

Review

The Interaction Between Lentiviral Integrase and LEDGF: Structural and Functional Insights

Stephen Hare and Peter Cherepanov *

Division of Medicine, Imperial College London, St. Mary's Campus, Norfolk Place,
London, W2 1PG, UK; E-mail: s.hare@imperial.ac.uk

* Author to whom correspondence should be addressed; E-mail: p.cherepanov@imperial.ac.uk;
Tel.: +44 20 759 43 655; Fax: +44 20 759 43 906.

Received: 12 October 2009; in revised form: 28 October 2009 / Accepted: 6 November 2009 /
Published: 6 November 2009

Abstract: Since its initial description as an HIV-1 integrase (IN) interactor seven years ago, LEDGF has become one of the best-characterized host factors involved in viral replication. Results of intensive studies in several laboratories indicated that the protein serves as a targeting factor for the lentiviral DNA integration machinery, and accounts for the characteristic preference of *Lentivirus* to integrate within active transcription units. The IN-LEDGF interaction has been put forward as a promising target for antiretroviral drug development and as a potential tool to improve safety of lentiviral vectors for use in gene therapy. Additionally, as a natural ligand of lentiviral IN proteins, LEDGF has been successfully used in structural biology studies of retroviral DNA integration. This review focuses on the structural aspects of the IN-LEDGF interaction and their functional consequences.

Keywords: retrovirus; integrase; host factors; protein-protein interaction

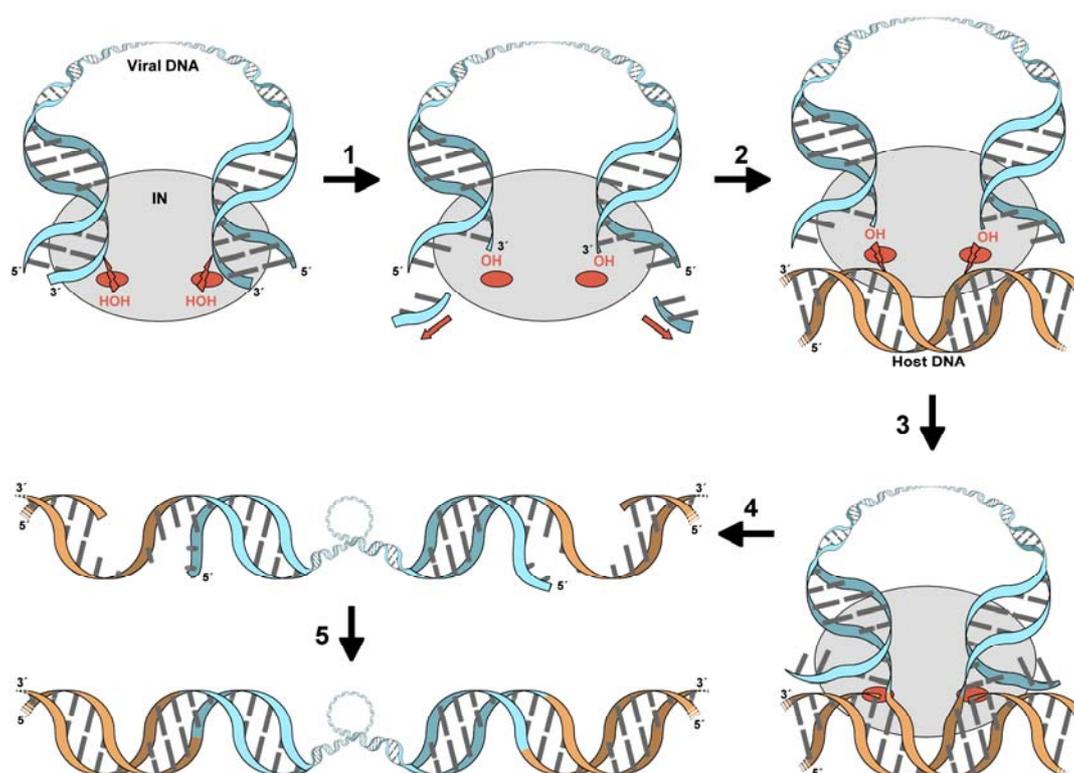
1. Introduction

For successful replication, HIV and other retroviruses depend on virally encoded IN enzymes to orchestrate insertion of their reverse transcribed genomes into host cell DNA (reviewed in [1,2]). The active site of retroviral IN catalyzes two distinct nucleophilic substitution (SN2) reactions during the integration process (Figure 1). Firstly, the 3' processing reaction takes place in the cytoplasm of the host cell in the context of a large nucleoprotein complex, termed the preintegration complex (PIC). In this step, IN removes two or three nucleotides from the 3' ends of the viral DNA, exposing the reactive 3' hydroxyl groups of the invariant CA dinucleotides. The second reaction, strand transfer, occurs in the nucleus and involves a pair of coordinated transesterification reactions that cut both strands of target DNA, simultaneously joining them to the 3'-ends of the viral DNA molecule. These concerted strand transfer events target a pair of phosphodiester bonds on the opposing strands of the target DNA, across its major groove. Consequently, following gap repair by host enzymes, the resulting provirus is flanked by short (4-6 bp, depending on the retroviral genus) duplications of the target DNA sequences.

Retroviral IN comprises three domains, an N-terminal domain (NTD) containing the invariant Zn²⁺-binding HHCC motif, a catalytic core domain (CCD) containing the active site, and a positively charged C-terminal domain (CTD) [3-6]. The IN active site contains three invariant acidic residues, forming the so-called D,DX₃₅E motif [6,7]. Based on analogy with distantly related polynucleotidyl transferases, most notably prokaryotic transposases and ribonuclease H, the IN catalytic triad carboxylates are expected to coordinate a pair of Mg²⁺ cations [8-12]. All three IN domains have been implicated in multimerization [13-16] and DNA binding [3,17-19]. In particular, the CCD of HIV-1 IN was shown to participate in sequence-specific recognition of viral DNA termini [20,21]. Solution NMR and X-ray crystallography have been used to determine the structures of the individual domains and two-domain fragments of retroviral INs [22-30] (reviewed in [31]). While each isolated domain is dimeric in solution, recombinant full-length retroviral INs exist in varying multimeric states and, with only rare exceptions [30], are highly prone to aggregation. The two active sites of the spherical CCD dimer are located on opposing faces, separated by ~40 Å [22]. Therefore, it follows that a tetramer of IN would be the minimal protomer to correctly position a pair of active sites for concerted strand transfer events targeting phosphodiester bonds across the major groove (~18 Å). A growing number of recent reports suggest that the tetrameric form of IN is indeed its functionally-relevant state [16,32-37].

Given a historical misnomer, lens epithelium derived growth factor (LEDGF) is a ubiquitous chromatin-associated protein with poorly characterized cellular functions. LEDGF is most notorious for its tight interaction with lentiviral INs and its role in HIV-1 replication. First implicated in virology as a cellular binding partner of ectopically expressed HIV-1 IN, LEDGF was also found to stimulate its enzymatic activity *in vitro* [38]. The protein-protein interaction was soon corroborated by two independent laboratories [39,40]. The functional aspects of this virus-host interaction were recently reviewed [41,42]; herein, we primarily focus on its structural details.

Figure 1. The retroviral DNA integration pathway. The PIC, formed following reverse transcription of the viral RNA genome in the host cell cytoplasm, contains viral DNA (light blue), IN (grey oval), along with other viral and host cell proteins (not shown). Within the PIC, the active sites of IN (red ovals) activate water molecules for nucleophilic attacks on the phosphodiester backbone close to the 3' ends of the viral DNA. This 3'-processing reaction (1) results in the removal of a di- or tri-nucleotide from both 3' ends of the viral DNA, exposing the reactive 3' hydroxyl groups attached to invariant CA dinucleotides. Following nuclear import, the PIC comes into contact with host chromosomal DNA (orange) (2). Whereupon, the IN active sites activate the hydroxyl groups at the 3' viral DNA ends to cut a pair of phosphodiester bonds in the opposing strands of chromosomal DNA, 4-6 bp apart (the exact separation depends on the retroviral genus, and equals 5 bp for lentiviruses) (3). The resulting intermediate (4) contains viral DNA joined at each 3' end to chromosomal DNA, flanked by short gaps and 5'-overhangs. The final DNA repair step (5) that joins the 5' viral DNA ends to the host DNA is presumably carried out by host proteins.



2. Domain organization of LEDGF

LEDGF belongs to the hepatoma derived growth factor (HDGF) related protein (HRP) family and is by far the most extensively studied IN binding partner. Predicted to be largely disordered, LEDGF contains two small structural domains [43]. One of these, the PWWP domain (LEDGF residues 1-91)

is present at the N-termini of all HRP family members [43-46]. Together with the nuclear localization signal (residues 148-156) and a pair of AT-hook motifs (residues 178-197), the PWWP domain is responsible for the tight association of LEDGF with chromatin [47-50]. The second structural domain is located within the C-terminal region of LEDGF (residues 347-429). Found responsible for the interaction with lentiviral INs, it was termed the integrase-binding domain (IBD) [43,48]. An alternative splice form of LEDGF, p52, lacks the IBD and concordantly fails to interact with HIV-1 IN or activate its enzymatic activity [43,51]. Of the five other human HRP family members, only HRP2 contains a conserved IBD within its C-terminal region, enabling it to interact with and stimulate the strand transfer activity of HIV-1 IN *in vitro* [43]. However, it is important to note that the affinity of HIV-1 IN for HRP2 is markedly lower than it is for LEDGF [43]; it is currently unclear whether HRP2 has any function in lentiviral replication.

The IBD was reported to mediate interactions between LEDGF and a number of cellular proteins: JPO2, a putative transcription factor [52,53]; Menin:MLL, a histone methyltransferase complex involved in transcriptional regulation and oncogenesis [54]; and PogZ, a protein of yet unknown function [55]. Intriguingly, JPO2 and Menin:MLL were shown to be tethered to chromatin by LEDGF [52,54]. Based on its domain organization, intracellular localization and characterized cellular binding partners, LEDGF is likely to play a role in regulation of gene expression, and/or as an adaptor protein tethering a plethora of cellular proteins to chromatin.

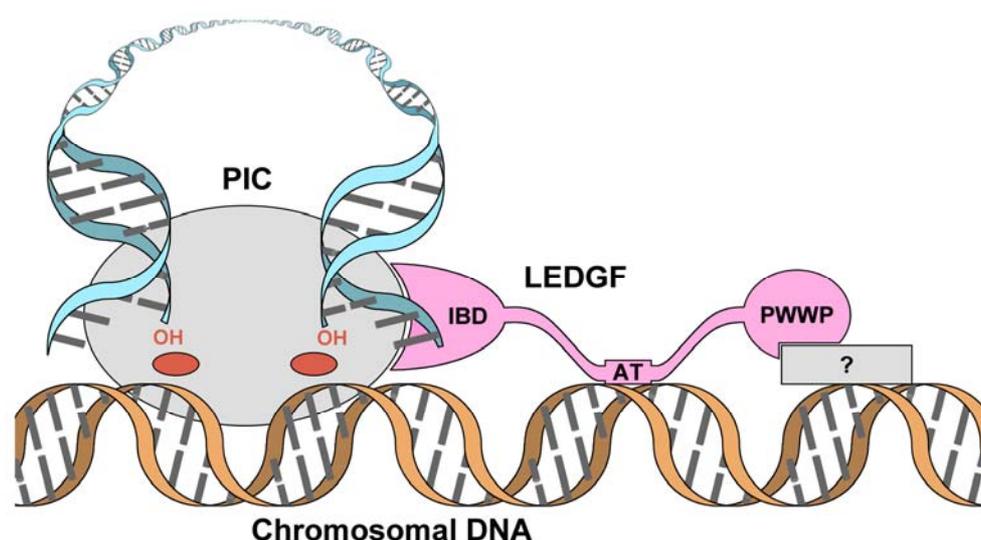
3. The role of LEDGF in lentiviral infection

The initial insight into possible role(s) of LEDGF in lentiviral replication came from early experiments that used ectopically-expressed IN. Unexpectedly, the chromatin binding activity of HIV-1 IN, initially thought to be intrinsic to this protein [56], was shown to be dependent on endogenous LEDGF [51,57]. Additionally, stability and nuclear accumulation of HIV-1 IN in human cells were drastically impaired by LEDGF depletion [51,57,58]. These results suggested that LEDGF might tether IN to host cell chromatin and that it may also be involved in its nuclear import and protection from proteasomal degradation. The early steps of viral infection are essentially single-molecule events and thus present substantial challenges to studies of ubiquitous host factors. Consequently, it required considerable efforts to generate cell lines with sufficient levels of LEDGF depletion or genetic knockout models to arrive at a consensus on the importance of this protein for lentiviral DNA integration. HIV-1 infection of cells depleted for or lacking LEDGF was substantially reduced relative to controls, due to a specific block at the integration step [59-61]. The infectivity could be restored by re-expression of full-length LEDGF, while over-expression of the isolated IBD lead to even more drastic suppression of HIV-1 integration [60,62]. So far no evidence has emerged to support a role of LEDGF in nuclear import or protection of the lentiviral PIC, although HIV-1 and feline immunodeficiency virus (FIV) PICs can be immunoprecipitated with anti-LEDGF antibodies [57].

The interaction with LEDGF is exclusive to INs from the retroviral genus of *Lentivirus*, while those from the members of *Alpha-*, *Beta-*, *Gamma-*, *Deltaretrovirus*, and *Spumavirus* genera do not bind LEDGF [57,63,64]. The characteristic features of lentiviruses include their marked propensity to integrate within active transcription units of the host cell genome, and their bias against insertion into promoters and CpG islands [65-70] (reviewed in [71]). Using LEDGF knockdown and knockout models, these properties were shown to depend on the IN-LEDGF interaction [61,72,73]. Intriguingly, with respect to genomic features, the integration site profiles of HIV-1 in the absence of LEDGF are reminiscent of non-lentiviral genera [61,73].

The current model for LEDGF function in lentiviral integration advocates that the N-terminal region of LEDGF interacts with the host chromatin at active transcription units and, via the connected IBD, brings the PIC into proximity with this preferred region of the genome for integration while stimulating IN strand transfer activity (Figure 2). Presumably, integration into transcriptionally-active genomic loci improves the efficiency of lentiviral gene expression, providing a sufficient evolutionary advantage [74].

Figure 2. The role of LEDGF in lentiviral biology. Representations of IN and viral and host DNA are conserved from Figure 1. LEDGF (pink) interacts with the PIC via its C-terminal IBD, with host DNA via its AT-hooks and with an unidentified component of the chromatin (grey rectangle) via its N-terminal PWWP domain, tethering the PIC to select loci of host cell chromatin.



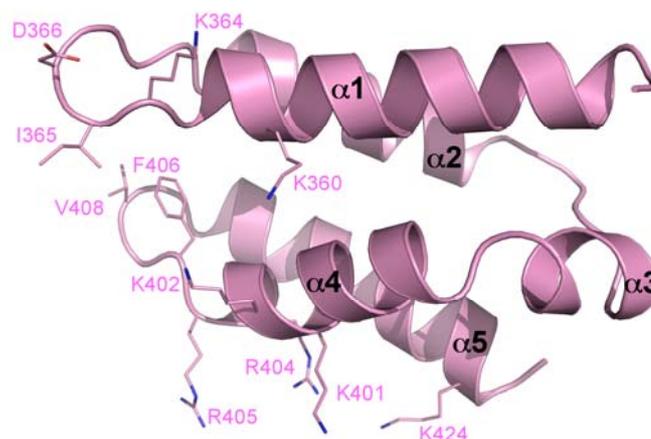
This model suggests an exciting possibility for engineering artificial LEDGF-like molecules for targeting integration of lentiviral gene therapy vectors. The concern over using integrating vectors is a tangible danger of insertional mutagenesis. Indeed, several unfortunate cases of leukemia associated with gammaretroviral vector insertions in the vicinity of a proto oncogene have been observed in two gene therapy trials [75,76]. An artificial targeting factor linking the IBD of LEDGF to a DNA binding

domain, able to selectively associate with a safe genomic locus, could in principle be used to direct lentiviral integration. The potential for this was demonstrated *in vitro* using a fusion construct comprising the IBD of LEDGF and the DNA binding domain of the λ phage repressor. The chimeric construct targeted a modest fraction of IN strand transfer events into the vicinity of λ repressor binding sites [77]. More recently, further validation of this approach was provided by Meehan *et al.* [78] who fused heterologous chromatin binding domains to the IBD and showed that the chimeras were able to rescue both IN chromatin tethering and HIV-1 integration under conditions of endogenous LEDGF depletion.

4. The primary IN:LEDGF interface

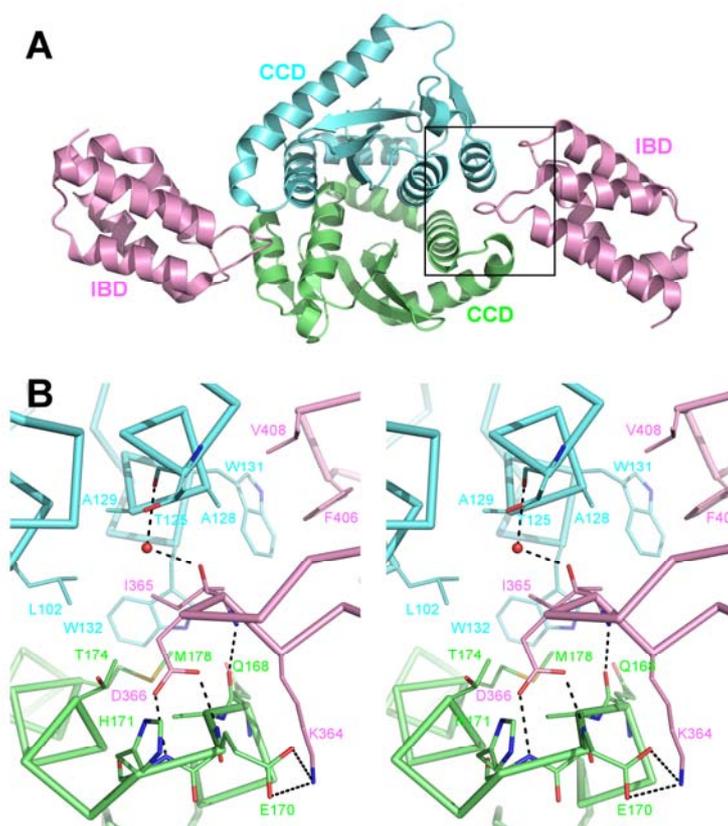
The solution structure of isolated LEDGF_{IBD} was determined by NMR spectroscopy (PDB ID 1z9e), revealing a bundle of four long α -helices ($\alpha 1$, $\alpha 2$, $\alpha 4$, and $\alpha 5$) (Figure 3) [79]. At one end of the bundle, hairpin turns connect $\alpha 1$ to $\alpha 2$ and $\alpha 4$ to $\alpha 5$. Meanwhile, at the other end of the structure, a shorter α -helix ($\alpha 3$) links $\alpha 2$ and $\alpha 4$. Several exposed hydrophobic side chains are located on the two adjacent hairpin turns and alanine scanning mutagenesis revealed three of these residues (Ile-365, Phe-406 and to a lesser extent Val-408) to be involved in the interaction with HIV-1 IN. The adjacent acidic residue Asp-366 was shown by mutation to Asn (D366N) to be essential for both the interaction with IN and for stimulating its enzymatic activity [79]. This loss-of-function mutation has since been widely used in studies of the roles of LEDGF in HIV-1 replication. Fortuitously, mutations of Asp-366 do not seem to affect the interaction of LEDGF with its known cellular binding partners [52,53,55].

Figure 3. The solution structure of the LEDGF IBD. The helical bundle is shown in cartoon representation, with individual helices labeled (PDB ID 1z9e). Side chains contributing to the hydrophobic area at the left side of the helical bundle as drawn and the positive face on the underside are shown as sticks and labeled.



The primary site of LEDGF interaction with IN is the CCD and the details of this interface were revealed in a co-crystal structure containing HIV-1 IN_{CCD} and LEDGF_{IBD} constructs (Figure 4) (PDB ID 2b4j) [80]. The loop connecting $\alpha 1$ and $\alpha 2$ of the IBD inserts into the CCD dimer interface and interacts with both monomers of the CCD dimer. Here, the side chain of Asp-366 makes dual hydrogen bonds with the main chain amino groups of IN chain A residues Glu-170 and His-171, located on the loop connecting CCD alpha helices $\alpha 4$ and $\alpha 5$ (known as $\alpha 4/5$ connector). The hydrophobic side chain of LEDGF Ile-365 makes Van der Waals interactions with the side chains of IN chain A residue Met-178 and chain B residues Leu-102, Ala-128, Ala-129, Trp-131 and Trp-132 (Figure 4B). The second IBD loop at this location contains the hydrophobic side chains of Phe-406 and Val-408, which pack against a hydrophobic patch on IN chain B formed by Ala-128 and Trp-131 (Figure 4B).

Figure 4. The primary IN:LEDGF interface. **A.** Cartoon representation of the co-crystal structure of the IN_{CCD}:LEDGF_{IBD} complex (PDB ID 2b4j). IN chains are colored green (chain A) and cyan (chain B) and a pair of LEDGF chains interacting at either end of the IN CCD dimer are pink. **B.** Stereo close-up view of the region enclosed by a black rectangle in **A**, showing details of the CCD:IBD interface. The protein backbone and side chains, shown in ribbon and stick representations, respectively, are colored by atom. Side chains of residues involved in interactions are shown, as well as a water molecule coordinated between main chain carbonyls of IN Thr-125 and LEDGF Ile-365.



As well as identifying LEDGF residues crucial for the interaction with IN, mutational analyses highlighted HIV-1 IN Val-165, Arg-166, Gln-168, Leu-172, and Lys-173 as being important for the interaction [39,79,81]. Interestingly, these residues do not form direct interactions with LEDGF. Mutations at these positions are thought to affect conformation of the α 4/5 connector, disrupting the surface complementarity between the CCD and the IBD [80]. Side chains of HIV-1 IN residues Ala-128, Ala-129, Trp-131 and Trp-132 directly interact with the IBD and were also found to have varying effects on the strength of the HIV-1 IN-LEDGF interaction [80-82]. Mutations within or close to the α 4/5 connector typically result in lethal viral phenotypes, initially attributed to a defect in nuclear import (V165A and R166A, [83]) or their inability to interact with LEDGF (Q168A, [39]). However, follow up analyses revealed a more complex picture with the mutant viruses displaying additional defects in reverse transcription [82,84,85], which cannot be explained merely by disruption of the virus-host interaction [60,61].

Intuitively, this primary IN-LEDGF interaction is of interest for the development of a novel class of antiretrovirals. The LEDGF binding cavity on the IN CCD dimer interface has the potential for both hydrogen bonding and hydrophobic interactions, suggesting a small molecule could specifically bind there with high affinity. A number of ligands have been observed binding there in crystal structures of IN fragments ([86]; S.H. and P.C. unpublished). Chemical library screening approaches and *in silico* pharmacophore modeling have been used to identify lead inhibitors of HIV-1 IN-LEDGF interaction [87-89].

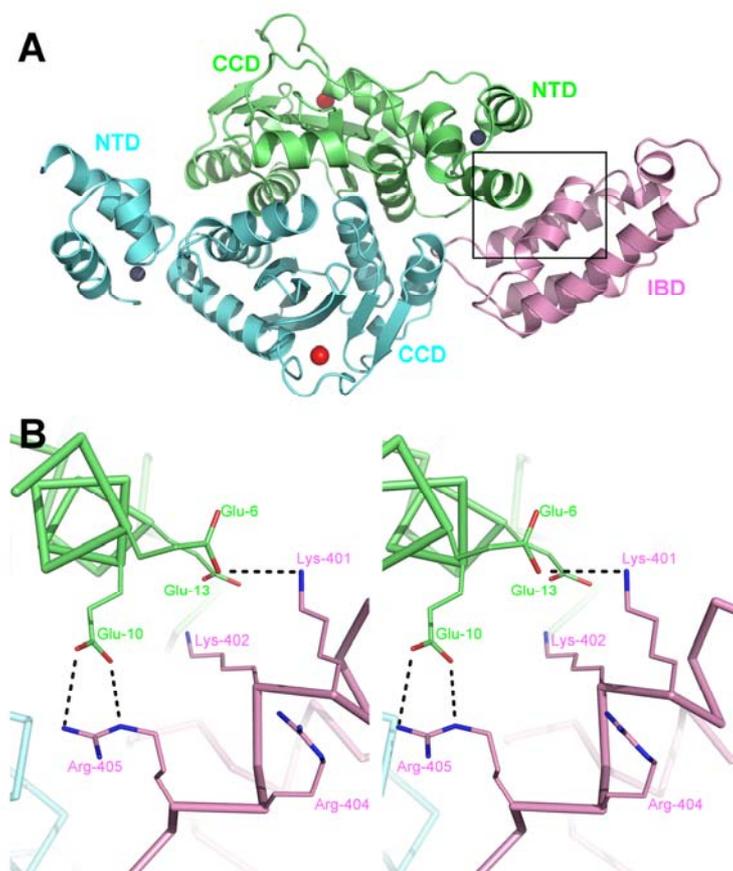
Recent structures of LEDGF_{IBD} bound to INs of HIV-2 and maedi-visna virus (MVV) have revealed the conservation of the co-factor binding at this same pocket [16,90]. Retroviruses are notorious for their high evolutionary rate; hence conservation of this interaction throughout the entire genus indicates its vital importance for lentiviral biology. HIV-2 and MVV INs share approximately 60% and 30% amino acid sequence identity with HIV-1 IN, respectively. While HIV-2 IN binds the IBD in a manner nearly identical to that of HIV-1 IN, the IBD shows a twist of $\sim 34^\circ$ when binding to the more divergent MVV IN. This is due to a slight change in the size and shape of the MVV CCD pocket, a result of local amino acid changes [16]. More drastic differences in the local structures of the analogous pockets at the CCD dimer interfaces of non-lentiviral INs (in particular the conformations of their α 4/5 connectors) explain the notable lentiviral specificity of LEDGF [30,80].

5. The IN NTD and the high affinity IN-LEDGF interaction

Although the CCD of HIV-1 IN is essential and minimally sufficient for the interaction with LEDGF, the NTD is required for high-affinity binding [51]. Thus, the HIV-1 IN H12N mutant, with disrupted NTD structure, was unable to interact with LEDGF in pull-down assays and required over-expression of LEDGF to associate with condensed chromatin in mitotic cells [51]. The crystal structure of LEDGF_{IBD} in complex with a two-domain construct of HIV-2 IN containing its NTD and CCD revealed the details of this interaction (PDB ID 3f9k) (Figure 5) [90]. In this structure, while the

previously identified IBD:CCD interaction is preserved, the positive face of the IBD makes additional contacts with a negatively charged surface of the NTD. Specifically, IBD residues Lys-401, Lys-402, Arg-404, and Arg-405 oppose HIV-2 IN Glu-6, Glu-10, and Glu-13 (Figure 5B). Pull-down and yeast two-hybrid experiments using mutants targeting this interface confirmed its importance for the protein-protein interaction. Additionally, it was observed that reversing the charges on both sides, *i.e.* making Lys/Arg to Glu mutations on LEDGF and Glu to Lys substitutions on IN, recuperated the interaction. *In vitro* concerted integration assays and single round HIV-1 infection experiments using reverse charge mutants showed the cofactor role of LEDGF also depends on this interface. These activity and infection assays also indicated that the charge-charge interface could be reversed, with mutant INs requiring LEDGF containing complementary mutations for activity. However, the reversed mutant combinations were significantly less active/infectious than the wild type [90].

Figure 5. The interaction between the IBD and IN NTD. **A.** Cartoon representation of the HIV-2 IN_{NTD+CCD} and LEDGF_{IBD} co-crystal structure (PDB ID 3f9k) with the IN dimer colored green (chain A) and cyan (chain B) and LEDGF colored pink. Red spheres represent magnesium ions in the IN active sites and dark grey spheres represent Zn²⁺ ions coordinated by the HHCC motif of the NTDs. **B.** Details of the NTD:IBD interface, showing the area enclosed by a black rectangle in **A**. The charge-charge interactions are shown as black dashed lines between the stick representations of the side chains involved.



6. A role for LEDGF in IN tetramerization

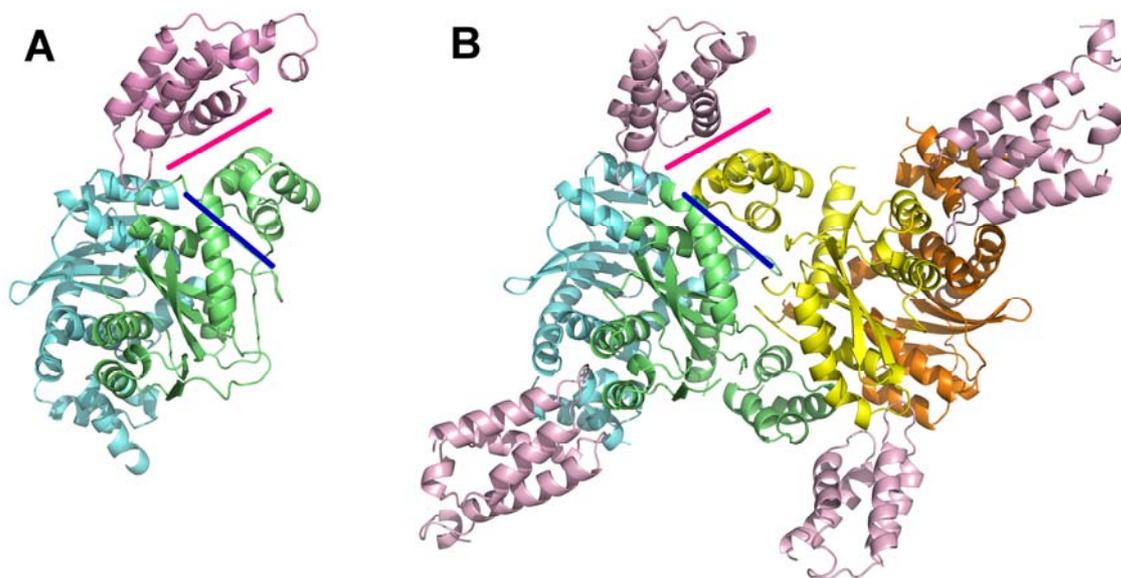
Structural information detailing protein-protein interfaces involved in PIC assembly and the organization of the IN active site is invaluable for the development of antiretroviral drugs. The odds of obtaining useful crystals can often be improved by using a natural or even artificial ligand for the target protein. Typically, complexes of lentiviral INs with full-length LEDGF or LEDGF_{IBD} are considerably more soluble than their free forms (unpublished observations). Furthermore, LEDGF can be expected to stabilize the functionally-relevant conformation(s) of divergent lentiviral INs. Crystallization and structure determination of two such complexes containing a two-domain fragments of HIV-2 and MVV IN_{NTD+CCD} helped to elucidate the mechanism of functional IN tetramerization [16,90]. Notably, while crystallization of unliganded HIV-1 IN fragments required presence of various solubilizing point mutations [22,25,26,91], the analogous changes were not necessary to obtain crystals of HIV-2 or MVV IN_{NTD+CCD} complexed with LEDGF_{IBD} [16,90].

The co-crystal structures of MVV IN_{NTD+CCD} with LEDGF_{IBD} (PDB IDs 3hpg and 3hph) revealed a series of IN tetrameric arrangements [16], while a similar HIV-2 - derived complex (PDB ID 3f9k) was captured in a dimeric form [90]. Importantly, the dimer-of-dimers tetrameric architecture observed in the crystals of the MVV complex is very similar to that reported earlier for the unliganded HIV-1 IN_{NTD+CCD} construct (PDB ID 1k6y) [26], despite less than 30% amino acid sequence identity between MVV and HIV-1 INs. The IN dimer-dimer interface is stabilized by an NTD of one dimer interacting with a CCD of the opposing dimer. The analogous NTD:CCD interface is also observed in the crystals of dimeric HIV-2 IN_{NTD+CCD}:LEDGF_{IBD}, although in this case the NTD interacts with its own CCD dimer [90]. These observations indicated the mechanism for IN tetramerization that involves swapping of a pair of NTDs between interacting dimers. Disruption of the NTD:CCD interface abrogates tetramerization and dramatically reduces the enzymatic activities of HIV-1 IN [16,36,92]. The long and flexible NTD-CCD linkers allow striking flexibility of the dimer-dimer IN interface within the two-domain IN constructs [16]. Based on the tetrameric IN_{NTD+CCD} structures, the LEDGF binding platform includes the CCDs from one IN dimer and an NTD from another (Figure 6). Concordantly, LEDGF binding dramatically stimulates tetramerization of HIV-1 IN *in vitro* [16,36]. Furthermore, the co-factor can partially rescue multimerization of HIV-1 IN mutants with defects in the NTD:CCD interface [16]. Intriguingly, Hayouka *et al.* [93], reported that peptides derived from the LEDGF IBD loops (residues 361-370 and 402-411) promoted HIV-1 IN multimerization *in vitro*. This observation suggests that engagement of the primary IBD:CCD interface alone, possibly through allosteric effects on the crucial CCD:NTD interface, could stabilize the tetramer.

The stoichiometry of the biologically-relevant IN:LEDGF complex is yet to be resolved. The crystal structures of the MVV IN_{NTD+CCD}:LEDGF_{IBD} complexes contained four LEDGF chains associated with each IN tetramer, forming both primary and secondary interfaces with IN [16]. This same 1:1 stoichiometry was observed earlier in crystals of the minimal HIV-1 IN_{CCD}:LEDGF_{IBD}

complex [80]. However, HIV-2 IN_{NTD+CCD} construct was co-crystallized with a single LEDGF chain per IN dimer [90], which could also be attributable to crystal packing forces. Therefore, the available structural data therefore support either 1:1 or 2:1 IN:LEDGF stoichiometry. Mass spectrometry analyses of full-length HIV-1 IN:LEDGF complex produced by co-expression in bacteria [37] and size exclusion chromatography of IN:LEDGF_{IBD} complexes assembled *in vitro* [36] argued for the latter stoichiometry, although more data are required to ascertain its biological relevance. While it is reasonable to speculate that the functional IN tetramer might possess as many as four high affinity binding sites for LEDGF, as the main proposed role of the host factor is tethering the PIC to chromatin, a single LEDGF molecule might well suffice.

Figure 6. The structural basis for IN tetramerization. **A.** The crystal structure of the dimeric HIV-2 IN_{NTD+CCD} construct in complex with LEDGF_{IBD} (PDB ID 3f9k). The intramolecular NTD:CCD interface is shown as a blue line and the IBD:NTD interface as a pink line. **B.** The crystal structure of tetrameric MVV IN_{NTD+CCD} in complex with LEDGF_{IBD} (PDB ID 3hph). The NTD-CCD interface (blue line) is conserved from the dimeric HIV-2 structure, although, as a consequence of NTD swapping, in the tetrameric structure this interface is intermolecular. Charge-charge interactions between the IN NTD and LEDGF IBD (pink line) act to stabilize the tetramer thereby explaining the observed effect of LEDGF on IN multimerization *in vitro* [16,36].



7. Effects of LEDGF on enzymatic activities of lentiviral IN

Retroviral DNA integration can be re-constituted *in vitro* using recombinant IN, viral DNA mimics (commonly referred to as donor DNA) and target DNA [94,95]. Typically, such reactions lead to

formation of abundant Y-shaped strand transfer products, resulting from the insertion of a single donor DNA end into one strand of a target DNA molecule. Under optimized conditions, it is possible to observe biologically-relevant concerted strand transfer products, arising from coordinated insertions of pairs of donor DNA molecules. The IN tetramer has been implicated as the basic catalytic unit for both 3'-processing and strand transfer reactions [16,32,96]. While residual 3'-processing and half-site integration can be carried out by IN mutants impaired for tetramerization, the tetramer is essential for concerted integration [16]. Accordingly, HIV-1 IN tetramers were observed within *in vitro* assembled nucleoprotein complexes competent for concerted integration [34].

In accordance with its role in lentiviral IN tetramerization, LEDGF robustly stimulates its 3'-processing and strand transfer activities *in vitro* [36,38,64,90,97-99]. Nevertheless, ambiguity persisted over the effect of LEDGF on concerted HIV-1 integration. Cryptically, depending on reaction conditions such as the order of addition and relative input ratios of HIV-1 IN and LEDGF, the type of DNA substrates, and the reaction buffer components, the co-factor can both stimulate and inhibit concerted HIV-1 integration *in vitro* [36,90,98,99]. Thus, Raghavendra and Engelman [99] observed that while LEDGF promoted overall levels of strand transfer activity of HIV-1 IN, it specifically inhibited formation of concerted integration products. Pandey *et al.* [98] went further and showed that, under similar conditions, while using excess LEDGF inhibited concerted integration, a modest (two- to three- fold) stimulation was observed when using equimolar or lower ratios of LEDGF to HIV-1 IN, results that were later replicated by an independent group [36]. Pandey *et al.* [98] also showed that, in order to observe the stimulative effect of LEDGF on HIV-1 concerted integration, it is important to add donor DNA substrate before LEDGF, suggesting that the host factor might prevent IN from forming a productive complex with donor DNA [98]. As LEDGF binding locks HIV-1 IN into a tetrameric state [36,93], perhaps there is insufficient flexibility to subsequently engage a pair of viral DNA ends. During infection this is unlikely to be an issue, as PIC assembly occurs in the cytoplasm, and LEDGF may not be encountered prior to nuclear entry. Consistent with this view, PIC assembly proceeds normally in LEDGF-null cells, and therefore does not depend on LEDGF [61]. More recently it was shown that in the presence of LEDGF and higher inputs of donor DNA, HIV-1 IN displays very robust concerted strand transfer activity, albeit significant levels of half-site strand transfer persist [90]. Under these conditions, the NTD:IBD interface was specifically important for stimulation of the concerted strand transfer activity of HIV-1 IN. It is important to note that the relative ratio of concerted to half-site strand transfer products greatly depends on the viral source of the IN used. For example, while HIV-1 IN even under most optimized conditions generates copious amounts of half-site products, the INs from equine infectious anemia virus (EIAV) and prototype foamy virus (PFV) promote predominantly concerted integration *in vitro* [30,64]. The reasons for these differences are currently unknown.

Although the IBD, the only region of LEDGF known to directly interact with IN, is sufficient to stabilize IN tetramers and to stimulate its 3'-processing activity [36], it is not sufficient to bolster

strand transfer [43]. In fact, isolated LEDGF_{IBD} can competitively inhibit LEDGF-dependent strand transfer activity of HIV-1 IN [43]. Concordantly, over-expression of GFP-LEDGF_{IBD} fusions can potently suppress HIV-1 integration in human cells [60,62]. Using naked DNA targets, Turlure *et al.* [49] observed that a fragment spanning residues 226-530 of LEDGF retained approximately 50% of full-length LEDGF strand transfer stimulative activity. Stimulation of EIAV IN strand transfer activity required the presence of the DNA-binding AT hook region in the LEDGF construct (P.C., unpublished observations). Similar experiments using reconstituted polynucleosomes as target DNA revealed that the N-terminal PWWP domain is required for stimulation of HIV-1 integration into chromatinized DNA [100]. Furthermore, combining the IBD of LEDGF with heterologous chromatin binding domains, Meehan *et al.* [78] were able to create functional co-factors that rescued HIV-1 integration in LEDGF deficient cells. Thus, it seems likely that the effect of LEDGF on lentiviral IN activities is a combination of (i) enhancement of biologically-relevant multimerization, and, specifically pertaining to strand transfer, (ii) tethering of the PIC to target DNA. In addition, (iii) allosteric regulation of IN active site function by LEDGF cannot be ruled out at this point.

8. Concluding remarks, remaining questions and perspectives

Since its identification as an HIV-1 IN binding partner seven years ago, LEDGF and its role in retroviral replication has been subject to intense investigations. Despite initial controversy, the cellular protein has been validated as a *bona fide* co-factor of HIV-1 (and generally lentiviral) DNA integration. Recent studies revealed much about the functional and structural aspects of the IN-LEDGF interaction. Using RNA interference mediated knockdown and genetic knockout model systems several independent groups have demonstrated that the protein is important albeit not absolutely essential for HIV-1 integration, playing a major role in directing the virus into active transcription units of the host cell genome [59-61,72,73,101]. The most notable gap in our knowledge of LEDGF is its natural function in the cell. Even the intrinsic distribution of LEDGF along cellular chromatin is yet to be reported. The discovery of several LEDGF binding partners that in some but not all cases use LEDGF for chromatin tethering has done little to further our understanding of its native functions.

HIV DNA integration is an important target for antiretroviral drug discovery, and inhibition of the IN-LEDGF interaction is widely expected to produce a novel class of drugs [102,103]. Additionally, as a natural ligand of lentiviral INs, LEDGF has already served as a useful tool in structural biology of retroviral DNA integration [16,90]. We hope that using LEDGF-derived constructs will eventually allow crystallization and structure determination of the functional lentiviral PIC, which in turn would greatly stimulate the development of integrase inhibitors.

The IBD-NTD interface may also prove useful for exploitation in the design of safer gene therapy vectors. The recent success of creating artificial HIV-1 co-factors by linking alternative chromatin binding modules to the LEDGF IBD strongly advocates this possibility [78]. An extension of this work hypothesizes novel IBD fusion protein could target vector integration to safe genetic loci (see section

3, above). However, an obvious impediment to this strategy is the presence of endogenous LEDGF in target cells. This could potentially be overcome by using a mutant IN, unable to recognize endogenous cellular LEDGF, and a complementary IBD variant. Although the reverse-charge mutations provided a proof of principle that such a system could be designed [90], a more robust IN:IBD mutant pair needs to be developed for it to be applicable under conditions of endogenous LEDGF expression.

Recent genome-wide screening experiments have uncovered scores of cellular proteins required for HIV-1 infectivity [104-108]. Of these, transportin-SR2, shown to also bind IN *in vitro* [109], is already receiving attention of many laboratories, although a direct link between its interaction with IN and HIV-1 infection is yet to be established. The next few years should yield a wealth of functional and structural information about LEDGF, transportin-SR2, as well as novel potential IN co-factors [110], which will hopefully open new therapeutic possibilities.

Acknowledgements

We thank Alan Engelman for critical reading of the manuscript. Our work is funded by the UK Medical Research Council grant G0600009 (to P.C.).

References

1. Craigie, R. Retroviral DNA Integration. In *Mobile DNA II*, Craig, N.L.; Craigie, R.; Gellert, M.; Lambowitz, A.M. Eds. ASM Press: Washington DC, 2002; pp 613-630.
2. Lewinski, M.K.; Bushman, F.D. Retroviral DNA integration--mechanism and consequences. *Adv. Genet.* **2005**, *55*, 147-181.
3. Khan, E.; Mack, J.P.; Katz, R.A.; Kulkosky, J.; Skalka, A.M. Retroviral integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Res.* **1991**, *19*, 851-860.
4. Bushman, F.D.; Engelman, A.; Palmer, I.; Wingfield, P.; Craigie, R. Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 3428-3432.
5. Engelman, A.; Bushman, F.D.; Craigie, R. Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. *EMBO J.* **1993**, *12*, 3269-3275.
6. Engelman, A.; Craigie, R. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function *in vitro*. *J. Virol.* **1992**, *66*, 6361-6369.
7. Kulkosky, J.; Jones, K.S.; Katz, R.A.; Mack, J.P.; Skalka, A.M. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. *Mol. Cell. Biol.* **1992**, *12*, 2331-2338.
8. Wlodawer, A. Crystal structures of catalytic core domains of retroviral integrases and role of divalent cations in enzymatic activity. *Adv. Virus Res.* **1999**, *52*, 335-350.
9. Lovell, S.; Goryshin, I.Y.; Reznikoff, W.R.; Rayment, I. Two-metal active site binding of a Tn5 transposase synaptic complex. *Nat. Struct. Biol.* **2002**, *9*, 278-281.

10. Nowotny, M.; Gaidamakov, S.A.; Crouch, R.J.; Yang, W. Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell* **2005**, *121*, 1005-1016.
11. Yang, W.; Lee, J.Y.; Nowotny, M. Making and breaking nucleic acids: two-Mg²⁺-ion catalysis and substrate specificity. *Mol. Cell* **2006**, *22*, 5-13.
12. Diamond, T.L.; Bushman, F.D. Role of metal ions in catalysis by HIV integrase analyzed using a quantitative PCR disintegration assay. *Nucleic Acids Res.* **2006**, *34*, 6116-6125.
13. Hickman, A.B.; Palmer, I.; Engelman, A.; Craigie, R.; Wingfield, P. Biophysical and enzymatic properties of the catalytic domain of HIV-1 integrase. *J. Biol. Chem.* **1994**, *269*, 29279-29287.
14. Jenkins, T.M.; Engelman, A.; Ghirlando, R.; Craigie, R. A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. *J. Biol. Chem.* **1996**, *271*, 7712-7718.
15. Zheng, R.; Jenkins, T.M.; Craigie, R. Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 13659-13664.
16. Hare, S.; Di Nunzio, F.; Labeja, A.; Wang, J.; Engelman, A.; Cherepanov, P. Structural basis for functional tetramerization of lentiviral integrase. *PLoS Pathog.* **2009**, *5*, e1000515.
17. Engelman, A.; Hickman, A.B.; Craigie, R. The core and carboxyl-terminal domains of the integrase protein of human immunodeficiency virus type 1 each contribute to nonspecific DNA binding. *J. Virol.* **1994**, *68*, 5911-5917.
18. van den Ent, F.M.; Vos, A.; Plasterk, R.H. Dissecting the role of the N-terminal domain of human immunodeficiency virus integrase by trans-complementation analysis. *J. Virol.* **1999**, *73*, 3176-3183.
19. Zhao, Z.; McKee, C.J.; Kessl, J.J.; Santos, W.L.; Daigle, J.E.; Engelman, A.; Verdine, G.; Kvaratskhelia, M. Subunit-specific protein footprinting reveals significant structural rearrangements and a role for N-terminal Lys-14 of HIV-1 Integrase during viral DNA binding. *J. Biol. Chem.* **2008**, *283*, 5632-5641.
20. Jenkins, T.M.; Esposito, D.; Engelman, A.; Craigie, R. Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking. *EMBO J.* **1997**, *16*, 6849-6859.
21. Esposito, D.; Craigie, R. Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein-DNA interaction. *EMBO J.* **1998**, *17*, 5832-5843.
22. Dyda, F.; Hickman, A.B.; Jenkins, T.M.; Engelman, A.; Craigie, R.; Davies, D.R. Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* **1994**, *266*, 1981-1986.
23. Eijkelenboom, A.P.; Lutzke, R.A.; Boelens, R.; Plasterk, R.H.; Kaptein, R.; Hard, K. The DNA-binding domain of HIV-1 integrase has an SH3-like fold. *Nat. Struct. Biol.* **1995**, *2*, 807-810.
24. Cai, M.; Zheng, R.; Caffrey, M.; Craigie, R.; Clore, G.M.; Gronenborn, A.M. Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. *Nat. Struct. Biol.* **1997**, *4*, 567-577.
25. Chen, J.C.; Krucinski, J.; Miercke, L.J.; Finer-Moore, J.S.; Tang, A.H.; Leavitt, A.D.; Stroud, R.M. Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 8233-8238.

26. Wang, J.Y.; Ling, H.; Yang, W.; Craigie, R. Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. *EMBO J.* **2001**, *20*, 7333-7343.
27. Bujacz, G.; Jaskolski, M.; Alexandratos, J.; Wlodawer, A.; Merkel, G.; Katz, R.A.; Skalka, A.M. The catalytic domain of avian sarcoma virus integrase: conformation of the active-site residues in the presence of divalent cations. *Structure* **1996**, *4*, 89-96.
28. Yang, Z.N.; Mueser, T.C.; Bushman, F.D.; Hyde, C.C. Crystal structure of an active two-domain derivative of Rous sarcoma virus integrase. *J. Mol. Biol.* **2000**, *296*, 535-548.
29. Chen, Z.; Yan, Y.; Munshi, S.; Zugay-Murphy, J.; Xu, B.; Witmer, M.; Felock, P.; Wolfe, A.; Sardana, V.; Emini, E.A.; Hazuda, D.; Kuo, L.C. X-ray structure of simian immunodeficiency virus integrase containing the core and C-terminal domain (residues 50-293)--an initial glance of the viral DNA binding platform. *J. Mol. Biol.* **2000**, *296*, 521-533.
30. Valkov, E.; Gupta, S.S.; Hare, S.; Helander, A.; Roversi, P.; McClure, M.; Cherepanov, P. Functional and structural characterization of the integrase from the prototype foamy virus. *Nucleic Acids Res.* **2009**, *37*, 243-255.
31. Jaskolski, M.; Alexandratos, J.N.; Bujacz, G.; Wlodawer, A. Piecing together the structure of retroviral integrase, an important target in AIDS therapy. *FEBS J.* **2009**, *276*, 2926-2946.
32. Bao, K.K.; Wang, H.; Miller, J.K.; Erie, D.A.; Skalka, A.M.; Wong, I. Functional oligomeric state of avian sarcoma virus integrase. *J. Biol. Chem.* **2003**, *278*, 1323-1327.
33. Faure, A.; Calmels, C.; Desjobert, C.; Castroviejo, M.; Caumont-Sarcos, A.; Tarrago-Litvak, L.; Litvak, S.; Parissi, V. HIV-1 integrase crosslinked oligomers are active *in vitro*. *Nucleic Acids Res.* **2005**, *33*, 977-986.
34. Li, M.; Mizuuchi, M.; Burke, T.R. Jr.; Craigie, R. Retroviral DNA integration: reaction pathway and critical intermediates. *EMBO J.* **2006**, *25*, 1295-1304.
35. Ren, G.; Gao, K.; Bushman, F.D.; Yeager, M. Single-particle image reconstruction of a tetramer of HIV integrase bound to DNA. *J. Mol. Biol.* **2007**, *366*, 286-294.
36. McKee, C.J.; Kessl, J.J.; Shkriabai, N.; Dar, M.J.; Engelman, A.; Kvaratskhelia, M. Dynamic modulation of HIV-1 integrase structure and function by cellular lens epithelium-derived growth factor (LEDGF) protein. *J. Biol. Chem.* **2008**, *283*, 31802-31812.
37. Michel, F.; Crucifix, C.; Granger, F.; Eiler, S.; Mouscadet, J.F.; Korolev, S.; Agapkina, J.; Ziganshin, R.; Gottikh, M.; Nazabal, A.; Emiliani, S.; Benarous, R.; Moras, D.; Schultz, P.; Ruff, M. Structural basis for HIV-1 DNA integration in the human genome, role of the LEDGF/P75 cofactor. *EMBO J.* **2009**, *28*, 980-991.
38. Cherepanov, P.; Maertens, G.; Proost, P.; Devreese, B.; Van Beeumen, J.; Engelborghs, Y.; De Clercq, E.; Debyser, Z. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J. Biol. Chem.* **2003**, *278*, 372-381.
39. Emiliani, S.; Mousnier, A.; Busschots, K.; Maroun, M.; Van Maele, B.; Tempé, D.; Vandekerckhove, L.; Moisant, F.; Ben-Slama, L.; Witvrouw, M.; Christ, F.; Rain, J.C.; Dargemont, C.; Debyser, Z.; Benarous, R. Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication. *J. Biol. Chem.* **2005**, *280*, 25517-25523.

40. Turlure, F.; Devroe, E.; Silver, P.A.; Engelman, A. Human cell proteins and human immunodeficiency virus DNA integration. *Front. Biosci.* **2004**, *9*, 3187-3208.
41. Engelman, A.; Cherepanov, P. The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. *PLoS Pathog.* **2008**, *4*, e1000046.
42. Poeschla, E.M. Integrase, LEDGF/p75 and HIV replication. *Cell. Mol. Life Sci.* **2008**, *65*, 1403-1424.
43. Cherepanov, P.; Devroe, E.; Silver, P.A.; Engelman, A. Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. *J. Biol. Chem.* **2004**, *279*, 48883-48892.
44. Ge, Y.Z.; Pu, M.T.; Gowher, H.; Wu, H.P.; Ding, J.P.; Jeltsch, A.; Xu, G.L. Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. *J. Biol. Chem.* **2004**, *279*, 25447-25454.
45. Stec, I.; Nagl, S.B.; van Ommen, G.J.; den Dunnen, J.T. The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? *FEBS Lett.* **2000**, *473*, 1-5.
46. Izumoto, Y.; Kuroda, T.; Harada, H.; Kishimoto, T.; Nakamura, H. Hepatoma-derived growth factor belongs to a gene family in mice showing significant homology in the amino terminus. *Biochem. Biophys. Res. Commun.* **1997**, *238*, 26-32.
47. Maertens, G.; Cherepanov, P.; Debyser, Z.; Engelborghs, Y.; Engelman, A. Identification and characterization of a functional nuclear localization signal in the HIV-1 integrase interactor LEDGF/p75. *J. Biol. Chem.* **2004**, *279*, 33421-33429.
48. Vanegas, M.; Llano, M.; Delgado, S.; Thompson, D.; Peretz, M.; Poeschla, E. Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering. *J. Cell Sci.* **2005**, *118*, 1733-1743.
49. Turlure, F.; Maertens, G.; Rahman, S.; Cherepanov, P.; Engelman, A. 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.* **2006**, *34*, 1653-1665.
50. Llano, M.; Vanegas, M.; Hutchins, N.; Thompson, D.; Delgado, S.; Poeschla, E.M. Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75. *J. Mol. Biol.* **2006**, *360*, 760-773.
51. Maertens, G.; Cherepanov, P.; Pluymers, W.; Busschots, K.; De Clercq, E.; Debyser, Z.; Engelborghs, Y. LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J. Biol. Chem.* **2003**, *278*, 33528-33539.
52. Maertens, G.N.; Cherepanov, P.; Engelman, A. Transcriptional co-activator p75 binds and tethers the Myc-interacting protein JPO2 to chromatin. *J. Cell Sci.* **2006**, *119*, 2563-2571.
53. Bartholomeeusen, K.; De Rijck, J.; Busschots, K.; Desender, L.; Gijssbers, R.; Emiliani, S.; Benarous, R.; Debyser, Z.; Christ, F. Differential interaction of HIV-1 integrase and JPO2 with the C terminus of LEDGF/p75. *J. Mol. Biol.* **2007**, *372*, 407-421.
54. Yokoyama, A.; Cleary, M.L. Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell* **2008**, *14*, 36-46.

55. Bartholomeeusen, K.; Christ, F.; Hendrix, J.; Rain, J.C.; Emiliani, S.; Benarous, R.; Debyser, Z.; Gijbbers, R.; De Rijck, J. Lens epithelium-derived growth factor/p75 interacts with the transposase-derived DDE domain of PoxZ. *J. Biol. Chem.* **2009**, *284*, 11467-11477.
56. Cherepanov, P.; Pluymers, W.; Claeys, A.; Proost, P.; De Clercq, E.; Debyser, Z. High-level expression of active HIV-1 integrase from a synthetic gene in human cells. *FASEB J.* **2000**, *14*, 1389-1399.
57. Llano, M.; Vanegas, M.; Fregoso, O.; Saenz, D.; Chung, S.; Peretz, M.; Poeschla, E.M. LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. *J. Virol.* **2004**, *78*, 9524-9537.
58. Llano, M.; Delgado, S.; Vanegas, M.; Poeschla, E. LEDGF/p75 prevents proteasomal degradation of HIV-1 integrase. *J. Biol. Chem.* **2004**, *279*, 55570-55577.
59. Vandekerckhove, L.; Christ, F.; Van Maele, B.; De Rijck, J.; Gijbbers, R.; Van den Haute, C.; Witvrouw, M.; Debyser, Z. Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. *J. Virol.* **2006**, *80*, 1886-1896.
60. Llano, M.; Saenz, D.T.; Meehan, A.; Wongthida, P.; Peretz, M.; Walker, W.H.; Teo, W.; Poeschla, E.M. An essential role for LEDGF/p75 in HIV integration. *Science* **2006**, *314*, 461-464.
61. Shun, M.C.; Raghavendra, N.K.; Vandegraaff, N.; Daigle, J.E.; Hughes, S.; Kellam, P.; Cherepanov, P.; Engelman, A. LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev.* **2007**, *21*, 1767-1778.
62. De Rijck, J.; Vandekerckhove, L.; Gijbbers, R.; Hombrouck, A.; Hendrix, J.; Vercammen, J.; Engelborghs, Y.; Christ, F.; Debyser, Z. Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. *J. Virol.* **2006**, *80*, 11498-11509.
63. Busschots, K.; Vercammen, J.; Emiliani, S.; Benarous, R.; Engelborghs, Y.; Christ, F.; Debyser, Z. The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. *J. Biol. Chem.* **2005**, *280*, 17841-17847.
64. Cherepanov, P. LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity *in vitro*. *Nucleic Acids Res.* **2007**, *35*, 113-124.
65. Schroder, A.R.; Shinn, P.; Chen, H.; Berry, C.; Ecker, J.R.; Bushman, F. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **2002**, *110*, 521-529.
66. Hematti, P.; Hong, B.K.; Ferguson, C.; Adler, R.; Hanawa, H.; Sellers, S.; Holt, I.E.; Eckfeldt, C.E.; Sharma, Y.; Schmidt, M.; von Kalle, C.; Persons, D.A.; Billings, E.M.; Verfaillie, C.M.; Nienhuis, A.W.; Wolfsberg, T.G.; Dunbar, C.E.; Calmels, B. Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. *PLoS Biol.* **2004**, *2*, e423.
67. Crise, B.; Li, Y.; Yuan, C.; Morcock, D.R.; Whitby, D.; Munroe, D.J.; Arthur, L.O.; Wu, X. Simian immunodeficiency virus integration preference is similar to that of human immunodeficiency virus type 1. *J. Virol.* **2005**, *79*, 12199-12204.
68. Kang, Y.; Moressi, C.J.; Scheetz, T.E.; Xie, L.; Tran, D.T.; Casavant, T.L.; Ak, P.; Benham, C.J.; Davidson, B.L.; McCray, P.B., Jr. Integration site choice of a feline immunodeficiency virus vector. *J. Virol.* **2006**, *80*, 8820-8823.

69. Hacker, C.V.; Vink, C.A.; Wardell, T.W.; Lee, S.; Treasure, P.; Kingsman, S.M.; Mitrophanous, K.A.; Miskin, J.E. The integration profile of EIAV-based vectors. *Mol Ther* **2006**, *14*, 536-545.
70. MacNeil, A.; Sankale, J.L.; Meloni, S.T.; Sarr, A.D.; Mboup, S.; Kanki, P. Genomic sites of human immunodeficiency virus type 2 (HIV-2) integration: similarities to HIV-1 *in vitro* and possible differences *in vivo*. *J. Virol.* **2006**, *80*, 7316-7321.
71. Bushman, F.; Lewinski, M.; Ciuffi, A.; Barr, S.; Leipzig, J.; Hannenhalli, S.; Hoffmann, C. Genome-wide analysis of retroviral DNA integration. *Nat. Rev. Microbiol.* **2005**, *3*, 848-858.
72. Ciuffi, A.; Llano, M.; Poeschla, E.; Hoffmann, C.; Leipzig, J.; Shinn, P.; Ecker, J.R.; Bushman, F. A role for LEDGF/p75 in targeting HIV DNA integration. *Nat. Med.* **2005**, *11*, 1287-1289.
73. Marshall, H.M.; Ronen, K.; Berry, C.; Llano, M.; Sutherland, H.; Saenz, D.; Bickmore, W.; Poeschla, E.; Bushman, F.D. Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. *PLoS One* **2007**, *2*, e1340.
74. Lewinski, M.K.; Bisgrove, D.; Shinn, P.; Chen, H.; Hoffmann, C.; Hannenhalli, S.; Verdin, E.; Berry, C.C.; Ecker, J.R.; Bushman, F.D. Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. *J. Virol.* **2005**, *79*, 6610-6619.
75. 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.; Sorensen, R.; Forster, A.; Fraser, P.; Cohen, J.I.; de Saint Basile, G.; Alexander, I.; Wintergerst, U.; Frebourg, T.; Aurias, A.; Stoppa-Lyonnet, D.; Romana, S.; Radford-Weiss, I.; Gross, F.; Valensi, F.; Delabesse, E.; Macintyre, E.; Sigaux, F.; Soulier, J.; Leiva, L.E.; Wissler, M.; Prinz, C.; Rabbitts, T.H.; Le Deist, F.; Fischer, A.; Cavazzana-Calvo, M. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **2003**, *302*, 415-419.
76. Howe, S.J.; Mansour, M.R.; Schwarzwaelder, K.; Bartholomae, C.; Hubank, M.; Kempinski, H.; Brugman, M.H.; Pike-Overzet, K.; Chatters, S.J.; de Ridder, D.; Gilmour, K.C.; Adams, S.; Thornhill, S.I.; Parsley, K.L.; Staal, F.J.; Gale, R.E.; Linch, D.C.; Bayford, J.; Brown, L.; Quaye, M.; Kinnon, C.; Ancliff, P.; Webb, D.K.; Schmidt, M.; von Kalle, C.; Gaspar, H.B.; Thrasher, A.J. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J. Clin. Invest.* **2008**, *118*, 3143-3150.
77. Ciuffi, A.; Diamond, T.L.; Hwang, Y.; Marshall, H.M.; Bushman, F.D. Modulating target site selection during human immunodeficiency virus DNA integration *in vitro* with an engineered tethering factor. *Hum. Gene Ther.* **2006**, *17*, 960-967.
78. Meehan, A.M.; Saenz, D.T.; Morrison, J.H.; Garcia-Rivera, J.A.; Peretz, M.; Llano, M.; Poeschla, E.M. LEDGF/p75 proteins with alternative chromatin tethers are functional HIV-1 cofactors. *PLoS Pathog.* **2009**, *5*, e1000522.
79. Cherepanov, P.; Sun, Z.Y.; Rahman, S.; Maertens, G.; Wagner, G.; Engelman, A. Solution structure of the HIV-1 integrase-binding domain in LEDGF/p75. *Nat. Struct. Mol. Biol.* **2005**, *12*, 526-532.
80. Cherepanov, P.; Ambrosio, A.L.; Rahman, S.; Ellenberger, T.; Engelman, A. Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 17308-17313.

81. Busschots, K.; Voet, A.; De Maeyer, M.; Rain, J.C.; Emiliani, S.; Benarous, R.; Desender, L.; Debyser, Z.; Christ, F. Identification of the LEDGF/p75 binding site in HIV-1 integrase. *J. Mol. Biol.* **2007**, *365*, 1480-1492.
82. Rahman, S.; Lu, R.; Vandegraaff, N.; Cherepanov, P.; Engelman, A. Structure-based mutagenesis of the integrase-LEDGF/p75 interface uncouples a strict correlation between *in vitro* protein binding and HIV-1 fitness. *Virology* **2007**, *357*, 79-90.
83. Bouyac-Bertoia, M.; Dvorin, J.D.; Fouchier, R.A.; Jenkins, Y.; Meyer, B.E.; Wu, L.I.; Emerman, M.; Malim, M.H. HIV-1 infection requires a functional integrase NLS. *Mol. Cell* **2001**, *7*, 1025-1035.
84. Limon, A.; Devroe, E.; Lu, R.; Ghory, H.Z.; Silver, P.A.; Engelman, A. 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.* **2002**, *76*, 10598-10607.
85. Lu, R.; Limon, A.; Devroe, E.; Silver, P.A.; Cherepanov, P.; Engelman, A. Class II integrase mutants with changes in putative nuclear localization signals are primarily blocked at a postnuclear entry step of human immunodeficiency virus type 1 replication. *J. Virol.* **2004**, *78*, 12735-12746.
86. Molteni, V.; Greenwald, J.; Rhodes, D.; Hwang, Y.; Kwiatkowski, W.; Bushman, F.D.; Siegel, J.S.; Choe, S. Identification of a small-molecule binding site at the dimer interface of the HIV integrase catalytic domain. *Acta Crystallogr. D. Biol. Crystallogr.* **2001**, *57*, 536-544.
87. Du, L.; Zhao, Y.; Chen, J.; Yang, L.; Zheng, Y.; Tang, Y.; Shen, X.; Jiang, H. D77, one benzoic acid derivative, functions as a novel anti-HIV-1 inhibitor targeting the interaction between integrase and cellular LEDGF/p75. *Biochem. Biophys. Res. Commun.* **2008**, *375*, 139-144.
88. Hou, Y.; McGuinness, D.E.; Prongay, A.J.; Feld, B.; Ingravallo, P.; Ogert, R.A.; Lunn, C.A.; Howe, J.A. Screening for antiviral inhibitors of the HIV integrase-LEDGF/p75 interaction using the AlphaScreen luminescent proximity assay. *J. Biomol. Screen* **2008**, *13*, 406-414.
89. De Luca, L.; Barreca, M.L.; Ferro, S.; Christ, F.; Iraci, N.; Gitto, R.; Monforte, A.M.; Debyser, Z.; Chimirri, A. Pharmacophore-based discovery of small-molecule inhibitors of protein-protein interactions between HIV-1 integrase and cellular cofactor LEDGF/p75. *ChemMedChem* **2009**, *4*, 1311-1316.
90. Hare, S.; Shun, M.C.; Gupta, S.S.; Valkov, E.; Engelman, A.; Cherepanov, P. A novel co-crystal structure affords the design of gain-of-function lentiviral integrase mutants in the presence of modified PSIP1/LEDGF/p75. *PLoS Pathog.* **2009**, *5*, e1000259.
91. Maignan, S.; Guilloteau, J.P.; Zhou-Liu, Q.; Clement-Mella, C.; Mikol, V. Crystal structures of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: high level of similarity of the active site with other viral integrases. *J. Mol. Biol.* **1998**, *282*, 359-368.
92. Berthoux, L.; Sebastian, S.; Muesing, M.A.; Luban, J. The role of lysine 186 in HIV-1 integrase multimerization. *Virology* **2007**, *364*, 227-236.
93. Hayouka, Z.; Rosenbluh, J.; Levin, A.; Loya, S.; Lebendiker, M.; Veprintsev, D.; Kotler, M.; Hizi, A.; Loyter, A.; Friedler, A. Inhibiting HIV-1 integrase by shifting its oligomerization equilibrium. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 8316-8321.

94. Li, M.; Craigie, R. Processing of viral DNA ends channels the HIV-1 integration reaction to concerted integration. *J. Biol. Chem.* **2005**, *280*, 29334-29339.
95. Sinha, S.; Pursley, M.H.; Grandgenett, D.P. Efficient concerted integration by recombinant human immunodeficiency virus type 1 integrase without cellular or viral cofactors. *J. Virol.* **2002**, *76*, 3105-3113.
96. Bosserman, M.A.; O'Quinn, D.F.; Wong, I. Loop202-208 in avian sarcoma virus integrase mediates tetramer assembly and processing activity. *Biochemistry* **2007**, *46*, 11231-11239.
97. Yu, F.; Jones, G.S.; Hung, M.; Wagner, A.H.; MacArthur, H.L.; Liu, X.; Leavitt, S.; McDermott, M.J.; Tsiang, M. HIV-1 integrase preassembled on donor DNA is refractory to activity stimulation by LEDGF/p75. *Biochemistry* **2007**, *46*, 2899-2908.
98. Pandey, K.K.; Sinha, S.; Grandgenett, D.P. Transcriptional coactivator LEDGF/p75 modulates human immunodeficiency virus type 1 integrase-mediated concerted integration. *J. Virol.* **2007**, *81*, 3969-3979.
99. Raghavendra, N.K.; Engelman, A. 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* **2007**, *360*, 1-5.
100. Botbol, Y.; Raghavendra, N.K.; Rahman, S.; Engelman, A.; Lavigne, M. Chromatinized templates reveal the requirement for the LEDGF/p75 PWWP domain during HIV-1 integration *in vitro*. *Nucleic Acids Res.* **2008**, *36*, 1237-1246.
101. Ciuffi, A.; Bushman, F.D. Retroviral DNA integration: HIV and the role of LEDGF/p75. *Trends Genet.* **2006**, *22*, 388-395.
102. Marchand, C.; Maddali, K.; Metifiot, M.; Pommier, Y. HIV-1 IN Inhibitors: 2010 Update and Perspectives. *Curr Top Med Chem* **2009**.
103. Walker, M.A. New approaches for inhibiting HIV integrase: a journey beyond the active site. *Curr Opin Investig Drugs* **2009**, *10*, 129-136.
104. Brass, A.L.; Dykxhoorn, D.M.; Benita, Y.; Yan, N.; Engelman, A.; Xavier, R.J.; Lieberman, J.; Elledge, S.J. Identification of host proteins required for HIV infection through a functional genomic screen. *Science* **2008**, *319*, 921-926.
105. König, R.; Zhou, Y.; Elleder, D.; Diamond, T.L.; Bonamy, G.M.; Irelan, J.T.; Chiang, C.Y.; Tu, B.P.; De Jesus, P.D.; Lilley, C.E.; Seidel, S.; Opaluch, A.M.; Caldwell, J.S.; Weitzman, M.D.; Kuhen, K.L.; Bandyopadhyay, S.; Ideker, T.; Orth, A.P.; Miraglia, L.J.; Bushman, F.D.; Young, J.A.; Chanda, S.K. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* **2008**, *135*, 49-60.
106. Zhou, H.; Xu, M.; Huang, Q.; Gates, A.T.; Zhang, X.D.; Castle, J.C.; Stec, E.; Ferrer, M.; Strulovici, B.; Hazuda, D.J.; Espeseth, A.S. Genome-scale RNAi screen for host factors required for HIV replication. *Cell Host Microbe* **2008**, *4*, 495-504.
107. Bushman, F.D.; Malani, N.; Fernandes, J.; D'Orso, I.; Cagney, G.; Diamond, T.L.; Zhou, H.; Hazuda, D.J.; Espeseth, A.S.; König, R.; Bandyopadhyay, S.; Ideker, T.; Goff, S.P.; Krogan, N.J.; Frankel, A.D.; Young, J.A.; Chanda, S.K. Host cell factors in HIV replication: meta-analysis of genome-wide studies. *PLoS Pathog.* **2009**, *5*, e1000437.

108. Yeung, M.L.; Houzet, L.; Yedavalli, V.S.; Jeang, K.T. A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication. *J. Biol. Chem.* **2009**, *284*, 19463-19473.
109. Christ, F.; Thys, W.; De Rijck, J.; Gijssbers, R.; Albanese, A.; Arosio, D.; Emiliani, S.; Rain, J.C.; Benarous, R.; Cereseto, A.; Debyser, Z. Transportin-SR2 imports HIV into the nucleus. *Curr. Biol.* **2008**, *18*, 1192-1202.
110. Studamire, B.; Goff, S.P. Host proteins interacting with the Moloney murine leukemia virus integrase: multiple transcriptional regulators and chromatin binding factors. *Retrovirology* **2008**, *5*, 48.

© 2009 by the authors; licensee Molecular Diversity Preservation International, Basel, Switzerland. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).