Cell. Mol. Life Sci. 65 (2008) 1403 – 1424
1420-682X/08/091403-22
DOI 10.1007/s00018-008-7540-5

Cellular and Molecular Life Sciences

Review

Integrase, LEDGF/p75 and HIV replication

E. M. Poeschla

Guggenheim 18, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, 55905 (USA),
Fax: +1 507 266 2122, e-mail: emp@mayo.edu

Received 25 November 2007; received after revision 7 January 2008; accepted 10 January 2008
Online First 12 February 2008

Abstract. HIV integrates a DNA copy of its genome into a host cell chromosome in each replication cycle. The essential DNA cleaving and joining chemistry of integration is known, but there is less understanding of the process as it occurs in a cell, where two complex and dynamic macromolecular entities are joined: the viral pre-integration complex and chromatin. Among implicated cellular factors, much recent attention has coalesced around LEDGF/p75, a nuclear protein that may act as a chromatin docking factor or receptor for lentiviral pre-integration complexes. LEDGF/p75 tethers HIV integrase to chromatin, protects it from degradation, and strongly influences the genome-wide pattern of HIV integration. Depleting the protein from cells and/or over-expressing its integrase-binding domain blocks viral replication. Current goals are to establish the underlying mechanisms and to determine whether this knowledge can be exploited for antiviral therapy or for targeting lentiviral vector integration in human gene therapy.

Keywords. Integrase, integration, LEDGF/p75, HIV, lentivirus, chromatin, PWWP domain, AT Hook, HRP-2.

Introduction: integration as a viral strategy

Viruses, viewed most fundamentally, are short segments of nucleic acid that have evolved myriad ways to propagate themselves by parasitizing the replicative metabolism of cells. It is no surprise that tracking their intracellular journeys has frequently provided important insights into cell biology. From the perspective of a cell biologist, retroviruses are the most subversive of viruses because they invade not only the cell but also its genome. Shortly after its core enters the host cell cytoplasm, a retrovirus copies its single-stranded RNA genome into a double-stranded cDNA, which is maneuvered to the nucleus by still unclear mechanisms and integrated into a chromosome. There the provirus behaves for the most part like any of the cell's other 25 000 or so Pol II-transcribed genes. For HIV and other lentiviruses, the only really fundamental post-integration peculiarity in the nucleus is a necessary mechanism to bypass the cell's splicing checkpoint for the viral genome RNA and some coding mRNAs. Perhaps only prions infiltrate their hosts more intimately.

The numerous survival advantages that follow from integration include acquisition by an RNA virus of the long-term stability of chromosomal DNA, the capacity to replicate through mitosis, and the ability to forego considerable genetic encumbrance by parasitizing the elaborate cellular transcriptional apparatus. Thus, while the exceptional ability of HIV-1 to evade and slowly destroy human immunity rests on many mechanisms, the most fundamental may be integration. A stably integrated provirus can occupy a spectrum of transcriptional states, allowing it to evade immune surveillance through latency while retaining the capacity to scale up transcription rapidly and initiate progeny production [1, 2]. Chromosomal archiving of drug-resistant HIV genomes in long-lived cells can also permanently foreclose the utility of antiretroviral drugs or entire drug classes, which often

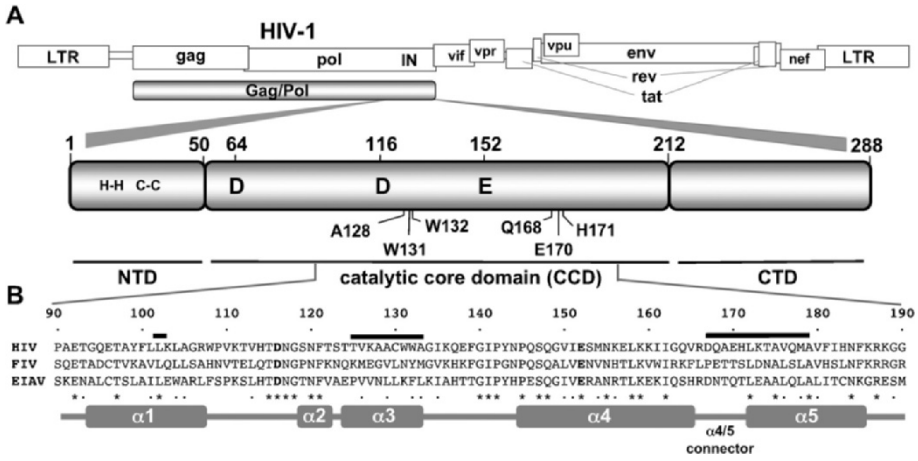

Figure 1. HIV-1 integrase: derivation, structure and key residues. (*A*) IN is derived from the Gag/Pol precursor by viral protease cleavage. The protein has three distinct structural and functional domains: an N-terminal domain (NTD, amino acids 1–50), which contains an HHCC zinc-binding motif, the core catalytic domain (CCD, residues 51–212, the dimer interface of which interacts with p75) and the C-terminal domain (CTD, residues 213–288). Of the three, the CTD is the least conserved between retroviral integrase proteins; it is involved in DNA binding, but all three domains likely contribute to this as well as to multimerization. D64, D16 and E152 coordinate magnesium ions in the catalytic center. They can be mutated to produce pure (class I) IN catalysis deficits without confounding (class II) effects on other Gag/Pol precursor functions. Class II effects are frequent, such that mutagenesis of the protein requires cautious assessment of multiple parts of the viral life cycle, including assembly. Some IN residues located in the pocket at the p75-IN dimer interface are highlighted at the bottom of the three-domain illustration, with those N-terminal to E152 in one monomer, and those C-terminal to E152 in the other monomer of the IN dimer. Although its exact oligomeric state and abundance in the PIC are still ambiguous, a wealth of *in vitro* enzymatic and virological complementation studies suggest that IN functions to integrate both long terminal repeat (LTR) ends into host DNA in a coupled (concerted) fashion, and that it acts as a multimer, with a tetramer likely [73, 154, 157, 158, 188–192]. *In vitro*, an IN dimer is sufficient for 3′ end processing, but a tetramer appears needed for DNA strand transfer activity [154, 157, 158]. In cells, p75 binds tetrameric IN [73]. (*B*) Alignment of the central part of the IN CCD for lentiviruses from the three main subgenera. Identity is indicated by asterisks and residues with conserved biochemical features by dots. IN alpha helices are indicated below, and the segments primarily involved in forming the IBD-binding pocket are indicated by heavy overlying black bars; see Figure 5 for placement of these protein elements in the IN-CCD interface. For an alignment that shows this region for additional ungulate lentiviruses and HIV-2 as well, see [100].

prevents achieving the durable viral suppression necessary for stable health. For these and other reasons, integration is not only the most defining feature of the retroviral life cycle, it has high priority as a therapeutic target. Blocking this step was in the past regarded as potentially problematic because the targeted process is a single pair of DNA cleaving and joining reactions per viral replication round (compared to about 20000 enzymatic cycles for reverse transcription, for example). However, a timely new class of antiretroviral drugs, at this point best exemplified by the HIV integrase strand transfer inhibitor raltegravir (FDA approved in October 2007) has proved that integration is a robust therapeutic target [3, 4].

Host cell factors and integration

The integration reaction is catalyzed by the virally encoded integrase (IN) protein (Fig. 1). The main DNA cleaving and joining steps are the initial removal of a terminal dinucleotide from each end of the reverse transcribed cDNA (3′ processing) and strand transfer (Fig. 2). The latter is a single-step transesterification reaction in which staggered cleavage of opposite target DNA strands provides the energy to join the newly recessed viral 3' ends to them [5–9]. In the third main step of integration, gap repair, extra nucleotides are trimmed from the 5' ends of the viral cDNA, and these are joined to host DNA 3' ends. This 'closing of the second joint' involves host cell DNA repair enzymes, but the full details remain to be elucidated [10]. Both 3' processing and strand transfer can be detected when the purified enzyme is combined with model DNAs, e.g. oligonucleotides, that serve as surrogate viral and host substrates [6, 11–14]. Such experiments have provided the framework for the goal of understanding HIV integration as it occurs in a cell, where two much more complex structures are joined, the viral pre-integration complex (PIC) and chromatin.

The PIC is a macromolecular assemblage that consists of the viral cDNA, a subset of viral proteins retained from the initial reverse transcription complex and probably host cell proteins [5, 15–19]. Similarly, chromatin is compositionally and structurally intricate. Furthermore, protein-protein and protein-DNA interactions are highly dynamic within chromatin [20, 21]. They are also likely to evolve within the PIC as it

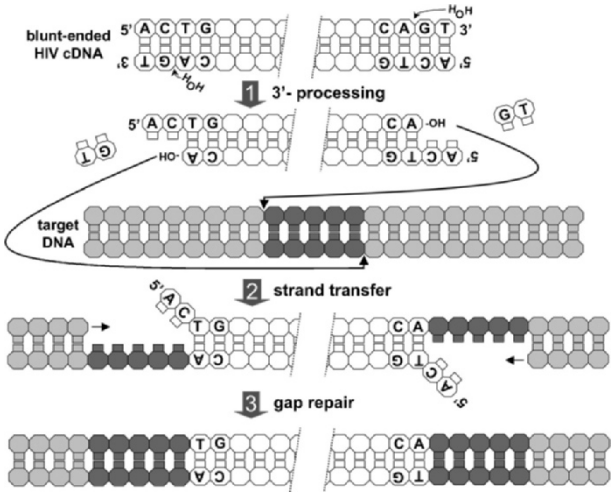

Figure 2. The integration reaction. A blunt-ended linear viral genome cDNA is the precursor to integration. 3′ processing occurs largely or entirely before nuclear entry for most retroviruses, including lentiviruses. Removal of a terminal dinucleotide from the linear cDNA molecule generates a recessed CA at the 3′ end. This first reaction step may serve to remove extra nucleotides occasionally added by reverse transcriptase [193] and promote stable complex formation [154, 194]. The CA is invariable in all retroviruses, while the removed terminal dinucleotide is not (it is TT for MLV, for example). Binding to host DNA is followed by a concerted transesterification reaction in which the 3'-OH groups of the newly recessed A nucleotides are used to attack opposite strand phosphodiester bonds separated by four to six bases in the 5′ direction (this interval is five for HIV). Therefore, the joinings occur on the same face of the double helix, flanking a major groove. Note that 3' processing and the subsequent strand transfer are really the same reaction chemically [7, 8]. In each case a hydroxyl group carries out a direct nucleophilic attack on an internucleotide phosphodiester bond. The only difference is that in 3' processing the attacking hydroxyl group is from a water molecule and in strand transfer it is the 3' hydroxyl group of the conserved, recessed A. In other words, this same 3'-OH is the leaving group in 3' processing and the attacking nucleophile in the strand transfer. An exogenous energy source is not required because this is supplied by the breaking of the chromosome DNA phosphodiester bond [5]. The process is completed by the third main step, gap repair, yielding a duplication of host DNA flanking the provirus (e.g., the five dark shaded bases shown here for HIV). Like the first two steps [6, 11, 13, 14, 195], gap repair can also be modeled *in vitro* [196, 197]. It appears to involve host cell DNA repair systems [10]. The N- and C-terminal domains of IN are essential for proper interaction with substrates. This aspect is illustrated by the reverse of the strand transfer reaction, which is termed 'disintegration' [198]. More orderly than it sounds, it is detected *in vitro* as cleavage of a substrate that mimics one viral DNA end joined to a target DNA. Disintegration can be carried out by the isolated catalytic core domain (CCD), since only in this instance are the viral DNA ends properly prepositioned; in contrast, the N- and C-terminal domains of HIV integrase are needed for 3' processing and strand transfer, presumably because without them the CCD cannot correctly position the viral cDNA termini at the active site [115]. See [8] for a detailed overview of integration.

transits the nucleus and completes the integration process [22]. Most aspects of the nuclear trajectory of the PIC have been persistently enigmatic: nuclear import, intra-nuclear trafficking, molecular evolution during transit and the engagement with chromatin [23]. Paralleling vibrant current interest in the roles cellular proteins play in other early retroviral life cycle events [24–29], the integration field is focusing on how cellular molecules participate either positively or negatively in these steps, in the integration reaction itself and in determining the site of integration within the host cell genome [10, 23, 30–33].

From this perspective, HIV research is poised to learn more about a substantial mystery: the nuclear life of the PIC. The numerous proteins that have been suspected to influence integration can be grouped into three main categories: those that (i) interact with IN directly; (ii) bind to the viral DNA to influence the process; or (iii) participate in the final gap repair step. Of equivalent interest is deciphering the roles cellular proteins are likely to play in genome-wide patterns of retroviral integration [34–46]. Experimental capabilities in this area changed with the availability of the human genome sequence, commensurately advancing bioinformatics and more recently, a second wave of high-throughput sequencing methodologies [47, 48]). Before such analyses, most evidence had suggested that retroviral integration was largely random except for highly local biases related to distortion of the double helix as it accommodates to nucleosome structure [49–54]. Exons account for only about 1.1% of human DNA. Since human exons average only 50 codons and are generally separated by much larger introns, transcription units comprise roughly 30–40% of the genome, depending on the particular gene catalogue used [55, 56]. Thus, random integration by HIV would be predicted to produce about this percentage of active transcription unit hits, whereas approximately twice this has been observed [34].

Moreover, each retrovirus genus appears to follow different targeting biases. Broadly sketched, lentiviruses favor integrating in active transcription units without favoring the promoter regions in particular [34, 35, 38, 39, 43, 45, 46], MLV favors the promoter regions of genes (transcription start sites and CpG islands) [35, 38], and ALV appears, in general, to show less specificity for transcription units [36, 37, 42]. MMTV shows none at all and is so far the most random integrator [57]. Foamy viruses are not found in excess within genes *per se*, but those that are tend to favor transcriptional starts and CpG islands [58, 59]. Within the overall patterns lie further mysteries, as some sites appear to be targeted as much as several hundredfold greater than random [34, 41, 54, 60]. HIV integration is also strongly disfavored in centromeric heterochromatin [34, 61], which is transcriptionally repressed and a poor substrate for viral transcription [62, 63]. Analyses that exploit deeper genome annotation have potential to reveal further mechanistic possibilities [48].

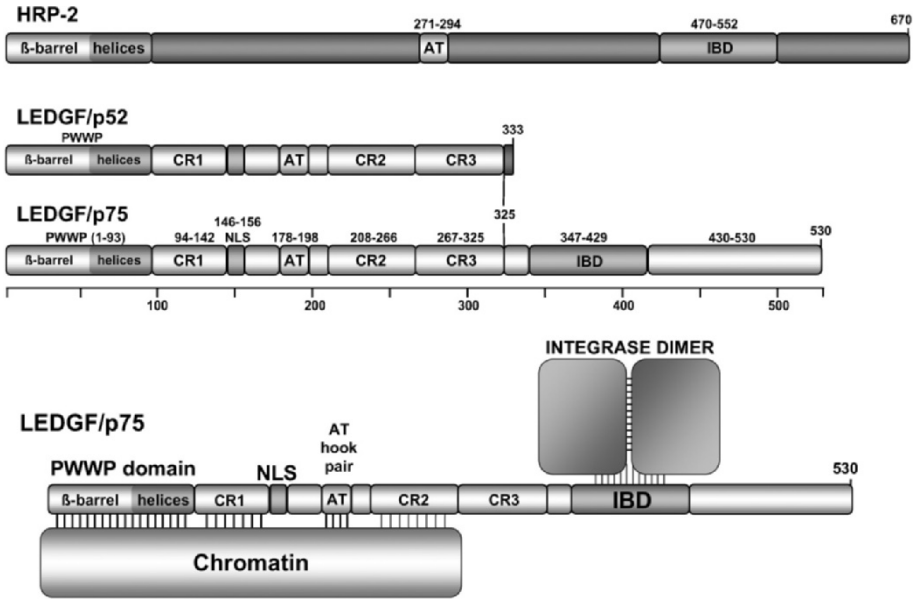

Figure 3. The modular basis for IN-to-chromatin tethering: molecular domain structure of LEDGF/p75. Lengths of protein elements are drawn to scale. The amino-terminal domain ensemble participating in chromatin binding is primarily the PWWP domain and the AT hook domain, with CR1-CR3 being relatively charged regions that influence chromatin binding to a lesser extent. However, PWWP-CR1 is necessary to transfer chromosome binding to GFP, and the N-terminus up to CR2 is required to transfer full chromatin avidity (see text). HRP-2 is shown at top. Outside of the PWWP domain, AT hooks and the integrase binding domain (IBD), HRP-2 is substantially divergent from p75. Numbers indicate the amino acids that comprise each domain. The IBD is strongly conserved in vertebrates [79–81]. See [81] for an alignment of the IBD across a variety of species.

p75: identification and putative cellular function of a lentiviral IN interactor

This remainder of this review will focus mostly on the host cell protein LEDGF/p75 (p75), which falls into the first of the three groups of host proteins noted above since it interacts directly with all lentiviral integrase proteins. I review the existing p75 literature [64–110], consider models for cofactor action, and discuss perspectives and opportunities for future research. 'HIV' denotes HIV-1 throughout. The main features of retroviral integration, including pertinent biochemical and structural considerations, as well as aspects of retrotransposon integration targeting, are also covered in Figure 1, 2, and 5 or at relevant points in each text section. For in-depth reviews of these areas and other cellular factors implicated in the integration process, see [8–10, 31–33, 111–117]; [29] provides a current overall review of host factors utilized by retroviruses.

p75 is a ubiquitously expressed 530 amino acid product of the gene *PSIP1* [64, 65]. A 333-amino acid splice variant, LEDGF/p52 (p52), shares p75's N-terminal 325 residues (Fig. 3). The two proteins have unique C-termini derived by alternative splicing, 8 amino acids in the case of p52 and 205 amino acids for p75 [64, 65]. p75 and p52 were first identified as proteins that copurified with PC4, a transcriptional coactivator [64]. Both p52 and p75 were shown to have transcriptional coactivator activity as well, with p52 being more active [64] and also having a more restricted intranuclear distribution [69]. This functional assignment remains provisional, since little subsequent work on the transcriptional role of either protein has followed. Transcriptional profiling in p75-deficient human cells revealed significant global changes (approximately 1000 mRNAs changed significantly in each direction), but Gene Ontology annotation did not suggest regulation of a coherent developmental or physiological program [85]. In –/– mouse embryonic fibroblasts (MEFs) fewer genes were up- or downregulated (<200), and again, a particular transcriptional network was not identified [101]. Studies focused on other biological questions have implicated p75 in modulating apoptosis and other cellular responses to stress and have suggested a role as an auto-antigen in certain disease states [66–68, 70–72, 84, 107]. However, p75 depletion does not disproportionately alter stress-responsive gene transcription, nor does HIV appear to preferentially target such genes [85, 101, 110].

Nearly coincident with the initial descriptions, vision science researchers isolated the cDNA from a lens epithelium library and coined 'lens epithelium-derived growth factor/p75 (LEDGF/p75)', which has entered common usage [66]. The acronym has the advantage (not trivial) of retrieving few if any confounding hits from biomedical databases, but this

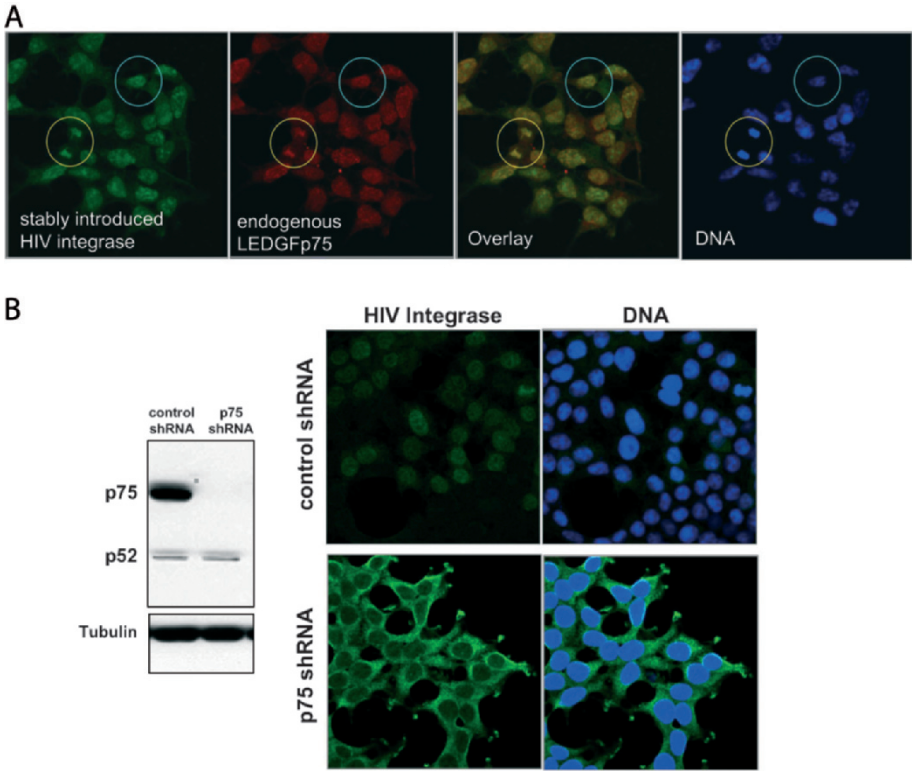

Figure 4. p75 and IN colocalization in cells and RNAi-induced untethering. (*A*) 293T cells stably expressing IN were double-stained for IN (green) and p75 (red) and analyzed by confocal microscopy. Colocalization is seen in uncondensed interphase chromatin and in condensed mitotic chromatin (see circled cells). (*B*) HIV IN is un-tethered from chromatin by p75 depletion. Western blotting of these paired cell lines expressing control and active shRNAs respectively is shown at left. The antibody recognizes an N-terminal region present in both p75 and p52. Note that only p75 levels were reduced. HIV IN was stably expressed by puromycin selection in these cells. Reproduced with permission from ref. [75].

nomenclature is otherwise somewhat unsatisfactory since the protein is neither lens-specific nor a growth factor. Rather, it is a ubiquitously expressed and constitutively nuclear transcription factor-type protein that is not secreted from the cell. Two different p75 knockout mice have had normal lenses [94, 101]. The knockouts also revealed that p75 is not essential for cell or organism survival. Knockout mice feed poorly in the perinatal period, leading to some mortality, but survivors display mild skeletal and neurobehavioral abnormalities and the derived MEFs grow normally [94]. In addition, stable knockdown is well tolerated in human cell lines, so far producing no discernible differences in growth rates, morphology or phenotypes other than lentivirus susceptibility [75, 80, 85, 95, 96, 110].

Interaction with lentiviral integrase proteins and the trafficking of integrase

The seminal finding for the HIV field was the isolation by Cherepanov et al. from 293T cells of an approximately 75 kDa protein by coprecipitation with HIV-1 IN, which mass spectrometry revealed to be LEDGF/p75 [73]. Confirmatory affinity-based proteome screens were subsequently reported [76, 86]. The interaction between p75 and IN is direct [73], and it is confined to only one of the seven retroviral genera, *lentivirinae* [75, 87, 100]. A number of laboratories established the role of p75 in chromatin tethering, mapped relevant functional domains and worked to answer the question of virological relevance [74–82, 85–91, 93, 95–97, 100–105, 108–110]. The first indication of the protein's significance was the clarification it provided about the vexing issue of the intracellular trafficking of retroviral integrase proteins. When HIV IN is expressed outside the viral context it is karyophilic [118–121] and chromatin-associated [73–75, 86, 122, 123]. This property is characteristic of the entire lentivirus genus since IN proteins from the two nonprimate (feline and ungulate) subgenera behave identically [75, 87, 100]. In contrast, gamma-retroviral (MLV) IN localizes to the cytoplasm [75]. A number of earlier studies had, quite logically, sought autonomous nuclear localization (NLS) signals in HIV IN itself (see [23, 124] for reviews). However, it is now clear that lentiviral IN proteins are cytoplasmic in the absence of p75. First, p75 and integrase were seen to precisely colocalize in association with condensed or uncondensed chromatin through all phases of the cell cycle (Fig. 4A) [73, 75]). Second, when p75 was depleted by RNAi, integrase shifted completely to the cytoplasm (Fig. 4B) [74, 75]). Third, IN mutants that are unable to interact with p75 lose chromatin association [86]. The consensus outcome was a tethering model in which p75 links IN to chromatin.

An experimental proviso that also emerged is that while fluorescent protein fusions are convenient (e.g., GFP-IN), they frequently produce ambiguous or artifactual phenotypes, with signal found in both nucleus and cytoplasm. In contrast, immunofluorescence studies with anti-IN antibodies or with small antigenic epitope-tagged versions of IN produce strict nuclear and cytoplasmic lentiviral IN distributions in the presence and absence of p75 respectively [75, 80]; see also [120], which first identified the nuclear location of epitope-tagged HIV IN. Further illustrating the point, fusions of IN to larger proteins, e.g., pyruvate kinase or beta-galactosidase, disrupt the p75-IN interaction entirely [75, 80, 123, 125]. Therefore, fluorescent protein-IN fusions are no longer preferred for subcellular trafficking experiments. To give one of a number of examples where this has been relevant, GFP-IN^{Q168A} does not segregate with mitotic chromosomes like GFP-IN, but it is indistinguishable from the latter in other cell cycle phases, with both proteins being virtually completely nuclear [86]. This phenotype suggested that p75 interaction is not required for nuclear localization of IN, whereas tracking IN without a GFP fusion partner makes it clear that p75 is necessary (e.g., Fig. 4B and [75, 80]). Nuclear import must also be distinguished from nuclear localization, since whether IN protein is trapped by p75 after import or is cotransported with it through the nucleopore is not yet clear, although studies with nuclei from digitonin-permeabilized cells suggest nuclear import can occur independently of p75 [86]. NLS-mutant p75 can sequester HIV IN in the cytoplasm [77, 80], and attachment of heterologous NLS sequences to the IBD of p75 can lead to nuclear import of IN [80].

The interest in the intracellular trafficking of IN has in part reflected longstanding interest in determining the mechanism by which HIV and other lentiviruses/lentiviral vectors infect non-dividing cells, e.g., macrophages and, in the gene therapy setting, many post-mitotic cell types [126, 127]. This now legendary problem remains unsolved [23, 124]. Despite the initially provocative correlation between intracellular trafficking of integrase proteins expressed outside the viral context and the differential capability of these retroviruses for non-dividing cell infection – HIV IN is nuclear, MLV IN is cytoplasmic – the weight of current evidence tilts against integrase or the integrase-p75 interaction being the determinant [75, 95, 96, 101, 121, 128–133]. As is discussed below, however, a role for p75 in PIC nuclear import is not definitively excluded.

Structure and the modular nature of the tether

The domain structure of p75 is illustrated in Figure 3. Discrete functional modules govern various activities. Nuclear localization of p75 is the outcome of two mechanisms, NLS-mediated nuclear import and chromatin trapping. A single transferable classical NLS is located in the N-terminal region (^{146}RRGRKRKAEKQ^{156}) [77, 80]. Four central basic residues, RKRK, are critical for NLS activity.

Chromatin linkage of IN is mediated by specific domains located at opposite ends of p75, hence the tether metaphor (Fig. 3). An 83-amino acid C-terminal integrase binding domain (IBD, residues 347–429) binds to IN [79, 80]. The crystal structure of the IN catalytic core domain (CCD)-IBD complex revealed critical contacts at the interface [82]. The structural data correlate with *in vitro* yeast two-hybrid and alanine scanning experiments that have mapped p75 interaction to the IN CCD [86, 102]. The IBD, which is highly conserved in vertebrates [79, 80], forms a right-handed alpha helix bundle that resembles so-called HEAT domains from other proteins (Fig. 5) [81]. A less-structured inter-helical loop of the IBD extends to bind in a pocket at the IN CCD dimer interface, especially residues 166–171 from one IN monomer, which connect IN alpha helices 4 and 5, and a hydrophobic patch composed principally of alpha helices 1 and 3 of the second IN monomer (Figs 1B, 5) [82]. Side chains of certain residues in the alpha-4/5 connector, e.g., Q168 and R166, do not interact obviously with IBD residues directly but rather form contacts within the same IN monomer (R166 salt-bridging with E69) or between opposite monomers (Q168 hydrogen bonding across the dimer interface with W132) (Fig. 5). Therefore, substitutions of alpha-4/5 connector amino acids are predicted to disrupt the structural integrity of IN and abrogate interaction with p75 either directly or indirectly [82]. Interpreting their viral phenotypes is complex, an issue that is discussed further below. The minimal IBD fragment that will interact with IN is not precisely determined. For example, residues 330–417 bound IN normally [80].

IBD residues identified in IN-p75 binding experiments are in good agreement with the crystal structure. For example, D366 is likely to hydrogen-bond with the backbone amides of E170 and H171 in the alpha-4/5 connector of one IN monomer (Fig. 5). These residues are not conserved in non-primate lentiviral IN proteins (Fig. 1B), which is consistent with the predicted interaction with the peptide backbone rather than the amino acid side chains [100]. The loss of a single negative charge through an Asp → Asn

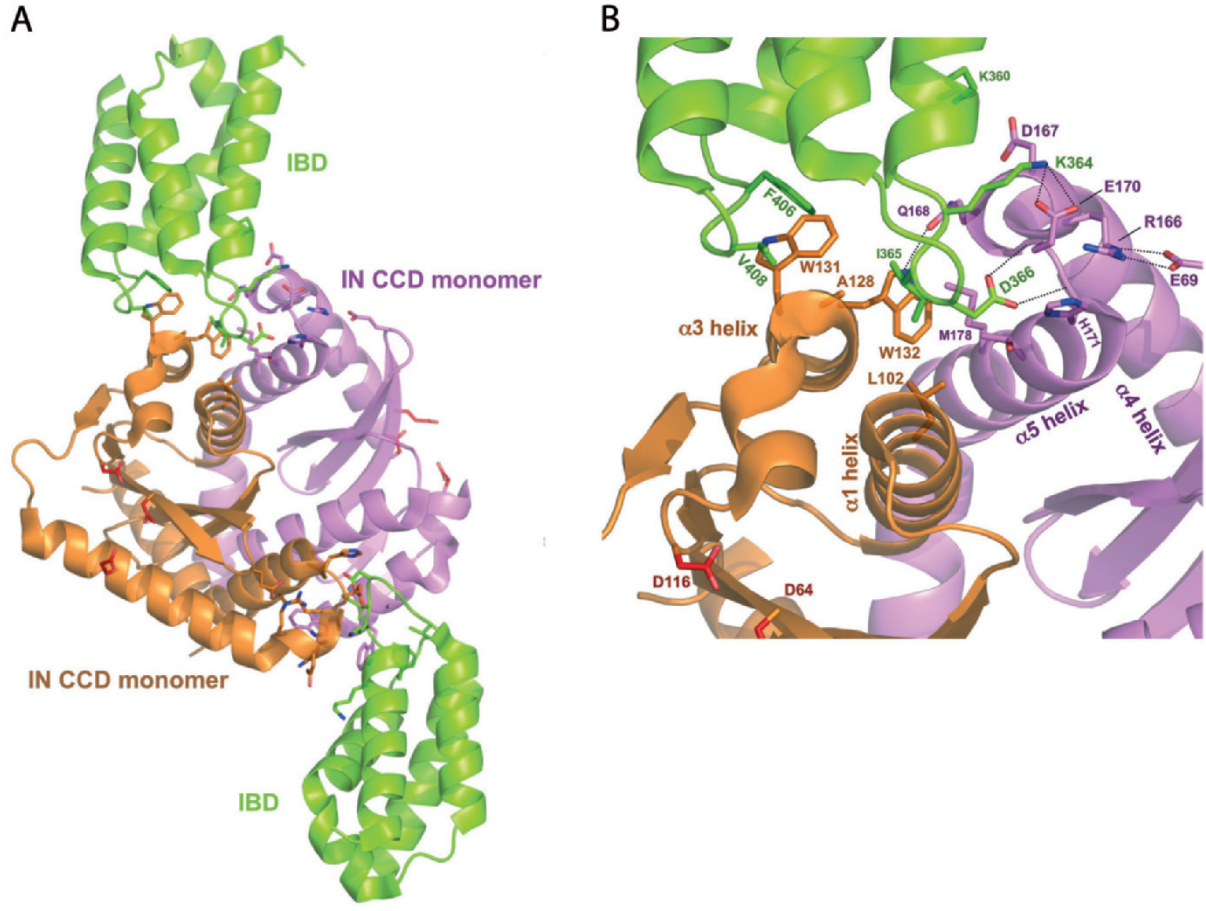

Figure 5. Structure of the HIV IN CCD-IBD interface. (*A*) An IN dimer interacting with the p75 IBD. (*B*) close-up view of the trimolecular interface. The figures were made with PyMOL, using Protein Data Bank crystal structure file 2BJ4 (www.pdb.org) from [82]. The IN CCD monomers are colored orange and pink and the IBD is colored green. Catalytic center residues D64, D116 and E152 are portrayed with red sticks. p75 contributes most of the amino acid side chains that make direct IN-p75 contacts, and mutagenesis of single IBD loop residues (e.g., I365, D366, F406) more definitively abrogates binding than do single amino acid changes in IN. For further details see text. Single-domain crystallographic or NMR (nuclear magnetic resonance) structures for IN were previously available for the CCD and the other two individual domains [199–202] as are two-domain structures for the CCD with either of the other two [203, 204]. A full-length IN structure is a long-sought but still unrealized goal. Ideally but ambitiously, this would reveal the protein complexed with LTR termini and possibly a chromatinized DNA target, with p75 a component also now to be considered. Solubility of IN has been a recurring challenge for structural biology and certain mutations are useful in this regard, in particular F185K [205]. See [206] for an updated review of HIV IN structural biology.

change at residue 366 (D366N) disrupts binding to IN. Mutants such as D366N and D366A have been useful for demonstrating specificity of effects in HIV replication experiments [90, 95–97, 101, 105, 109]. Another significant interaction is a salt bridge between K364 and E170 (Fig. 5), although mutation of the lysine alone is not sufficient to disrupt the interaction [79]. On the other IN monomer, multiple hydrophobic interactions predominate (e.g., W131 with p75 residues V408 and F406). IBD loop residue I365, mutation of which abrogates IN binding, is situated in a hydrophobic pocket formed by W132 and some other local resides, such as L102, A128 and M178 [82]. The lack of conservation among primate and non-primate lentiviral IN proteins at this interface extends to more than the alpha-4/5 connector residues noted above; other key contacts such as A128, W131, W132 and Q168 are also poorly conserved (Fig. 1B). Therefore, structure begets deep insight into function, since as noted by Cherepanov [100] the preservation of p75 interaction by all lentiviral INs despite these sequence differences is remarkable, implying a major functional advantage and a conserved role in the evolution of the lentiviral genus as a whole rather than, e.g., fortuitous binding to a common structural motif.

On the amino-terminal end of the p75 tether, the situation is more complex, but the first 198 amino acids of the protein contain a modular chromatin-binding ensemble [93] (Fig. 3). There are two dominant elements, a 'PWWP domain' (residues 1–93) and an AT-

Hook pair ([178]PKRGRPAATEVKIPKPRGRPK[198]). The PWWP domain is found in numerous other nuclear proteins [134, 135]. Deletion of this domain or just its beta-barrel subdomain from p75 disrupts microscopically detectable tethering to condensed chromosomes [93]. There is some disagreement, however, since Turlure et al. reported chromosome tethering to be fully preserved after PWWP domain deletion [89]. Analyzed biochemically, deletion of the domain impairs high affinity (Triton X-100-refractory) chromatin binding, and additional deletion of the AT Hook domain abolishes it [93]. Similarly, the PWWP domain in conjunction with an adjacent segment of 49 relatively charged amino acids (CR1) is sufficient to tether GFP to condensed chromosomes, but transfer of Triton-resistant binding to GFP is negligible without addition of the AT hook pair [93]. Two additional charged regions (CR2 and CR3) in the N-terminal region also enhance the dominant activity of the PWWP and AT hook domains, with full Triton-resistance only achieved when the N-terminal donor to GFP extends all the way through CR2 [93]. The p75 NLS has been reported to influence p75 chromatin interaction, perhaps through non-specific DNA binding properties [89]. However, the NLS-mutant protein (p75.NLS$^-$) is cytoplasmic when transiently expressed but becomes efficiently chromatin-trapped as soon as cells transit mitosis; when stably expressed p75.NLS$^-$ displays the same constitutively chromatin-bound phenotype as the normal protein, suggesting that the significant role of the NLS is rather to mediate nuclear entry [80]. The stable nuclear residence of p75.NLS$^-$ thus appears to result from efficient capture by chromatin during the mingling of nuclear and cytosolic contents at cell division. If the PWWP domain is removed from p75.NLS$^-$, it loses this chromatin-trapped property entirely, confirming the central role of this domain [93]. Detailed PWWP domain mutagenesis is indicated to pinpoint the amino acids required for chromatin interaction of p75 and other PWWP domain proteins.

The specific chromatin ligands of p75 are unknown. The relative contributions of binding to DNA versus protein in chromatin are not known either, for this protein or for PWWP domain proteins in general. While sequence-specific *in vitro* DNA binding of p75 to certain short stress response and heat shock elements was reported [68], this was not verified in a later study [89]. Non-specific *in vitro* DNA binding activity was detected for the NLS-AT hook region, but deletion of both these elements (or even the PWWP domain in addition) had only minor effects on *in vitro* integration [89], highlighting further the disconnect between such assays and the true *in vivo* situation. However, there is much about p75 modulation of targeting in relation to local DNA composition that is interesting and incompletely explained. Lentiviral integration sites have an average G/C content significantly lower than other retroviruses, an observation that stands in intriguing paradox to the overall tendency to integrate into transcription units, which are in general G/C-rich [136]. It can be reasonably speculated that the p75 AT hook pair participates in this relative A/T targeting, since an increase in HIV integration into G/C-rich regions of chromatin has been seen repeatedly in p75-deficient cells [85, 101, 110].

Other members of the HDGF family

The IBD is found in only one other human protein, hepatoma-derived growth factor (HDGF)-related protein (HRP-2). p75 and HRP-2 belong to the HDGF family, of which six human members are known (p52, p75, HDGF, HRP-1, HRP-2 and HRP-3). All six share the conserved N-terminal PWWP domain, but only p75 and HRP-2 contain the IBD and only they interact with lentiviral integrases [79, 80, 137, 138]. Outside the PWWP domain and IBD, HRP-2 and LEDGF/p75 are strongly divergent (Fig. 3). Note that HRP-2 possesses virtually the same PWWP and AT Hook domains but lacks CR1-3, and it is not tethered to chromatin constitutively [80]. HRP-2's most recognizably distinctive compositional feature is its repetitive polyserine-acidic residue motifs. Clustered serine, aspartate and glutamate residues comprise 33 % of the protein. Similar motifs, which seem likely to be relatively unstructured regions linking functional domains, have been identified in some other eukaryotic proteins, all of which are nuclear; generally, they are cellular and viral transcription factors [139–144].

p75 and integrase protein degradation in cells

In addition to chromatin tethering, and mechanistically separable from it, there is a second outcome of the p75-IN interaction: a striking increase in IN stability [78]. HIV IN is ubiquitylated and is subject to degradation in the proteasome [78, 123, 145]. Nevertheless, it is readily expressed from the IN cDNA at substantial levels provided Rev-dependence is taken into account [75]. In contrast, in p75-deficient cell lines IN was barely detectable despite IN mRNA levels equivalent to non-deficient cells [78]. p75 re-expression rescued IN expression, while neither p52 nor IBD-deleted p75 were effective [78]. Moreover, it is clear that the location of the IN-p75 complex in the

cell is irrelevant, since a nuclear localization signal-mutant version p75, which complexed with IN in the cytoplasm, still provides robust protection from degradation [78, 80]. How the interaction of the proteins is protective – whether it shields IN from a specific E3 ubiquitin ligase, for example – is not clear, but the effect is consistent with the concept of a stable p75-IN complex that is not perceived as structurally aberrant by protein recycling machinery. It could have direct virological relevance. For example, PIC-associated IN could become vulnerable to the proteasome at some point in its nuclear trajectory unless complexed with p75. It has also been noted that inhibition of the proteasome can enhance HIV infection under certain conditions, suggesting that the proteasome may target incoming virions or host proteins involved in the early steps of infection [146, 147]. Establishing a clear connection between the latter effect, the IN-stabilizing effect of p75 and the viral cofactor role requires further investigation. In addition, a cellular protein, von Hippel Lindau binding protein 1 (VBP1), has recently been suggested to foster transition of the nascently integrated PIC to transcriptional competence by facilitating IN degradation [148]. Whether and how VBP1-mediated IN degradation intersects with p75 protection of IN [78] or with the post-strand transfer gap repair step are unknown.

Lentiviral specificity

Each property discussed above – integrase interaction, chromatin tethering, proteasome protection – occurs only with the integrase proteins of lentiviruses [75, 87, 100]. As will be examined further below, the viral cofactor role is also lentivirus-specific. This has the convenient effect, aesthetically congenial to retrovirologists, of allowing gamma-retroviral vectors to be used for efficiently re-expressing p75 in RNAi- or knockout-depleted cell lines or as genus specificity controls in virus challenge experiments.

Viral cofactor role

Despite the increasingly provocative data from IN protein expression studies, the significance of p75 in the HIV life cycle was not initially clear. First, it should be emphasized that ectopic overexpression of free IN protein fails to recapitulate its normal context in important ways. IN exists as a discrete protein only after it is cleaved by the viral protease from the Gag/Pol precursor (this polyprotein includes Matrix, Capsid, Nucleocapsid, Protease, and Reverse Transcriptase as well as IN). Cleavage into the individual virion proteins occurs during late stages of assembly and IN remains PIC-associated post-entry. There is no evidence that p75 interacts with IN prior to cleavage of the latter from Gag/Pol. No producer cell interaction or virion incorporation has been detectable and no virological effects assignable to events in the producer cell have been found [75, 86, 90, 95, 96]. Thus, all p75-IN interactions are currently believed to occur with the PIC post-entry. Whether these can happen before and after nuclear import is not clearly resolved at present, but immunoprecipitation experiments have detected p75 within integration-competent PICs [75]. When PIC-associated IN engages p75 is an important question because it bears directly on mechanistic scenarios.

Evidence for an HIV integration cofactor role was first presented by Vandekerckhove et al., who reported delayed HIV replication in RNAi-depleted HeLa-P4 cells [95]. The inhibition observed was relatively small in virological terms, with slight delays detected at low virus inputs, e.g., approximately 70 versus 130 ng of HIV p24 produced at nine days at MOI = 0.002 infection, but it was rescued by p75 re-expression [95]. Also in favor of a cofactor role, knockdown of p75 changed the genomic distribution of HIV integration [85]. Three main effects were seen: the lentiviral bias for integrating into active genes was reduced about a third to halfway towards random, AT-rich DNA targeting decreased, conceivably due to the loss of p75 AT hook domain effects, and p75-regulated genes were relatively disfavored as well [85]. These integration pattern results have been confirmed and extended in two new studies that are described in more detail below [101, 110]. Since IN is the viral determinant of retroviral targeting specificity [149], the data are consistent with the exclusive interaction of p75 with lentiviral IN proteins [75, 87, 100]. *In vitro,* chimeric lambda repressor-IBD proteins were also found to increase recombinant IN-mediated strand transfer near lambda repressor-binding sites in a DNA target [150].

However, uncertainties and perplexities persisted for several years after the identification of the p75-IN interaction. In the same knocked-down cell lines in which IN protein was found to be entirely cytoplasmic and in which integration patterns were altered (e.g., the cells used in Fig. 4B, although other cells produced the same phenotypes) [75, 85], and in similar experiments with siRNAs [90], changes in overall levels of HIV integration and replication were either entirely absent [75, 85, 90], or modest (two-fold) [91]. Puzzlingly, no laboratory could detect significant effects in single-cycle infection analyses with HIV reporter viruses or vectors, whether the experiments were siRNA- or shRNA-based, or whether they were conducted in adherent fibroepithelial cell lines or in

CD4+ T cell lines [75, 85, 90, 95]. In Ciuffi et al., two different laboratories independently tested a variety of p75-depleted knockdown lines from various sources and found no viral replication impairment [85]. Vandegraaff et al. came to similar conclusions using siRNAs [90].

Another line of potential evidence for a cofactor role came from mutant virus experiments. For example, Emiliani et al. found that a mutation of HIV IN residue Q168, which they elegantly identified by two-hybrid screening of a high-complexity random IN fragment library, disrupted the IN-p75 interaction *in vitro* while preserving viral 3′ processing and strand transfer activity; in cells the mutation untethered IN protein from chromatin [86]. HIV engineered to have the mutation was also incapable of replication (this virus was dead, representing the most severe p75-attributed block reported to date). The Q168 mutant virus defect was assigned to an integration block by PCR tracking of post-entry viral cDNA forms [86]. Here it is worth noting that mutagenesis of IN is notorious for producing poorly understood pleiotropic effects on other life cycle steps (particle assembly, release, reverse transcription) by virtue of secondary ('class II') effects on other functions of the Gag/Pol precursor and its cleavage products (see [151] for a review). In this regard, the Q168A mutation was subsequently reported to only partially diminish IN binding to p75 [102, 105], and also to have some class II properties when placed in the virus, e.g., a defect in reverse transcription [105]. As such, it is presently uncertain whether the profound replication block is solely or even primarily attributable to abrogation of p75 interaction rather than perturbation of IN (or Gag/Pol) structural integrity or monomer interactions (see also the above discussion of crystal structure-informed considerations for Q168). Characterization of a number of other relevant IN mutations at the same time revealed that p75-IBD binding affinity correlates inconsistently with viral fitness, with a number of viral deficits likely the result of confounding class II effects related to perturbed structural integrity in and around the binding cleft rather than solely to loss of p75 interaction [102, 105]. Such studies have also shown that the ability to correlate viral fitness with *in vitro* binding of IN to a host factor is frequently constrained by the unpredictable dependence of binding on various assay particularities such as the pull-down domains used (e.g., GST vs. His_6) or their C- vs. N-terminal placement [102, 105]. In summary, analysis of IN mutants defective for p75 interaction has been essential for defining critical contacts to the host factor but has produced largely ambiguous evidence concerning the viral cofactor role. The example of p75 attests to the comprehensive array of approaches – from *in vitro* biochemistry to viral life cycle dissections – that are needed to fully characterize any given retroviral IN mutation.

Appreciating the strong chromatin avidity of p75 was important for understanding its viral cofactor role. As noted above, when cell lines were engineered to stably express a nuclear localization mutant of p75, it was found only in nuclei, where it was always detected as strictly chromatin-bound throughout the cell cycle [80]. The implication was that p75 is very efficiently chromatin-trapped. Consistent with this, the majority of the protein is refractory to Triton X-100 extraction and requires treatment with DNase and salt to mobilize it [93]. These and other experiments led to the hypothesis that even in apparently drastically knocked down cells, a virologically significant residuum of p75 might be confounding clear interpretation.

Definitive studies: three avenues to clarity

Faced with these ambiguities, investigators developed three basic experimental approaches: intensified RNAi [96], p75 knockout mice [101]) and overexpression of the IBD [96, 97, 109]. Prior RNAi experiments had used either siRNA transfection or plasmid-based shRNAs. Lentiviral vectors that coencoded sortable fluorescent markers and shRNAs were used to intensify RNAi in relevant HIV targets, i.e., human CD4+ T cell lines [96]. Controls included equivalently expressed control shRNAs, repeated derivation of stable lines, challenges with another lentivirus (FIV) and with a non-lentiviral retrovirus (MLV), and re-expression of shRNA-resistant p75 alleles to validate specificity [96]. Triton-resistant chromatin-bound p75 [93] was specifically correlated with infectivity of HIV-1 [96]. Human CD4+ T cells with no immunoblot-detectable p75 in whole-cell lysates but a scant detectable residue of p75 in the Triton-resistant, DNAse I- and salt-extractable chromatin fraction (S2 fraction-positive cells) were fully HIV-susceptible. In telling contrast, in cells in which the RNAi had resulted in undetectable p75 in the S2 fraction, infection was reduced 30-fold in multiple independently derived cell lines [96]. The required RNAi intensification was also evident at the mRNA level: only cells in which mRNA was reduced to less than 3% of baseline levels were S2 fraction-negative and resistant to infection. Thus, for HIV, p75 is functionally super-abundant, and must be stripped from the DNAse- and salt-extractable chromatin fraction to produce a substantial block to HIV integration [96]. Note that this model does not imply a separate chromatin-localized reservoir of the protein that is not in exchange with the total cellular pool, but rather

the opposite. Appreciating the strengths, limitations and proper controls [152, 153] for RNAi in its various forms has been and remains important in unraveling the cofactor role. p75 is a clear example of how apparently effective RNAi (see Fig. 4B) can be inadequate to reveal a protein phenotype because fractionally small residua retain potent biological activity. This has added significance when the target protein is spatially concentrated in a functionally critical way, as in this case, where it is sequestered with the cellular component of the key reaction.

The cofactor role of p75 is lentiviral-specific, applies to both single round infection and replicating virus challenge at high MOIs, and maps to integration [96]. Intra-nuclear circularization of HIV genomes (2-LTR circles) is also increased by knockdown, as expected with an integration block [95, 96]. Targeting HRP-2 did not produce similar effects, and HRP-2 overexpression did not rescue the p75 defect. Rescue required as well that the re-expressed p75 protein have functionally intact modules at both ends of the tether. For example, p75 lacking the PWWP and AT hook domains retains the function of protecting IN from proteasomal degradation but loses HIV infection-rescuing function [96].

Overexpression of chimeric GFP-IBD fusion proteins has also resulted in substantial inhibition of viral replication, which is specifically abrogated by mutating Asp366 at the heart of the CCD-IBD interface [96, 97, 109]. (Overexpressing the IBD alone has not proved feasible, because of instability that N-terminal fusion partners appear to mitigate.) An important observation is that the effects of knockdown and GFP-IBD expression appear to be multiplicative, inducing over 500-fold defects in single round infectivity [96], which has now been extended to several thousandfold in subsequent experiments [our unpublished data]. It is likely that this synergy reflects the abundance of endogenous p75, such that without the knockdown, the GFP-IBD protein is in limiting competition with wild type protein. Note that, *in vitro*, the disembodied IBD does not stimulate IN catalysis [79].

Most recently, comprehensive HIV life cycle analyses and genome-wide integration site mapping in embryonic fibroblasts from p75 knockout mice have provided clear additional evidence for the viral cofactor role [101]. The data from this study are on the one hand corroborative of the effects of stringent p75 knockdown on viral replication as well as multiple aspects of genome-wide targeting. Both ends of the tether were again shown to be required to rescue HIV infection. However, the knockout mice studies added the intriguing new information that HIV displays an MLV-resembling increase in promoter region and CpG island targeting in –/– MEFs [101]. This latter aspect was recently confirmed when MEFs from an independently derived knockout [94] and S2 fraction-negative human T cells [96] were studied [110]. The effect on genome-wide integration patterns in the human T cell lines was also greater than in the previously analyzed [85] human cells with weaker knockdowns [110]. The notion these two studies [101, 110] suggest, that removing p75 uncovers a default gamma-retroviral targeting pattern, is interesting since it might imply action of other tethering factors or targeting mechanisms that the ubiquitous presence of p75 obscures for lentiviruses (Fig. 6, factor 'X'). In addition, analysis of 15 lentiviral integration site data sets from 10 un-manipulated primary cell types and cell lines showed a linear correlation between the frequency of integration in transcription units and the cell-type-specific level of p75 expression [110]. Considered with the intensified RNAi data [96] and the prior targeting data [85], the latter result indicates that genome-wide integration patterns are affected at a range of p75 levels that are higher than the severely reduced (chromatin-stripping) levels needed to block to integration *per se*, and that the two effects might be functionally separable.

While integration-competent HIV and FIV PICs isolated from cytoplasmic fractions will coimmunoprecipitate with p75 [75], PICs isolated from either the cytoplasm or nuclei of +/+ and –/– MEFs have equivalent *in vitro* integration activities with naked DNA targets [101]. Both results are consistent with the present hypothesis that it is integration into chromatin that is the decisive variable. The MEF results also support the current consensus that earlier effects on viral cDNA integrity, e.g., completing reverse transcription properly, are not involved.

In contrast to the human T cell line studies, HRP-2 overexpression appeared to rescue the defect in p75 –/– MEFs [101]. Previously, over-expressing this protein in human cells did not rescue HIV replication, and knockdown was not additive to the inhibitory effect of p75 knockdown [96]. The virological relevance of HRP-2 is thus unclear at present, but it deserves further investigation. Its IBD displays about 50% identity and 70% similarity to the p75 IBD. Its interaction with IN is probably of lower affinity [79, 80]. HRP-2 also has distinct chromatin interaction properties. When overexpressed, it can rescue nuclear targeting of IN in p75-deficient cells, but unlike p75 it is Triton-extractable, and it does not constitutively tether IN to chromatin [80]. Fivefold knockdown did not influence HIV integration patterns [85]. As noted, regions other than the PWWP domain and IBD are quite divergent between these two HDGF family members (Fig. 3). Double p75/HRP-2 knockout mice may help clarify the role of HRP-2.

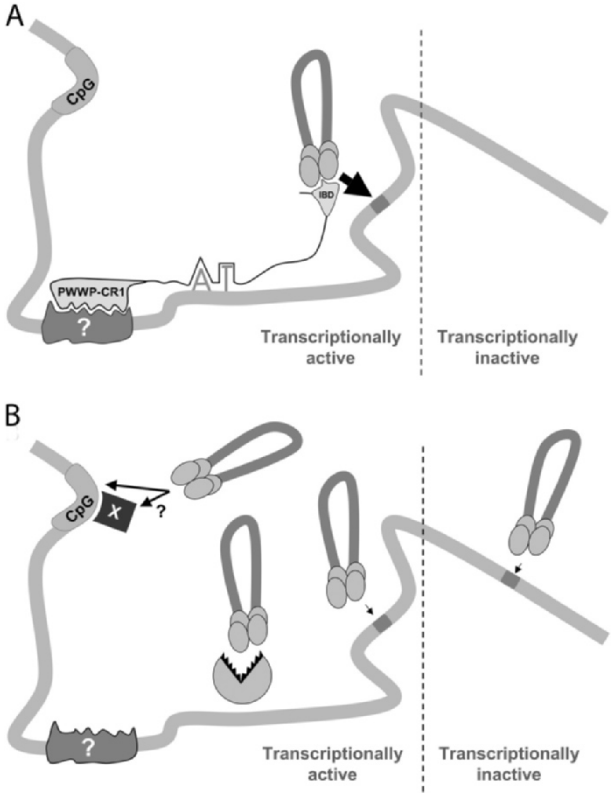

Figure 6. Basic models for p75 function in HIV integration. The chromatin fiber is depicted as a gray strand, without nucleosome or solenoid detail. Components are schematic and not drawn to scale. (*A*) p75 engages chromatin via the PWWP-CR1 elements and AT hook domains primarily. The NLS, CR2 and CR3 are not shown, but the latter two contribute to chromatin interaction. The chromatin element indicated with a question mark represents the unknown p75 ligand(s), which may include elements of the general transcription machinery. The relative contribution of direct DNA binding versus protein binding is unknown. In the presence of p75, integration is favored into active transcription units (large arrow). The four ovals represent a tetramer of integrase. (*B*) In the absence of p75, integration is reduced (arrows are smaller) and active transcription units are not favored. The IBD interaction may also protect IN from degradation. In addition, two recent papers indicate that promoter regions and CpG islands – the latter genomic feature is enriched in the unusual CpG dinucleotide and is associated with transcription start sites and regulatory regions – are relatively favored by HIV in the absence of p75 [101, 110], suggesting that removing p75 uncovers a gamma-retroviral-like targeting proclivity. The speculative black box ('X') indicates this could be mediated by a tethering factor or factors analogous to p75. More complicated scenarios, such as interaction with p75 prior to chromatin engagement, are not depicted, nor are p75-independent palindrome-like local sequence preferences illustrated.

p75 also interacts with a cellular c-Myc interactor, JPO2, perhaps reflecting the role of p75 in cellular transcription [98, 99]. JPO2 is in a narrow sense a 'cellular counterpart' of IN since it binds at or near the IBD. There are several additional parallels. The protein is protected from degradation by p75, it is tethered to chromatin and it competes with IN for p75 binding. However, while the binding regions overlap, they are different, since p75 proteins with IBD mutations that completely abrogate IN binding still bind JPO2 [99]. In addition, when overexpression of JPO2 was examined, it produced only a trivial diminution in HIV infectivity; on the other side of the life cycle, a small increase in HIV-1 Tat reporter assay activity was seen as well [99]. Careful JPO2 knockdowns remain to be reported, but a significant role for this protein in HIV replication is not apparent so far. Rather, studies of p75 and JP02 buttress the view that p75 is a general adaptor between chromatin and proteins or nucleic acid-protein complexes that must be brought into the proximity of chromatin to function, such as those involved in transcription (e.g., in c-Myc regulatory processes) or integration (lentiviral PICs). Considering further the issue of unknown cellular proteins that may interact with the p75 C-terminus, it can be deduced from domain deletion and exchange experiments that the C-terminal segment of p75 (amino acids 325–530) imparts some Triton-resistant chromatin binding to the PWWP-deleted N-terminal segment (i.e., to residues 94–325, which display no chromatin interaction when fused instead to GFP) [93]. This might suggest that, in addition to the dominant N-terminal ensemble, some component of p75 chromatin association may be mediated by bridging via proteins like JPO2.

The most convincing evidence for specificity of an antiviral effect is the evolution of mechanistically coherent mutations at the predicted binding site in the targeted viral protein. Hombrouck et al. have now reported such counterevolution of HIV [109]. After repeated passage of HIV in cells engineered to express a GFP-IBD chimeric, informative mutations arose in CCD residues that participate in the interface with p75 (E170G and A128T, which lie in opposite IN monomers; see Fig. 5). The induced mutations, which conferred reduced IBD binding affinity [109], can be reconciled well with the structural data: E170 forms a salt bridge with K364 of p75 and A128 is part of the hydrophobic pocket that engages I365 of p75 (Fig. 5). A virus constructed to have both mutations was able to replicate in the presence of the GFP-IBD chimeric but remained as impaired as wild-type HIV by p75 knockdown, consistent with additional evidence from Hombrouck et al. that the mutations disrupt GFP-IBD interaction preferentially [109]. These forced evolution studies suggest that escape is constrained by the need to retain affinity for endogenous p75.

Modeling p75 function in the HIV life cycle

Table 1 lists a number of central findings. Main conclusions for which there is now consensus are that

p75 determines IN trafficking in cells, is essential for efficient viral replication and is the first cellular protein demonstrated to influence the genome-wide integration distribution of a retrovirus. Taken together, current evidence suggests a working model in which p75 acts as a chromatin docking factor, and potentially as a receptor for the lentiviral PIC prior to chromatin engagement. The net effect in all variants of this model is to enhance the probability of the PIC completing its intranuclear journey, up to and including integration. Three processes may be involved: enhancing PIC capture by chromatin, enhancing integrase catalysis and reducing IN or PIC vulnerability to degradation by cellular pathways. There is some evidence to support each and they are not mutually exclusive (Fig. 6).

p75 knockdowns and knockout studies are in agreement, showing that in the absence of p75 infection becomes much less efficient (10- to 30-fold). Addition of a dominant-interfering GFP-IBD protein produces three or more logs of inhibition. Nevertheless, up to 10% of PICs complete integration in the absence of the cofactor. Several laboratories have also seen a 3- to 5-fold greater effect of p75 depletion on single-round luciferase reporter virus expression than is accounted for quantifying the integration deficit by techniques such as Alu-PCR or Southern blotting. Most likely this reflects the confluence of the two main effects: impaired integration and a shift of integration into chromosomal sites that are on average less conducive to HIV transcription. It should be noted that re-expressing p75 in S2 fraction-negative cells after HIV integration had occurred (or conversely, knocking it down in infected p75 wild-type cells) did not alter viral expression [96], supporting the idea that the salient variable is genomic location rather than transcriptional modulation by the protein. The lack of an absolute requirement of p75 for integration is compatible with the known ability of purified IN alone to carry out 3′ processing and strand transfer with oligonucleotide substrates *in vitro*. A probabilistic process is also compatible with the gene targeting effects, as this docking factor is a transcriptional coactivator expected to be relatively increased in abundance in the vicinity of genes. An optimally tuned bias for integrating into transcriptionally active genes may be advantageous for the virus since integration of a portion of PICs into sites where transcription is relatively repressed may foster establishing a diverse latent viral reservoir [1, 2].

It now seems robustly established that shifting integration site distributions requires a less stringent knockdown than does inhibiting integration [85, 96, 110]. These two effects might therefore be functionally separable. Conceivably, transcription unit targeting could involve PIC-chaperoning effects of p75 prior to chromatin engagement. More speculatively, PIC-p75 interactions could also be reversible. Alternatively, the chromatin-bound pool may be the only functionally critical one, which is consistent with the most straightforward PIC-to-chromatin tethering scenario in which the PIC is captured and activated by chromatin-resident p75 (Fig. 6). This would also fit a more passive scenario where the gene targeting and

Table 1. Properties of LEDGF/p75

- Transcriptional coactivator.
- Ubiquitously expressed.
- Abundant in cell.
- Nuclear and chromatin-associated throughout the cell cycle.
- Tethers lentiviral IN proteins to chromatin.
- Protects lentiviral IN proteins from proteasomal degradation.
- Enhances lentiviral IN strand transfer with *in vitro* DNA targets under most conditions.
- Primary determinants of chromatin binding: PWWP-CR1 and a tandemly situated pair of AT hooks, with ancillary effects of flanking segments.
- Nuclear import: mediated by a single classical NLS.
- C-terminal IBD mediates binding of p75 to IN.
- IBD is present in one other human protein, the HDGF family comember HRP-2.
- IBD is well conserved in vertebrates.
- Structure of IN CCD-IBD interface is established at 2 ångström resolution.
- Single amino acid mutations in the IBD (e.g., D366N or D366A) abrogate IN interaction, HIV cofactor activity.
- Detected in lentiviral PICs. PICs from p75 –/– MEFs showed no defect in 3′ processing or strand transfer into a DNA target *in vitro*.
- Knockdown or knockout reduces lentiviral targeting of transcription units. This correlates with decreased AT-rich DNA targeting and also with emergence of gamma-retrovirus-resembling CpG island/promoter region favoring.
- The shifts in integration site distributions occur with less-intensive depletion than is needed to impair viral replication/integration.
- Inhibition of viral infection requires RNAi effective enough to eradicate detectable chromatin-bound p75 (shown in human cells) or a knockout (shown in mouse cells).
- Overexpressed GFP-IBD chimeric proteins inhibit HIV infection.
- Combination of RNAi and GFP-IBD expression leads to over 3 logs of inhibition of HIV infection.
- Viral replication in presence of GFP-IBD forces evolution of specific IN resistance mutations at the IN dimer-IBD interface.
- Interaction with IN is lentivirus-specific. Consistent with this, infection by other viruses from other retroviral genera is unaffected by p75 depletion.
- JP02, a cellular protein, interacts with the p75 at the IBD, although the binding interface differs from that of IN. A virological role is not yet evident.

integration-fostering effects are not mechanistically dissociable, if all segments of chromatin are approached by PICs with statistical equivalence and gene targeting is simply a matter of rescue of those that 'land' near a p75-replete chromatin segment.

In terms of augmenting IN catalysis, p75 binding may simply bring the PIC and chromosome DNA into closer proximity or it may augment intrinsic catalytic activity, acting as an 'IN coactivator' in analogy to its role as a coactivator for the Pol II complex. *In vitro* evidence supports an enzymatic cofactor role, since recombinant p75 can augment purified IN protein catalysis (the strand transfer component) with oligonucleotide substrates [73, 79, 87, 89]. The equivalent 3' processing and strand transfer capacity of PICs isolated from +/+ and –/– MEFs [101] is consistent with the primary effect requiring engagement with chromatin-resident p75. However, the caveat is that a naked DNA template rather than chromatin is the target in these kinds of PIC integration assays, so that even if the PICs do 'pre-load' with functionally significant p75 in +/+ cells, they might not score differently than –/– cell-derived PICs. There is also intriguing evidence that recombinant p75 can promote the functionally coupled insertion of both LTR ends into a DNA target *in vitro* [100]. The literature variously refers to this as coupled, concerted or full-site integration. The two covalent joinings are not necessarily simultaneous [154]. The full-site product has the gapped, 5′ staggered structure of the bona fide strand transfer product shown in Figure 2. This essential aspect of the *in vivo* reaction has typically been difficult to recapitulate *in vitro* with the purified enzyme, although recent adjustments to assay parameters have boosted efficiency [155, 156]. The effect of p75 differed sharply between different recombinant lentiviral IN proteins: its addition enhanced full-site integration for EIAV IN, while HIV IN was instead directed virtually exclusively to catalyze uncoupled (half-site) integration in which only one viral DNA end is joined to one strand of target DNA [100]. Preferential promotion of an HIV IN half-site reaction by p75 was corroborated in a separate study [104]. Thus, the results could suggest that an additional host or viral factor may be required specifically for HIV IN to achieve the true concerted reaction that must occur *in vivo* [100]. Alternatively, in the second study, p75 appeared to interfere with higher-order oligomerization of recombinant HIV IN [104]. Since achieving the full-site reaction *in vitro* is thought to require the tetrameric form of the enzyme while the dimer mediates half-site integration [154, 157, 158], properly coupled joining may require that p75 binding occur after assembly of a mature 'synaptic complex' in which tetrameric HIV-1 IN is complexed to the viral cDNA ends [104]. A note of complexity was added by Pandey et al., who found that full site integration by HIV IN was promoted at lower p75 concentrations but inhibited at higher concentrations [103].

Consistent with a mechanism localized to chromatin, the breakpoint for significantly impairing lentiviral infection is achieving S2 chromatin fraction negativity for p75 [96]. It cannot be entirely excluded at this point that the latter is a surrogate for the total cellular deficit and that the PIC acquires p75 earlier, perhaps in interchromatin space [159] or even prior to nuclear entry as was suggested by PIC coimmunoprecipitation [75]. The lack of detail on this question parallels the larger ignorance about whether retroviral PICs negotiate intranuclear transit by specific directional processes or by random diffusion. For example, evidence has been presented that the nuclear lamina protein Emerin, which the PIC would presumably encounter immediately after nuclear entry, is required for the PIC to orient properly to the nuclear landscape and access chromatin [160]. However, this result was challenged by subsequent studies in Emerin-knockout MEFs, in which HIV infection was unimpaired; introduction of human Emerin into these cells was also without effect [161]. Finally, the reality for p75 may be dynamically complex. Many chromatin-associated proteins are highly mobile, such that they rapidly and continuously exchange among chromatin binding sites in a stop-and-go manner [20, 21]. The inferred competition between endogenous p75 and IBD fusion proteins would be consistent with a situation whereby p75 interacts in such a manner with both the PIC and chromatin [96, 97].

Completion of the integrated proviral state is a process of survival. The susceptibility of HIV reverse transcription complexes and HIV IN to the proteasome in tandem with the striking protective effect of p75 for IN [78] suggests that vulnerability of IN to cellular degradative pathways could increase during the late window of strand transfer and gap repair, perhaps as it becomes more exposed in relation to other PIC constituents. The nucleus is likely primed to recognize the threat of a foreign DNA-protein complex in a variety of ways that terminate in the ubiquitin-proteasome pathway (Fig. 6B). p75 defense from such mechanisms could result by shielding of the preintegration synaptic complex from nuclear ligands – i.e., 'nuclear restriction factors' – that channel it into degradative or restrictive pathways. Perhaps the underlying process is primarily kinetic, i.e., without p75 the PIC might simply reside longer in the nucleus in an un-integrated state, increasing the likelihood it will be interdicted or lethally modified. However, there is some evidence against p75-dependent increases in wholesale lentiviral PIC degradation *per se*.

Differential decay rates for functional nuclear PICs in the presence and absence of p75 are not established. But in p75-depleted cells in which integration is impaired, the level of total late viral cDNAs is not significantly decreased within the first 36–48 h after infection and there is an increase in dead-end circularized genomes [95, 96, 101]. This is the same pathway to which class I mutant (and p75-binding) IN viruses are routed. In addition, integration-capable nuclear PICs were isolated from p75 –/– mouse cells [101]. Although harder to address experimentally, it has not been ruled out that the ensemble of host cell enzymes that carries out gap repair is not properly entrained in the absence of p75 and the semi-joined provirus is 'repaired' by its removal; however, the increases in 2-LTR circles that have been seen in p75-deficient cells weigh against this. Note also that p75 knockout did not affect the well-established tendency of retroviral integrations to occur at weakly conserved yet virus and/or genus-specific DNA sequences with palindrome-approximating symmetry [48, 61, 116, 136, 162, 163], indicating cofactor-independent, presumably IN-autonomous selectivity that depends on the local DNA sequence irrespective of larger-scale genomic features or the G/C content of the DNA [101].

p75 and future directions in HIV research and human gene therapy

Four years of investigation have produced substantial progress. p75 is the dominant IN-interacting factor in cells. It explains IN chromatin association, is essential for efficient HIV replication and is the first cellular factor found to regulate a genome-wide retroviral integration pattern. A satisfactory explanation for IN-to-chromatin tethering by p75 has been deduced from the modular organization of the protein, and the structure of the IBD-IN interface is defined. A hypothesis that p75 was the key to lentiviral PIC nuclear import drew strong initial interest, but evidence to date does not point to a significant role. This has not been definitively excluded, however, and further analyses in specific cell types, e.g., macrophages, and under conditions of growth arrest could be revealing. It should also be noted in this context that HIV-specific integration patterns are the same in dividing and nondividing cells [44]. Whether similar proteins contribute to effects observed for p75, most notably HRP-2, will also be important to clarify. A more general experimental lesson learned is that cellular factors that engage the incoming HIV substructure can have virological efficacy at fractionally minute cellular levels. Proteins with significant roles may be missed by RNAi screens of lower intensity. It seems doubtful, for example, that p75 would have been identified in a broad RNAi library screen. Gene therapy to reduce cellular p75 levels also seems unlikely to be an effective antiviral strategy.

Targeting the p75-IBD interaction

Two main possibilities for exploiting the p75-IN interaction come readily to mind. First, is it druggable? Targeting protein-protein interactions with small molecule drugs has traditionally been viewed with skepticism. Compared with enzyme active sites the surfaces involved are generally much larger (and flatter), known reaction substrates are not available to template antagonist molecule design, and the higher affinity of protein-protein interactions may prevent drug molecule competition [164]. See [165] for a recent reconsideration of these obstacles and a description of successful precedents, e.g., [166, 167]. Optimistically, the IBD inserts into a relatively small and deep cleft at the IN dimer interface (Fig. 5). In fact, prior to the recognition of this region as the p75-binding site, there was evidence it could contain hotspots for small-molecule binding. Molteni et al. [168] reported the structures of two related small molecules (tetraphenylarsonium and dihydroxyphenyltriphenylarsonium) bound to the IN CCD, where they were found to occupy what is now recognized as the IBD-binding cleft (compare panel B of Fig. 5 with Figure 3b of Molteni et al., which shows a similarly rotated view of the CCD). Binding specificity appeared to depend on a charge-charge interaction with the carbonyl O atom of Q168. The second of these two compounds also inhibited 3′ processing and strand transfer *in vitro*, albeit weakly ($IC_{50} = 150$ μM). The compounds consist of a core of arsenic surrounded by four aromatic groups; replacing the arsenic center of dihydroxyphenyltriphenylarsonium with phosphorous reduced the IC_{50} to 13.5 μM. Potency in disrupting IBD-IN binding and the numerous other parameters relevant to realistic lead compound development have not been reported.

More recently, the binding site for an IN inhibitor with similar *in vitro* activity (a coumarin derivative) was found to be IN residues ^{128}AACWWAGIK^{136} [169]. This peptide overlaps with IBD contact residues A128, A129, W131 and W132 (Fig. 5B). The inhibitor appears to disrupt IN oligomerization, with mutations of C130 and W132 conferring resistance to this effect [169]. In addition, peptides derived from the p75 IBD (fragments of the loop formed by residues 353–411) have been reported to block HIV integration *in vitro* and in cultured cells through a hypothesized mecha-

nism involving a shift of IN to a tetrameric state defective for the 3' processing step [106]. It is therefore possible that appropriate library screens will identify lead compounds that can that disrupt the p75-IN interaction. Challenges will include the relatively high affinity of the interaction and devising high-throughput methods that surmount the relative insolubility of IN.

Targeting lentiviral vectors

A second opportunity relates to the lentiviral integration-targeting properties of p75. Lentiviral vectors derived from HIV-1 [170] and other lentiviruses [171, 172] have shown preclinical promise as the only gene therapy vectors that reliably combine two properties: integration (permanence) and non-dividing cell transduction [173]. The largest safety concern with such vectors is now believed to stem not from their derivation from pathogenic lentiviruses, but rather from the risk of oncogenic insertional mutagenesis. Although little evidence for this problem has been specifically attributable to lentiviral vectors *per se*, its potential for any retroviral vector has long been recognized from preclinical experiments [174]. It became a reality after recent human gene therapy carried out with gamma-retroviral vectors [175]. There are many caveats about generalizing from the specifics of this trial to other vectors and settings [176], and the relative disinclination of lentiviruses to target promoters is reassuring. Nevertheless, an ability to target lentiviral vectors to specific genome sites would be a major advance. Such highly focused targeting may prove difficult, but even developing the capacity to generally shift lentiviral integration site profiles more broadly – a feat achieved for the first time by p75 depletion [85, 101, 110] – could enhance safety and possibly foster other desirable gene therapy properties.

For targeting, the modular separation of p75's integrase binding function from its chromatin interaction domains suggests opportunities. As noted above, when the IBD was fused to the DNA-binding domain of phage lambda repressor, increased strand transfer by purified integrase was observed near lambda repressor-binding sites of a small naked DNA target [150]. Moving from *in vitro* targets like this to lentiviral vectors negotiating the human genome is obviously a challenge. Two basic tests of principle can be envisioned: randomizing integration by replacing the p75 N-terminal domains with protein modules that associate with chromatin promiscuously and narrowing integration with ones having different gradations of specificity. The initial questions to be addressed will be whether such proteins can rescue p75 cofactor function, followed by the question of whether genome-wide integration patterns are shifted and, more ambitiously, how narrowly targeting can be constrained.

For targeting lentiviral vectors, interesting success has been achieved with chimeric fusions of IN directly to DNA-binding proteins, although efficiency is not yet sufficient for practical use [177, 178]. As noted by Bushman et al., retroviral vector targeting strategies are all challenged by the overwhelming excess of non-specific competitor DNA, where sequence-specific DNA-binding proteins can accumulate when not bound to their specific sites [41]. However, more sharply focused targeting of integration through tethering of a viral integrase by a chromatin-associated protein has a striking natural precedent in yeast LTR retrotransposons, which are under strong selection pressure to integrate benignly because they can transpose in haploid cells and they cannot spread horizontally, so that their genetic fates are bound to that of the particular cell in which they reside [179]. Consider that the preference of HIV for active transcription units is an approximately twofold effect, albeit one established with incontrovertible rigor due to the large number of independent events in the samples. In contrast, fully 90–95 % of Ty5 retrotransposon insertions occur within confined heterochromatic regions at yeast telomeres or at the silent mating loci [180]. This process is mediated by a direct tethering interaction between Ty5 IN and a heterochromatin-associated protein, Sir4p [181]. In contrast to the complex HIV IN dimer interface, a six-amino acid peptide near the Ty5 IN C-terminus interacts with Sir4p. This targeting domain (TD) must be phosphorylated at a single serine to function, and preventing that post-translational modification or substituting other single amino acids in the TD can randomize Ty5 integration patterns [182].

Moreover, Ty5 can be retargeted by tethering of Sir4p to ectopic DNA sites or by replacing the TD with peptides that recognize alternative chromatin proteins [183, 184]. There is even evidence that yeast decontrol Ty5 integration in times of environmental stress by reducing TD serine phosphorylation [182]. By unleashing mutagenic genome-wide transposition in this way, it was hypothesized that yeast enlarge the mutation pool available for adaptive natural selection [182]. The scenario is consistent with Barbara McClintock's proposal that cells activate transposable element replication to restructure their genomes in response to stress [185]. From this perspective, Ty5 and its *Saccharomyces* host are mutualists [182]. Commingling of retroelement and host cell genetic fates through deep evolutionary time is the rule in

many other genomes, including our own, almost half of which is recognizably derived from retroelements (8% from endogenous retroviruses) [55, 56, 186, 187]. Perhaps prions do not more intimately infiltrate their hosts.

Finally, p75's chromatin ligands are unknown. We do not know either the versatility of the p75 tether. Lentiviral INs and JP02 may be first members of what will turn out to be a larger a group of p75-tethered proteins. Illuminating how p75 engages the Pol II transcription machinery and/or other nuclear molecules holds additional promise for understanding the nuclear life of the PIC.

Acknowledgements. My laboratory has received support from NIH grants (AI47536, EY14411, AI77433), a Pfizer Scholars Grant, the Helen C. Leavitt Foundation and the John H. Tietze Foundation Research Trust. I thank numerous laboratory colleagues who contributed to the studies discussed, and Emeric Wasielewski for generously teaching me the intricacies of PyMOL scripts and helping with Figure 5.

1 Han, Y., Wind-Rotolo, M., Yang, H. C., Siliciano, J. D. and Siliciano, R. F. (2007). Experimental approaches to the study of HIV-1 latency. Nat. Rev. Microbiol. 5, 95–106.
2 Bisgrove, D., Lewinski, M., Bushman, F. and Verdin, E. (2005). Molecular mechanisms of HIV-1 proviral latency. Expert Rev. Anti. Infect. Ther. 3, 805–814.
3 Hazuda, D. J., Young, S. D., Guare, J. P., Anthony, N. J., Gomez, R. P., Wai, J. S., Vacca, J. P., Handt, L., Motzel, S. L., Klein, H. J., Dornadula, G., Danovich, R. M., Witmer, M. V., Wilson, K. A., Tussey, L., Schleif, W. A., Gabryelski, L. S., Jin, L., Miller, M. D., Casimiro, D. R., Emini, E. A. and Shiver, J. W. (2004). Integrase inhibitors and cellular immunity suppress retroviral replication in rhesus macaques. Science 305, 528–532.
4 Grinsztejn, B., Nguyen, B. Y., Katlama, C., Gatell, J. M., Lazzarin, A., Vittecoq, D., Gonzalez, C. J., Chen, J., Harvey, C. M. and Isaacs, R. D. (2007). Safety and efficacy of the HIV-1 integrase inhibitor raltegravir (MK-0518) in treatment-experienced patients with multidrug-resistant virus: a phase II randomised controlled trial. Lancet 369, 1261–1269.
5 Brown, P. O., Bowerman, B., Varmus, H. E. and Bishop, J. M. (1987). Correct integration of retroviral DNA in vitro. Cell 49, 347–356.
6 Craigie, R., Fujiwara, T. and 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.
7 Engelman, A., Mizuuchi, K. and Craigie, R. (1991). HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell 67, 1211–1221.
8 Brown, P. O. (1999) Integration. In: Retroviruses Coffin, J. M., Hughes, S. H. and Varmus, H. E., (eds.), pp. 161–203, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
9 Craigie, R. (2001). HIV integrase, a brief overview from chemistry to therapeutics. J. Biol. Chem. 276, 23213–23216.
10 Skalka, A. M. and Katz, R. A. (2005). Retroviral DNA integration and the DNA damage response. Cell Death Differ. 12 Suppl. 1, 971–978.
11 Bushman, F. D., Fujiwara, T. and Craigie, R. (1990). Retroviral DNA integration directed by HIV integration protein in vitro. Science 249, 1555–1558.
12 Katz, R. A., Merkel, G., Kulkosky, J., Leis, J. and Skalka, A. M. (1990). The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. Cell 63, 87–95.
13 Vora, A. C., Fitzgerald, M. L. and Grandgenett, D. P. (1990). Removal of 3'-OH-terminal nucleotides from blunt-ended long terminal repeat termini by the avian retrovirus integration protein. J. Virol. 64, 5656–5659.
14 Bushman, F. D. and 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.
15 Bowerman, B., Brown, P. O., Bishop, J. M. and Varmus, H. E. (1989). A nucleoprotein complex mediates the integration of retroviral DNA. Genes Dev. 3, 469–478.
16 Ellison, V., Abrams, H., Roe, T., Lifson, J. and Brown, P. (1990). Human immunodeficiency virus integration in a cell-free system. J. Virol. 64, 2711–2715.
17 Farnet, C. M. and Haseltine, W. A. (1990). Integration of human immunodeficiency virus type 1 DNA in vitro. Proc. Natl. Acad. Sci. USA 87, 4164–4168.
18 Bukrinsky, M. I., Sharova, N., McDonald, T. L., Pushkarskaya, T., Tarpley, W. G. and Stevenson, M. (1993). Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. Proc. Nat. Acad. Sci. USA 90, 6125–6129.
19 Miller, M. D., Farnet, C. M. and Bushman, F. D. (1997). Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. J. Virol. 71, 5382–5390.
20 Phair, R. D., Scaffidi, P., Elbi, C., Vecerova, J., Dey, A., Ozato, K., Brown, D. T., Hager, G., Bustin, M. and Misteli, T. (2004). Global nature of dynamic protein-chromatin interactions in vivo: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. Mol. Cell. Biol. 24, 6393–6402.
21 Catez, F., Ueda, T. and Bustin, M. (2006). Determinants of histone H1 mobility and chromatin binding in living cells. Nat. Struct. Mol. Biol. 13, 305–310.
22 Fassati, A. and Goff, S. P. (2001). Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. J. Virol. 75, 3626–3635.
23 Suzuki, Y. and Craigie, R. (2007). The road to chromatin – nuclear entry of retroviruses. Nat. Rev. Microbiol. 5, 187–196.
24 Sheehy, A. M., Gaddis, N. C., Choi, J. D. and Malim, M. H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature 418, 646–650.
25 Towers, G. J., Hatziioannou, T., Cowan, S., Goff, S. P., Luban, J. and Bieniasz, P. D. (2003). Cyclophilin A modulates the sensitivity of HIV-1 to host restriction factors. Nat. Med. 9, 1138–1143.
26 Stremlau, M., Owens, C. M., Perron, M. J., Kiessling, M., Autissier, P. and Sodroski, J. (2004). The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. Nature 427, 848–853.
27 Bieniasz, P. D. (2004). Intrinsic immunity: a front-line defense against viral attack. Nat. Immunol. 5, 1109–1115.
28 Sokolskaja, E. and Luban, J. (2006). Cyclophilin, TRIM5, and innate immunity to HIV-1. Curr. Opin. Microbiol. 9, 404–408.
29 Goff, S. P. (2007). Host factors exploited by retroviruses. Nat. Rev. Microbiol. 5, 253–263.
30 Bushman, F. D. (1999). Host proteins in retroviral cDNA integration. Adv. Virus Res. 52, 301–317.
31 Engelman, A. (2003). The roles of cellular factors in retroviral integration. Curr. Top. Microbiol. Immunol. 281, 209–238.
32 Sorin, M. and Kalpana, G. V. (2006). Dynamics of virus-host interplay in HIV-1 replication. Curr. HIV Res. 4, 117–130.
33 Van Maele, B., Busschots, K., Vandekerckhove, L., Christ, F. and Debyser, Z. (2006). Cellular cofactors of HIV-1 integration. Trends Biochem. Sci. 31, 98–105.
34 Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R. and Bushman, F. (2002). HIV-1 integration in the human

genome favors active genes and local hotspots. Cell 110, 521–529.
35 Wu, X., Li, Y., Crise, B. and Burgess, S. M. (2003). Transcription start regions in the human genome are favored targets for MLV integration. Science 300, 1749–1751.
36 Narezkina, A., Taganov, K. D., Litwin, S., Stoyanova, R., Hayashi, J., Seeger, C., Skalka, A. M. and Katz, R. A. (2004). Genome-wide analyses of avian sarcoma virus integration sites. J. Virol. 78, 11656–11663.
37 Mitchell, R. S., Beitzel, B. F., Schroder, A. R., Shinn, P., Chen, H., Berry, C. C., Ecker, J. R. and Bushman, F. D. (2004). Retroviral DNA Integration: ASLV, HIV, and MLV Show Distinct Target Site Preferences. PLoS Biol. 2, E234.
38 Hematti, P., Hong, B. K., Ferguson, C., Adler, R., Hanawa, H., Sellers, S., Holt, I. E., Eckfeldt, C. E., Sharma, Y., Schmidt, M. et al. (2004). Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. PLoS Biol. 2, e423.
39 Crise, B., Li, Y., Yuan, C., Morcock, D. R., Whitby, D., Munroe, D. J., Arthur, L. O. and Wu, X. (2005). Simian immunodeficiency virus integration preference is similar to that of human immunodeficiency virus type 1. J. Virol. 79, 12199–12204.
40 Maxfield, L. F., Fraize, C. D. and Coffin, J. M. (2005). Relationship between retroviral DNA-integration-site selection and host cell transcription. Proc. Natl. Acad. Sci. USA 102, 1436–1441.
41 Bushman, F., Lewinski, M., Ciuffi, A., Barr, S., Leipzig, J., Hannenhalli, S. and Hoffmann, C. (2005). Genome-wide analysis of retroviral DNA integration. Nat. Rev. Microbiol. 3, 848–858.
42 Barr, S. D., Leipzig, J., Shinn, P., Ecker, J. R. and Bushman, F. D. (2005). Integration targeting by avian sarcoma-leukosis virus and human immunodeficiency virus in the chicken genome. J. Virol. 79, 12035–12044.
43 Kang, Y., Moressi, C. J., Scheetz, T. E., Xie, L., Tran, D. T., Casavant, T. L., Ak, P., Benham, C. J., Davidson, B. L. and McCray, P. B., Jr. (2006). Integration site choice of a feline immunodeficiency virus vector. J. Virol. 80, 8820–8823.
44 Ciuffi, A., Mitchell, R. S., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J. R. and Bushman, F. D. (2006). Integration site selection by HIV-based vectors in dividing and growth-arrested IMR-90 lung fibroblasts. Mol. Ther. 13, 366–373.
45 MacNeil, A., Sankale, J. L., Meloni, S. T., Sarr, A. D., Mboup, S. and Kanki, P. (2006). Genomic sites of human immunodeficiency virus type 2 (HIV-2) integration: similarities to HIV-1 in vitro and possible differences in vivo. J. Virol. 80, 7316–7321.
46 Hacker, C. V., Vink, C. A., Wardell, T. W., Lee, S., Treasure, P., Kingsman, S. M., Mitrophanous, K. A. and Miskin, J. E. (2006). The integration profile of EIAV-based vectors. Mol. Ther. 14, 536–545.
47 Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z. et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376–380.
48 Wang, G. P., Ciuffi, A., Leipzig, J., Berry, C. C. and Bushman, F. D. (2007). HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. Genome Res. Published online before print, 1 June.
49 Pryciak, P. M., Sil, A. and Varmus, H. E. (1992). Retroviral integration into minichromosomes in vitro. EMBO J. 11, 291–303.
50 Pryciak, P. M. and Varmus, H. E. (1992). Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. Cell 69, 769–780.
51 Pruss, D., Bushman, F. D. and Wolffe, A. P. (1994). Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. Proc. Nat. Acad. Sci. USA 91, 5913–5917.
52 Pruss, D., Reeves, R., Bushman, F. D. and Wolffe, A. P. (1994). The influence of DNA and nucleosome structure on integration events directed by HIV integrase. J. Biol. Chem. 269, 25031–25041.
53 Muller, H. P. and Varmus, H. E. (1994). DNA bending creates favored sites for retroviral integration: an explanation for preferred insertion sites in nucleosomes. EMBO J. 13, 4704–4714.
54 Withers-Ward, E. S., Kitamura, Y., Barnes, J. P. and Coffin, J. M. (1994). Distribution of targets for avian retrovirus DNA integration in vivo. Genes Dev. 8, 1473–1487.
55 Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. et al. (2001). Initial sequencing and analysis of the human genome. Nature 409, 860–921.
56 Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A. et al. (2001). The sequence of the human genome. Science 291, 1304–1351.
57 Faschinger, A., Rouault, F., Sollner, J., Lukas, A., Salmons, B., Gunzburg, W. H. and Indik, S. (2008). Mouse mammary tumor virus integration site selection in human and mouse genomes. J. Virol. 82, 1360–1367.
58 Nowrouzi, A., Dittrich, M., Klanke, C., Heinkelein, M., Rammling, M., Dandekar, T., von Kalle, C. and Rethwilm, A. (2006). Genome-wide mapping of foamy virus vector integrations into a human cell line. J. Gen. Virol. 87, 1339–1347.
59 Trobridge, G. D., Miller, D. G., Jacobs, M. A., Allen, J. M., Kiem, H. P., Kaul, R. and Russell, D. W. (2006). Foamy virus vector integration sites in normal human cells. Proc. Natl. Acad. Sci. USA 103, 1498–1503.
60 Bushman, F. D. (2002). Integration site selection by lentiviruses: biology and possible control. Curr. Top. Microbiol. Immunol. 261, 165–177.
61 Carteau, S., Hoffmann, C. and Bushman, F. (1998). Chromosome structure and human immunodeficiency virus type 1 cDNA integration: centromeric alphoid repeats are a disfavored target. J. Virol. 72, 4005–4014.
62 Jordan, A., Bisgrove, D. and Verdin, E. (2003). HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. EMBO J. 22, 1868–1877.
63 Lewinski, M. K., Bisgrove, D., Shinn, P., Chen, H., Hoffmann, C., Hannenhalli, S., Verdin, E., Berry, C. C., Ecker, J. R. and Bushman, F. D. (2005). Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. J. Virol. 79, 6610–6619.
64 Ge, H., Si, Y. and Roeder, R. G. (1998). Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. EMBO J. 17, 6723–6729.
65 Ge, H., Si, Y. and Wolffe, A. P. (1998). A novel transcriptional coactivator, p52, functionally interacts with the essential splicing factor ASF/SF2. Mol. Cell 2, 751–759.
66 Singh, D. P., Ohguro, N., Chylack, L. T., Jr. and Shinohara, T. (1999). Lens epithelium-derived growth factor: increased resistance to thermal and oxidative stresses. Invest. Ophthalmol. Vis. Sci. 40, 1444–1451.
67 Sharma, P., Singh, D. P., Fatma, N., Chylack, L. T., Jr. and Shinohara, T. (2000). Activation of LEDGF gene by thermal- and oxidative-stresses. Biochem. Biophys. Res. Commun. 276, 1320–1324.
68 Singh, D. P., Fatma, N., Kimura, A., Chylack, L. T., Jr. and Shinohara, T. (2001). LEDGF binds to heat shock and stress-related element to activate the expression of stress-related genes. Biochem. Biophys. Res. Commun. 283, 943–955.
69 Nishizawa, Y., Usukura, J., Singh, D. P., Chylack, L. T., Jr. and 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.
70 Fatma, N., Singh, D. P., Shinohara, T. and Chylack, L. T., Jr. (2001). Transcriptional regulation of the antioxidant protein 2

gene, a thiol-specific antioxidant, by lens epithelium-derived growth factor to protect cells from oxidative stress. J. Biol. Chem. 276, 48899–48907.
71 Wu, X., Daniels, T., Molinaro, C., Lilly, M. B. and Casiano, C. A. (2002). Caspase cleavage of the nuclear autoantigen LEDGF/p75 abrogates its pro-survival function: implications for autoimmunity in atopic disorders. Cell Death Differ. 9, 915–925.
72 Ganapathy, V., Daniels, T. and Casiano, C. A. (2003). LEDGF/p75: a novel nuclear autoantigen at the crossroads of cell survival and apoptosis. Autoimmun. Rev. 2, 290–297.
73 Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E. and Debyser, Z. (2003). HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. J. Biol. Chem. 278, 372–381.
74 Maertens, G., Cherepanov, P., Pluymers, W., Busschots, K., De Clercq, E., Debyser, Z. and 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.
75 Llano, M., Vanegas, M., Fregoso, O., Saenz, D. T., Chung, S., Peretz, M. and Poeschla, E. M. (2004). LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral pre-integration complexes. J. Virol. 78, 9524–9537.
76 Turlure, F., Devroe, E., Silver, P. A. and Engelman, A. (2004). Human cell proteins and human immunodeficiency virus DNA integration. Front. Biosci. 9, 3187–3208.
77 Maertens, G., Cherepanov, P., Debyser, Z., Engelborghs, Y. and Engelman, A. (2004). Identification and characterization of a functional nuclear localization signal in the HIV-1 integrase (IN) interactor LEDGF/p75. J. Biol. Chem. 279, 33421–33429.
78 Llano, M., Delgado, S., Vanegas, M. and Poeschla, E. M. (2004). LEDGF/p75 prevents proteasomal degradation of HIV-1 integrase. J. Biol. Chem. 279, 55570–55577.
79 Cherepanov, P., Devroe, E., Silver, P. A. and Engelman, A. (2004). Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional coactivator p75 (LEDGF/p75) that binds HIV-1 integrase. J. Biol. Chem. 279, 48883–48892.
80 Vanegas, M., Llano, M., Delgado, S., Thompson, D., Peretz, M. and Poeschla, E. (2005). Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering. J. Cell Sci. 118, 1733–1743.
81 Cherepanov, P., Sun, Z. Y., Rahman, S., Maertens, G., Wagner, G. and Engelman, A. (2005). Solution structure of the HIV-1 integrase-binding domain in LEDGF/p75. Nat. Struct. Mol. Biol. 12, 526–532.
82 Cherepanov, P., Ambrosio, A. L., Rahman, S., Ellenberger, T. and Engelman, A. (2005). Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. Proc. Natl. Acad. Sci. USA 102, 17308–17313.
83 Bradley, C. M. and Craigie, R. (2005). Seeing is believing: structure of the catalytic domain of HIV-1 integrase in complex with human LEDGF/p75. Proc. Natl. Acad. Sci. USA 102, 17543–17544.
84 Daniels, T., Zhang, J., Gutierrez, I., Elliot, M. L., Yamada, B., Heeb, M. J., Sheets, S. M., Wu, X. and Casiano, C. A. (2005). Antinuclear autoantibodies in prostate cancer: immunity to LEDGF/p75, a survival protein highly expressed in prostate tumors and cleaved during apoptosis. Prostate 62, 14–26.
85 Ciuffi, A., Llano, M., Poeschla, E., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J. R. and Bushman, F. (2005). A role for LEDGF/p75 in targeting HIV DNA integration. Nat. Med. 11, 1287–1289.
86 Emiliani, S., Mousnier, A., Busschots, K., Maroun, M., Van Maele, B., Tempe, D., Vandekerckhove, L., Moisant, F., Ben-Slama, L., Witvrouw, M. et al. (2005). Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication. J. Biol. Chem. 280, 25517–25523.
87 Busschots, K., Vercammen, J., Emiliani, S., Benarous, R., Engelborghs, Y., Christ, F. and Debyser, Z. (2005). The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. J. Biol. Chem. 280, 17841–17847.
88 Maertens, G., Vercammen, J., Debyser, Z. and Engelborghs, Y. (2005). Measuring protein-protein interactions inside living cells using single color fluorescence correlation spectroscopy. Application to human immunodeficiency virus type 1 integrase and LEDGF/p75. FASEB J. 19, 1039–1041.
89 Turlure, F., Maertens, G., Rahman, S., Cherepanov, P. and 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, 1663–1675.
90 Vandegraaff, N., Devroe, E., Turlure, F., Silver, P. A. and 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.
91 Zielske, S. and Stevenson, M. (2006). Modest but reproducible inhibition of human immunodeficiency virus type 1 infection in macrophages following LEDGFp75 silencing. J. Virol. 80, 7275–7280.
92 Ciuffi, A. and Bushman, F. D. (2006). Retroviral DNA integration: HIV and the role of LEDGF/p75. Trends Genet. 22, 388–395.
93 Llano, M., Vanegas, M., Hutchins, N., Thompson, D., Delgado, S. and Poeschla, E. M. (2006). Identification and characterization of the chromatin binding domains of the HIV-1 integrase interactor LEDGF/p75. J. Mol. Biol. 360, 760–773.
94 Sutherland, H. G., Newton, K., Brownstein, D. G., Holmes, M. C., Kress, C., Semple, C. A. and Bickmore, W. A. (2006). Disruption of Ledgf/Psip1 results in perinatal mortality and homeotic skeletal transformations. Mol. Cell. Biol. 26, 7201–7210.
95 Vandekerckhove, L., Christ, F., Van Maele, B., De Rijck, J., Gijsbers, R., Van den Haute, C., Witvrouw, M. and 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.
96 Llano, M., Saenz, D. T., Meehan, A., Wongthida, P., Peretz, M., Walker, W. H., Teo, W. and Poeschla, E. M. (2006). An Essential Role for LEDGF/p75 in HIV Integration. Science 314, 461–464.
97 De Rijck, J., Vandekerckhove, L., Gijsbers, R., Hombrouck, A., Hendrix, J., Vercammen, J., Engelborghs, Y., Christ, F. and 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.
98 Maertens, G. N., Cherepanov, P. and Engelman, A. (2006). Transcriptional coactivator p75 binds and tethers the Myc-interacting protein JPO2 to chromatin. J. Cell Sci. 119, 2563–2571.
99 Bartholomeeusen, K., De Rijck, J., Busschots, K., Desender, L., Gijsbers, R., Emiliani, S., Benarous, R., Debyser, Z. and Christ, F. (2007). Differential Interaction of HIV-1 Integrase and JPO2 with the C Terminus of LEDGF/p75. J. Mol. Biol. 372, 407–421.
100 Cherepanov, P. (2007). LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. Nucleic Acids Res. 35, 113–124.
101 Shun, M. C., Raghavendra, N. K., Vandegraaff, N., Daigle, J. E., Hughes, S., Kellam, P., Cherepanov, P. and Engelman, A. (2007). LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. Genes Dev. 21, 1767–1778.

102 Busschots, K., Voet, A., De Maeyer, M., Rain, J. C., Emiliani, S., Benarous, R., Desender, L., Debyser, Z. and Christ, F. (2007). Identification of the LEDGF/p75 binding site in HIV-1 integrase. J. Mol. Biol. 365, 1480–1492.

103 Pandey, K. K., Sinha, S. and Grandgenett, D. P. (2007). Transcriptional coactivator LEDGF/p75 modulates human immunodeficiency virus type 1 integrase-mediated concerted integration. J. Virol. 81, 3969–3979.

104 Raghavendra, N. K. and 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.

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

106 Hayouka, Z., Rosenbluh, J., Levin, A., Loya, S., Lebendiker, M., Veprintsev, D., Kotler, M., Hizi, A., Loyter, A. and Friedler, A. (2007). Inhibiting HIV-1 integrase by shifting its oligomerization equilibrium. Proc. Natl. Acad. Sci. USA 104, 8316–8321.

107 Huang, T. S., Myklebust, L. M., Kjarland, E., Gjertsen, B. T., Pendino, F., Bruserud, O., Doskeland, S. O. and Lillehaug, J. R. (2007). LEDGF/p75 has increased expression in blasts from chemotherapy-resistant human acute myelogenic leukemia patients and protects leukemia cells from apoptosis in vitro. Mol. Cancer 6, 31.

108 Yu, F., Jones, G. S., Hung, M., Wagner, A. H., MacArthur, H. L., Liu, X., Leavitt, S., McDermott, M. J. and Tsiang, M. (2007). HIV-1 integrase preassembled on donor DNA is refractory to activity stimulation by LEDGF/p75. Biochemistry 46, 2899–2908.

109 Hombrouck, A., De Rijck, J., Hendrix, J., Vandekerckhove, L., Voet, A., Maeyer, M. D., Witvrouw, M., Engelborghs, Y., Christ, F., Gijsbers, R. et al. (2007). Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV. PLoS Pathog. 3, e47.

110 Marshall, H. M., Ronen, K., Berry, C., Llano, M., Sutherland, H., Saenz, D., Bickmore, W., Poeschla, E. and Bushman, F. D. (2007). Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. PLoS ONE 2, e1340.

111 Kulkosky, J. and Skalka, A. M. (1994). Molecular mechanism of retroviral DNA integration. Pharmacol. Ther. 61, 185–203.

112 Andrake, M. D. and Skalka, A. M. (1996). Retroviral integrase, putting the pieces together. J. Biol. Chem. 271, 19633–19636.

113 Esposito, D. and Craigie, R. (1999). HIV integrase structure and function. Adv. Virus Res. 52, 319–333.

114 Asante-Appiah, E. and Skalka, A. M. (1999). HIV-1 integrase: structural organization, conformational changes, and catalysis. Adv. Virus Res. 52, 351–369.

115 Chiu, T. K. and Davies, D. R. (2004). Structure and function of HIV-1 integrase. Curr. Top. Med. Chem. 4, 965–977.

116 Grandgenett, D. P. (2005). Symmetrical recognition of cellular DNA target sequences during retroviral integration. Proc. Natl. Acad. Sci. USA 102, 5903–5904.

117 Lewinski, M. K. and Bushman, F. D. (2005). Retroviral DNA integration – mechanism and consequences. Adv. Genet. 55, 147–181.

118 Gallay, P., Hope, T., Chin, D. and Trono, D. (1997). HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. Proc. Natl. Acad. Sci. USA 94, 9825–9830.

119 Pluymers, W., Cherepanov, P., Schols, D., De Clercq, E. and Debyser, Z. (1999). Nuclear localization of human immunodeficiency virus type 1 integrase expressed as a fusion protein with green fluorescent protein. Virology 258, 327–332.

120 Petit, C., Schwartz, O. and Mammano, F. (1999). Oligomerization within virions and subcellular localization of human immunodeficiency virus type 1 integrase. J. Virol. 73, 5079–5088.

121 Petit, C., Schwartz, O. and Mammano, F. (2000). The karyophilic properties of human immunodeficiency virus type 1 integrase are not required for nuclear import of proviral DNA. J. Virol. 74, 7119–7126.

122 Cherepanov, P., Pluymers, W., Claeys, A., Proost, P., De Clercq, E. and Debyser, Z. (2000). High-level expression of active HIV-1 integrase from a synthetic gene in human cells. FASEB J. 14, 1389–1399.

123 Devroe, E., Engelman, A. and Silver, P. A. (2003). Intracellular transport of human immunodeficiency virus type 1 integrase. J. Cell Sci. 116, 4401–4408.

124 Fassati, A. (2006). HIV infection of non-dividing cells: a divisive problem. Retrovirology 3, 74.

125 Kukolj, G., Jones, K. S. and Skalka, A. M. (1997). Subcellular localization of avian sarcoma virus and human immunodeficiency virus type 1 integrases. J. Virol. 71, 843–847.

126 Roe, T., Reynolds, T. C., Yu, G. and Brown, P. O. (1993). Integration of murine leukemia virus DNA depends on mitosis. EMBO J. 12, 2099–2108.

127 Lewis, P. F. and Emerman, M. (1994). Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. J. Virol. 68, 510–516.

128 Dvorin, J. D., Bell, P., Maul, G. G., Yamashita, M., Emerman, M. and Malim, M. H. (2002). Reassessment of the roles of integrase and the central DNA flap in human immunodeficiency virus type 1 nuclear import. J. Virol. 76, 12087–12096.

129 Limon, A., Devroe, E., Lu, R., Ghory, H. Z., Silver, P. A. and Engelman, A. (2002). Nuclear localization of human immunodeficiency virus type 1 preintegration complexes (PICs): V165A and R166A are pleiotropic integrase mutants primarily defective for integration, not PIC nuclear import. J. Virol. 76, 10598–10607.

130 Yamashita, M. and Emerman, M. (2004). Capsid is a dominant determinant of retrovirus infectivity in nondividing cells. J. Virol. 78, 5670–5678.

131 Yamashita, M. and Emerman, M. (2005). The cell cycle independence of HIV infections is not determined by known karyophilic viral elements. PLoS Pathog. 1, e18.

132 Dismuke, D. J. and Aiken, C. (2006). Evidence for a functional link between uncoating of the human immunodeficiency virus type 1 core and nuclear import of the viral preintegration complex. J. Virol. 80, 3712–3720.

133 Yamashita, M., Perez, O., Hope, T. J. and Emerman, M. (2007). Evidence for direct involvement of the capsid protein in HIV infection of nondividing cells. PLoS Pathog. 3, 1502–1510.

134 Stec, I., Nagl, S. B., van Ommen, G. J. and den Dunnen, J. T. (2000). The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? FEBS Lett. 473, 1–5.

135 Qiu, C., Sawada, K., Zhang, X. and Cheng, X. (2002). The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. Nat. Struct. Biol. 9, 217–224.

136 Berry, C., Hannenhalli, S., Leipzig, J. and Bushman, F. D. (2006). Selection of target sites for mobile DNA integration in the human genome. PLoS Comput. Biol. 2, e157.

137 Izumoto, Y., Kuroda, T., Harada, H., Kishimoto, T. and 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.

138 Ikegame, K., Yamamoto, M., Kishima, Y., Enomoto, H., Yoshida, K., Suemura, M., Kishimoto, T. and Nakamura, H. (1999). A new member of a hepatoma-derived growth factor gene family can translocate to the nucleus. Biochem. Biophys. Res. Commun. 266, 81–87.

139 van de Wetering, M., Oosterwegel, M., van Norren, K. and Clevers, H. (1993). Sox-4, an Sry-like HMG box protein, is a

transcriptional activator in lymphocytes. EMBO J. 12, 3847–3854.
140 Smith, C. A., Bates, P., Rivera-Gonzalez, R., Gu, B. and DeLuca, N. A. (1993). ICP4, the major transcriptional regulatory protein of herpes simplex virus type 1, forms a tripartite complex with TATA-binding protein and TFIIB. J. Virol. 67, 4676–4687.
141 Wu, C. L. and Wilcox, K. W. (1991). The conserved DNA-binding domains encoded by the herpes simplex virus type 1 ICP4, pseudorabies virus IE180, and varicella-zoster virus ORF62 genes recognize similar sites in the corresponding promoters. J. Virol. 65, 1149–1159.
142 Perera, L. P., Mosca, J. D., Ruyechan, W. T., Hayward, G. S., Straus, S. E. and Hay, J. (1993). A major transactivator of varicella-zoster virus, the immediate-early protein IE62, contains a potent N-terminal activation domain. J. Virol. 67, 4474–4483.
143 Meier, U. T. and Blobel, G. (1992). Nopp140 shuttles on tracks between nucleolus and cytoplasm. Cell 70, 127–138.
144 Ge, H. and Roeder, R. G. (1994). Purification, cloning, and characterization of a human coactivator, PC4, that mediates transcriptional activation of class II genes. Cell 78, 513–523.
145 Mulder, L. C. and Muesing, M. A. (2000). Degradation of HIV-1 integrase by the N-end rule pathway. J. Biol. Chem. 275, 29749–29753.
146 Schwartz, O., Marechal, V., Friguet, B., Arenzana-Seisdedos, F. and Heard, J. M. (1998). Antiviral activity of the proteasome on incoming human immunodeficiency virus type 1. J. Virol. 72, 3845–3850.
147 Butler, S. L., Johnson, E. P. and Bushman, F. D. (2002). Human immunodeficiency virus cDNA metabolism: notable stability of two- long terminal repeat circles. J. Virol. 76, 3739–3747.
148 Mousnier, A., Kubat, N., Massias-Simon, A., Segeral, E., Rain, J. C., Benarous, R., Emiliani, S. and Dargemont, C. (2007). von Hippel Lindau binding protein 1-mediated degradation of integrase affects HIV-1 gene expression at a postintegration step. Proc. Natl. Acad. Sci. USA 104, 13615–13620.
149 Lewinski, M. K., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., Collins, F., Shinn, P., Leipzig, J., Hannenhalli, S. et al. (2006). Retroviral DNA integration: viral and cellular determinants of target-site selection. PLoS Pathog. 2, e60.
150 Ciuffi, A., Diamond, T. L., Hwang, Y., Marshall, H. M. and Bushman, F. D. (2006). Modulating target site selection during human immunodeficiency virus DNA integration in vitro with an engineered tethering factor. Hum. Gene Ther. 17, 960–967.
151 Engelman, A. (1999). In vivo analysis of retroviral integrase structure and function. Adv. Virus Res. 52, 411–426.
152 (2003). Whither RNAi? Nat. Cell. Biol. 5, 489–490.
153 Llano, M., Gaznick, N. and Poeschla, E. M. (2007) Rapid, controlled and intensive lentiviral vector-based RNAi. In: HIV Protocols, 2nd edn Kalpana, G. and Prasad, V. (eds.), Humana Press, Totowa, NJ, in press.
154 Li, M., Mizuuchi, M., Burke, T. R., Jr. and Craigie, R. (2006). Retroviral DNA integration: reaction pathway and critical intermediates. EMBO J. 25, 1295–1304.
155 Sinha, S. and Grandgenett, D. P. (2005). Recombinant human immunodeficiency virus type 1 integrase exhibits a capacity for full-site integration in vitro that is comparable to that of purified preintegration complexes from virus-infected cells. J. Virol. 79, 8208–8216.
156 Li, M. and Craigie, R. (2005). Processing of viral DNA ends channels the HIV-1 integration reaction to concerted integration. J. Biol. Chem. 280, 29334–29339.
157 Faure, A., Calmels, C., Desjobert, C., Castroviejo, M., Caumont-Sarcos, A., Tarrago-Litvak, L., Litvak, S. and Parissi, V. (2005). HIV-1 integrase crosslinked oligomers are active in vitro. Nucleic Acids Res. 33, 977–986.
158 Guiot, E., Carayon, K., Delelis, O., Simon, F., Tauc, P., Zubin, E., Gottikh, M., Mouscadet, J. F., Brochon, J. C. and Deprez, E. (2006). Relationship between the oligomeric status of HIV-1 integrase on DNA and enzymatic activity. J. Biol. Chem. 281, 22707–22719.
159 Cremer, T., Cremer, M., Dietzel, S., Muller, S., Solovei, I. and Fakan, S. (2006). Chromosome territories – a functional nuclear landscape. Curr. Opin. Cell Biol. 18, 307–316.
160 Jacque, J. M. and Stevenson, M. (2006). The inner-nuclear-envelope protein emerin regulates HIV-1 infectivity. Nature 441, 641–645.
161 Shun, M. C., Daigle, J. E., Vandegraaff, N. and Engelman, A. (2007). Wild-type levels of human immunodeficiency virus type 1 infectivity in the absence of cellular emerin protein. J. Virol. 81, 166–172.
162 Holman, A. G. and Coffin, J. M. (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.
163 Wu, X., Li, Y., Crise, B., Burgess, S. M. and Munroe, D. J. (2005). Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. J. Virol. 79, 5211–5214.
164 Arkin, M. R. and Wells, J. A. (2004). Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. Nat. Rev. Drug Discov. 3, 301–317.
165 Wells, J. A. and McClendon, C. L. (2007). Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature 450, 1001–1009.
166 Vassilev, L. T., Vu, B. T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C. et al. (2004). In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science 303, 844–848.
167 He, M. M., Smith, A. S., Oslob, J. D., Flanagan, W. M., Braisted, A. C., Whitty, A., Cancilla, M. T., Wang, J., Lugovskoy, A. A., Yoburn, J. C. et al. (2005). Small-molecule inhibition of TNF-alpha. Science 310, 1022–1025.
168 Molteni, V., Greenwald, J., Rhodes, D., Hwang, Y., Kwiatkowski, W., Bushman, F. D., Siegel, J. S. and Choe, S. (2001). Identification of a small-molecule binding site at the dimer interface of the HIV integrase catalytic domain. Acta Crystallogr. D Biol. Crystallogr. 57, 536–544.
169 Al-Mawsawi, L. Q., Fikkert, V., Dayam, R., Witvrouw, M., Burke, T. R., Jr., Borchers, C. H. and Neamati, N. (2006). Discovery of a small-molecule HIV-1 integrase inhibitor-binding site. Proc. Natl. Acad. Sci. USA 103, 10080–10085.
170 Naldini, L., Bloemer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F. H., Verma, I. M. and Trono, D. (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science 272, 263–267.
171 Poeschla, E., Wong-Staal, F. and Looney, D. (1998). Efficient transduction of nondividing cells by feline immunodeficiency virus lentiviral vectors. Nat. Med. 4, 354–357.
172 Olsen, J. C. (1998). Gene transfer vectors derived from equine infectious anemia virus. Gene Ther. 5, 1481–1487.
173 Loewen, N. and Poeschla, E. M. (2005). Lentiviral vectors. Adv. Biochem. Eng. Biotechnol. 99, 169–191.
174 Donahue, R. E., Kessler, S. W., Bodine, D., McDonagh, K., Dunbar, C., Goodman, S., Agricola, B., Byrne, E., Raffeld, M., Moen, R. et al. (1992). Helper virus induced T cell lymphoma in nonhuman primates after retroviral mediated gene transfer. J. Exp. Med. 176, 1125–1135.
175 Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M. P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C. S., Pawliuk, R., Morillon, E. et al. (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science 302, 415–419.
176 Pike-Overzet, K., van der Burg, M., Wagemaker, G., van Dongen, J. J. and Staal, F. J. (2007). New insights and unresolved issues regarding insertional mutagenesis in X-linked SCID gene therapy. Mol. Ther. 15, 1910–1916.
177 Holmes-Son, M. L. and Chow, S. A. (2002). Correct integration mediated by integrase-LexA fusion proteins incorporated into HIV-1. Mol. Ther. 5, 360–370.

178 Tan, W., Dong, Z., Wilkinson, T. A., Barbas, C. F., 3rd and Chow, S. A. (2006). Human immunodeficiency virus type 1 incorporated with fusion proteins consisting of integrase and the designed polydactyl zinc finger protein E2C can bias integration of viral DNA into a predetermined chromosomal region in human cells. J. Virol. 80, 1939–1948.
179 Craigie, R. (1992). Hotspots and warm spots: integration specificity of retroelements. Trends Genet. 8, 187–190.
180 Zou, S., Ke, N., Kim, J. M. and Voytas, D. F. (1996). The Saccharomyces retrotransposon Ty5 integrates preferentially into regions of silent chromatin at the telomeres and mating loci. Genes Dev. 10, 634–645.
181 Xie, W., Gai, X., Zhu, Y., Zappulla, D. C., Sternglanz, R. and Voytas, D. F. (2001). Targeting of the yeast Ty5 retrotransposon to silent chromatin is mediated by interactions between integrase and Sir4p. Mol. Cell. Biol. 21, 6606–6614.
182 Dai, J., Xie, W., Brady, T. L., Gao, J. and Voytas, D. F. (2007). Phosphorylation regulates integration of the yeast Ty5 retrotransposon into heterochromatin. Mol. Cell 27, 289–299.
183 Zhu, Y., Dai, J., Fuerst, P. G. and Voytas, D. F. (2003). Controlling integration specificity of a yeast retrotransposon. Proc. Natl. Acad. Sci. USA 100, 5891–5895.
184 Sandmeyer, S. (2003). Integration by design. Proc. Natl. Acad. Sci. USA 100, 5586–5588.
185 McClintock, B. (1984). The significance of responses of the genome to challenge. Science 226, 792–801.
186 Deininger, P. L. and Batzer, M. A. (2002). Mammalian retroelements. Genome Res. 12, 1455–1465.
187 Han, K., Konkel, M. K., Xing, J., Wang, H., Lee, J., Meyer, T. J., Huang, C. T., Sandifer, E., Hebert, K., Barnes, E. W. et al. (2007). Mobile DNA in Old World monkeys: a glimpse through the rhesus macaque genome. Science 316, 238–240.
188 Heuer, T. S. and Brown, P. O. (1998). Photo-cross-linking studies suggest a model for the architecture of an active human immunodeficiency virus type 1 integrase-DNA complex. Biochemistry 37, 6667–6678.
189 Hindmarsh, P., Ridky, T., Reeves, R., Andrake, M., Skalka, A. M. and Leis, J. (1999). HMG protein family members stimulate human immunodeficiency virus type 1 and avian sarcoma virus concerted DNA integration in vitro. J. Virol. 73, 2994–3003.
190 Sinha, S., Pursley, M. H. and Grandgenett, D. P. (2002). Efficient concerted integration by recombinant human immunodeficiency virus type 1 integrase without cellular or viral cofactors. J. Virol. 76, 3105–3113.
191 Fletcher, T. M., 3rd, Soares, M. A., McPhearson, S., Hui, H., Wiskerchen, M., Muesing, M. A., Shaw, G. M., Leavitt, A. D., Boeke, J. D. and Hahn, B. H. (1997). Complementation of integrase function in HIV-1 virions. EMBO J. 16, 5123–5138.
192 Wu, X., Liu, H., Xiao, H., Conway, J. A., Hunter, E. and Kappes, J. C. (1997). Functional RT and IN incorporated into HIV-1 particles independently of the Gag/Pol precursor protein. EMBO J. 16, 5113–5122.
193 Patel, P. H. and Preston, B. D. (1994). Marked infidelity of human immunodeficiency virus type 1 reverse transcriptase at RNA and DNA template ends. Proc. Natl. Acad. Sci. USA 91, 549–553.
194 Ellison, V. and Brown, P. O. (1994). A stable complex between integrase and viral DNA ends mediates human immunodeficiency virus integration in vitro. Proc. Natl. Acad. Sci. USA 91, 7316–7320.
195 Vora, A. C. and Grandgenett, D. P. (1995). Assembly and catalytic properties of retrovirus integrase-DNA complexes capable of efficiently performing concerted integration. J. Virol. 69, 7483–7488.
196 Brin, E., Yi, J., Skalka, A. M. and Leis, J. (2000). Modeling the late steps in HIV-1 retroviral integrase-catalyzed DNA integration. J. Biol. Chem. 275, 39287–39295.
197 Yoder, K. E. and Bushman, F. D. (2000). Repair of gaps in retroviral DNA integration intermediates. J. Virol. 74, 11191–11200.
198 Chow, S. A., Vincent, K. A., Ellison, V. and Brown, P. O. (1992). Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. Science 255, 723–726.
199 Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R. and Davies, D. R. (1994). Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. Science 266, 1981–1986.
200 Eijkelenboom, A. P., Lutzke, R. A., Boelens, R., Plasterk, R. H., Kaptein, R. and Hard, K. (1995). The DNA-binding domain of HIV-1 integrase has an SH3-like fold. Nat. Struct. Biol. 2, 807–810.
201 Cai, M., Zheng, R., Caffrey, M., Craigie, R., Clore, G. M. and Gronenborn, A. M. (1997). Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. Nat. Struct. Biol. 4, 567–577.
202 Goldgur, Y., Dyda, F., Hickman, A. B., Jenkins, T. M., Craigie, R. and Davies, D. R. (1998). Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. Proc. Natl. Acad. Sci. USA 95, 9150–9154.
203 Chen, J. C., Krucinski, J., Miercke, L. J., Finer-Moore, J. S., Tang, A. H., Leavitt, A. D. and Stroud, R. M. (2000). Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. Proc. Natl. Acad. Sci. USA 97, 8233–8238.
204 Wang, J. Y., Ling, H., Yang, W. and Craigie, R. (2001). Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. EMBO J. 20, 7333–7343.
205 Jenkins, T. M., Hickman, A. B., Dyda, F., Ghirlando, R., Davies, D. R. and Craigie, R. (1995). Catalytic domain of human immunodeficiency virus type 1 integrase: identification of a soluble mutant by systematic replacement of hydrophobic residues. Proc. Natl. Acad. Sci. USA 92, 6057–6061.
206 Chiu, T. K. and Davies, D. R. (2007). Structure and function of HIV-1 integrase: an update. Fron. Med. Chem. 3, 1–20.