Cellular and Molecular Life Sciences (2018) 75:2491–2507
https://doi.org/10.1007/s00018-018-2772-5

REVIEW

Cellular and molecular mechanisms of HIV-1 integration targeting

Alan N. Engelman[1,2] · **Parmit K. Singh**[1,2]

Received: 16 November 2017 / Revised: 23 January 2018 / Accepted: 1 February 2018 / Published online: 7 February 2018

Abstract

Integration is central to HIV-1 replication and helps mold the reservoir of cells that persists in AIDS patients. HIV-1 interacts with specific cellular factors to target integration to interior regions of transcriptionally active genes within gene-dense regions of chromatin. The viral capsid interacts with several proteins that are additionally implicated in virus nuclear import, including cleavage and polyadenylation specificity factor 6, to suppress integration into heterochromatin. The viral integrase protein interacts with transcriptional co-activator lens epithelium-derived growth factor p75 to principally position integration within gene bodies. The integrase additionally senses target DNA distortion and nucleotide sequence to help fine-tune the specific phosphodiester bonds that are cleaved at integration sites. Research into virus–host interactions that underlie HIV-1 integration targeting has aided the development of a novel class of integrase inhibitors and may help to improve the safety of viral-based gene therapy vectors.

Keywords HIV/AIDS · Integrase · LEDGF/p75 · Capsid · CPSF6 · Virus–host interaction · Viral DNA integration · Latency

Abbreviations

PIC	Preintegration complex
MoMLV	Moloney murine leukemia virus
LEDGF	Lens epithelium-derived growth factor
CPSF6	Cleavage and polyadenylation specificity factor 6
RNAi	RNA interference
MLL	Mixed-lineage leukemia
HDGF	Hepatoma-derived growth factor
HRP	HDGF-related protein
HDGFL	HDGF like
CR	Charged regions
IBD	Integrase-binding domain
PHAT	Pseudo HEAT repeat analogous topology
NTD	N-terminal domain
CCD	Catalytic core domain
CTD	C-terminal domain
PHD	Plant homeodomain
SMARCB1	SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily B, member 1
INI1	Integrase interactor 1
NPC	Nuclear pore complex
NUP	Nucleoporin
CYPA	Cyclophilin A
RNP	Ribonucleoprotein
RRM	RNA recognition motif
PRD	Pro-rich domain
RSLD	RS-like domain
IN	Integrase
GFP	Green fluorescent protein
NLS	Nuclear localization signal
CFIm	Cleavage factor I mammalian
ChIP-Seq	Chromatin-immunoprecipitation sequencing
Y	Pyrimidine
R	Purine
LEDGIN	LEDGF/p75-integrase interaction site
ALLINI	Allosteric integrase inhibitor
NCINI	Non-catalytic site integrase inhibitor
INLAI	Integrase-LEDGF allosteric inhibitor
MDM	Monocyte-derived macrophages
LRA	Latency-reversing agent
PDB	Protein database
TNPO1	Transportin 1

✉ Alan N. Engelman
alan_engelman@dfci.harvard.edu

[1] Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, 450 Brookline Avenue, CLS-1010, Boston, MA 02215, USA

[2] Department of Medicine, Harvard Medical School, A-111, 25 Shattuck Street, Boston, MA 02115, USA

TNPO3 Transportin 3
CA Capsid

Introduction

The lentivirus HIV-1, like all retroviruses, integrates the DNA copy of its RNA genome into a host cell chromosome. Integration provides a favorable environment for viral gene expression and ensures that both daughter cells receive a copy of the virus following cell division. Integration is mediated by the viral integrase protein, a specialized DNA recombinase that is incorporated into the virus during particle biogenesis from an infected cell. In the ensuing round of infection, reverse transcription and integration occur within the context of large subviral nucleoprotein structures that are termed the reverse transcription complex and preintegration complex (PIC), respectively [1, 2]. Within the PIC, the ends of the linear viral reverse transcript are held together by a multimer of integrase in a complex that is referred to as the intasome [3–6]. The active sites of two integrase protomers within the intasome interact with the viral DNA ends intimately, and it is these two active sites that promote the chemical steps of DNA recombination. The number of integrase protomers that comprise the intasome differs among the different retroviruses, from a minimum of four for the spumaviruses [6] to as many as 16 molecules for the lentiviruses [7, 8] (see [9] for a recent review).

Retroviral integrases possess two distinct catalytic activities, 3′ processing and strand transfer, which are required for integration. Integrase processes the viral DNA 3′ ends adjacent to conserved CA sequences, which most usually removes a dinucleotide from each end [10–13]. Integrase uses the resulting 3′-hydroxyl groups as nucleophiles to cut chromosomal DNA in staggered fashion, which at the same time joins the viral DNA CA_{OH}-3′ ends to target DNA 5′-phosphate groups [14]. The resulting DNA recombination intermediate is repaired by the host cell to yield a short 4–6 bp duplication flanking the integrated provirus, the sequence of which is derived from the staggered DNA cut (Fig. 1). Clinical integrase strand transfer inhibitors engage the intasome and displace the 3′-hydroxyl nucleophiles from the enzyme active site, disarming the nucleoprotein complex

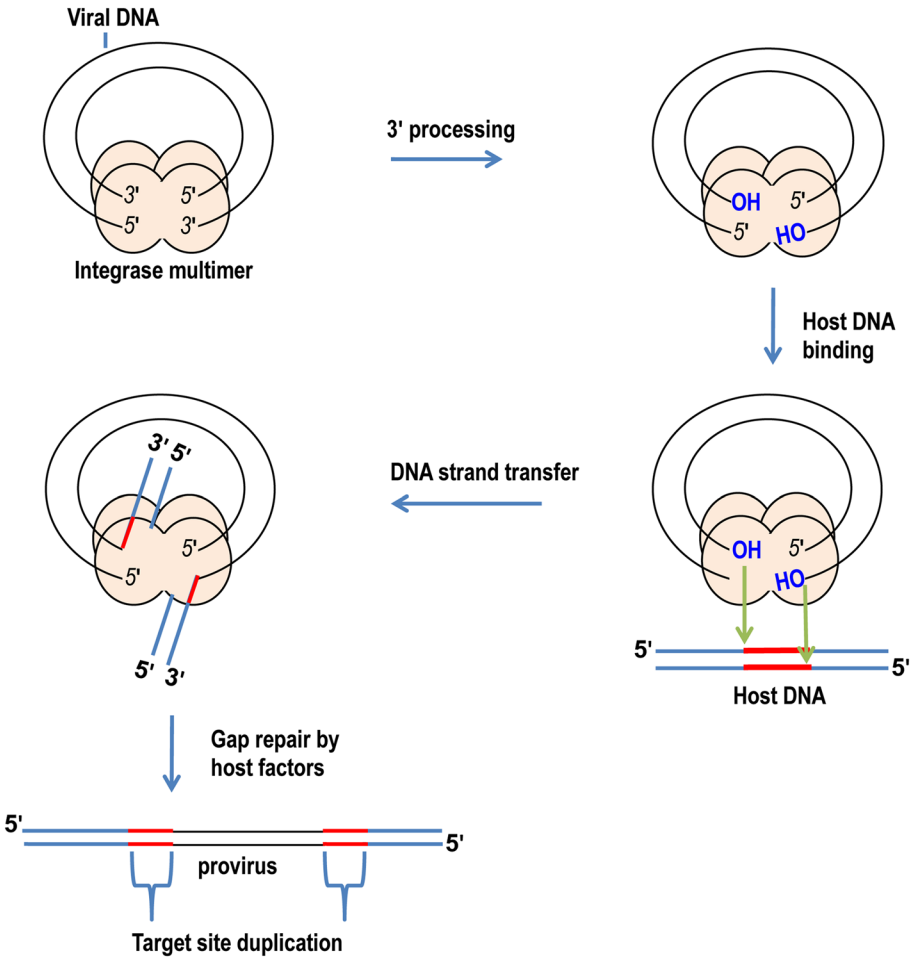

Fig. 1 Mechanism of retroviral DNA integration. A multimer of integrase (light orange) engages the ends of linear viral DNA (black lines), synapsing them together within the intasome complex. Integrase 3′ processing activity yields chemically reactive 3′-hydroxyl groups at the viral DNA ends. Following the binding of host DNA (blue lines) in the nucleus, integrase promotes the strand transfer or DNA joining reaction (staggered cut site in red). Repair of the DNA recombination intermediate yields a short duplication of host DNA (red lines) flanking the integrated provirus

[6, 15]. See [16] for a recent review of retroviral integration and its inhibition.

Integration in the host genome is non-random, with aspects of chromatin structure/function that span from DNA sequence to nuclear architecture influencing the sites of DNA recombination. Different types of retroviruses display distinctive preferences for transcription units, promoter/enhancer regions, histone modifications, and transcriptional activity (see [17–19] for recent reviews). HIV-1 in particular favors the interior regions of genes that reside in relatively gene-dense, transcriptionally active regions of chromatin [20]. Gammaretroviruses, typified by Moloney murine leukemia virus (MoMLV), in contrast favor promoter/enhancer regions [21–23] with more modest preferences for transcriptional activity than HIV-1 [24]. Recent research has highlighted that HIV-1 integration targeting is in large part mediated by the interaction of two virus proteins with cognate cellular receptors [25]. One of these interactions is between integrase and lens epithelium-derived growth factor (LEDGF)/p75 [26], while the other is between the viral capsid protein and cleavage and polyadenylation specificity factor 6 (CPSF6) [27]. On the local level, nucleosomes [28–33] in addition to target DNA sequence [34–38] can influence the choice of integration site. Globally, HIV-1 preferentially integrates into chromatin located in the peripheral region of the nucleus [39–41] in association with nuclear pore complexes (NPCs) [42, 43]. Herein we review in detail the cellular and molecular mechanisms of HIV-1 integration targeting, focusing on what is known alongside questions for future research.

Integrase—a primary determinant of integration targeting

HIV-1 integrase purified following its expression in *Escherichia coli* is enzymatically active in vitro, revealing that human cellular proteins are dispensable for integrase catalytic function [13, 44]. However, because viruses from different genera of Retroviridae display distinct genome targeting preferences, it seemed reasonable to assume that integrase-binding cell factors played a role in integration targeting [45, 46]. Detailed review of the published literature [47–64] alongside the HIV-1 Human Interaction Database that is maintained at the National Center for Biotechnology Information [65] reveals that several hundred cellular proteins reportedly interact with HIV-1 integrase. Clearly, not all of these interactions can play an important role in HIV-1 infection, and a central tenet of viral interactome research is to determine which interactions are key to virus replication. Among the reported HIV-1 integrase interactors, LEDGF/p75 has been shown to play an important role in virus infection and integration targeting.

LEDGF/p75

From initial reports, it was unclear if LEDGF/p75 played an important role in HIV-1 biology. Efficient knockdown of LEDGF/p75 via RNA interference (RNAi) yielded approximate two- to fivefold infection defects in some studies [66, 67] where others failed to detect a specific replication defect [68, 69]. The estimated cellular content of LEDGF/p75 is several 100,000 molecules [69], which seemingly greatly outweighs the level required for the efficient integration of any given PIC. Consistent with this interpretation, deep knockdown of LEDGF/p75 by short-hairpin RNA [70] or disruption of the LEDGF/p75-encoding *PSIP1* gene, which theoretically depleted all traces of the factor from the cell, yielded ~ tenfold or greater reductions in HIV-1 integration and virus infection [71–75]. CRISPR-Cas9 mediated knockout of *PSIP1* in primary human $CD4^+$ T cells yielded a relatively mild ~ twofold infection defect [76], a level consistent with early RNAi studies [66, 67]. This similarity could be due to residual LEDGF/p75 levels across studies, as Cas9 nucleoprotein complexes were introduced by electroporation, which may not transfect all cells in the population [76]. A significant reduction in the level at which HIV-1 targeted active genes during integration was observed even when knockdown cells supported normal levels of virus infection [25, 77], effects that were exacerbated via *PSIP1* knockout [25, 71, 72, 78, 79].

LEDGF/p75 and its shorter p52 splice variant (collectively referred to as LEDGF) were described initially as transcriptional co-activators [80], and both interact with RNA splicing factors [79, 81, 82]. LEDGF/p75 plays a role in homologous DNA recombination [83] and mixed-lineage leukemia (MLL) [84–87]. LEDGF belongs to the hepatoma-derived growth factor (HDGF)-related protein (HRP) family that also includes HDGF and HDGF like (HDGFL) 1–3 [88]. The most common sequence element among these proteins is an N-terminal PWWP domain [88], which is a type of chromatin reader [89]; the LEDGF PWWP domain has a preference for histone H3 tri-methylated on Lys36 [90–92]. LEDGF/p75 at steady-state binds chromatin [26, 93, 94], which is mediated by the PWWP domain and additional charged elements within the N-terminal portion of the protein [95, 96]. LEDGF/p75 also displays affinity for supercoiled DNA [97] (Fig. 2a).

The interaction of LEDGF/p75 with integrase is specific to the lentiviridae genus of Retroviridae [68, 98, 99]. An evolutionarily conserved domain in the C-terminal portion of LEDGF/p75, which is missing from LEDGF/p52, was termed the integrase-binding domain (IBD) because it was necessary and sufficient for binding to HIV-1 integrase in vitro [100] (Fig. 2a). The LEDGF/p75 IBD is a PHAT

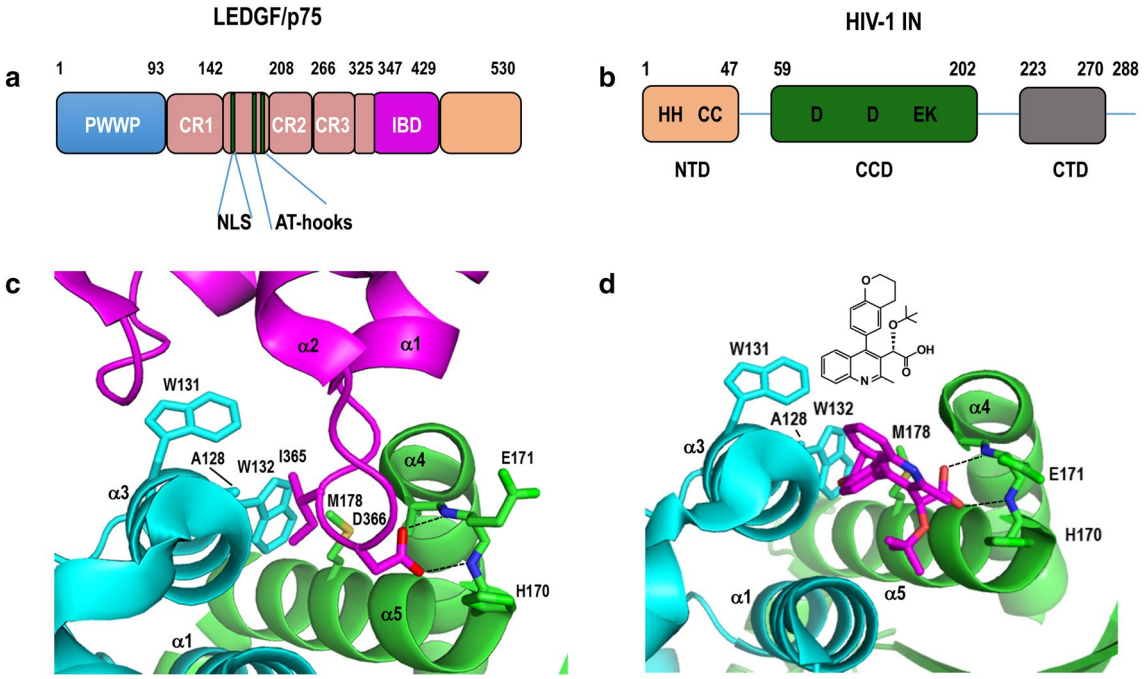

Fig. 2 LEDGF/p75 and ALLINI structures and binding to integrase. **a** Schematic diagram of LEDGF/p75, highlighting different protein regions/domains. Chromatin binding is mediated by the PWWP domain, charged regions (CR) 1–3, two copies of an AT-hook DNA binding motif, and basic NLS [95, 96]. CR 2 and 3 confer binding to supercoiled DNA [97]. Numbers refer to domain boundary positions. **b** Schematic of HIV-1 integrase; numbers demarcate domain boundaries. Amino acids invariant across Retroviridae are indicated by single-letter code. **c** X-ray structure of the LEDGF/p75 IBD (magenta) bound at the HIV-1 integrase CCD dimer (cyan and green). Shown in sticks are LEDGF/p75 hotspot residues Ile365 and Asp366 as well as integrase residues that help mold the binding pocket. Dashed lines, hydrogen bonds. Blue, red, and yellow denote nitrogen, oxygen, and sulfur, respectively (protein database (PDB) accession code 2B4J [102]). **d** X-ray structure of ALLINI BI-D (magenta, with chemical structure shown above) bound to the integrase CCD dimer (PDB code 4ID1 [190]), oriented as in **c**. The carboxylic acid attached to position 3 of the quinoline ring via the *tert*-butoxy group makes the same hydrogen bond contacts with integrase (dashed lines) as LEDGF/p75 residue Asp366 (compare with **c**). Other labeling is as in **c**

domain (for pseudo HEAT repeat analogous topology) composed of two helix–hairpin–helix motifs, and evolutionarily conserved hotspot residues Ile365 and Asp366 in the N-terminal hairpin were critical for integrase binding [101]. Retroviral integrase proteins harbor three common domains that are referred to as the N-terminal domain (NTD), catalytic core domain (CCD), and C-terminal domain (CTD) (reviewed in [9]) (Fig. 2b). The HIV-1 integrase CCD comprised the primary LEDGF/p75-binding determinant, while the NTD was required for high affinity binding [94]. An X-ray crystal structure of the IBD in complex with the HIV-1 integrase CCD revealed that the host factor engaged the CCD dimerization interface. The side chain carboxylic acid of Asp366 contacted backbone amides of residues Glu170 and His171 from one integrase monomer, while the adjacent LEDGF/p75 Ile365 residue mediated hydrophobic interactions principally with the other CCD molecule [102] (Fig. 2c). A subsequent structure with the 2-domain NTD-CCD construct from HIV-2 integrase revealed that electronegative side chains on one face of NTD helix 1 engaged conserved electropositive residues of the IBD [103]. Interestingly, reverse-charge LEDGF/p75 mutants restored partial infectivity to otherwise non-infectious HIV-1 integrase mutant viruses that harbored complementary reverse charge substitutions in NTD helix 1 [103, 104]. All-in-all, the structural basis of the LEDGF/p75-integrase interaction is fairly well understood. Additional structures that include the full-length host factor with the HIV-1 intasome [8] and nucleosomes would shed significant new light into the structural basis of integration targeting.

Both chromatin and integrase-binding activities are critical for LEDGF/p75 to function as an HIV-1 integration co-factor [70, 72, 105]. LEDGF/p75 normally functions to bring cellular IBD partners such as cell division cycle associated 7 like (a.k.a. JPO2) and MLL in proximity to chromatin [106], and thus HIV-1 hijacks this ancient chromatin-associated molecular beacon to fulfill its nefarious needs. Hybrid LEDGF/p75 constructs that swapped the N-terminal chromatin-binding elements for heterologous chromatin readers supported HIV-1 infection [107] and redirected integration to novel positions in the genome that were consistent with the known chromatin-binding properties of the substituted domains [108–111]. The plasticity of this

approach was rather remarkable, as both promoter-proximal readers such as PHD (plant homeodomain) fingers as well as heterochromatin protein modules such as CBX1 and HP1α similarly supported HIV-1 infection. Genotoxic side effects of retroviral gene therapy that are linked to sites of MoMLV vector integration have led to fatal cases of childhood leukemia [112], so the ability to reprogram HIV-1 integration to near random via fusion proteins such as CBX1-LEDGF suggested possible novel approaches to safer integrating vectors. A major hurdle in such approaches is how to effectively introduce the hybrid LEDGF/p75 construct into the cell. LEDGF/p75 is inefficiently packaged into HIV-1 particles and the scant amount that is packaged is cleaved by the viral protease [113]; so the potential to hitchhike retargeting LEDGF/p75 into the cell as a virus vector component is an apparent non-starter. The need to introduce a hybrid construct into patient cells prior to a therapeutic lentiviral vector necessarily complicates the clinical utility of retargeting LEDGF/p75 constructs in human gene therapy [103, 104, 111, 114].

Other integrase-binding proteins

Although *PSIP1* knockout significantly reduced the extent to which HIV-1 targeted genes during integration, the preference to integrate into genes remained much greater than would be expected based on random chance [71, 72, 78, 79]. Such observations suggested that other integrase-binding proteins might also play a role in HIV-1 integration targeting. SMARCB1 (SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily B, member 1) was the first cellular interactor of HIV-1 integrase to be published in the literature, which at the time was called INI1 for integrase interactor 1 [115]. As implicated by its name, SMARCB1 is a component of the ATP-dependent SWI/SNF chromatin remodeling complex [116], so it was from the get-go an obvious candidate for an integration targeting cofactor. Although SMARCB1 can influence the choice of HIV-1 integration site in in vitro reactions [117], it remains unclear if it plays a role in integration targeting during HIV-1 infection.

HRP protein family member HDGFL2 in addition to LEDGF/p75 harbors a functional IBD [100]; so it too was an obvious candidate for an integrase targeting cofactor. HDGFL2 was depleted from cells using RNAi [69, 70, 73] or gene knockout [74]. None of these manipulations significantly affected the level of HIV-1 infection or the selection of integration sites. Depleting HDGFL2 from *PSIP1* knockout cells additionally reduced HIV-1 infection by ~ twofold, and further reduced integration into genes from the levels observed by sole *PSIP1* knockout [74, 78]. However, since the sole depletion of HDGFL2 failed to significantly affect integration targeting, it seems unlikely that this integrase-binding factor plays a role in integration targeting under normal conditions of HIV-1 infection. Additional research is required to determine if any integrase-binding partner other than LEDGF/p75 plays a significant role in determining the profile of HIV-1 integration targeting.

HIV-1 capsid and integration targeting

A key biological parameter that distinguishes the lentiviruses from the gammaretroviruses is the requirement of cell cycling for virus infection. While HIV-1 can productively infect terminally differentiated, non-cycling cells, MoMLV infection is dependent upon cell division [118–121] (reviewed in [122]). This difference maps to PIC biology: HIV-1 PICs are efficiently transported through the cellular NPC in an energy-dependent manner [123], while MoMLV PICs require nuclear envelop breakdown such that their p12 Gag protein can latch onto chromatin for subsequent integration [124, 125]. The mechanistic details that underlie HIV-1 PIC nuclear import are still being worked out and exceed the scope of this review. In short, seminal work from Yamashita and Emerman took advantage of the biological distinction of HIV-1 versus MoMLV nuclear import mechanisms to construct chimeric viruses, which mapped the viral determinant of HIV-1 nuclear import to the capsid protein [126]. The use of such constructs first indicated a role for HIV-1 Gag, the polyprotein that includes capsid, in integration targeting [127]. HIV-1 capsid interacts directly with several host factors implicated in HIV-1 PIC nuclear import including nucleoporin (NUP) 153 [128], NUP358 [129], CPSF6 [130], and cyclophilin A (CYPA) [131–133], and each of these factors has additionally been shown to influence the choice of integration site [25, 129, 134–136]. Such observations highlight a potential mechanistic link between HIV-1 PIC nuclear import and integration targeting [137, 138].

HIV-1 capsid is a 2-domain protein composed of an NTD and CTD separated by a flexible linker. Capsid monomers assemble into hexameric or pentameric ring-like structures, which are the building blocks of the virion capsid shell [139, 140]. The capsid shell houses the ribonucleoprotein (RNP) guts of the virus, composed of viral RNA and viral proteins nucleocapsid, reverse transcriptase, and integrase (reviewed in [141]).

Capsid and integrase-binding host factors influence HIV-1 integration targeting in different ways. Integration in cells depleted for NUP358 was preferentially decreased in gene-dense regions of chromosomes as compared to gene bodies [134]. Similarly, capsid mutant viruses N74D and A77V, which are defective for binding to CPSF6 [27, 142], lost the targeting preference for gene-dense regions while retaining partial preference for genes [25, 129, 136, 142].

While LEDGF/p75 depletion via RNAi reduced integration into genes [25, 77], an initial report indicated this did not significantly affect the targeting of gene-dense regions [134]. Subsequent work indicated this too was likely due to residual LEDGF/p75 levels, as *PSIP1* knockout reduced the targeting of HIV-1 to gene-dense regions of chromosomes as well as gene bodies [25, 136]. Thus, although there is some phenotypic overlap between capsid and integrase-binding proteins, cofactors that bind capsid principally influence the targeting of gene-dense regions, while LEDGF/p75 primarily influences intragenic targeting.

While loss of CPSF6 or NUP358 binding to capsid preferentially reduced the targeting of gene-dense regions, loss of CYPA binding yielded the opposite phenotype of enhanced gene-dense region targeting [129]. This phenomenon is not well understood, as NUP358 and CYPA bind the same loop region in the capsid NTD [132, 143, 144] (reviewed in [145]) (Fig. 3). While CYPA is composed of the sole cyclophilin domain, NUP358 is a relatively large protein composed of numerous domains including its C-terminal cyclophilin homology domain [146, 147]. One possibility is that other parts of NUP358 could interact with the capsid/PIC to account for the differential effect it versus CYPA has on HIV-1 integration targeting. Consistent with this interpretation, the N74D mutation in the CPSF6-binding pocket of capsid, which is distal from the CYPA-binding loop (Fig. 3), disrupted the interaction of the PIC with NUP358 in the cytoplasm of infected cells [148].

CPSF6

The roles of LEDGF/p75 and CPSF6 in integration targeting have been scrutinized through side-by-side analysis of isogenic sets of cells knocked down or knocked out for each or both factors [25]. In such studies, random targeting is calculated based on known annotation distribution within the human genome. While LEDGF/p75 knockdown in U2OS cells marginally reduced integration into genes, gene-tropic integration was impressively reduced to near random via CPSF6 knockdown. As residual levels of LEDGF/p75 can readily mask its role in integration [69, 70], the factors were subsequently knocked out alone or in tandem in HEK293T cells. While *PSIP1* knockout reduced integration into genes from ~ 83 to ~ 63%, *CPSF6* knockout yielded 57% gene targeting. Dual factor knockout reduced integration into genes

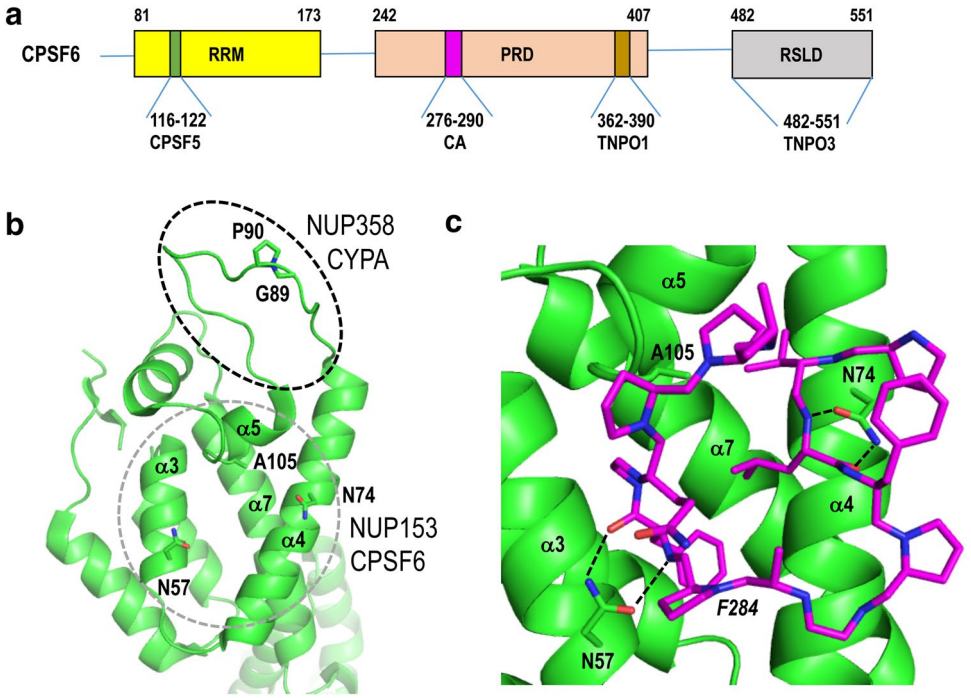

Fig. 3 CPSF6 structure and interaction with capsid. **a** CPSF6 isoform 1, which lacks internal sequences encoded by exon 6. Regions that facilitate binding to cellular CPSF5, transportin 1 (TNPO1), transportin 3 (TNPO3), as well as HIV-1 capsid (CA), are indicated. RRM, RNA recognition motif; PRD, Pro-rich domain; RSLD, RS-like domain; numbers, domain boundaries. **b** X-ray structure of the HIV-1 capsid NTD highlighting approximate positions of NUP358 and CYPA (black dotted line) versus NUP153 and CPSF6 (gray dotted line) binding (PDB code 4XFY [223]). Shown in sticks are capsid residues that help mediate binding to the different host factors. Blue and red, nitrogen and oxygen, respectively. **c** Close-up view of CPSF6 residues 276–289 (magenta) bound to hexameric HIV-1 capsid (PDB code 4U0B [165]) highlighting hotspot residue Phe284 (italic type). Dashed lines represent hydrogen bonds. Other labeling is as in **b**

to ~ 48%, just a few percentage points above the random value of 44.7%. In terms of gene density, each Mb of human chromatin harbors on average 8.7 genes, whereas HIV-1 targets regions that on average contain ~ 21 genes/Mb. *PSIP1* knockout reduced this preference to ~ 14 genes/Mb. Impressively, *CPSF6* knockout reduced it to 5.8 genes, well below the level expected by pure chance. Inspection of promoter proximal integration patterns further distinguished the roles of the integrase versus capsid-binding host factors. Consistent with prior reports [71, 72], *PSIP1* knockout increased promoter proximal integration, from ~ 4 to 10% within a 5 kb window surrounding transcriptional start sites, while *CPSF6* knockout reduced this metric to 1.8%, well below the random 4.0% value [25].

A number of additional genomic annotations were analyzed to further tease out unique contributions of LEDGF/p75 versus CPSF6 in HIV-1 integration targeting [25]. One of the more telltale was association with histone post-translational modifications. HIV-1 integration favors regions in the vicinity of activating epigenetic marks and disfavors regions nearby repressive marks [31, 149]. While *PSIP1* knockout weakened each of these tendencies, both preferences were nevertheless maintained when compared to random chance values. In contrast, *CPSF6* knockout flipped both preferences such that the virus now preferred repressive regions and disfavored regions nearby activating marks. Another telltale metric was genome-wide averaging of intragenic integration sites, which revealed a dominant role for LEDGF/p75 in positioning HIV-1 integration along gene bodies [25, 79]. We accordingly concluded that the primary role of the capsid–CPSF6 interaction was to steer the PIC away from heterochromatic regions of chromatin, while the LEDGF/p75–integrase interaction primarily positioned integration along gene bodies (Fig. 4) [25]. Such observations invoke a model whereby during its journey the PIC sequentially interacts with CPSF6 and then LEDGF/p75, but we currently do not know the order of binding events. The PIC may very well engage CPSF6 in the cytoplasm to enable its import through the NPC [27, 150–152] (Fig. 4), though this model is clouded by the fact that cytoplasmic accumulation of CPSF6 can potently restrict HIV-1 infection [27, 153–155]. The PIC is susceptible to LEDGF/p75 binding in the cytoplasm [156], but even here studies differ as to whether LEDGF/p75 is a component of the PIC [68, 157, 158]. Recombinant LEDGF/p75 protein can significantly stimulate the strand transfer activity of purified HIV-1 integrase [26, 96, 100, 103, 159, 160], yet PICs extracted from *PSIP1* knockout cells notably support the wild-type level of strand transfer activity in vitro [72, 74]. Based on these observations, we favor a model whereby the HIV-1 PIC engages CPSF6 during or shortly after nuclear import, but need not engage LEDGF/p75 until it scans chromatin for integration sites (Fig. 4). Other scenarios that invoke

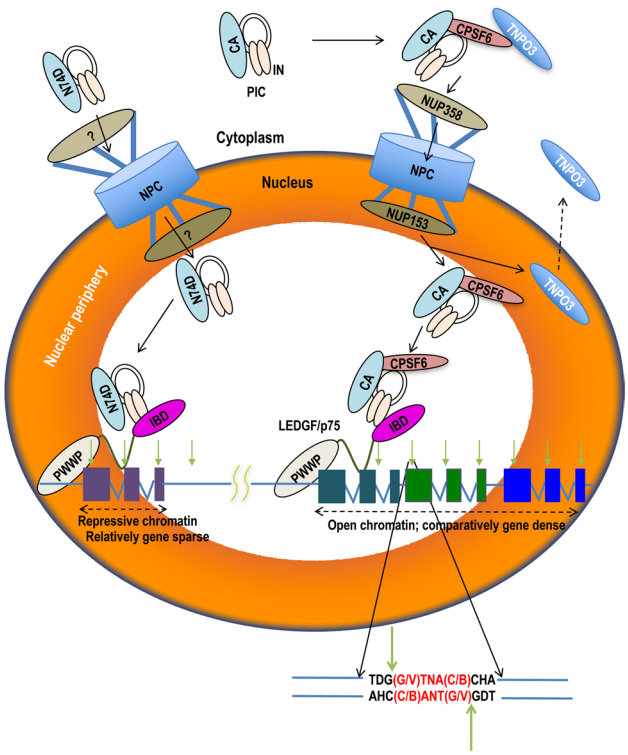

Fig. 4 Cellular and molecular mechanisms of HIV-1 integration targeting. The HIV-1 PIC harbors viral DNA (black lines), integrase (IN, light orange ovals), and capsid (CA). CPSF6 may bind CA in the cytoplasm to facilitate PIC nuclear import in a NUP358 and NUP153-dependent manner; CPSF6 is shown in association with transportin 3 (TNPO3), one of its known β-karyopherin-binding partners. Following nuclear import, TNPO3 will recycle to engage additional transport substrates in the cytoplasm. The PIC utilizes both CPSF6 and LEDGF/p75 to target integration to active genes (represented as three colored exons separated by introns) preferentially located within the peripheral region of the nucleus (orange shade). A typical integration site (downward light green arrow) is enlarged below to indicate nucleotide sequence preference in International Union of Biochemistry base code (B: G, C, or T; D: G, A, or T; H: C, A, or T; N: G, C, A, or T; V: G, C, or A); opposing green arrows denote scissile phosphodiester bonds. PICs that cannot properly engage CPSF6, as represented by the N74D change in CA, enter the nucleus via an alternate route (marked ?) that may require NUP155 [27]. Such PICs prefer gene sparse regions and hyper target the peripheral region of the nucleus for integration [25, 129, 136, 142, 152]

cytoplasmic engagement or integrase–LEDGF/p75 binding in the nucleoplasm are also plausible. The integrase hexadecamer in the lentiviral intasome could be modeled to bind up to 16 LEDGF/p75 molecules, which may help the PIC target regions relatively enriched in LEDGF/p75 content [7]. Recent evidence suggests that HIV-1 preferentially integrates into genes that regulate T-cell activation, but the roles of LEDGF/p75 and/or CPSF6 in this process have yet to be addressed [161].

CPSF6 is an SR-like protein composed of an N-terminal RNA recognition motif (RRM), a central Pro-rich domain

(PRD), and a C-terminal RS-like domain (RSLD) [162] (Fig. 3a). Binding to capsid is mediated via the PRD [130, 163] and although a 15-mer peptide could bind the isolated HIV-1 capsid NTD, it displayed significantly higher affinity to capsid hexamers in vitro [164, 165]. The form of the capsid that engages CPSF6 or other nuclear transport factors during HIV-1 infection is unknown, but it seems likely that CPSF6 will minimally engage a hexamer. Additional work to determine the form of the capsid that is bound by CPSF6 during HIV-1 infection is warranted.

CPSF6 at steady state is nuclear, and a fusion protein composed of the RSLD and green fluorescent protein (GFP) accumulated in cell nuclei, revealing that the RSLD harbors a functional nuclear localization signal (NLS) [162]. An internal deletion mutant that lacked the PRD but retained the RSLD localized to both the cytoplasm and nucleus, indicating that sequences outside of the RSLD may also contribute to CPSF6 nuclear localization [162]. Consistent with this interpretation, two different β-karyopherin proteins, transportin 1 and transportin 3, were shown to bind sequences within the PRD and RSLD, respectively [150, 166]. It will be informative to determine if these CPSF6 sequences actually comprise functional NLSs.

CPSF6 is part of the cleavage factor I mammalian (CFIm) complex that regulates positions of polyadenylation in the 3′ untranslated regions of mRNAs [167–169]. CFIm is a heterotetramer composed of two copies of CPSF5 and two copies of either CPSF6 or CPSF7 [167, 168, 170]. Unlike LEDGF/p75, CPSF6-binding partners important for integration targeting that function downstream from its interaction with the PIC are unknown, but one obvious candidate was CPSF5. However, expression of a CPSF6 mutant that is defective for CPSF5 binding in *CPSF6* knockout cells in large part restored HIV-1 integration targeting, indicating that the function of CPSF6 in integration targeting is independent from its role in polyadenylation regulation [171]. Consistent with this interpretation, CPSF6 did not preferentially direct integration into the genes for which it regulated polyadenylation [25]. CPSF6 resides predominantly in nuclear paraspeckles [162, 172] and is not known to directly bind chromatin. Nevertheless, a chromatin-immunoprecipitation sequencing (ChIP-Seq) dataset is available [173], and these sequences correlated with regions of CPSF6-dependent integration [25]. We accordingly expect that factor(s) that apparently bridge CPSF6 to chromatin, as evident by the ChIP-Seq dataset, could play a role in integration targeting. Additional work is required to clarify the mechanism of CPSF6 action in HIV-1 integration targeting.

Global integration targeting and nuclear architecture

Imaging-based studies have indicated that HIV-1 prefers to integrate into chromatin that resides within the peripheral region of the nucleus [39–41, 43, 152] (reviewed in [174]). LEDGF/p75 and NUP153 were reportedly both required for peripheral integration targeting, as each knockdown resulted in shifts of the viral DNA away from the periphery toward the central region of the nucleus [43, 175]. Results of two other studies, however, fail to support a significant role for LEDGF/p75 in the intranuclear localization of HIV-1 DNA [176, 177]. CPSF6 knockdown, or infection with binding defective capsid mutant viruses such as N74A or A105T, by contrast increased the accumulation of peripherally located proviruses [152]. The binding sites for NUP153 and CPSF6 on the capsid overlap [130, 164, 165], but in each case only relatively short peptides of the host factors have been crystallized with capsid (Fig. 3b, c). Thus, as previously mentioned for NUP358, regions of NUP153 or CPSF6 outside of the visualized peptides could potentially interact with capsid or PICs in unknown ways to effect the different fates of integration within the structure of the nucleus. Additional work is clearly required to assess the roles of the different integration targeting cofactors in HIV-1 PIC trafficking to architecturally distinct regions of the nucleus for integration.

Local chromatin features in integration targeting

Retroviral integration favors particular nucleobases at integration sites, although these preferences are rather subtle and thus only become evident through the alignment of multiple sequences [34, 35]. Spumavirus integrase cuts target DNA with a 4 bp stagger, and X-ray crystal structures of the spumaviral intasome with target DNA revealed a severe kink at the center of the 4 bp region [178]. The center of spumaviral integration sites is enriched in pyrimidine (Y)-purine (R) dinucleotides, which, due to base stacking, are naturally more flexible than YY and RR, or the most rigid RY dinucleotide [179]. Thus, the spumaviral intasome preferentially selects for target DNA sequences that can best bend to fit into the relatively confined space between two opposing integrase active sites [178]. Expanded preferential target DNA sequence analyses indicated that inherent bendability contributed to the integration site preferences of most retroviruses, and that viruses that cut target DNA with a 6 bp stagger more routinely utilized less bendable DNA than viruses that cut

with 4 or 5 bp staggers [36, 37]. Substitutions of integrase residues that interact with target DNA in crystal structures [178] or molecular models [36, 180] not only altered nucleotide selectivity at integration sites [36, 178, 180], but also could retarget integration toward gene-sparse regions of chromatin [180].

Nucleosomes are favored over matched naked DNA templates for HIV-1 integration in vitro, with preferred sites mapping to outward regions of the distorted nucleosomal DNA major groove [29, 30]. Analysis of large numbers of integration sites indicated that HIV-1 is likely to retain this targeting preference during virus infection [31]. The cryo-electron microscopy structure of the spumaviral intasome bound to a mononucleosome revealed that the intasome induced the same target DNA bend that was observed previously with naked target DNA through the lifting of the DNA from the surface of the underlying histones [181]. Although these data suggest that retroviruses might universally integrate into preferred target DNA sequences on the surface of nucleosomes, other work has revealed differential effects of chromatin compaction on the in vitro activities of retroviral integrase proteins [32]. HIV-1 integrase in particular favored regions of lower nucleosome density under conditions where MoMLV and spumaviral integrase favored compact chromatin [32]. Thus, the HIV-1 intasome may require chromatin remodeling complexes such as SWI/SNF [115, 117] or FACT [64] to effectively access nucleosomal target DNAs during virus infection. Additional research that maps the integration sites of a variety of retroviruses such as HIV-1, MoMLV, and spumavirus in cells depleted for remodeling complexes should inform the extent that such complexes play a role in viral integration targeting.

Allosteric inhibitors of HIV-1 integrase activity

Over-expression of mutant LEDGF/p75 constructs that retained the IBD but substituted the N-terminal chromatin-binding elements for heterologous sequences such as GFP blocked HIV-1 infection at the integration step [70, 182]. Moreover, combining this approach with RNAi-mediated knockdown of LEDGF/p75 could inhibit HIV-1 infection > 500-fold [156]. These observations highlighted that small molecule inhibitors of the LEDGF/p75–integrase interaction could potentially possess antiviral activity.

Two different approaches led to the identification of what turned out to be highly similar inhibitors of the interaction of LEDGF/p75 with HIV-1 integrase. Whereas one utilized a high-throughput screen for integrase 3′ processing activity [183], the other used an in silico approach to identify molecules that mimicked the positions of key LEDGF/p75 interacting residues at the CCD dimer-binding site [184] (Fig. 2c, d). In both cases, micromolar lead compounds were developed into low nM inhibitors, with Boehringer Ingelheim compound BI 224436 evaluated in human clinical trials [185]. Such compounds go by various names in the literature. The Debyser laboratory coined LEDGINs for LEDGF/p75-integrase interaction site [184], whereas we prefer ALLINIs for allosteric-integrase inhibitors to represent the fact that such compounds bind far from the integrase active site and thus allosterically inhibit catalytic function [186]. Other terms include NCINIs for non-catalytic site integrase inhibitors [183, 185, 187] and INLAIs for integrase-LEDGF allosteric inhibitors [188].

ALLINIs possess a remarkable antiviral mechanism of action that is linked to their ability to hyper-multimerize integrase [189–192]. HIV-1 integrase in solution adopts numerous forms including monomer, dimer, tetramer, and higher-order multimer, the details of which in part depend on protein and salt concentration [160, 190, 193–199]. HIV-1 intasome assembly, which may occur via DNA-mediated tetramerization of the solution integrase tetramer [7], is exquisitely sensitive to perturbation of integrase multimerization such as that incurred via premature LEDGF/p75 binding [160, 200, 201]. ALLINI binding induces the formation of large integrase aggregates [191, 192, 202] through the bridging of neighboring protein molecules [203, 204]. Aggregation accounts for their anti-integrase activity, as the aggregates are unable to assemble with viral DNA to form functional integrase–DNA complexes in vitro [186]. And, when added to susceptible target cells, the compounds act as integrase inhibitors that specifically block the integration step of HIV-1 infection [184, 187–190, 205]. However, the compounds are across the board more potent inhibitors of HIV-1 particle maturation [187–190, 192, 206]. Particle maturation is particularly sensitive to changes in integrase, leading to the definition of a subclass of HIV-1 integrase mutant viruses, the class II mutants, to distinguish them from the class I mutants that mature normally and thus are specifically defective for the integration step of HIV-1 replication (see [207, 208] for reviews). ALLINI treatment phenocopies the maturation defect associated with certain class II integrase mutant viruses, which is highlighted by the mis-localization of the RNP complex to a region of the virion particle outside of the conical shell [187, 190, 206, 209, 210]. Such viral RNA is susceptible to degradation following infection [211], highlighting the common reverse transcription defect of ALLINI-treated and class II integrase mutant viruses [187–190, 192, 206–208]. HIV-1 integrase binds virion RNA in the virus, which is disrupted similarly by ALLINIs and class II mutations [212]. Thus, integrase binding to RNA may play a critical role in RNP localization during HIV-1 particle maturation [212], though the biochemical mechanism of how integrase–RNA binding orchestrates capsid shell morphogenesis around the RNP

remains to be elucidated. It will also be telltale to ascertain if integrase plays a similar role in the morphogenesis of other retroviruses.

In the long run, it appears that inhibition of the LEDGF/p75–integrase interaction may have little to do with the mechanism of ALLINI action. Sites of HIV-1 integration are altered when susceptible target cells are treated with ALLINIs, highlighting that the compounds are likely effective inhibitors of the LEDGF/p75–integrase interaction under this infection condition [175, 192, 202]. However, ALLINI potency actually increases during the afferent arm of HIV-1 replication when LEDGF/p75 is depleted, indicating that LEDGF/p75 protects the intasome from the aggregation inducing effects of the molecules [74, 187–190, 192]. By contrast, LEDGF/p75 has no apparent influence on compound potency under conditions of virus assembly [75, 187–190, 192]. As LEDGF/p75 is not effectively incorporated into virions [113], it seems that the inability for LEDGF/p75 to bind integrase during viral late events unleashes the full potency of the compounds.

Conclusions and perspectives

Recent research has highlighted roles for both integrase- and capsid-binding proteins in HIV-1 integration targeting. The brunt of this data was derived from transformed cell lines such as U2OS and HEK293T due to the sensitivity of these cells to RNAi and CRISPR-Cas9 genetic manipulation. It therefore is important to extend this line of research to primary cells such as monocyte-derived macrophages (MDM) and $CD4^+$ T cells to ascertain the extents that LEDGF/p75 and CPSF6 influence integration targeting under physiologically relevant conditions. As discussed, LEDGF/p75 has been depleted from primary T cells via electroporation of Cas9-guide RNA complexes [76], so mapping HIV-1 integration sites in these cells should be informative. Treating MDM with short-interfering RNA against CPSF6 or LEDGF/p75 yielded fairly modest knockdowns and reduced HIV-1 integration into genes, yet the targeting defects were, perhaps expectedly, much more modest than those seen by knocking down CPSF6 in U2OS cells or knocking out either factor [25]. Additional approaches to increase the potency of protein depletion in primary cells are warranted [76]. The A77V change in capsid, which disrupts binding to CPSF6, importantly ablated integration into gene-dense regions of chromatin in both transformed cells and MDM [142].

Patients on antiretroviral therapy harbor a reservoir of latently infected cells that rekindle systemic viremia upon cessation of therapy [213] (see [214] for review). Analysis of AIDS patients integration sites over time revealed the persistence of certain proviruses as evident by the clonal expansion of infected cells [215, 216]. As the brunt of viral DNA in patients is defective [217], it may be expected that many of the proviruses in clonally expanded cells are also defective [218], though in at least one patient a dominant cell clone was responsible for much if not all of the circulating virus [219]. What is currently unknown is whether proviruses, defective or otherwise, drive clonal expansion due to subtle growth advantages instilled by the particular site of integration, though this is surely a popular working model (see [220] for review). Significant research is currently dedicated to HIV cure, where cells that harbor latent proviruses would either be permanently locked down for viral expression or transcriptionally activated and then purged from the body (reviewed in [221]). The site of HIV-1 integration can influence both entry into latency and transcriptional reactivation by latency-reversing agents (LRAs) [222]. Interestingly, suppression of LEDGF/p75-mediated integration targeting by RNAi knockdown or ALLINI treatment increased the number of latent viruses and decreased the extent of LRA reactivation [175]. As CPSF6 shields the HIV-1 PIC from targeting heterochromatin [25], it will be of interest to determine the latent state of proviruses formed in the absence of this targeting pathway and how such viruses respond to LRAs. The cellular and molecular mechanisms that drive HIV-1 integration targeting are inextricably linked to the persistence of particular proviruses in AIDS patient reservoirs and research into HIV cure.

Acknowledgements Work in the corresponding author's laboratory is funded by Grants AI039394 and AI052014 from the US National Institutes of Health. The authors thank Vasudevan Achuthan, Gregory Bedwell, and Sooin Jang for their critical review of the manuscript.

References

1. Fassati A, Goff SP (1999) Characterization of intracellular reverse transcription complexes of Moloney murine leukemia virus. J Virol 73:8919–8925
2. Bowerman B, Brown PO, Bishop JM, Varmus HE (1989) A nucleoprotein complex mediates the integration of retroviral DNA. Genes Dev 3:469–478
3. Miller MD, Farnet CM, Bushman FD (1997) Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. J Virol 71:5382–5390
4. Wei SQ, Mizuuchi K, Craigie R (1997) A large nucleoprotein assembly at the ends of the viral DNA mediates retroviral DNA integration. EMBO J 16:7511–7520
5. Chen H, Wei S-Q, Engelman A (1999) Multiple integrase functions are required to form the native structure of the human immunodeficiency virus type I intasome. J Biol Chem 274:17358–17364
6. Hare S, Gupta SS, Valkov E, Engelman A, Cherepanov P (2010) Retroviral intasome assembly and inhibition of DNA strand transfer. Nature 464:232–236
7. Ballandras-Colas A, Maskell DP, Serrao E, Locke J, Swuec P, Jonsson SR, Kotecha A, Cook NJ, Pye VE, Taylor IA, Andresdottir V, Engelman AN, Costa A, Cherepanov P (2017) A

supramolecular assembly mediates lentiviral DNA integration. Science 355:93–95
8. Passos DO, Li M, Yang R, Regensburg S, Ghirlando R, Jeon Y, Kvaratskhelia M, Craigie R, Lyumkis D (2017) CryoEM structures and atomic model of the HIV-1 strand transfer complex intasome. Science 355:89–92
9. Engelman AN, Cherepanov P (2017) Retroviral intasomes arising. Curr Opin Struct Biol 47:23–29
10. Katzman M, Katz RA, Skalka AM, Leis J (1989) The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the in vivo sites of integration. J Virol 63:5319–5327
11. Roth MJ, Schwartzberg PL, Goff SP (1989) Structure of the termini of DNA intermediates in the integration of retroviral DNA: dependence on IN function and terminal DNA sequence. Cell 58:47–54
12. Craigie R, Fujiwara T, Bushman F (1990) The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. Cell 62:829–837
13. Sherman PA, Fyfe JA (1990) Human immunodeficiency virus integration protein expressed in *Escherichia coli* possesses selective DNA cleaving activity. Proc Natl Acad Sci USA 87:5119–5123
14. Engelman A, Mizuuchi K, Craigie R (1991) HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell 67:1211–1221
15. Hare S, Vos AM, Clayton RF, Thuring JW, Cummings MD, Cherepanov P (2010) Molecular mechanisms of retroviral integrase inhibition and the evolution of viral resistance. Proc Natl Acad Sci USA 107:20057–20062
16. Lesbats P, Engelman AN, Cherepanov P (2016) Retroviral DNA integration. Chem Rev 116:12730–12757
17. Kvaratskhelia M, Sharma A, Larue RC, Serrao E, Engelman A (2014) Molecular mechanisms of retroviral integration site selection. Nucleic Acids Res 42:10209–10225
18. Craigie R, Bushman FD (2014) Host factors in retroviral integration and the selection of integration target sites. Microbiol Spectr 2:6
19. Demeulemeester J, Rijck JD, Gijsbers R, Debyser Z (2015) Retroviral integration: site matters: mechanisms and consequences of retroviral integration site selection. BioEssays 37:1202–1214
20. Schroder ARW, 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
21. Wu X, Li Y, Crise B, Burgess SM (2003) Transcription start regions in the human genome are favored targets for MLV integration. Science 300:1749–1751
22. LaFave MC, Varshney GK, Gildea DE, Wolfsberg TG, Baxevanis AD, Burgess SM (2014) MLV integration site selection is driven by strong enhancers and active promoters. Nucleic Acids Res 42:4257–4269
23. De Ravin SS, Su L, Theobald N, Choi U, Macpherson JL, Poidinger M, Symonds G, Pond SM, Ferris AL, Hughes SH, Malech HL, Wu X (2014) Enhancers are major targets for murine leukemia virus vector integration. J Virol 88:4504–4513
24. Mitchell RS, Beitzel BF, Schroder AR, Shinn P, Chen H, Berry CC, Ecker JR, Bushman FD (2004) Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol 2:E234
25. Sowd GA, Serrao E, Wang H, Wang W, Fadel HJ, Poeschla EM, Engelman AN (2016) A critical role for alternative polyadenylation factor CPSF6 in targeting HIV-1 integration to transcriptionally active chromatin. Proc Natl Acad Sci USA 113:E1054–E1063
26. Cherepanov P, Maertens G, Proost P, Devreese B, Van Beeumen J, Engelborghs Y, De Clercq E, Debyser Z (2003) HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. J Biol Chem 278:372–381
27. Lee K, Ambrose Z, Martin TD, Oztop I, Mulky A, Julias JG, Vandegraaff N, Baumann JG, Wang R, Yuen W, Takemura T, Shelton K, Taniuchi I, Li Y, Sodroski J, Littman DR, Coffin JM, Hughes SH, Unutmaz D, Engelman A, KewalRamani VN (2010) Flexible use of nuclear import pathways by HIV-1. Cell Host Microbe 7:221–233
28. Pryciak PM, Varmus HE (1992) Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. Cell 69:769–780
29. Pruss D, Bushman FD, Wolffe AP (1994) Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. Proc Natl Acad Sci USA 91:5913–5917
30. Pruss D, Reeves R, Bushman FD, Wolffe AP (1994) The influence of DNA and nucleosome structure on integration events directed by HIV integrase. J Biol Chem 269:25031–25041
31. Wang GP, Ciuffi A, Leipzig J, Berry CC, Bushman FD (2007) HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. Genome Res 17:1186–1194
32. Benleulmi MS, Matysiak J, Henriquez DR, Vaillant C, Lesbats P, Calmels C, Naughtin M, Leon O, Skalka AM, Ruff M, Lavigne M, Andreola ML, Parissi V (2015) Intasome architecture and chromatin density modulate retroviral integration into nucleosome. Retrovirology 12:13
33. Naughtin M, Haftek-Terreau Z, Xavier J, Meyer S, Silvain M, Jaszczyszyn Y, Levy N, Miele V, Benleulmi MS, Ruff M, Parissi V, Vaillant C, Lavigne M (2015) DNA physical properties and nucleosome positions are major determinants of HIV-1 integrase selectivity. PLoS One 10:e0129427
34. Holman AG, Coffin JM (2005) Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites. Proc Natl Acad Sci USA 102:6103–6107
35. Wu X, Li Y, Crise B, Burgess SM, Munroe DJ (2005) Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. J Virol 79:5211–5214
36. Serrao E, Krishnan L, Shun MC, Li X, Cherepanov P, Engelman A, Maertens GN (2014) Integrase residues that determine nucleotide preferences at sites of HIV-1 integration: implications for the mechanism of target DNA binding. Nucleic Acids Res 42:5164–5176
37. Serrao E, Ballandras-Colas A, Cherepanov P, Maertens GN, Engelman AN (2015) Key determinants of target DNA recognition by retroviral intasomes. Retrovirology 12:39
38. Kirk PDW, Huvet M, Melamed A, Maertens GN, Bangham CRM (2016) Retroviruses integrate into a shared, non-palindromic DNA motif. Nat Microbiol 2:16212
39. Albanese A, Arosio D, Terreni M, Cereseto A (2008) HIV-1 preintegration complexes selectively target decondensed chromatin in the nuclear periphery. PLoS One 3:e2413
40. Burdick RC, Hu W-S, Pathak VK (2013) Nuclear import of APOBEC3F-labeled HIV-1 preintegration complexes. Proc Natl Acad Sci USA 110:E4780–E4789
41. Di Primio C, Quercioli V, Allouch A, Gijsbers R, Christ F, Debyser Z, Arosio D, Cereseto A (2013) Single-cell imaging of HIV-1 provirus (SCIP). Proc Natl Acad Sci USA 110:5636–5641
42. Lelek M, Casartelli N, Pellin D, Rizzi E, Souque P, Severgnini M, Di Serio C, Fricke T, Diaz-Griffero F, Zimmer C, Charneau P, Di Nunzio F (2015) Chromatin organization at the nuclear pore favours HIV replication. Nat Commun 6:6483
43. Marini B, Kertesz-Farkas A, Ali H, Lucic B, Lisek K, Manganaro L, Pongor S, Luzzati R, Recchia A, Mavilio F, Giacca M,

Lusic M (2015) Nuclear architecture dictates HIV-1 integration site selection. Nature 521:227–231
44. Bushman FD, Craigie R (1991) Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. Proc Natl Acad Sci USA 88:1339–1343
45. Bushman F, Lewinski M, Ciuffi A, Barr S, Leipzig J, Hannenhalli S, Hoffmann C (2005) Genome-wide analysis of retroviral DNA integration. Nat Rev Microbiol 3:848–858
46. Engelman A (2005) The ups and downs of gene expression and retroviral DNA integration. Proc Natl Acad Sci USA 102:1275–1276
47. Engelman A (2007) Host cell factors and HIV-1 integration. Future HIV Ther 1:415–426
48. Christ F, Thys W, De Rijck J, Gijsbers R, Albanese A, Arosio D, Emiliani S, Rain JC, Benarous R, Cereseto A, Debyser Z (2008) Transportin-SR2 imports HIV into the nucleus. Curr Biol 18:1192–1202
49. Huang L, G-l Xu, J-q Zhang, Tian L, J-l Xue, J-z Chen, Jia W (2008) Daxx interacts with HIV-1 integrase and inhibits lentiviral gene expression. Biochem Biophys Res Commun 373:241–245
50. Studamire B, Goff SP (2008) Host proteins interacting with the Moloney murine leukemia virus integrase: multiple transcriptional regulators and chromatin binding factors. Retrovirology 5:48
51. Woodward CL, Prakobwanakit S, Mosessian S, Chow SA (2009) Integrase interacts with nucleoporin NUP153 to mediate the nuclear import of human immunodeficiency virus type 1. J Virol 83:6522–6533
52. J-q Zhang, J-j Wang, W-j Li, Huang L, Tian L, J-l Xue, J-z Chen, Jia W (2009) Cellular protein TTRAP interacts with HIV-1 integrase to facilitate viral integration. Biochem Biophys Res Commun 387:256–260
53. Ao Z, Danappa Jayappa K, Wang B, Zheng Y, Kung S, Rassart E, Depping R, Kohler M, Cohen EA, Yao X (2010) Importin α3 interacts with HIV-1 integrase and contributes to HIV-1 nuclear import and replication. J Virol 84:8650–8663
54. Manganaro L, Lusic M, Gutierrez MI, Cereseto A, Del Sal G, Giacca M (2010) Concerted action of cellular JNK and Pin1 restricts HIV-1 genome integration to activated $CD4^+$ T lymphocytes. Nat Med 16:329–333
55. Terreni M, Valentini P, Liverani V, Gutierrez MI, Di Primio C, Di Fenza A, Tozzini V, Allouch A, Albanese A, Giacca M, Cereseto A (2010) GCN5-dependent acetylation of HIV-1 integrase enhances viral integration. Retrovirology 7:18
56. Allouch A, Cereseto A (2011) Identification of cellular factors binding to acetylated HIV-1 integrase. Amino Acids 41:1137–1145
57. Allouch A, Di Primio C, Alpi E, Lusic M, Arosio D, Giacca M, Cereseto A (2011) The TRIM family protein KAP1 inhibits HIV-1 integration. Cell Host Microbe 9:484–495
58. Kobbi L, Octobre G, Dias J, Comisso M, Mirande M (2011) Association of mitochondrial lysyl-tRNA synthetase with HIV-1 GagPol involves catalytic domain of the synthetase and transframe and integrase domains of Pol. J Mol Biol 410:875–886
59. Sorin M, Cano J, Das S, Mathew S, Wu X, Davies KP, Shi X, Cheng SW, Ott D, Kalpana GV (2011) Recruitment of a SAP18-HDAC1 complex into HIV-1 virions and its requirement for viral replication. PLoS Pathog 5:e1000463
60. Yamamoto SP, Okawa K, Nakano T, Sano K, Ogawa K, Masuda T, Morikawa Y, Koyanagi Y, Suzuki Y (2011) Huwe1, a novel cellular interactor of Gag-Pol through integrase binding, negatively influences HIV-1 infectivity. Microbes Infect 13:339–349
61. Zheng Y, Ao Z, Wang B, Jayappa KD, Yao X (2011) Host protein Ku70 binds and protects HIV-1 integrase from proteasomal degradation and is required for HIV replication. J Biol Chem 286:17722–17735
62. Ao Z, Jayappa KD, Wang B, Zheng Y, Wang X, Peng J, Yao X (2012) Contribution of host nucleoporin 62 in HIV-1 integrase chromatin association and viral DNA integration. J Biol Chem 287:10544–10555
63. Jager S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, Clarke SC, Shales M, Mercenne G, Pache L, Li K, Hernandez H, Jang GM, Roth SL, Akiva E, Marlett J, Stephens M, D'Orso I, Fernandes J, Fahey M, Mahon C, O'Donoghue AJ, Todorovic A, Morris JH, Maltby DA, Alber T, Cagney G, Bushman FD, Young JA, Chanda SK, Sundquist WI, Kortemme T, Hernandez RD, Craik CS, Burlingame A, Sali A, Frankel AD, Krogan NJ (2012) Global landscape of HIV-human protein complexes. Nature 481:365–370
64. Matysiak J, Lesbats P, Mauro E, Lapaillerie D, Dupuy J-W, Lopez AP, Benleulmi MS, Calmels C, Andreola M-L, Ruff M, Llano M, Delelis O, Lavigne M, Parissi V (2017) Modulation of chromatin structure by the FACT histone chaperone complex regulates HIV-1 integration. Retrovirology 14:39
65. Ako-Adjei D, Fu W, Wallin C, Katz KS, Song G, Darji D, Brister JR, Ptak RG, Pruitt KD (2015) HIV-1, human interaction database: current status and new features. Nucleic Acids Res 43:D566–D570
66. Vandekerckhove L, Christ F, Van Maele B, De Rijck J, Gijsbers R, Van den Haute C, Witvrouw M, Debyser Z (2006) Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. J Virol 80:1886–1896
67. Zielske SP, Stevenson M (2006) Modest but reproducible inhibition of human immunodeficiency virus type 1 infection in macrophages following LEDGFp75 silencing. J Virol 80:7275–7280
68. Llano M, Vanegas M, Fregoso O, Saenz D, Chung S, Peretz M (2004) LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. J Virol 78:9524–9537
69. Vandegraaff N, Devroe E, Turlure F, Silver PA, Engelman A (2006) Biochemical and genetic analyses of integrase-interacting proteins lens epithelium-derived growth factor (LEDGF)/p75 and hepatoma-derived growth factor related protein 2 (HRP2) in preintegration complex function and HIV-1 replication. Virology 346:415–426
70. Llano M, Saenz DT, Meehan A, Wongthida P, Peretz M, Walker WH, Teo W, Poeschla EM (2006) An essential role for LEDGF/p75 in HIV integration. Science 314:461–464
71. Marshall HM, Ronen K, Berry C, Llano M, Sutherland H, Saenz D, Bickmore W, Poeschla E, Bushman FD (2007) Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. PLoS One 2:e1340
72. 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
73. Schrijvers R, De Rijck J, Demeulemeester J, Adachi N, Vets S, Ronen K, Christ F, Bushman FD, Debyser Z, Gijsbers R (2012) LEDGF/p75-independent HIV-1 replication demonstrates a role for HRP-2 and remains sensitive to inhibition by LEDGINs. PLoS Pathog 8:e1002558
74. 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

75. Fadel HJ, Morrison JH, Saenz DT, Fuchs JR, Kvaratskhelia M, Ekker SC, Poeschla EM (2014) TALEN knockout of the PSIP1 gene in human cells: analyses of HIV-1 replication and allosteric integrase inhibitor mechanism. J Virol 88:9704–9717
76. Hultquist JF, Schumann K, Woo JM, Manganaro L, McGregor MJ, Doudna J, Simon V, Krogan NJ, Marson A (2016) A Cas9 ribonucleoprotein platform for functional genetic studies of HIV-host interactions in primary human T cells. Cell Rep 17:1438–1452
77. 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
78. Schrijvers R, Vets S, De Rijck J, Malani N, Bushman FD, Debyser Z, Gijsbers R (2012) HRP-2 determines HIV-1 integration site selection in LEDGF/p75 depleted cells. Retrovirology 9:84
79. Singh PK, Plumb MR, Ferris AL, Iben JR, Wu X, Fadel HJ, Luke BT, Esnault C, Poeschla EM, Hughes SH, Kvaratskhelia M, Levin HL (2015) LEDGF/p75 interacts with mRNA splicing factors and targets HIV-1 integration to highly spliced genes. Genes Dev 29:2287–2297
80. Ge H, Si Y, Roeder RG (1998) Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. EMBO J 17:6723–6729
81. Ge H, Si Y, Wolffe AP (1998) A novel transcriptional coactivator, p52, functionally interacts with the essential splicing factor ASF/SF2. Mol Cell 2:751–759
82. Morchikh M, Naughtin M, Di Nunzio F, Xavier J, Charneau P, Jacob Y, Lavigne M (2013) TOX4 and NOVA1 proteins are partners of the LEDGF PWWP domain and affect HIV-1 replication. PLoS One 7:e1001280
83. Daugaard M, Baude A, Fugger K, Povlsen LK, Beck H, Sorensen CS, Petersen NH, Sorensen PH, Lukas C, Bartek J, Lukas J, Rohde M, Jaattela M (2012) LEDGF (p75) promotes DNA-end resection and homologous recombination. Nat Struct Mol Biol 19:803–810
84. Yokoyama A, Cleary ML (2008) Menin critically links MLL proteins with LEDGF on cancer-associated target genes. Cancer Cell 14:36–46
85. Čermáková K, Tesina P, Demeulemeester J, El Ashkar S, Méreau H, Schwaller J, Řezáčová P, Veverka V, De Rijck J (2014) Validation and structural characterization of the LEDGF/p75–MLL interface as a new target for the treatment of MLL-dependent leukemia. Cancer Res 74:5139–5151
86. Murai MJ, Pollock J, He S, Miao H, Purohit T, Yokom A, Hess JL, Muntean AG, Grembecka J, Cierpicki T (2014) The same site on the integrase-binding domain of lens epithelium–derived growth factor is a therapeutic target for MLL leukemia and HIV. Blood 124:3730–3737
87. El Ashkar S, Schwaller J, Pieters T, Goossens S, Demeulemeester J, Christ F, Van Belle S, Juge S, Boeckx N, Engelman A, Van Vlierberghe P, Debyser Z, De Rijck J (2018) LEDGF/p75 is dispensable for hematopoiesis but essential for MLL-rearranged leukemogenesis. Blood 131:95–107
88. Izumoto Y, Kuroda T, Harada H, Kishimoto T, Nakamura H (1997) Hepatoma-derived growth factor belongs to a gene family in mice showing significant homology in the amino terminus. Biochem Biophys Res Commun 238:26–32
89. Qin S, Min J (2014) Structure and function of the nucleosome-binding PWWP domain. Trends Biochem Sci 39:536–547
90. Pradeepa MM, Sutherland HG, Ule J, Grimes GR, Bickmore WA (2012) Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. PLoS Genet 8:e1002717
91. Eidahl JO, Crowe BL, North JA, McKee CJ, Shkriabai N, Feng L, Plumb M, Graham RL, Gorelick RJ, Hess S, Poirier MG, Foster MP, Kvaratskhelia M (2013) Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. Nucleic Acids Res 41:3924–3936
92. van Nuland R, van Schaik FM, Simonis M, van Heesch S, Cuppen E, Boelens R, Timmers HM, van Ingen H (2013) Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain. Epigenetics Chromatin 6:12
93. Nishizawa Y, Usukura J, Singh DP, Chylack LTJ, Shinohara T (2001) Spatial and temporal dynamics of two alternatively spliced regulatory factors, lens epithelium-derived growth factor (ledgf/p75) and p52, in the nucleus. Cell Tissue Res 305:107–114
94. Maertens G, Cherepanov P, Pluymers W, Busschots K, De Clercq E, Debyser Z, Engelborghs Y (2003) LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. J Biol Chem 278:33528–33539
95. Llano M, Vanegas M, Hutchins N, Thompson D, Delgado S, Poeschla EM (2006) Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75. J Mol Biol 360:760–773
96. Turlure F, Maertens G, Rahman S, Cherepanov P, Engelman A (2006) A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. Nucleic Acids Res 34:1653–1665
97. Tsutsui KM, Sano K, Hosoya O, Miyamoto T, Tsutsui K (2011) Nuclear protein LEDGF/p75 recognizes supercoiled DNA by a novel DNA-binding domain. Nucleic Acids Res 39:5067–5081
98. Busschots K, Vercammen J, Emiliani S, Benarous R, Engelborghs Y, Christ F, Debyser Z (2005) The interaction of LEDGF/p75 with integrase Is lentivirus-specific and promotes DNA binding. J Biol Chem 280:17841–17847
99. Cherepanov P (2007) LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. Nucleic Acids Res 35:113–124
100. Cherepanov P, Devroe E, Silver PA, Engelman A (2004) Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. J Biol Chem 279:48883–48892
101. Cherepanov P, Sun Z-YJ, Rahman S, Maertens G, Wagner G, Engelman A (2005) Solution structure of the HIV-1 integrase-binding domain in LEDGF/p75. Nat Struct Mol Biol 12:526–532
102. Cherepanov P, Ambrosio AL, Rahman S, Ellenberger T, Engelman A (2005) Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. Proc Natl Acad Sci USA 102:17308–17313
103. Hare S, Shun MC, Gupta SS, Valkov E, Engelman A, Cherepanov P (2009) A novel co-crystal structure affords the design of gain-of-function lentiviral integrase mutants in the presence of modified PSIP1/LEDGF/p75. PLoS Pathog 5:e1000259
104. Wang H, Shun MC, Li X, Di Nunzio F, Hare S, Cherepanov P, Engelman A (2014) Efficient transduction of LEDGF/p75 mutant cells by gain-of-function HIV-1 integrase mutant viruses. Mol Ther Methods Clin Dev 1:2
105. Shun M-C, Botbol Y, Li X, Di Nunzio F, Daigle JE, Yan N, Lieberman J, Lavigne M, Engelman A (2008) Identification and characterization of PWWP domain residues critical for LEDGF/p75 chromatin binding and human immunodeficiency virus type 1 infectivity. J Virol 82:11555–11567
106. Tesina P, Čermáková K, Hořejší M, Procházková K, Fábry M, Sharma S, Christ F, Demeulemeester J, Debyser Z, Rijck JD, Veverka V, Řezáčová P (2015) Multiple cellular proteins

interact with LEDGF/p75 through a conserved unstructured consensus motif. Nat Commun 6:7968
107. Meehan AM, Saenz DT, Morrison JH, Garcia-Rivera JA, Peretz M, Llano M, Poeschla EM (2009) LEDGF/p75 proteins with alternative chromatin tethers are functional HIV-1 cofactors. PLoS Pathog 5:e1000522
108. Ferris AL, Wu X, Hughes CM, Stewart C, Smith SJ, Milne TA, Wang GG, Shun MC, Allis CD, Engelman A, Hughes SH (2010) Lens epithelium-derived growth factor fusion proteins redirect HIV-1 DNA integration. Proc Natl Acad Sci USA 107:3135–3140
109. Gijsbers R, Ronen K, Vets S, Malani N, De Rijck J, McNeely M, Bushman FD, Debyser Z (2010) LEDGF hybrids efficiently retarget lentiviral integration into heterochromatin. Mol Ther 18:552–560
110. Silvers RM, Smith JA, Schowalter M, Litwin S, Liang Z, Geary K, Daniel R (2010) Modification of integration site preferences of an HIV-1-based vector by expression of a novel synthetic protein. Hum Gene Ther 21:337–349
111. Vranckx LS, Demeulemeester J, Debyser Z, Gijsbers R (2016) Towards a safer, more randomized lentiviral vector integration profile exploring artificial LEDGF chimeras. PLoS One 11:e0164167
112. Hacein-Bey-Abina S, Garrigue A, Wang GP, Soulier J, Lim A, Morillon E, Clappier E, Caccavelli L, Delabesse E, Beldjord K, Asnafi V, MacIntyre E, Dal Cortivo L, Radford I, Brousse N, Sigaux F, Moshous D, Hauer J, Borkhardt A, Belohradsky BH, Wintergerst U, Velez MC, Leiva L, Sorensen R, Wulffraat N, Blanche S, Bushman FD, Fischer A, Cavazzana-Calvo M (2008) Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J Clin Invest 118:3132–3142
113. Desimmie BA, Weydert C, Schrijvers R, Vets S, Demeulemeester J, Proost P, Paron I, De Rijck J, Mast J, Bannert N, Gijsbers R, Christ F, Debyser Z (2015) HIV-1 IN/Pol recruits LEDGF/p75 into viral particles. Retrovirology 12:16
114. Vets S, De Rijck J, Brendel C, Grez M, Bushman F, Debyser Z, Gijsbers R (2013) Transient expression of an LEDGF/p75 chimera retargets lentivector integration and functionally rescues in a model for X-CGD. Mol Ther Nucleic Acids 2:e77
115. Kalpana G, Marmon S, Wang W, Crabtree G, Goff S (1994) Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. Science 266:2002–2006
116. Roberts CWM, Orkin SH (2004) The SWI/SNF complex—chromatin and cancer. Nat Rev Cancer 4:133–142
117. Lesbats P, Botbol Y, Chevereau G, Vaillant C, Calmels C, Arneodo A, Andreola ML, Lavigne M, Parissi V (2011) Functional coupling between HIV-1 integrase and the SWI/SNF chromatin remodeling complex for efficient in vitro integration into stable nucleosomes. PLoS Pathog 7:e1001280
118. Weinberg JB, Matthews TJ, Cullen BR, Malim MH (1991) Productive human immunodeficiency virus type 1 (HIV-1) infection of nonproliferating human monocytes. J Exp Med 174:1477–1482
119. Lewis P, Hensel M, Emerman M (1992) Human immunodeficiency virus infection of cells arrested in the cell cycle. EMBO J 11:3053–3058
120. Roe T, Reynolds TC, Yu G, Brown PO (1993) Integration of murine leukemia virus DNA depends on mitosis. EMBO J 12:2099–2108
121. Lewis PF, Emerman M (1994) Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. J Virol 68:510–516
122. Matreyek KA, Engelman A (2013) Viral and cellular requirements for the nuclear entry of retroviral preintegration nucleoprotein complexes. Viruses 5:2483–2511
123. Bukrinsky MI, Sharova N, Dempsey MP, Stanwick TL, Bukrinskaya AG, Haggerty S, Stevenson M (1992) Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. Proc Natl Acad Sci USA 89:6580–6584
124. Elis E, Ehrlich M, Prizan-Ravid A, Laham-Karam N, Bacharach E (2012) p12 tethers the murine leukemia virus pre-integration complex to mitotic chromosomes. PLoS Pathog 8:e1003103
125. Schneider WM, Brzezinski JD, Aiyer S, Malani N, Gyuricza M, Bushman FD, Roth MJ (2013) Viral DNA tethering domains complement replication-defective mutations in the p12 protein of MuLV Gag. Proc Natl Acad Sci USA 110:9487–9492
126. Yamashita M, Emerman M (2004) Capsid is a dominant determinant of retrovirus infectivity in nondividing cells. J Virol 78:5670–5678
127. Lewinski MK, Yamashita M, Emerman M, Ciuffi A, Marshall H, Crawford G, Collins F, Shinn P, Leipzig J, Hannenhalli S, Berry CC, Ecker JR, Bushman FD (2006) Retroviral DNA integration: viral and cellular determinants of target-site selection. PLoS Pathog 2:e60
128. Matreyek KA, Yucel SS, Li X, Engelman A (2013) Nucleoporin NUP153 phenylalanine-glycine motifs engage a common binding pocket within the HIV-1 capsid protein to mediate lentiviral infectivity. PLoS Pathog 9:e1003693
129. Schaller T, Ocwieja KE, Rasaiyaah J, Price AJ, Brady TL, Roth SL, Hue S, Fletcher AJ, Lee K, KewalRamani VN, Noursadeghi M, Jenner RG, James LC, Bushman FD, Towers GJ (2011) HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. PLoS Pathog 7:e1002439
130. Price AJ, Fletcher AJ, Schaller T, Elliott T, Lee K, KewalRamani VN, Chin JW, Towers GJ, James LC (2012) CPSF6 defines a conserved capsid interface that modulates HIV-1 replication. PLoS Pathog 8:e1002896
131. Luban J, Bossolt KL, Franke EK, Kalpana GV, Goff SP (1993) Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. Cell 73:1067–1078
132. Gamble TR, Vajdos FF, Yoo S, Worthylake DK, Houseweart M, Sundquist WI, Hill CP (1996) Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell 87:1285–1294
133. De Iaco A, Luban J (2014) Cyclophilin A promotes HIV-1 reverse transcription but its effect on transduction correlates best with its effect on nuclear entry of viral cDNA. Retrovirology 11:11
134. Ocwieja KE, Brady TL, Ronen K, Huegel A, Roth SL, Schaller T, James LC, Towers GJ, Young JA, Chanda SK, Konig R, Malani N, Berry CC, Bushman FD (2011) HIV integration targeting: a pathway involving Transportin-3 and the nuclear pore protein RanBP2. PLoS Pathog 7:e1001313
135. Di Nunzio F, Fricke T, Miccio A, Valle-Casuso JC, Perez P, Souque P, Rizzi E, Severgnini M, Mavilio F, Charneau P, Diaz-Griffero F (2013) Nup153 and Nup98 bind the HIV-1 core and contribute to the early steps of HIV-1 replication. Virology 440:8–18
136. Koh Y, Wu X, Ferris AL, Matreyek KA, Smith SJ, Lee K, KewalRamani VN, Hughes SH, Engelman A (2013) Differential effects of human immunodeficiency virus type 1 capsid and cellular factors nucleoporin 153 and LEDGF/p75 on the efficiency and specificity of viral DNA integration. J Virol 87:648–658
137. Konig R, Zhou Y, Elleder D, Diamond TL, Bonamy GM, Irelan JT, Chiang CY, Tu BP, De Jesus PD, Lilley CE (2008) Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. Cell 135:49–60
138. Di Nunzio F (2013) New insights in the role of nucleoporins: a bridge leading to concerted steps from HIV-1 nuclear entry until integration. Virus Res 178:187–196

139. Pornillos O, Ganser-Pornillos BK, Kelly BN, Hua Y, Whitby FG, Stout CD, Sundquist WI, Hill CP, Yeager M (2009) X-ray structures of the hexameric building block of the HIV capsid. Cell 137:1282–1292
140. Pornillos O, Ganser-Pornillos BK, Yeager M (2011) Atomic-level modelling of the HIV capsid. Nature 469:424–427
141. Sundquist WI, Kräusslich H-G (2012) HIV-1 assembly, budding, and maturation. Cold Spring Harb Perspect Med 2:a006924
142. Saito A, Henning MS, Serrao E, Dubose BN, Teng S, Huang J, Li X, Saito N, Roy SP, Siddiqui MA, Ahn J, Tsuji M, Hatziioannou T, Engelman AN, Yamashita M (2016) Capsid-CPSF6 interaction is dispensable for HIV-1 replication in primary cells but is selected during virus passage in vivo. J Virol 90:6918–6935
143. Bichel K, Price AJ, Schaller T, Towers GJ, Freund SM, James LC (2013) HIV-1 capsid undergoes coupled binding and isomerization by the nuclear pore protein NUP358. Retrovirology 10:81
144. Lin DH, Zimmermann S, Stuwe T, Stuwe E, Hoelz A (2013) Structural and functional analysis of the C-terminal domain of Nup358/RanBP2. J Mol Biol 425:1318–1329
145. Yamashita M, Engelman AN (2017) Capsid-dependent host factors in HIV-1 infection. Trends Microbiol 25:741–755
146. Yokoyama N, Hayashi N, Seki T, Pante N, Ohba T, Nishii K, Kuma K, Hayashida T, Miyata T, Aebi U (1995) A giant nucleopore protein that binds Ran/TC4. Nature 376:184–188
147. Wu J, Matunis MJ, Kraemer D, Blobel G, Coutavas E (1995) Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP binding sites, zinc fingers, a cyclophilin A homologous domain, and a leucine-rich region. J Biol Chem 270:14209–14213
148. Dharan A, Talley S, Tripathi A, Mamede JI, Majetschak M, Hope TJ, Campbell EM (2016) KIF5B and Nup358 cooperatively mediate the nuclear import of HIV-1 during infection. PLoS Pathog 12:e1005700
149. Roth SL, Malani N, Bushman FD (2011) Gammaretroviral integration into nucleosomal target DNA in vivo. J Virol 85:7393–7401
150. Maertens GN, Cook NJ, Wang W, Hare S, Gupta SS, Öztop I, Lee K, Pye VE, Cosnefroy O, Snijders AP, KewalRamani VN, Fassati A, Engelman A, Cherepanov P (2014) Structural basis for nuclear import of splicing factors by human Transportin 3. Proc Natl Acad Sci USA 111:2728–2733
151. Peng K, Muranyi W, Glass B, Laketa V, Yant SR, Tsai L, Cihlar T, Müller B, Kräusslich H-G (2014) Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid. eLife 3:e04114
152. Chin CR, Perreira JM, Savidis G, Portmann JM, Aker AM, Feeley EM, Smith MC, Brass AL (2015) Direct visualization of HIV-1 replication intermediates shows that capsid and CPSF6 modulate HIV-1 intra-nuclear invasion and integration. Cell Rep 13:1717–1731
153. De Iaco A, Santoni F, Vannier A, Guipponi M, Antonarakis S, Luban J (2013) TNPO3 protects HIV-1 replication from CPSF6-mediated capsid stabilization in the host cell cytoplasm. Retrovirology 10:20
154. Fricke T, Valle-Casuso JC, White TE, Brandariz-Nuñez A, Bosche WJ, Reszka N, Gorelick R, Diaz-Griffero F (2013) The ability of TNPO3-depleted cells to inhibit HIV-1 infection requires CPSF6. Retrovirology 10:46
155. Hori T, Takeuchi H, Saito H, Sakuma R, Inagaki Y, Yamaoka S (2013) A carboxy-terminally truncated human CPSF6 lacking residues encoded by exon 6 inhibits HIV-1 cDNA synthesis and promotes capsid disassembly. J Virol 87:7726–7736
156. Meehan AM, Saenz DT, Morrison J, Hu C, Peretz M, Poeschla EM (2011) LEDGF dominant interference proteins demonstrate prenuclear exposure of HIV-1 integrase and synergize with LEDGF depletion to destroy viral infectivity. J Virol 85:3570–3583
157. Yan N, Cherepanov P, Daigle JE, Engelman A, Lieberman J (2009) The SET complex acts as a barrier to autointegration of HIV-1. PLoS Pathog 5:e1000327
158. Gérard A, Soler N, Ségéral E, Belshan M, Emiliani S (2013) Identification of low molecular weight nuclear complexes containing integrase during the early stages of HIV-1 infection. Retrovirology 10:13
159. Botbol Y, Raghavendra NK, Rahman S, Engelman A, Lavigne M (2008) Chromatinized templates reveal the requirement for the LEDGF/p75 PWWP domain during HIV-1 integration in vitro. Nucleic Acids Res 36:1237–1246
160. 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
161. Zhyvoloup A, Melamed A, Anderson I, Planas D, Lee CH, Kriston-Vizi J, Ketteler R, Merritt A, Routy JP, Ancuta P, Bangham CRM, Fassati A (2017) Digoxin reveals a functional connection between HIV-1 integration preference and T-cell activation. PLoS Pathog 13:e1006460
162. Dettwiler S, Aringhieri C, Cardinale S, Keller W, Barabino SML (2004) Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein-protein interactions, and subcellular localization. J Biol Chem 279:35788–35797
163. Lee K, Mulky A, Yuen W, Martin TD, Meyerson NR, Choi L, Yu H, Sawyer SL, Kewalramani VN (2012) HIV-1 capsid-targeting domain of cleavage and polyadenylation specificity factor 6. J Virol 86:3851–3860
164. Bhattacharya A, Alam SL, Fricke T, Zadrozny K, Sedzicki J, Taylor AB, Demeler B, Pornillos O, Ganser-Pornillos BK, Diaz-Griffero F, Ivanov DN, Yeager M (2014) Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. Proc Natl Acad Sci USA 111:18625–18630
165. Price AJ, Jacques DA, McEwan WA, Fletcher AJ, Essig S, Chin JW, Halambage UD, Aiken C, James LC (2014) Host cofactors and pharmacologic ligands share an essential interface in HIV-1 capsid that is lost upon disassembly. PLoS Pathog 10:e1004459
166. Lee BJ, Cansizoglu AE, Süel KE, Louis TH, Zhang Z, Chook YM (2006) Rules for nuclear localization sequence recognition by karyopherin beta2. Cell 126:543–558
167. Rüegsegger U, Beyer K, Keller W (1996) Purification and characterization of human cleavage factor I involved in the 3′ end processing of messenger RNA precursors. J Biol Chem 271:6107–6113
168. Gruber AR, Martin G, Keller W, Zavolan M (2012) Cleavage factor Im is a key regulator of 3′ UTR length. RNA Biol 9:1405–1412
169. Martin G, Gruber AR, Keller W, Zavolan M (2012) Genome-wide analysis of pre-mRNA 3′ end processing reveals a decisive role of human cleavage factor I in the regulation of 3′ UTR length. Cell Rep 1:753–763
170. Yang Q, Coseno M, Gilmartin GM, Doublié S (2011) Crystal structure of a human cleavage factor CFIm25/CFIm68/RNA complex provides an insight into poly(A) site recognition and RNA looping. Structure 19:368–377
171. Rasheedi S, Shun M-C, Serrao E, Sowd GA, Qian J, Hao C, Dasgupta T, Engelman AN, Skowronski J (2016) The cleavage and polyadenylation specificity factor 6 (CPSF6) subunit of the capsid-recruited pre-messenger RNA cleavage factor I (CFIm) complex mediates HIV-1 integration into genes. J Biol Chem 291:11809–11819
172. Cardinale S, Cisterna B, Bonetti P, Aringhieri C, Biggiogera M, Barabino SML (2007) Subnuclear localization and dynamics of

the pre-mRNA 3′ end processing factor mammalian cleavage factor I 68-kDa subunit. Mol Biol Cell 18:1282–1292
173. Katahira J, Okuzaki D, Inoue H, Yoneda Y, Maehara K, Ohkawa Y (2013) Human TREX component Thoc5 affects alternative polyadenylation site choice by recruiting mammalian cleavage factor I. Nucleic Acids Res 41:7060–7072
174. Lusic M, Siliciano RF (2016) Nuclear landscape of HIV-1 infection and integration. Nat Rev Microbiol 15:69–82
175. Vranckx LS, Demeulemeester J, Saleh S, Boll A, Vansant G, Schrijvers R, Weydert C, Battivelli E, Verdin E, Cereseto A, Christ F, Gijsbers R, Debyser Z (2016) LEDGIN-mediated inhibition of integrase–LEDGF/p75 interaction reduces reactivation of residual latent HIV. EBioMedicine 8:248–264
176. Quercioli V, Di Primio C, Casini A, Mulder LCF, Vranckx LS, Borrenberghs D, Gijsbers R, Debyser Z, Cereseto A (2016) Comparative analysis of HIV-1 and murine leukemia virus three-dimensional nuclear distributions. J Virol 90:5205–5209
177. Burdick RC, Delviks-Frankenberry KA, Chen J, Janaka SK, Sastri J, Hu WS, Pathak VK (2017) Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. PLoS Pathog 13:e1006570
178. Maertens GN, Hare S, Cherepanov P (2010) The mechanism of retroviral integration from X-ray structures of its key intermediates. Nature 468:326–329
179. Johnson RC, Stella S, Heiss JK (2008) Bending and compaction of DNA by proteins. In: Rice PA, Correll CC (eds) Protein-nucleic acid interactions. RCS Publishing, London, pp 176–220
180. Demeulemeester J, Vets S, Schrijvers R, Madlala P, De Maeyer M, De Rijck J, Ndung'u T, Debyser Z, Gijsbers R (2014) HIV-1 integrase variants retarget viral integration and are associated with disease progression in a chronic infection cohort. Cell Host Microbe 16:651–662
181. Maskell DP, Renault L, Serrao E, Lesbats P, Matadeen R, Hare S, Lindemann D, Engelman AN, Costa A, Cherepanov P (2015) Structural basis for retroviral integration into nucleosomes. Nature 523:366–369
182. De Rijck J, Vandekerckhove L, Gijsbers R, Hombrouck A, Hendrix J, Vercammen J, Engelborghs Y, Christ F, Debyser Z (2006) Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. J Virol 80:11498–11509
183. Fader LD, Malenfant E, Parisien M, Carson R, Bilodeau F, Landry S, Pesant M, Brochu C, Morin S, Chabot C, Halmos T, Bousquet Y, Bailey MD, Kawai SH, Coulombe R, LaPlante S, Jakalian A, Bhardwaj PK, Wernic D, Schroeder P, Amad M, Edwards P, Garneau M, Duan J, Cordingley M, Bethell R, Mason SW, Bös M, Bonneau P, Poupart MA, Faucher AM, Simoneau B, Fenwick C, Yoakim C, Tsantrizos Y (2014) Discovery of BI 224436, a noncatalytic site integrase inhibitor (NCINI) of HIV-1. ACS Med Chem Lett 5:422–427
184. 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
185. Fenwick C, Amad M, Bailey MD, Bethell R, Bös M, Bonneau P, Cordingley M, Coulombe R, Duan J, Edwards P, Fader LD, Faucher AM, Garneau M, Jakalian A, Kawai S, Lamorte L, LaPlante S, Luo L, Mason S, Poupart MA, Rioux N, Schroeder P, Simoneau B, Tremblay S, Tsantrizos Y, Witvrouw M, Yoakim C (2014) Preclinical profile of BI 224436, a novel HIV-1 non-catalytic-site integrase inhibitor. Antimicrob Agents Chemother 58:3233–3244
186. Kessl JJ, Jena N, Koh Y, Taskent-Sezgin H, Slaughter A, Feng L, de Silva S, Wu L, Le Grice SF, Engelman A, Fuchs JR, Kvaratskhelia M (2012) Multimode, cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. J Biol Chem 287:16801–16811
187. 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
188. 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
189. 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
190. 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 USA 110:8690–8695
191. 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
192. 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
193. van Gent DC, Elgersma Y, Bolk MWJ, Vink C, Plasterk RHA (1991) DNA binding properties of the integrase proteins of human immunodeficiency viruses types 1 and 2. Nucleic Acids Res 19:3821–3827
194. Engelman A, Bushman FD, Craigie R (1993) Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. EMBO J 12:3269–3275
195. Vincent KA, Ellison V, Chow SA, Brown PO (1993) Characterization of human immunodeficiency virus type 1 integrase expressed in *Escherichia coli* and analysis of variants with amino-terminal mutations. J Virol 67:425–437
196. Taddeo B, Carlini F, Verani P, Engelman A (1996) Reversion of a human immunodeficiency virus type 1 integrase mutant at a second site restores enzyme function and virus infectivity. J Virol 70:8277–8284
197. Hayouka Z, Rosenbluh J, Levin A, Loya S, Lebendiker M, Veprintsev D, Kotler M, Hizi A, Loyter A, Friedler A (2007) Inhibiting HIV-1 integrase by shifting its oligomerization equilibrium. Proc Natl Acad Sci USA 104:8316–8321
198. Hare S, Di Nunzio F, Labeja A, Wang J, Engelman A, Cherepanov P (2009) Structural basis for functional tetramerization of lentiviral integrase. PLoS Pathog 5:e1000515
199. Pandey KK, Bera S, Grandgenett DP (2011) The HIV-1 integrase monomer induces a specific interaction with LTR DNA for concerted integration. Biochemistry 50:9788–9796
200. Raghavendra NK, Engelman A (2007) LEDGF/p75 interferes with the formation of synaptic nucleoprotein complexes that catalyze full-site HIV-1 DNA integration in vitro: implications for the mechanism of viral cDNA integration. Virology 360:1–5

201. Kessl JJ, Li M, Ignatov M, Shkriabai N, Eidahl JO, Feng L, Musier-Forsyth K, Craigie R, Kvaratskhelia M (2011) FRET analysis reveals distinct conformations of IN tetramers in the presence of viral DNA or LEDGF/p75. Nucleic Acids Res 39:9009–9022
202. Feng L, Dharmarajan V, Serrao E, Hoyte A, Larue RC, Slaughter A, Sharma A, Plumb MR, Kessl JJ, Fuchs JR, Bushman FD, Engelman AN, Griffin PR, Kvaratskhelia M (2016) The competitive interplay between allosteric HIV-1 integrase inhibitor BI/D and LEDGF/p75 during the early stage of HIV-1 replication adversely affects inhibitor potency. ACS Chem Biol 11:1313–1321
203. Deng N, Hoyte A, Mansour YE, Mohamed MS, Fuchs JR, Engelman AN, Kvaratskhelia M, Levy R (2016) Allosteric HIV-1 integrase inhibitors promote aberrant protein multimerization by directly mediating inter-subunit interactions: structural and thermodynamic modeling studies. Protein Sci 25:1911–1917
204. 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
205. 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
206. 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
207. Engelman A (1999) In vivo analysis of retroviral integrase structure and function. Adv Virus Res 52:411–426
208. Engelman A (2011) Pleiotropic nature of HIV-1 integrase mutations. In: Neamati N (ed) HIV-1 integrase: mechanism and inhibitor design. Wiley, Hoboken, pp 67–81
209. 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
210. Johnson BC, Métifiot 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
211. Madison MK, Lawson DQ, Elliott J, Ozantürk AN, Koneru PC, Townsend D, Errando M, Kvaratskhelia M, Kutluay SB (2017) Allosteric HIV-1 integrase inhibitors lead to premature degradation of the viral RNA genome and integrase in target cells. J Virol 91:e00821–00817
212. 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):e1212
213. Chun T-W, Stuyver L, Mizell SB, Ehler LA, Mican JAM, Baseler M, Lloyd AL, Nowak MA, Fauci AS (1997) Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. Proc Natl Acad Sci USA 94:13193–13197
214. Chun TW, Fauci AS (2012) HIV reservoirs: pathogenesis and obstacles to viral eradication and cure. AIDS 26:1261–1268
215. Maldarelli F, Wu X, Su L, Simonetti FR, Shao W, Hill S, Spindler J, Ferris AL, Mellors JW, Kearney MF, Coffin JM, Hughes SH (2014) Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. Science 345:179–183
216. Wagner TA, McLaughlin S, Garg K, Cheung CY, Larsen BB, Styrchak S, Huang HC, Edlefsen PT, Mullins JI, Frenkel LM (2014) Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. Science 345:570–573
217. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DI, Lai J, Blankson JN, Siliciano JD, Siliciano RF (2013) Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell 155:540–551
218. Cohn LB, Silva IT, Oliveira TY, Rosales RA, Parrish EH, Learn GH, Hahn BH, Czartoski JL, McElrath MJ, Lehmann C, Klein F, Caskey M, Walker BD, Siliciano JD, Siliciano RF, Jankovic M, Nussenzweig MC (2015) HIV-1 integration landscape during latent and active infection. Cell 160:420–432
219. Simonetti FR, Sobolewski MD, Fyne E, Shao W, Spindler J, Hattori J, Anderson EM, Watters SA, Hill S, Wu X, Wells D, Su L, Luke BT, Halvas EK, Besson G, Penrose KJ, Yang Z, Kwan RW, Van Waes C, Uldrick T, Citrin DE, Kovacs J, Polis MA, Rehm CA, Gorelick R, Piatak M, Keele BF, Kearney MF, Coffin JM, Hughes SH, Mellors JW, Maldarelli F (2016) Clonally expanded $CD4^+$ T cells can produce infectious HIV-1 in vivo. Proc Natl Acad Sci USA 113:1883–1888
220. Hughes SH, Coffin JM (2016) What integration sites tell us about HIV persistence. Cell Host Microbe 19:588–598
221. Darcis G, Van Driessche B, Van Lint C (2017) HIV latency: should we shock or lock? Trends Immunol 38:217–228
222. Chen HC, Martinez JP, Zorita E, Meyerhans A, Filion GJ (2017) Position effects influence HIV latency reversal. Nat Struct Mol Biol 24:47–54
223. Gres AT, Kirby KA, KewalRamani VN, Tanner JJ, Pornillos O, Sarafianos SG (2015) X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. Science 349:99–103