LEDGF Hybrids Efficiently Retarget Lentiviral Integration Into Heterochromatin

Rik Gijsbers[1], Keshet Ronen[2], Sofie Vets[1], Nirav Malani[2], Jan De Rijck[1], Melissa McNeely[1], Frederic D Bushman[2] and Zeger Debyser[1]

[1]Division of Molecular Medicine, Katholieke Universiteit Leuven, Leuven, Belgium; [2]Department of Microbiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA

Correction of genetic diseases requires integration of the therapeutic gene copy into the genome of patient cells. Retroviruses are commonly used as delivery vehicles because of their precise integration mechanism, but their use has led to adverse events in which vector integration activated proto-oncogenes and contributed to leukemogenesis. Here, we show that integration by lentiviral vectors can be targeted away from genes using an artificial tethering factor. During normal lentivirus infection, the host cell–encoded transcriptional coactivator lens epithelium–derived growth factor/p75 (LEDGF/p75) binds lentiviral integrase (IN), thereby targeting integration to active transcription units and increasing the efficiency of infection. We replaced the LEDGF/p75 chromatin interaction–binding domain with CBX1. CBX1 binds histone H3 di- or trimethylated on K9, which is associated with pericentric heterochromatin and intergenic regions. The chimeric protein supported efficient transduction of lentiviral vectors and directed the integration outside of genes, near bound CBX1. Despite integration in regions rich in epigenetic marks associated with gene silencing, lentiviral vector expression remained efficient. Thus, engineered LEDGF/p75 chimeras provide technology for controlling integration site selection by lentiviral vectors.

Received 23 September 2009; accepted 6 December 2009; published online 5 January 2010. doi:10.1038/mt.2010.36

INTRODUCTION

Lens epithelium–derived growth factor/p75 (LEDGF/p75) is a transcriptional coactivator[1,2] that colocalizes with chromatin[3] and interacts with the integrase (IN) of the human immunodeficiency virus type-1 (HIV-1) and other lentivirinae[4–7] (**Figure 1**). RNA interference (RNAi)–mediated depletion of LEDGF/p75 results in the relocalization of IN to the cytoplasm and blocks HIV replication at the integration step of the viral life cycle.[8–11] In addition, LEDGF/p75 depletion alters the genomic distribution of lentiviral integration sites.[12–14] Lentiviruses preferentially integrate in active transcription units and disfavor promoter regions and locations within 1 kb of CpG islands.[13–17] For both HIV and EIAV (equine infectious anemia virus), integration in LEDGF/p75-depleted cells is reduced in transcription units, but enriched in CpG islands and at the 5′-end of genes, together with an increased GC content of regions surrounding integration sites. A model has therefore been proposed in which LEDGF/p75 functions as a molecular tether, bridging IN in the viral preintegration complex and host chromatin.[11,18]

Overexpression of a C-terminal fragment of LEDGF/p75 (amino acid 325–530; $LEDGF_{325–530}$) or the IN-binding domain alone does not mediate chromatin binding, but relocates HIV-IN to the cytoplasm and blocks HIV replication.[18,19] The mechanism of chromatin association is poorly understood, but elements in the N-terminal portion of LEDGF/p75 have been shown to be necessary. These include a PWWP domain, which contains a Pro-Trp-Trp-Pro signature related to the Tudor domain "Royal Family",[20] a nuclear localization signal, and two AT hooks[11,21–23] (**Figure 1**).

Meehan *et al.* recently showed that LEDGF proteins bearing H1.1, H1.5, and LANA in place of LEDGF's first 199 amino acids are functional HIV-1 cofactors.[24] Here, we used the LEDGF–IN interaction to retarget lentiviral integration to alternative regions of the genome. We engineered artificial chromatin tethers by fusing the C-terminal IN-binding fragment of LEDGF/p75 to alternative chromatin-binding proteins, expressed these in LEDGF/p75-depleted cells, and asked whether (i) infection was rescued and (ii) integration was retargeted to the regions bound by the chimeric protein. In a previous study, Ciuffi *et al.* created fusions of LEDGF/p75 IN-binding domain and the λ repressor DNA–binding domain and found increased *in vitro* strand transfer activity near λ repressor–binding sites.[25] However, this approach has not yet been used to redirect viral integration in cells.

We compared integration targeting for many hybrids between chromatin-binding proteins and LEDGF/p75, with particular focus on domains with binding specificities that might be useful during human gene therapy. The heterochromatin protein 1β (CBX1, formerly HP1β) binds to sites enriched in histone H3K9 di- and trimethylation at centromeric heterochromatin and transcriptionally silent regions—this provides a chromosomal target present at high copy in gene-sparse regions. We found that a fusion in which CBX1 replaced the chromatin-interaction domain of LEDGF/p75 rescued the infection block in LEDGF/p75-depleted cells. We characterized proviral integration sites

Correspondence: *Zeger Debyser, Division of Molecular Medicine, Katholieke Universiteit Leuven, Kapucijnenvoer 33, VCTB+5, Leuven, B-3000, Flanders, Belgium. E-mail: zeger.debyser@med.kuleuven.be*

using 454 pyrosequencing and found integration to be retargeted in the presence of the fusion to genomic sites bound by CBX1. These regions are low in gene expression and normally disfavored for lentiviral integration, but transgene expression from the vector was nevertheless efficient. These findings open possibilities for targeting of gene therapy vectors by using the LEDGF/p75–IN interaction, potentially to gene-poor regions where their genotoxic potential may be reduced.

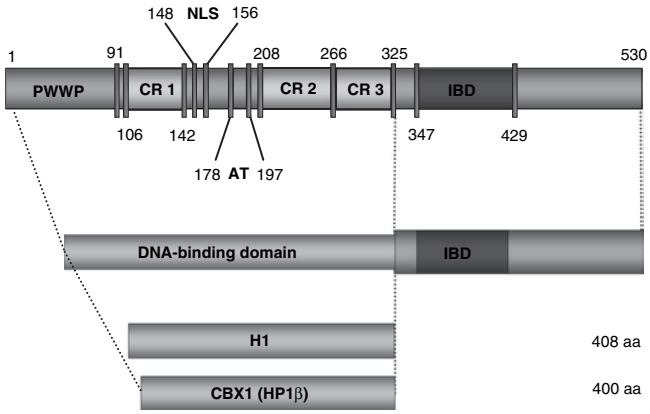

Figure 1 Domain structure of LEDGF/p75 and schematic representation of LEDGF$_{325–530}$ fusions. LEDGF/p75 contains an integrase-binding domain (IBD) in the C-terminus and a combination of chromatin-interacting modules located in the N-terminal end, most notably the PWWP domain, the AT-hook domain, and three relatively charged regions (CR1–CR3) influence chromatin binding. In the lower panel the DNA-binding domain fusions with LEDGF$_{325–530}$ are depicted, H1-LEDGF$_{325–530}$, and CBX1-LEDGF$_{325–530}$, respectively. Protein elements are drawn to scale. Numbers indicate amino acids of each domain. CBX1, heterochromatin protein 1β (formerly HP1β); H1, histone H1; LEDGF, lens epithelium–derived growth factor; NLS, nuclear localization signal.

RESULTS

Generation of cell lines depleted for LEDGF/p75

To study the role of LEDGF/p75 in tethering and targeting, we generated potent knockdown (KD) cell lines using murine leukemia virus (MLV)– based retroviral vectors encoding two miRNA-based short-hairpin RNAs[26] and a zeocin resistance cassette (**Supplementary Figure S1a**). Transduction of HeLaP4-CCR5 cells and subsequent selection resulted in a polyclonal cell line suppressing LEDGF/p75 mRNA levels to 7% of parental HeLaP4-CCR5 cells (**Supplementary Figure S1b**). Monoclonal lines were established, selecting the most potent LEDGF KD cells. Three monoclonal lines were isolated that contain <4% of the wild-type (WT) LEDGF/p75 mRNA, referred to as A3, B5, and D11 (96.7, 97, and 97.6% KD, respectively) (**Supplementary Figure S1b**). LEDGF/p75 protein was undetectable by western blot analysis in either the polyclonal or the monoclonal cell lines (**Supplementary Figure S1c**). Whereas immunocytochemistry showed that LEDGF/p75 was not depleted from all nuclei in the polyclonal cell line, the protein was undetectable in the monoclonal lines (**Supplementary Figure S1d**).

Generation of LEDGF hybrids

To retarget lentiviral integration, we substituted the LEDGF/p75 chromatin-binding region (amino acid 1–324, **Figure 1**) by alternative DNA-binding proteins. LEDGF$_{325–530}$ was fused to linker histone 1 (H1; histone 1, H1F0) and heterochromatin protein 1β (CBX1, formerly HP1β). H1F0 binds to nucleosomes without apparent preference for the underlying DNA sequence,[27] continuously shuttling among chromatin-binding sites.[28] CBX1 is associated with pericentric heterochromatin. CBX1 has a single N-terminal chromodomain, which recognizes histone tails via methylated lysine residues, for example trimethylated histone

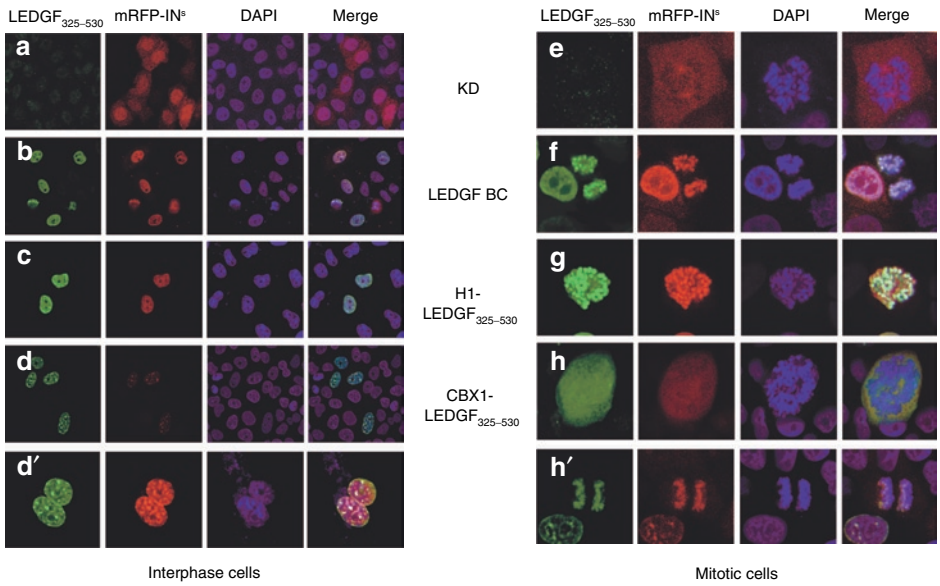

Figure 2 Subcellular localization and HIV-IN interaction of stably expressed LEDGF hybrids in interphase and mitotic cells. Stable KD cell lines were complemented with LEDGF$_{325–530}$ fusion proteins and transfected with mRFP-IN^{s}. Laser scanning microscopy (LSM) images of cells stained with anti-LEDGF$_{325–530}$ antibody are shown (green). Nuclei were stained with DAPI (4′,6-diamidino-2-phenylindole; blue). The respective constructs are indicated. Mitotic cells are displayed in the right panel. The data are representative for the vast majority of the imaged cells. CBX1, heterochromatin protein 1β (formerly HP1β); H1, histone H1; HIV, human immunodeficiency virus; IN, integrase; KD, knockdown; LEDGF, lens epithelium–derived growth factor; mRFP, monomeric red fluorescent protein; NLS, nuclear localization signal.

H3 at lysine 9 (H3K9me3) (ref. 29). Both constructs, referred to as H1-LEDGF$_{325–530}$ and CBX1-LEDGF$_{325–530}$, were introduced in LEDGF/p75-depleted cell lines (A3, B5, and D11, respectively) using MLV-based viral vectors and selected with blasticidin. In parallel, control cell lines complemented with MLV-based vectors encoding RNAi-resistant LEDGF/p75 (LEDGF BC) or eGFP-LEDGF$_{325–530}$ were generated.

Viability of the selected cell lines was similar to the parental HeLaP4-CCR5 cell line (data not shown). Expression of the fusion proteins in the A3 KD cell line was verified by western blot (**Supplementary Figure S2**). Although not all proteins were expressed to the same extent, the migration of all fusion proteins corresponded with their predicted molecular weight. Comparable data were obtained for the B5 and D11-derived cell lines (data not shown).

LEDGF hybrids colocalize with HIV-IN in the nucleus

Endogenous LEDGF/p75 displays a characteristic pattern of dense fine speckles in the interphase nucleoplasm and localizes to condensed chromosomes during mitosis.[3,30] Immunocytochemistry using antibodies recognizing the C-terminal portion of LEDGF/p75 showed no fluorescence in the KD cells (**Figure 2a**) corroborating depletion of LEDGF/p75. In contrast, KD cells complemented with RNAi-resistant LEDGF/p75 (LEDGF BC) displayed the typical dense fine speckled pattern of LEDGF/p75 (refs. 8,9,31) (**Figure 2b**). Complementation of KD cells with the H1-LEDGF$_{325–530}$ fusion resulted in a nuclear distribution (**Figure 2c**). In contrast, CBX1-LEDGF$_{325–530}$ was distributed in multiple irregularly shaped foci over the nuclear area during interphase (**Figure 2d,d′**), a pattern paralleling that of WT CBX1 (refs. 32,33). Similar to LEDGF/p75, the H1-LEDGF$_{325–530}$ fusion colocalized with condensed chromatin during mitosis (**Figure 2g**). In contrast, CBX1-LEDGF$_{325–530}$ diffused throughout the cytoplasm during late prophase (**Figure 2h**) and metaphase (not shown), but colocalized with condensed chromatin during anaphase (**Figure 2h′**), resembling the distribution pattern of WT CBX1 (refs. 32,33).

In order to reconstitute LEDGF function, LEDGF$_{325–530}$ fusions should, in addition to nuclear localization, support chromatin tethering of HIV-IN. In accordance with previous data,[3,4] transient expression of IN fused to the monomeric red fluorescent protein (mRFP-IN^s) in KD cells resulted in a diffuse fluorescent signal throughout the cytoplasm throughout the cell cycle (**Figure 2a,e**). Complementation with LEDGF/p75 relocated mRFP-IN^s to the nucleus and condensed chromatin (**Figure 2b,f**). Fusion of LEDGF$_{325–530}$ to the linker histone H1 rescued the nuclear phenotype of mRFP-IN^s and the binding to condensed chromatin (**Figure 2c,g**). CBX1-LEDGF$_{325–530}$ relocated mRFP-IN^s to the nucleus (**Figure 2d,d′**), in accordance with the distribution of WT CBX1 during late prophase (diffuse) and anaphase (condensed) (**Figure 2h,h′**).[33] B5 and D11 KD lines complemented in parallel with the same viral vector constructs resulted in similar cellular phenotypes (data not shown).

LEDGF hybrids rescue lentiviral transduction

After demonstrating that the fusions were capable of interacting with HIV-1 IN and tethering IN to chromatin, we asked whether they could support efficient lentiviral transduction. We infected the engineered cell lines with an HIV vector expressing firefly luciferase (fLuc).[34] Transduction efficiency was evaluated by assaying fLuc activity (relative light units/μg total protein). Transduction efficiency of KD cells was sevenfold lower than that of the parental cells (**Figure 3a**, WT). Back-complementation of KD with RNAi-resistant LEDGF/p75 (LEDGF BC) rescued lentiviral vector transduction to WT levels. Fusion of LEDGF$_{325–530}$ to either the linker histone H1 or CBX1 partially rescued viral vector transduction (36.3 and 47.5%, respectively). Fusion of enhanced green fluorescent protein (eGFP) to LEDGF$_{325–530}$ did not rescue transduction above the levels seen in KD cells. Similar data were obtained when evaluating transduction using eGFP as a reporter (**Supplementary Figure S4**) or using a near-complete HIV construct (HIV$_{NL4-3}$.fLuc virus) in which fLuc is driven by the proviral long terminal repeat promoter (**Supplementary Figure S5**). In parallel, we quantified the number of integrated proviral vector copies in the complemented KD lines following transduction

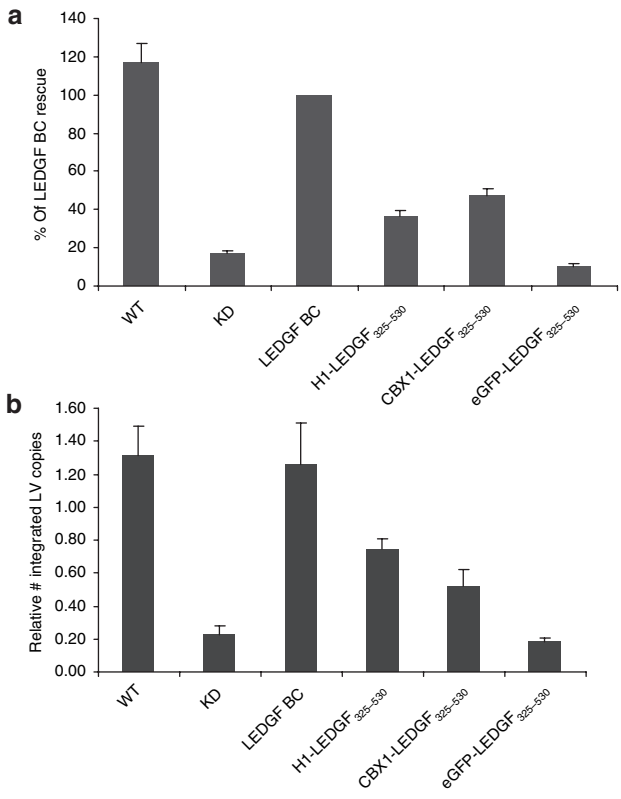

Figure 3 Rescue of HIV-based lentiviral vector transduction by LEDGF$_{325–530}$ hybrids. WT and LEDGF BC cells are used as controls. (**a**) Relative luciferase activity (RLU/μg protein) following HIV-based vector transduction (LV CMV eGFP-T2A-fLuc). Data are compiled at least six independent experiments and are expressed as percentages relative to LEDGF BC cells (mean ± SD). Complemented B5 and D11 cells resulted in similar results (**Supplementary Figure S3**). (**b**) Vector integration measured by Q-PCR in WT cells and LEDGF BC cells and LEDGF$_{325–530}$ fusions, data are represented as mean ± SD. CBX1, heterochromatin protein 1β (formerly HP1β); CMV, cytomegalovirus; eGFP, enhanced green fluorescent protein; H1, histone H1; HIV, human immunodeficiency virus; IN, integrase; KD, knockdown; LEDGF, lens epithelium–derived growth factor; LV, lentivirus; Q-PCR, quantitative PCR; RLU, relative light units; WT, wild type.

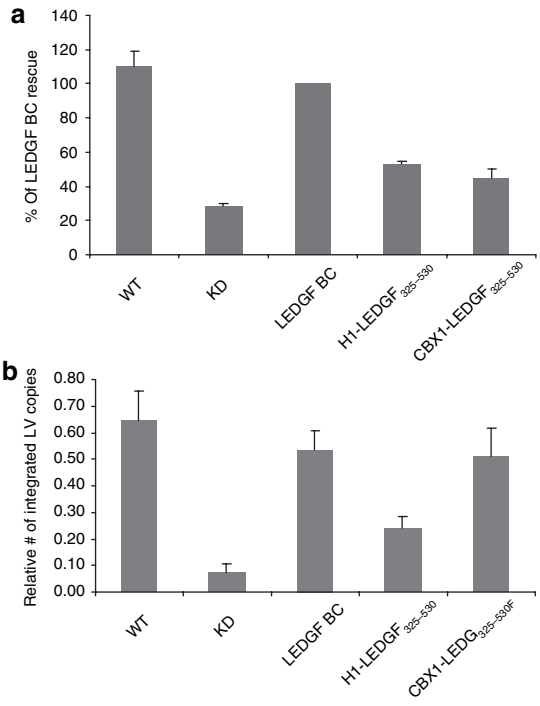

Figure 4 Rescue of EIAV-based vector transduction by LEDGF hybrids. $LEDGF_{325–530}$ fusions are tested for their ability to rescue EIAV-eGFP transduction. (**a**) eGFP fluorescence was evaluated by fluorescence-activated cell-sorting analysis. Data are compiled at least six independent experiments and are expressed relative to LEDGF BC complemented cells (mean ± SD). (**b**) EIAV integration measured by Q-PCR. CBX1, heterochromatin protein 1β (formerly HP1β); eGFP, enhanced green fluorescent protein; EIAV, equine infectious anemia virus; H1, histone H1; KD, knockdown; LEDGF, lens epithelium–derived growth factor; LV, lentivirus; Q-PCR, quantitative PCR; WT, wild type.

Table 1 Integration frequency near mapped genomic features in the human genome

	Cell line	# sites	% in RefSeq genes	% <2-kb CpG island
EIAV sites	WT	717	67.2***	1.3
	KD	213	51.2***	5.6**
	LEDGF BC	862	70.2***	1.9
	$H1\text{-}LEDGF_{325–530}$	449	46.1**	3.3
	$CBX1\text{-}LEDGF_{325–530}$	528	32.6*	1.1
MRC sites	MRC WT	2,151	37.3	2.8
	MRC KD	639	36.5	1.9
	MRC BC	2,586	36.9	2.1
	MRC H1	1,347	36.8	2.2
	MRC CBX1	1,584	37.8	2.1

Abbreviations: CBX1, heterochromatin protein 1β (formerly HP1β); EIAV, equine infectious anemia virus; H1, histone H1; KD, knockdown; LEDGF, lens epithelium–derived growth factor; MRC, matched random controls; WT, wild type.
Integration sets generated in this study and their genomic distributions.
Significant deviation from MRC in the Fisher's exact test is denoted by $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.0001$.

with HIV-based lentiviral vectors (**Figure 3b**). In KD cells, a 5.8-fold decrease in integrated copies was shown compared with WT cells, which was rescued completely upon back-complementation. Expression of fusion proteins partially rescued integration (60 and 41% of LEDGF BC integration for H1- and $CBX1\text{-}LEDGF_{325–530}$, respectively).

LEDGF/p75 is known to interact with other lentiviral INs in addition to HIV-IN.[4,5,7] To investigate complementation of other lentiviruses, fusion-containing and control cells were transduced with an EIAV vector, engineered to encode eGFP. $LEDGF_{325–530}$-hybrids were capable of rescuing EIAV transduction in the LEDGF/p75-depleted cells (**Figure 4**), paralleling the results obtained with HIV-1 based vectors. Whereas back-complementation of KD cells rescued vector transduction to WT levels (**Figure 4a**), fusion of $LEDGF_{325–530}$ to linker histone H1 or CBX1 partially rescued viral vector transduction (53 and 45.1%, respectively). The number of integrated copies in KD cells was decreased 8.8-fold compared to WT cells (**Figure 4b**). Complementing the KD cells with $H1\text{-}LEDGF_{325–530}$ resulted in a partial rescue of vector integration (3.3-fold increase compared to KD), and expression of $CBX1\text{-}LEDGF_{325–530}$ resulted in a 6.9-fold increase. Thus, expression of the chimeric proteins rescued EIAV as well as HIV transduction.

Sequencing of proviral integration sites

We next asked whether the $LEDGF_{325–530}$ fusions retargeted integration to genomic sites bound by the fusion partner. As HeLaP4 cells contain integrated HIV long terminal repeats that would interfere with the isolation of HIV provirus, we used the EIAV vector for distribution analysis. Both EIAV and HIV-IN interact with the LEDGF/p75 IN-binding domain[7] and show the same integration site preferences in WT[35] or LEDGF/p75-depleted[13] cells. Integration sites were analyzed as described previously,[13] yielding a total of 2,769 integration sites. Random control sites were generated computationally, and matched to experimental sites with respect to the distance to the nearest *Mse*I cleavage site (matched random control, MRC). In the analyses that follow, the distribution of experimental EIAV sites is normalized to that of the MRC sites, as a control for recovery bias due to cleavage by restriction enzymes.[36,37]

Retroviral INs show weak but detectable target sequence specificity at the local site of integration. In line with previous reports,[13,14] LEDGF/p75 depletion did not affect the consensus sequence flanking the integration site (**Supplementary Figure S6**). Likewise, expression of $LEDGF_{325–530}$ fusions did not alter the consensus sequence, consistent with the idea that IN binding to local target DNA determines the sequence preference, independent of the tethering mechanism.

CBX1 fusion directs integration to intergenic regions

Lentiviruses favor integration in transcription units and gene-dense regions.[15,35] In the absence of LEDGF/p75, this preference is reduced, and a preference for CpG islands and gene 5′-ends emerges.[12–14] As an initial survey of the proviral integration site distribution, we examined the frequency of integration in these features. In KD cells, a reduction in the integration frequency in RefSeq transcription units from 67.2 to 51.2% was observed (**Table 1**), as previously reported for LEDGF/p75-depleted cells.[13] Although this reduction was statistically significant ($P < 0.0001$, Fisher's exact test), integration events in the KD cells were still significantly favored in transcription units over random ($P = 4.8 \times 10^{-6}$). In accordance

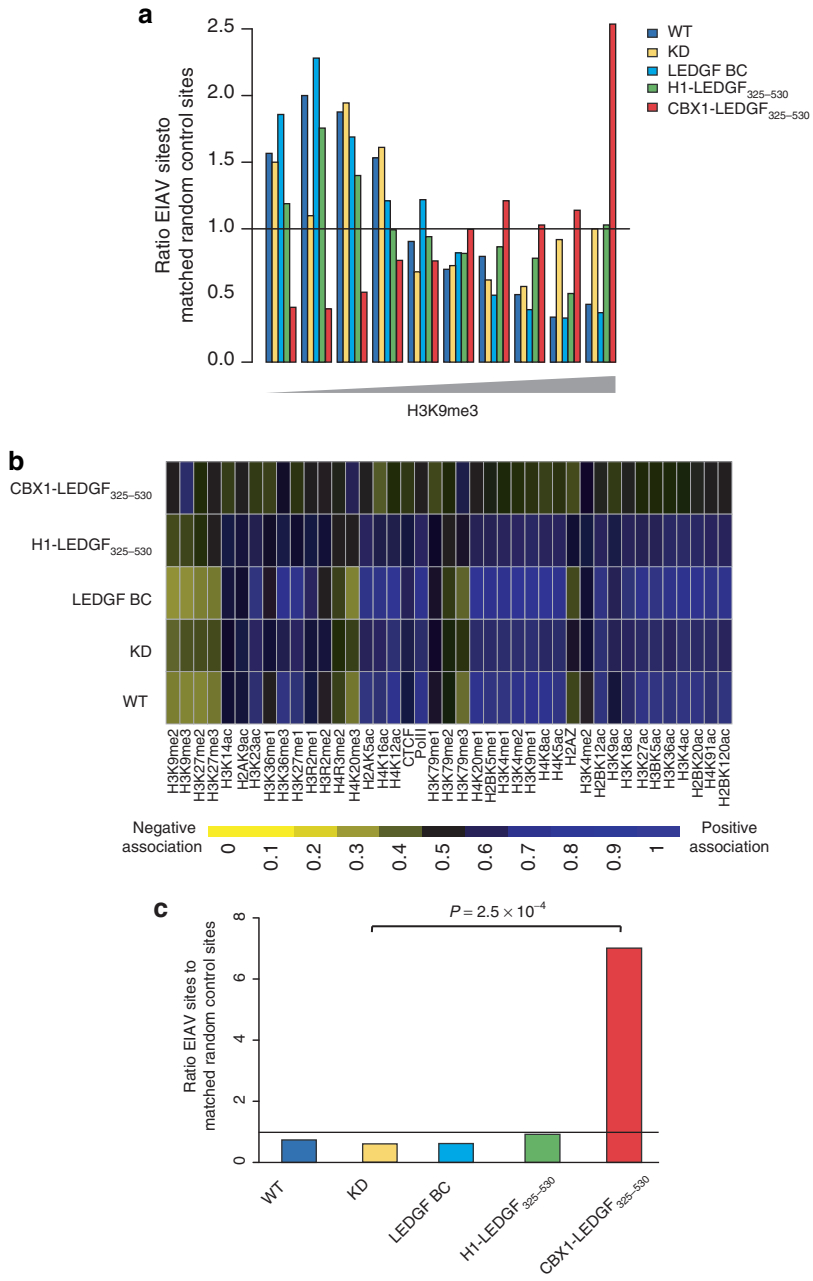

Figure 5 Expression of the CBX1-LEDGF$_{325–530}$ retargets EIAV integration into CBX1-rich heterochromatin regions. (**a**) Relationship of integration frequency to sites of H3K9me3 in the human genome. (**b**) Integration frequency relative to density of histone methylation and acetylation. (**c**) Integration frequency in human chromosome 19 near CBX1-binding sites. CBX1, heterochromatin protein 1β (formerly HP1β); EIAV, equine infectious anemia virus; H1, histone H1; KD, knockdown; LEDGF, lens epithelium–derived growth factor; WT, wild type.

with previous reports, we found that integration sites in the KD cells were favored near CpG islands. Both trends were reversed by LEDGF/p75 back-complementation. In contrast, expression of H1-LEDGF$_{325–530}$ did not rescue integration in transcription units. However, upon expression of CBX1-LEDGF$_{325–530}$, integration was significantly disfavored in transcription units compared with random (P = 0.026, Fisher's exact test), consistent with the distribution pattern of CBX1 in heterochromatic regions, which are generally gene-poor. Correlation of integration sites with the expression level of the targeted genes showed a slight shift towards genes with lower expression ($P < 0.0001$, χ^2-test to trend comparing CBX1-LEDGF$_{325–530}$ and WT cells) (**Supplementary Figure S7**). In addition, we analyzed the distribution of integration sites relative to transcription start sites (7.5-kb window around 5′-end of gene; **Supplementary Figure S8**). No significant differences were found between the positioning of integration sites in the different cell lines (χ^2-test to trend).

CBX1 is known to bind H3 di- or trimethylated at K9 (H3K9me2 and H3K9me3, respectively) via its chromodomain,[29,38,39] so we investigated integration near the sites of these histone modifications.[40,41] The H3K9me3 density near the sites of EIAV integration is summarized in **Figure 5a**. In WT cells, integration

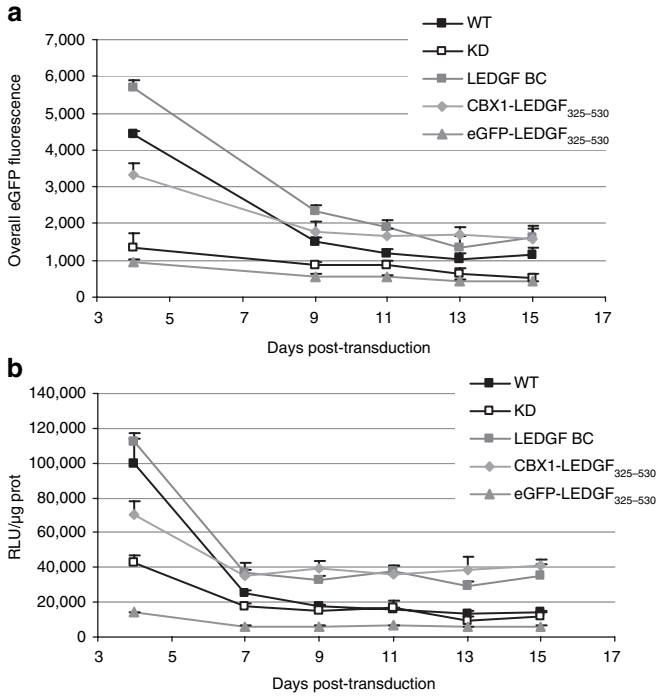

Figure 6 Effect of retargeting by CBX1-LEDGF$_{325–530}$ on transgene expression over time. WT and KD cells, together with LEDGF BC and eGFP-LEDGF$_{325–530}$ cells were used as controls. All cells were transduced with an HIV-based vector carrying both eGFP and fLuc as reporter genes (LV CMV eGFP-T2A-fLuc) as in **Supplementary Figure S3**. Reporter activity was determined at the indicated time point following HIV-based vector (LV CMV eGFP-T2A-fLuc) transduction (days post-transduction). (**a**) Overall eGFP fluorescence over time is calculated as MFI × % gated cells and displayed as mean ± SD (n = 6). (**b**) Relative luciferase activity (RLU/µg protein) over time; data are expressed as mean ± SD (n = 6). CBX1, heterochromatin protein 1β (formerly HP1β); CMV, cytomegalovirus; eGFP, enhanced green fluorescent protein; fLuc, firefly luciferase; H1, histone H1; HIV, human immunodeficiency virus; KD, knockdown; LEDGF, lens epithelium–derived growth factor; MFI, mean fluorescence intensity; RLU, relative light units; WT, wild type.

was disfavored in the areas high in H3K9me3 ($P = 2.9 \times 10^{-29}$), consistent with the role of H3K9me3 in transcriptional repression and establishment of silent heterochromatin, features generally disfavored by lentiviral integration. In the KD cells, the same negative correlation remained, though its magnitude was reduced ($P = 0.0012$). Complementation with LEDGF/p75 restored the negative effect of H3K9me3 to WT levels. Integration site distribution in H1-LEDGF$_{325–530}$ cells paralleled that seen in KD cells. In cells expressing CBX1-LEDGF$_{325–530}$, however, the correlation was reversed, with integration sites showing a clear preference for regions denser in H3K9me3 ($P = 1.3 \times 10^{-13}$).

We carried out the same analysis using genome-wide ChIP-seq data for a panel of 39 histone modifications.[41] **Figure 5b** shows correlations between integration sites and the density of these modifications. Each correlation is represented as a tile on the heat map, with the color denoting the strength and direction of the correlation. Histone modifications are grouped into clusters, reported to colocalize and associate with classes of functional genomic elements.[41] In WT cells, EIAV sites positively correlated with histone modifications generally associated with active transcription, such as all acetylations, and some histone methylations (shown in blue). Integration sites in WT cells negatively correlated (shown in yellow) with H3K9me3 and other markers reported to be associated with transcriptionally silent regions (*e.g.*, H3K27me3) and heterochromatin (*e.g.*, H4K20me3 and H3K79me3) (refs. 40,42,43). In KD cells, most of the correlations persisted, though they were less pronounced. Complementation with LEDGF restored correlations to WT levels. In cells expressing CBX1-LEDGF$_{325–530}$, however, most of the correlations were reversed, suggesting a dramatic redistribution of integration sites. In addition to H3K9me3, the modification bound by CBX1, regions high in H4K20me3 and H3K79me3 became favored for EIAV integration. The latter two modifications have also been associated with pericentric heterochromatin.

As CBX1 is enriched around centromeres, we compared the frequency of integration sites in pericentric regions. Integration sites in WT, KD, or LEDGF BC cells did not differ from random (**Supplementary Figure S9**). In contrast, in cells expressing CBX1-LEDGF$_{325–530}$, these regions contained 2.7-fold as many integration sites as MRC sites ($P = 0.0052$, Fisher's exact test), significantly higher than KD cells ($P = 0.0236$, Fisher's exact test). Sites from H1-LEDGF$_{325–530}$ cells also showed a preference for these regions, but this was not significantly higher than in KD cells ($P = 0.0851$, Fisher's exact test).

Finally, we used CBX1-binding sites mapped by DamID (ref. 44) to calculate the average number of CBX1-binding sites around integration sites. CBX1 occupancy around EIAV integration sites on chromosome 19 did not differ from random in WT cells and KD cells, and was not altered in H1-LEDGF$_{325–530}$-expressing cells (**Figure 5c**). However, in cells complemented with the CBX1 fusion, 10-kb windows around integration sites contain seven times as many CBX1-binding sites as random ($P = 2.5 \times 10^{-4}$). The same pattern held when integration sites across the genome were compared to CBX1-binding sites mapped genome-wide ($P = 0.015$, not shown). Thus, the CBX1-LEDGF/p75 fusion redirected integration to sites known to bind CBX1 and a collection of associated features.

Reporter gene expression remains efficient over time

Having shown that the CBX1-LEDGF$_{325–530}$ fusion retargets lentiviral integration to sites bound by CBX1, we wondered whether gene expression from the vector remained efficient, despite integration in regions rich in epigenetic marks associated with gene silencing. Vector encoded reporter activity was assessed over time following transduction of CBX1-LEDGF$_{325–530}$ cells and compared to WT cells, KD cells, or KD cells complemented with LEDGF BC or eGFP- LEDGF$_{325–530}$. Engineered cell lines were transduced with an HIV-based vector expressing both eGFP and fLuc (multiplicity of infection < 1) (ref. 34), and reporter expression was measured in cells over 2 weeks. Mean fluorescence intensity gradually decreased in all cell lines over time (**Figure 6a**). The relative difference in overall eGFP fluorescence (fold difference to the first measurement at day 4) reached 3.7- and 3.3-fold in WT cells or LEDGF BC cells, respectively, and 2.1-fold in the KD or eGFP-LEDGF$_{325–530}$ cells. Surprisingly, eGFP reporter activity also decreased only twofold in the CBX1-LEDGF$_{325–530}$ cells. In parallel, the same cells were analyzed for luciferase activity. Likewise, fLuc reporter activity decreased during the first week post-transduction and remained more or less constant thereafter

(**Figure 6b**). The relative difference in luciferase activity (relative to day 4) showed the most prominent effect in the WT cells (6.1-fold), followed by the KD cells, the LEGDF BC cells and eGFP-LEDGF$_{325–530}$ cells (3.1-, 3.3-, and 2.4-fold, respectively), whereas KD cells complemented with CBX1-LEDGF$_{325–530}$ showed a 1.9-fold decrease. Taken together, these data demonstrate that despite retargeting to CBX1-binding regions transgene expression from HIV-based vectors remains efficient.

DISCUSSION

In this study, we present evidence that LEDGF/p75 can be engineered to target lentiviral integration to new positions in the genome. Alternative chromatin-binding domains (linker histone H1 or the heterochromatin protein 1β, CBX1, were fused to the C-terminal portion of LEDGF/p75 (amino acid 325–530, LEDGF$_{325–530}$). CBX1 was selected to target sites of H3K9 di- and trimethylation, which are mapped in the genome and usually disfavored for lentiviral integration, so retargeting would be readily identifiable. H1 was used as a control, as it has no known preference for the underlying DNA sequence. Fusing a new chromatin-binding module to LEDGF$_{325–530}$ changed the behavior of this protein from an integration-inhibitor into an efficient cofactor. Upon challenge by lentiviral vectors, LEDGF$_{325–530}$-fusions supported efficient lentiviral transduction and integration compared to KD cells. Similar data were recently reported by Meehan *et al.*[24]—albeit using LEDGF hybrids that only lack the PWWP- and AT-hook domain (amino acid 1–199).

In addition, we characterized proviral integration sites using 454 pyrosequencing. Analysis of the EIAV integration distribution demonstrated that the CBX1 fusion retargeted lentiviral integration away from RefSeq genes (**Table 1**), to regions high in H3K9me3 (**Figure 5a**) and CBX1 binding (**Figure 5c**). The observation that integration can be retargeted away from genes and into heterochromatin using LEDGF hybrids raises hope for the development of safer lentiviral vectors for gene therapy. Before this study, attempts to retarget HIV integration employed fusions of IN with DNA-binding proteins.[45–48] Some of these showed retargeting as purified enzymes, but until now this approach had limited effect on the distribution of integration sites in cells.

The CBX1 hybrid provides the first example of global redistribution of lentiviral integration sites in the cellular genome, and the first instance of manipulation of a host tethering factor to do so. The success of the CBX1 fusion may be due to the abundance in the genome of its target ligand compared with site-specific DNA-binding domains previously employed, or perhaps its level of occupancy. Even though integration is targeted toward regions in the genome that are generally associated with gene silencing, transgene expression remained efficient over time (**Figure 6**). Whether new classes of genes are activated as a result is at present unknown and remains to be investigated.

Our findings open possibilities to engineer viral vectors that incorporate LEDGF hybrids to target integration into safe landing sites, thereby reducing the risk of insertional mutagenesis. Hare *et al.* have recently reported[49] a set of amino-acid substitutions in HIV-IN that abolish LEDGF/p75 binding, together with mutations in the LEDGF/p75 protein that restore binding. Gene delivery vectors could thus use an altered IN/LEDGF pair to direct integration, even in the presence of WT LEDGF/p75. To date, the altered IN does not show WT integration activity, but this may be improved with further engineering.

Our data also address issues in HIV biology. Our findings strengthen the idea that LEDGF/p75 is the dominant tether for lentiviral integration. Moreover, we show that chromatin-binding proteins with multiple specificities can successfully replace the LEDGF/p75 DNA-binding elements and rescue HIV infection in an LEDGF/p75 KD model. Still, the hybrids did not mediate rescue to WT levels, which leaves open the question of whether some portions of the N-terminus of LEDGF/p75 absent from our fusions stimulate IN activity or reporter gene expression. The fact that integration can be retargeted to genomic regions usually disfavored for integration indicates that integration in these areas in WT cells is disfavored due to the lack of a tether, rather than to an inherent integration barrier such as steric hindrance resulting from the condensed chromatin structure.

In conclusion, these results establish that LEDGF/p75 is the dominant targeting factor for lentiviral integration and that its interaction with lentiviral INs can be exploited to develop safe and target-specific lentiviral vectors for gene therapy.

MATERIALS AND METHODS

Generation of LEDGF/p75 depleted cell lines. Stable KD cells were generated using MLV-based vectors encoding miRNA-based short-hairpin RNAs[26] against LEDGF/p75 with a zeocin resistance cassette (**Supplementary Figure S1a**). HeLaP4-CCR5 cells (gift from P. Charneau, Institut Pasteur, Paris, France) were transduced and selected with zeocin (200 μg/ml; Invitrogen, Merelbeke, Belgium), resulting in polyclonal cells with 93% suppression of LEDGF/p75 mRNA compared to WT cells (**Supplementary Figure S1b**). Monoclonal cells were selected having <3% of WT LEDGF/p75 mRNA (referred to as A3, B5, and D11 in **Supplementary Figure S1b**). LEDGF protein was undetectable in polyclonal or monoclonal cells by western blot (**Supplementary Figure S1c**). Whereas LEDGF/p75 was not depleted from all nuclei in the polyclonal cell line, no protein could be detected in the monoclonals (**Supplementary Figure S1d**).

Construction of MLV-based retroviral vectors. MLV vector constructs were cloned in pLNCX (Clontech, Saint-Germain-en-Laye, France). pLNC-2x miRNA_L3 ZeoR was constructed by cloning zeocin cDNA and an artificial miR30-based shRNA-dimer into pLNC_MCS (primer sequences are included in **Supplementary Table S1**). Artificial miR30-based hairpin structures were cloned as described by Sun *et al.*[26] Briefly, Hs LEDGF/p75_L3_miRNA was amplified using UNI_miRNAi_s and UNI_miRNAi_as primers and a template with the LEDGF/p75 specific L3 siRNA sequence.[3,10,26] The product was cloned twice in peGFP-N3 (Clontech), resulting in peGFP-N3_2x L3mir.

pLNC_LEDGF BC-Ires-Bsd was constructed by digesting LEDGF BC-Ires-Bsd from pCHMWS-LEDGF BC-Ires-Bsd with *BamH*I–*Mlu*I and cloning in pLNC_MCS digested with *Bgl*II–*Mlu*I. pCHMWS-LEDGF BC-Ires-Bsd was constructed by replacing eGFP in pCHMWS-eGFP-Ires-Bsd with LEDGF BC (kind gift from M. Llano[4]) using LEDGF_KZ and LEDGF_as *Sal*I primers. The androgen receptor DNA-binding domain (AR-DBD) (gift from F. Claessens) was amplified using Flag_s *Bgl*II and AR DBD.CTE15_as primers and cloned in pCHMWS_eGFP-LEDGF$_{325–530}$ BC-Ires-Puro.[19] Next, flag AR-DBD-LEDGF$_{325–530}$ was amplified (Flag_s *Age*I, *Stu*I 325_as primers), generating pLNC_flag AR-DBD-LEDGF$_{325–530}$-Ires-Bsd. The latter plasmid was used to generate all fusions. pLNC_H1-LEDGF$_{325–530}$-Ires-Bsd was constructed by amplifying human H1F0

(NM_005318) from a HeLa cDNA. The PCR fragment was *Xma*I–*Xho*I digested to replace flag AR-DBD in pLNC_flag AR-DBD-LEDGF$_{325–530}$-Ires-Bsd. For pLNC_CBX1-LEDGF$_{325–530}$-Ires-Bsd, human CBX1 (NM_006807) was amplified with HsCBX1_s *Age*I and HsCBX1_as *Xho*I primers from pLgwCbx1-V5-EcoDam (gift from B. Van Steensel[44]).

Retroviral vector production and transduction. Lentiviral vector production was performed as described earlier.[34,50] Briefly, vesicular stomatitis virus glycoprotein pseudotyped HIV-based particles were produced by PEI transfection using pCHMWS_eGFP-T2A-fLuc as a transfer plasmid.[34] EIAV-vector particles were produced likewise using p6.1G3CeGFPw (M. Patel and J. Olsen, unpublished results, University of North Carolina, Chapel Hill) and pEV53B and vesicular stomatitis virus glycoprotein encoding pMD.G. HIV_NL4-3.fLuc single round virus was prepared by transient transfection with pNL4-3.LucR$^-$E$^-$ (National Institutes of Health AIDS Research and Reference Reagent Program) and pMD.G. MLV-based viral vectors were essentially produced as described for HIV-based vectors,[50] except that 293T producer cells were transfected with pCMVgagpol, the respective transfer plasmids (see higher) and pMD.G in a 12.5/32/7 ratio.

For lentiviral transduction experiments, cells were typically plated at 20,000 cells/well in a 96-well plate and transduced overnight. After 73 hours, 90% of cells were reseeded into two plates (FACS analysis and Luc-assay). The remainder was cultured for quantitative PCR or integration site analysis for at least 20 days to eliminate nonintegrated DNA. Stable cell lines were generated by transduction of the monoclonal LEDGF/p75 KD cells with retroviral vectors and subsequent selection with blasticidin (3 μg/ml; Invitrogen, Merelbeke, Belgium).

Quantitative PCR. Integrated proviral copies were quantified by real-time quantitative PCR on genomic DNA as reported earlier.[13] To determine LEDGF mRNA levels, total RNA was used for reverse transcription using the High-Capacity cDNA Archive kit (Applied Biosystems, Nieuwerkerk a/d. Ijssel, the Netherlands). Samples corresponding to 400 ng RNA were used for analysis using iQ5 Multicolor RT PCR detection system (BioRad, Nazareth, Belgium). Each reaction contains 12.5 μl 2× iQ Supermix (Biorad, Nazareth, Belgium), 40 nmol/l of forward and reverse primer, and 40 nmol/l of probe in a final volume of 25 μl. LEDGF/p75 primer/probe set: LEDGF Fwd4, 5′-GAA CTT GCT TCA CTT CAG GTC-3′, LEDGF Rev4, 5′-TCG CCG TAT TTT TTT CAG TGT-3′, LEDGF probe4, 5′-FAM-TGC AAC AAG CTC AGA AAC ACA CAG AGA TGA-TAMRA-3′. In all cases, RNaseP was used as endogenous house-keeping control (TaqMan RnaseP Control Reagent; Applied Biosystems). All samples were run in quadruplet for 3 minutes at 95 °C followed by 50 cycles of 10 seconds at 95 °C and 30 seconds at 55 °C. Data were analyzed with iQ5 Optical System Software (BioRad, Nazareth, Belgium).

Luciferase activity assay. Cells were lysed with 70 μl of lysis buffer (50 mmol/l Tris pH 7.5, 200 mmol/l NaCl, 0.2% NP40, 10% glycerol). The lysate was assayed according to the manufacturer's protocol (ONE-Glo; Promega, Madison, WI). Luciferase activity was normalized for total protein (BCA; Pierce, Rockford, IL). All conditions were run at least in triplicate in each experiment.

Western blot analysis. SDS (1%) protein extracts were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis. LEDGF fusions were detected using A300-848a antibody (1:2,000 dilution; Bethyl Laboratories, Montgomery, TX) and visualized by chemiluminescence (ECL$^+$; Amersham Biosciences, Uppsala, Sweden). Equal loading was verified with β-tubulin (1:4,000 dilution, T-4026; Sigma-Aldrich, St Louis, MO).

Laser scanning microscopy. Cells were transfected with pmRFP-IN^s and visualized as described earlier.[19] LEDGF$_{325–530}$ fusion proteins were detected as described previously using a polyclonal rabbit LEDGF$_{325–530}$-specific antibody (A300-848a, 1/500 dilution; Bethyl Laboratories) and Alexa-488-labeled goat anti-rabbit secondary antibodies (Molecular Probes, Invitrogen, Merelbeke, Belgium). All images were acquired using an LSM 510 META imaging unit (Carl Zeiss, Zaventem, Belgium). Alexa-488 was excited at 488 nm (AI laser), mRFP at 543 nm (HeNe laser) and DAPI (4′,6-diamidino-2-phenylindole) at 790 nm (Spectra-Physics Mai Tai laser; Spectra Physics, Mountain View, CA). After the main beam splitter (HFT KP 700/543 for mRFP, HFT UV/488/543/633 for eGFP, and HFT KP650 for DAPI) the fluorescence signal was divided by a secondary dichroic beam splitter (NFT 490 for eGFP, NFT 545 for mRFP).

Integration site amplification. Integration sites were amplified by linker-mediated PCR as described previously.[13] Genomic DNA was digested using *Mse*I and linkers were ligated (**Supplementary Table S2**). Proviral–host junctions were amplified by nested PCR using barcoded primers. This enabled pooling of PCR products into one sequencing reaction. Products were gel-purified and sequenced on the 454 GS-FLX instrument at the University of Pennsylvania.

Bioinformatic analysis. For integration sites to be authentic, sequences needed a best unique hit when aligned to the human genome (hg18 draft) using BLAT, the alignment began within 3 bp of the viral long terminal repeat end, and had >98% sequence identity. Statistical methods are detailed in Berry *et al.*[37] Integration site counts were compared with MRCs by a Fisher's exact test (where stated), or by multiple regression models for integration intensity and a c-logit test for significance.[37] Analysis was carried out using R (http://www.r-project.org). Histone modification data from Barski *et al.*[40] and Wang *et al.*[41] were used. The number of sequence tags from the ChIP-Solexa data sets in a defined window around each EIAV integration site or MRC, was calculated. CBX1-binding sites were analyzed using data from Vogel *et al.*[44] For each DamID probe set available, probes were aligned onto the hg18 draft using BLAT, and their associated $\log_2$-binding ratios used to select the top 5% of sites. For each integration site or MRC, the average number of high-affinity probes within a defined window around the site was calculated. Pericentric regions were defined as 1-Mb upstream or downstream of the unsequenced gap on each chromosome.

SUPPLEMENTARY MATERIAL

Figure S1. Generation of LEDGF/p75 depleted HeLaP4 cell lines.
Figure S2. Western blot analysis of LEDGF hybrids.
Figure S3. Rescue of HIV-based LV transduction by LEDGF hybrids in B5 and D11 monoclonal LEDGF/p75 KD lines.
Figure S4. Rescue of HIV-based LV transduction by LEDGF hybrids using eGFP fluorescence as a read-out.
Figure S5. Rescue of HIV$_{NL4.3}$-fLuc infection by LEDGF hybrids.
Figure S6. Alignment of 20 bp surrounding EIAV integration sites from each cell type.
Figure S7. Distribution of EIAV integration sites relative to gene expression level.
Figure S8. Distribution of EIAV integration sites relative to transcription start sites.
Figure S9. Integration frequency of EIAV near centromeres.
Table S1. Sequences of primers used in this study to clone hybrid LEDGF constructs.
Table S2. Sequences of primers and linkers used in this study to isolate EIAV integration sites.

ACKNOWLEDGMENTS

We thank the K.U. Leuven Cell Imaging Core for use of the confocal microscope. R.G. is a postdoctoral fellow of the Flemish Fund for Scientific Research (FWO Vlaanderen). K.R. was supported by National Institutes of Health (NIH) training grant T32 AI-07324-17. S.V. is funded by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). J.D. has a Mathilde-Krim postdoctoral fellowship from Amfar. M.M. is funded by the Affiliation Française contre les Maladies Musculaire. Research was funded by the

CellCoVir SBO grant (60813) of the IWT; FWO grant G.0530.08; EC grant THINC (HEALTH-F3-2008-201032) to Z.D. and by NIH grant AI52845, the University of Pennsylvania Center for AIDS Research, and the Penn Genome Frontiers Institute with a grant with the Pennsylvania Department of Health to F.D.B.

REFERENCES

1. Dietz, F, Franken, S, Yoshida, K, Nakamura, H, Kappler, J and Gieselmann, V (2002). The family of hepatoma-derived growth factor proteins: characterization of a new member HRP-4 and classification of its subfamilies. *Biochem J* **366**(Pt 2): 491–500.
2. Shinohara, T, Singh, DP and Fatma, N (2002). LEDGF, a survival factor, activates stress-related genes. *Prog Retin Eye Res* **21**: 341–358.
3. Maertens, G, Cherepanov, P, Pluymers, W, Busschots, K, De Clercq, E, Debyser, Z *et al.* (2003). LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J Biol Chem* **278**: 33528–33539.
4. Llano, M, Vanegas, M, Fregoso, O, Saenz, D, Chung, S, Peretz, M *et al.* (2004). LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. *J Virol* **78**: 9524–9537.
5. Busschots, K, Vercammen, J, Emiliani, S, Benarous, R, Engelborghs, Y, Christ, F *et al.* (2005). The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. *J Biol Chem* **280**: 17841–17847.
6. Cherepanov, P, Devroe, E, Silver, PA and Engelman, A (2004). Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. *J Biol Chem* **279**: 48883–48892.
7. Cherepanov, P (2007). LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity *in vitro*. *Nucleic Acids Res* **35**: 113–124.
8. 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 interactor LEDGF/p75. *J Biol Chem* **279**: 33421–33429.
9. 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**(Pt 8): 1733–1743.
10. Vandekerckhove, L, Christ, F, Van Maele, B, De Rijck, J, Gijsbers, R, Van den Haute, C *et al.* (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.
11. Llano, M, Saenz, DT, Meehan, A, Wongthida, P, Peretz, M, Walker, WH *et al.* (2006). An essential role for LEDGF/p75 in HIV integration. *Science* **314**: 461–464.
12. Ciuffi, A, Llano, M, Poeschla, E, Hoffmann, C, Leipzig, J, Shinn, P *et al.* (2005). A role for LEDGF/p75 in targeting HIV DNA integration. *Nat Med* **11**: 1287–1289.
13. Marshall, HM, Ronen, K, Berry, C, Llano, M, Sutherland, H, Saenz, D *et al.* (2007). Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. *PLoS ONE* **2**: e1340.
14. Shun, MC, Raghavendra, NK, Vandegraaff, N, Daigle, JE, Hughes, S, Kellam, P *et al.* (2007). LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev* **21**: 1767–1778.
15. Schröder, AR, Shinn, P, Chen, H, Berry, C, Ecker, JR and Bushman, F (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**: 521–529.
16. Ciuffi, A and Bushman, FD (2006). Retroviral DNA integration: HIV and the role of LEDGF/p75. *Trends Genet* **22**: 388–395.
17. Wang, GP, Ciuffi, A, Leipzig, J, Berry, CC and Bushman, FD (2007). HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res* **17**: 1186–1194.
18. Hombrouck, A, De Rijck, J, Hendrix, J, Vandekerckhove, L, Voet, A, De Maeyer, M *et al.* (2007). Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV. *PLoS Pathog* **3**: e47.
19. De Rijck, J, Vandekerckhove, L, Gijsbers, R, Hombrouck, A, Hendrix, J, Vercammen, J *et al.* (2006). Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. *J Virol* **80**: 11498–11509.
20. Maurer-Stroh, S, Dickens, NJ, Hughes-Davies, L, Kouzarides, T, Eisenhaber, F and Ponting, CP (2003). The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem Sci* **28**: 69–74.
21. 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**: 1653–1665.
22. Llano, M, Vanegas, M, Hutchins, N, Thompson, D, Delgado, S and Poeschla, EM (2006). Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75. *J Mol Biol* **360**: 760–773.
23. Shun, MC, Botbol, Y, Li, X, Di Nunzio, F, Daigle, JE, Yan, N *et al.* (2008). Identification and characterization of PWWP domain residues critical for LEDGF/p75 chromatin binding and human immunodeficiency virus type 1 infectivity. *J Virol* **82**: 11555–11567.
24. Meehan, AM, Saenz, DT, Morrison, JH, Garcia-Rivera, JA, Peretz, M, Llano, M *et al.* (2009). LEDGF/p75 proteins with alternative chromatin tethers are functional HIV-1 cofactors. *PLoS Pathog* **5**: e1000522.
25. Ciuffi, A, Diamond, TL, Hwang, Y, Marshall, HM and Bushman, FD (2006). Modulating target site selection during human immunodeficiency virus DNA integration *in vitro* with an engineered tethering factor. *Hum Gene Ther* **17**: 960–967.
26. Sun, D, Melegari, M, Sridhar, S, Rogler, CE and Zhu, L (2006). Multi-miRNA hairpin method that improves gene knockdown efficiency and provides linked multi-gene knockdown. *BioTechniques* **41**: 59–63.
27. Bustin, M, Catez, F and Lim, JH (2005). The dynamics of histone H1 function in chromatin. *Mol Cell* **17**: 617–620.
28. 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.
29. Maison, C and Almouzni, G (2004). HP1 and the dynamics of heterochromatin maintenance. *Nat Rev Mol Cell Biol* **5**: 296–304.
30. Ochs, RL, Muro, Y, Si, Y, Ge, H, Chan, EK and Tan, EM (2000). Autoantibodies to DFS 70 kd/transcription coactivator p75 in atopic dermatitis and other conditions. *J Allergy Clin Immunol* **105**(6 Pt 1): 1211–1220.
31. Llano, M, Delgado, S, Vanegas, M and Poeschla, EM (2004). Lens epithelium-derived growth factor/p75 prevents proteasomal degradation of HIV-1 integrase. *J Biol Chem* **279**: 55570–55577.
32. Minc, E, Allory, Y, Courvalin, JC and Buendia, B (2001). Immunolocalization of HP1 proteins in metaphasic mammalian chromosomes. *Methods Cell Sci* **23**: 171–174.
33. Hayakawa, T, Haraguchi, T, Masumoto, H and Hiraoka, Y (2003). Cell cycle behavior of human HP1 subtypes: distinct molecular domains of HP1 are required for their centromeric localization during interphase and metaphase. *J Cell Sci* **116**: 3327–3338.
34. Ibrahimi, A, Vande Velde, G, Reumers, V, Toelen, J, Thiry, I, Vandeputte, C *et al.* (2009). Highly efficient multicistronic lentiviral vectors with peptide 2A sequences. *Hum Gene Ther* **20**: 845–860.
35. Hacker, CV, Vink, CA, Wardell, TW, Lee, S, Treasure, P, Kingsman, SM *et al.* (2006). The integration profile of EIAV-based vectors. *Mol Ther* **14**: 536–545.
36. Mitchell, RS, Beitzel, BF, Schroder, AR, Shinn, P, Chen, H, Berry, CC *et al.* (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol* **2**: E234.
37. Berry, C, Hannenhalli, S, Leipzig, J and Bushman, FD (2006). Selection of target sites for mobile DNA integration in the human genome. *PLoS Comput Biol* **2**: e157.
38. Lachner, M, O'Carroll, D, Rea, S, Mechtler, K and Jenuwein, T (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**: 116–120.
39. Bannister, AJ, Zegerman, P, Partridge, JF, Miska, EA, Thomas, JO, Allshire, RC *et al.* (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**: 120–124.
40. Barski, A, Cuddapah, S, Cui, K, Roh, TY, Schones, DE, Wang, Z *et al.* (2007). High-resolution profiling of histone methylations in the human genome. *Cell* **129**: 823–837.
41. Wang, Z, Zang, C, Rosenfeld, JA, Schones, DE, Barski, A, Cuddapah, S *et al.* (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* **40**: 897–903.
42. Schotta, G, Lachner, M, Sarma, K, Ebert, A, Sengupta, R, Reuter, G *et al.* (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* **18**: 1251–1262.
43. Jones, B, Su, H, Bhat, A, Lei, H, Bajko, J, Hevi, S *et al.* (2008). The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet* **4**: e1000190.
44. Vogel, MJ, Guelen, L, de Wit, E, Peric-Hupkes, D, Lodén, M, Talhout, W *et al.* (2006). Human heterochromatin proteins form large domains containing KRAB-ZNF genes. *Genome Res* **16**: 1493–1504.
45. Bushman, FD (1994). Tethering human immunodeficiency virus 1 integrase to a DNA site directs integration to nearby sequences. *Proc Natl Acad Sci USA* **91**: 9233–9237.
46. Holmes-Son, ML and Chow, SA (2002). Correct integration mediated by integrase-LexA fusion proteins incorporated into HIV-1. *Mol Ther* **5**: 360–370.
47. Bushman, FD and Miller, MD (1997). Tethering human immunodeficiency virus type 1 preintegration complexes to target DNA promotes integration at nearby sites. *J Virol* **71**: 458–464.
48. Tan, W, Dong, Z, Wilkinson, TA, Barbas, CF 3rd and Chow, SA (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.
49. Hare, S, Shun, MC, Gupta, SS, Valkov, E, Engelman, A and Cherepanov, P (2009). A novel co-crystal structure affords the design of gain-of-function lentiviral integrase mutants in the presence of modified PSIP1/LEDGF/p75. *PLoS Pathog* **5**: e1000259.
50. Geraerts, M, Michiels, M, Baekelandt, V, Debyser, Z and Gijsbers, R (2005). Upscaling of lentiviral vector production by tangential flow filtration. *J Gene Med* **7**: 1299–1310.