HUMAN GENE THERAPY 17:960–967 (September 2006)
© Mary Ann Liebert, Inc.

# Brief Report

# Modulating Target Site Selection During Human Immunodeficiency Virus DNA Integration *In Vitro* with an Engineered Tethering Factor

ANGELA CIUFFI, TRACY L. DIAMOND, YOUNG HWANG, HEATHER M. MARSHALL, and FREDERIC D. BUSHMAN

## ABSTRACT

**The mechanisms controlling retroviral integration have been the topic of intense interest, in part because of adverse clinical events that occurred during retrovirus-mediated human gene therapy. Here we investigate the use of artificial tethering interactions to constrain retroviral integration site selection in an *in vitro* model. During normal infection, HIV DNA integration is favored in active cellular transcription units. One component of the targeting mechanism is the cellular LEDGF/p75 protein. LEDGF/p75 binds tightly to HIV integrase (IN) protein, and depletion of LEDGF/p75 from target cells results in reduced integration in transcription units, suggesting integration targeting by a tethering mechanism. We constructed and analyzed fusions of LEDGF/p75 or its IN-binding domain (IBD) to the DNA-binding domain of phage λ repressor protein (λR). In the presence of the λR-LEDGF/p75 fusions, increased strand transfer by IN was seen in target DNA near λR-binding sites *in vitro*. These data support the idea that a direct interaction between LEDGF/p75 and IN can mediate targeting via a tethering mechanism, and provide proof of concept for the idea that protein–protein interactions might be engineered to constrain integration site selection during human gene therapy.**

## OVERVIEW SUMMARY

A series of studies has focused on methods for controlling integration site selection by retroviruses. Some success has been achieved with fusions of retroviral integrase proteins (INs) to sequence-specific DNA-binding domains, which promote integration near cognate DNA recognition sites. However, in experiments performed so far, vectors containing such fusions show only slight specificity and reduced titer. Here we provide an *in vitro* model for another approach: linking INs to target DNA via protein–protein interactions that tether IN near favored sites. Our data support the idea that this mechanism operates during normal HIV infection, and provide the first example of engineering target site specificity by this means.

## INTRODUCTION

RETROVIRAL VECTORS are one of the most popular gene delivery vehicles used in human gene therapy because of their ability to precisely integrate new sequences into the human genome (Fig. 1A). Unfortunately, the fact that integration is only weakly influenced by primary sequence at the point of integration (Stevens and Griffith, 1996; Carteau *et al.*, 1998; Holman and Coffin, 2005; Wu *et al.*, 2005) means that integration can take place at many genomic locations, including sites that activate oncogenes and result in transformation. Such oncogene activation has now been seen during human gene therapy trials (Hacein-Bey-Abina *et al.*, 2003a,b). This has focused intense interest on the mechanisms driving integration site selection,

Department of Microbiology, University of Pennsylvania School of Medicine, Philadelphia, PA 19104.

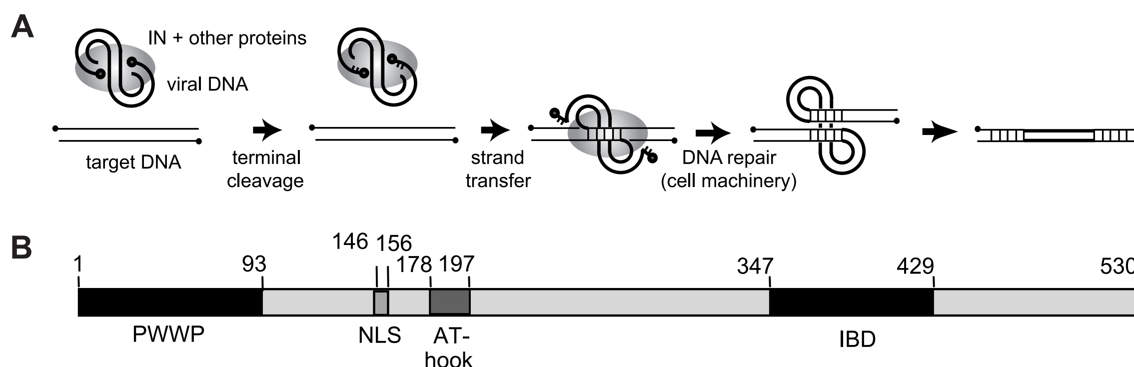


**FIG. 1.** Early steps of HIV DNA integration, and diagram of the LEDGF/p75 domain structure. (**A**) The initial DNA-breaking and joining reactions that mediate retroviral DNA integration are carried out by the virus-encoded IN protein (Katzman *et al.*, 1989; Bushman and Craigie, 1990; Bushman *et al.*, 1990; Craigie *et al.*, 1990; Katz *et al.*, 1990; Sherman and Fyfe, 1990; Bushman and Craigie, 1991). IN first cleaves the ends of the viral DNA (terminal cleavage) and then joins the exposed 3′ ends to target DNA on one strand at each host–virus DNA junction (strand transfer). The joining of the second DNA strand at each end and removal of the 5′ overhang is probably carried out by the host cell enzymes that normally repair DNA gaps (Yoder and Bushman, 2000). (**B**) The name LEDGF/p75 derives from early findings that (1) LEDGF/p75 was important in lens cells (hence "lens epithelium-derived growth factor") (Shinohara *et al.*, 2002) and (2) acts as a transcriptional coactivator and was named for its molecular weight ("p75") (Ge *et al.*, 1998). The name notwithstanding, LEDGF/p75 is expressed in most cell types and not just lens cells (Ge *et al.*, 1998). IBD, integrin-binding domain; NLS, nuclear localization signal.

with the long-term hope of increasing the safety of human gene therapy.

Previous studies *in vitro* with engineered retroviral integrases (INs) have shown that tethering via fused sequence-specific DNA-binding domains could constrain integration site selection. Combinations tested included human immunodeficiency virus (HIV) IN linked to the DNA-binding domain of phage $\lambda$ repressor protein ($\lambda$R) (Bushman, 1994), HIV IN linked to LexA (Goulaouic and Chow, 1996), HIV IN linked to zinc finger proteins (Bushman and Miller, 1997; Tan *et al.*, 2004), and avian sarcoma-leukosis virus (ASLV) IN linked to LexA (Katz *et al.*, 1996). Each of these fusions programmed integration near the binding site for the linked DNA-binding domain *in vitro*. In an important step, Tan *et al*. reported that such fusions *in vivo* can increase integration near the binding site for a fusion partner 10-fold (Tan *et al.*, 2006). However, most integration was still elsewhere in the genome and vector titer was reduced. Here we investigate whether a protein–protein interaction between an IN protein and a tethering protein engineered to bind both IN and specific DNA sites can target integration, using an *in vitro* model.

The IN-binding protein lens epithelium-derived growth factor (LEDGF)/p75 has been implicated previously in binding tightly to HIV IN and targeting HIV integration, providing one of the reagents used in our study (Cherepanov *et al.*, 2003; Maertens *et al.*, 2003; Llano *et al.*, 2004a,b; Turlure *et al.*, 2004; Bushman *et al.*, 2005; Ciuffi *et al.*, 2005; Vandegraaff *et al.*, 2006). Although LEDGF/p75 binds HIV IN tightly, tests of HIV titer after infection of target cells depleted for LEDGF/p75 showed only modest effects on viral titer (Boese *et al.*, 2004; Llano *et al.*, 2004b; Vandegraaff *et al.*, 2006; and A. Ciuffi and F.D. Bushman, unpublished data; but see Emiliani *et al.* [2005]

TABLE 1. OLIGONUCLEOTIDES USED IN THIS STUDY

| *Oligonucleotides used in:* | *Name* | *Sequence (5′ → 3′)* |
|---|---|---|
| DNA constructions | LEDGF-F/NdeI[a] | CGTGCATCTGTGAATCATATGGATTTCAAACC |
| | LEDGF-R/stop/BamHI[a,b] | CTCTATATTCCAGGTATGGATCC**CTA**GTTATCTAG |
| | IBD-F/NdeI[a] | GGAACAAATGGAACATATGCAGCAGAATAAAG |
| | LEDGF-R/KpnI[a] | CTATATTCCAGGTATGTCGGTACCGTTATCTAGTGTAG |
| | pFB257-F/NdeI[a] | CACAGGAAACAGCATATGAGCACAA |
| | pFB257-R/NdeI[a] | GCCGGATCCTCATATGACGTTCACC |
| | IN-F/NdeI[a] | TTTTTTCATATGTTTCTAGACGGTATCGATA |
| | IN-R/BsrGI[a] | TTTTTTTGTACAATCTTCGTCCTGACGAG |
| *In vitro* integration assay | FB64 | ACTGCTAGAGATTTTCCACACGGATCCTAGGC |
| | FB65-2 | GCCTAGGATCCGTGTGGAAAATCTCTAGCA |
| | FB66 | GCCTAGGATCCGTGTGGAAAATC |
| | FB182 | GACAGATTCCTGGGATAAGCCAAG |
| | FB183 | CGCGCTTTGATATACGCCGAGATC |

[a]Restriction sites introduced in the primers to facilitate cloning are underlined.
[b]Boldface characters indicate additional mutations to introduce a stop codon.

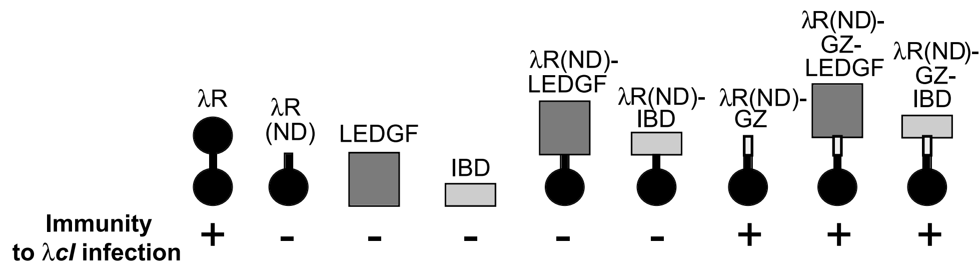


**FIG. 2.** DNA binding by λR-LEDGF/p75 fusion proteins in *E. coli*, assayed by immunity to phage λ*cI* infection. Fusions of λR(ND) to LEDGF or IBD confer high-level immunity to λ phage infection only when the GCN4 dimerization domain (GZ) is also included. Immunity (+) was assayed by cross-streaking *E. coli* expressing the indicated proteins versus phage λ*cI*.

and Vandekerckhove *et al.* [2006] for an opposing view), thus we have focused here on involvement of LEDGF/p75 in integration target site selection. One study supported a role for LEDGF/p75 in integration targeting *in vivo* (Ciuffi *et al.*, 2005). In human cells, HIV is known to favor integration in active transcription units (abbreviated TUs) (Schroder *et al.*, 2002; Wu *et al.*, 2003; Mitchell *et al.*, 2004). LEDGF/p75 was previously implicated in gene regulation (Ge *et al.*, 1998), making it attractive as a factor mediating HIV integration in active TUs. Three human cell lines were depleted for LEDGF/p75, using short hairpin RNAs (shRNAs) against LEDGF/p75, and more than 4000 HIV integration site sequences were determined in knockdown cells and controls. This revealed several effects on HIV integration site selection, including a significant reduction in the frequency of integration in TUs in the knockdown cells (Ciuffi *et al.*, 2005). These data supported the idea that LEDGF/p75 promoted HIV integration in transcription units by tethering integration complexes at these sites.

Here we used assays *in vitro* to investigate HIV integration targeting in the presence of fusion proteins composed of LEDGF/p75 linked to a sequence-specific DNA-binding domain. Initial studies showed that LEDGF/p75 bound DNA sequence nonspecifically and so did not contribute a confounding DNA-binding specificity of its own. We prepared and analyzed fusions of LEDGF/p75 to the λR. We found that the λR-LEDGF/p75 fusions directed favored integration *in vitro* near repressor-binding sites in integration target DNA, and that both IN-binding and DNA-binding functions were necessary for targeting. These studies *in vitro* support a tethering model for the mechanism of HIV integration targeting by LEDGF/p75, and demonstrate proof of concept for the idea that a protein–protein interaction can constrain integration site selection.

## MATERIALS AND METHODS

### *DNA constructions*

A plasmid encoding the N-terminal domain of the λR alone or fused to the leucine zipper GCN4 (pFB257) (Hu *et al.*, 1990; Bushman, 1994) was used as a cloning vector to introduce LEDGF/p75 or the integrase-binding domain (IBD), in frame, through *Nde*I and *Bam*HI restriction sites. Primers (Table 1) containing the same restriction sites were used to amplify LEDGF/p75 or IBD by polymerase chain reaction (PCR), in order to obtain constructs encoding λR(ND)-LEDGF, λR(ND)-IBD, λR(ND)-GZ-LEDGF, and λR(ND)-GZ-IBD. The pETDuet-Mxe vector (a modified version of pETDuet-1 from Novagen [Madison, WI] with insertion of an intein and a chitin-binding domain [CBD] from pTwin1 from New England BioLabs [Beverly, MA], a kind gift from G. Van Duyne, University of Pennsylvania School of Medicine, Philadelphia, PA) was used. *LEDGF* or *IBD* was joined to DNA encoding an intein, a CBD, and a histidine tag, using *Nde*I and *Kpn*I sites. These constructs encode LEDGF-intein-CBD-His and IBD-intein-CBD-His. IN from pNL4-3 was amplified with primers containing *Nde*I and *Bsr*GI restriction sites and introduced into pETDuet-Mxe previously digested with *Nde*I and *Acc*65I. An *Nde*I restriction site was added by site-directed mutagenesis (Stratagene, La Jolla, CA) at the start codon for λR(ND)-GZ from pFB257, in order to subclone this fragment into the pETDuet-Mxe-LEDGF plasmids, generating constructs encoding λR(ND)-GZ-LEDGF-in-

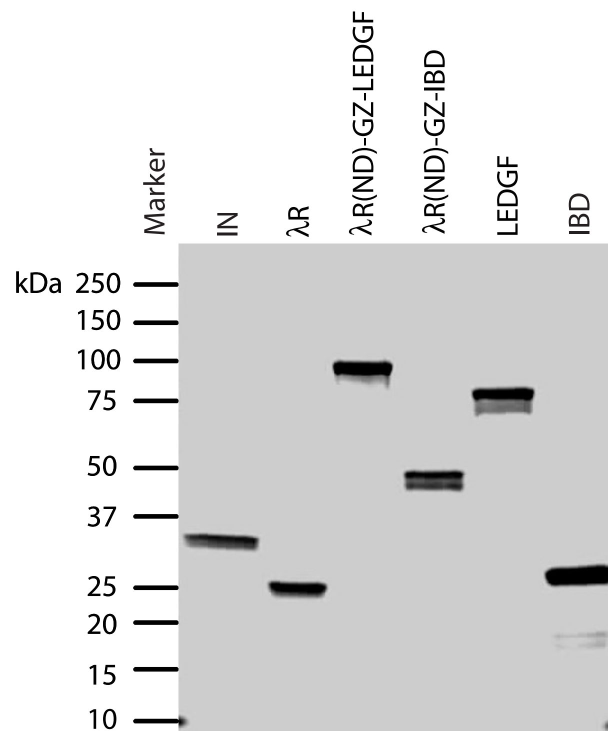


**FIG. 3.** Protein purification. SDS–PAGE analysis of each purified protein is shown after Coomassie staining.

tein-CBD-His and λR(ND)-GZ-IBD-intein-CBD-His. Immunity phenotypes of *Escherichia coli* containing λR fusions were determined by cross-streak assay versus λ*cI* (λ phages mutant in the *cI* gene, which encodes λR) (Hendrix *et al.*, 1983).

### *Protein purification*

Protein purification was performed essentially as in Diamond and Bushman (2005). DNA plasmids encoding the various fusion proteins were transformed into *E. coli* BL21(DE3) and grown in 1 liter of SuperBroth (3.2% tryptone, 2% yeast extract, 0.5% NaCl, pH to 7.5 with 1 *N* NaOH. Add 100 ml/L of pH 7.5 phosphate buffer) to a minimal optical density of 0.5. Expression of the fusion proteins was induced by adding 500 $\mu M$ isopropyl-β-D-1-thiogalactopyranoside (IPTG) for 3 hr. Bacteria were then pelleted, resuspended in Ni-lysis buffer (50 n*M* Na,K $PO_4$: pH 8.0, 300 n*M* NaCl), supplemented with protease inhibitors (Sigma), and sonicated. Disrupted bacteria were pelleted and the

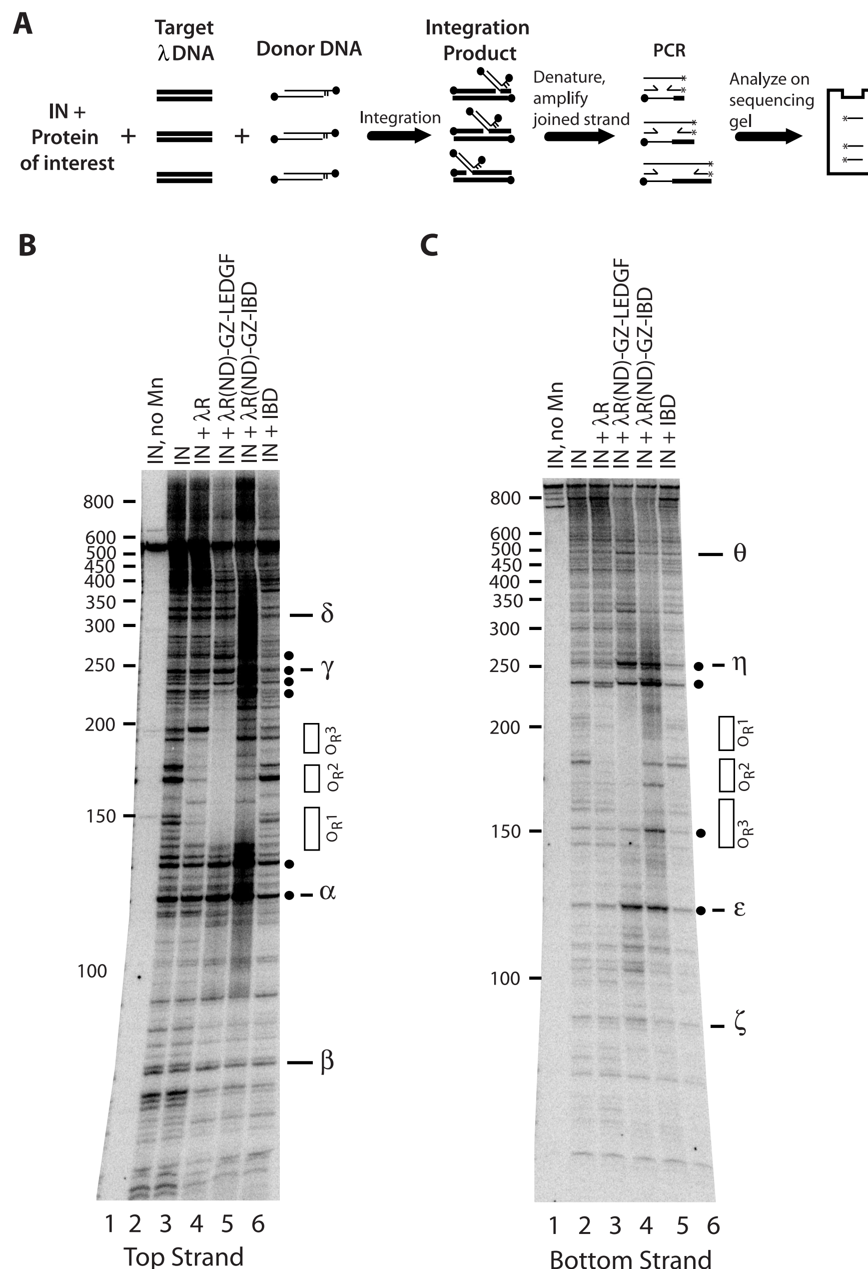


**FIG. 4.** Tethering HIV IN to target DNA via λR-LEDGF fusions results in local integration hotspots. (**A**) Diagram of the PCR method used for analysis. "Donor DNA" refers to the LTR duplex oligonucleotide. The arrows indicate PCR primers, the asterisk (*) indicates a $^{32}P$-labeled DNA 5′ end. (**B**) Integration site selection on the top DNA strand at λ $O_R$ in the presence of LEDGF or fusion proteins. Greek letters indicated on the gel mark bands used for statistical analysis. (**C**) Integration site selection on the bottom DNA strand at λ $O_R$ in the presence of LEDGF or fusion proteins. Dots indicate positions of increased integration when the λR-LEDGF fusions are present. Boxes indicate positions of the λ operators ($O_R$).

clear supernatant was fractionated on Ni-NTA agarose (Qiagen, Valencia, CA) to achieve purification of histidine-tagged proteins. The eluate from the Ni-NTA column was applied to chitin beads (New England BioLabs) to allow further purification of the proteins of interest through the chitin-binding domain. Finally, the fusion proteins were released by intein-mediated self-cleavage induced with 50 m*M* dithiothreitol (DTT). The proteins were further concentrated on an SP Sepharose Fast Flow column (GE Healthcare Life Sciences, Piscataway, NJ) according to manufacturer's instructions and eluted with increasing NaCl concentrations. Purification of the viral IN was done similarly, but without the SP Sepharose step and in the presence of 10 m*M* 3-[(3-cholamidopropyl)dimethylammonio]-1-propane sulfonic acid (CHAPS) throughout the purification procedure to prevent precipitation.

### *Assay of integration* in vitro

*In vitro* integration assays were performed essentially as described in Bushman (1994) and Diamond and Bushman (2005), with some modifications. Briefly, 30 pmol of purifed IN was incubated with various amounts of λR, λR(ND)-GZ-LEDGF, λR-(ND)-GZ-IBD, LEDGF, or IBD for 20 min on ice. The target DNA (λ DNA cleaved with *Hin*dIII, 3 μg) was added to the protein mix and incubated for a further 20 min on ice. The *in vitro* integration reaction was started by addition of 3 pmol of the donor DNA (constructed by annealing FB64 and FB65-2; Table 1) and incubating the mixture for 30 min at 37°C in 25 m*M* KCl, 10 m*M* 2-mercaptoethanol, 30 m*M* 2-(*N*-morpholino)ethanesulfonic acid (MES, pH 6.7), 15 m*M* $MnCl_2$, 10% glycerol, bovine serum albumin (BSA, 100 μg/ml). The reaction was stopped by adding proteinase K (50 μg/ml) and 0.5% sodium dodecyl sulfate (SDS) for 30 min at 37°C. The DNA was purified with a QIAquick PCR purification kit (Qiagen). The PCR amplification, using FB66 as a primer in the donor DNA region and radiolabeled FB182 (bottom strand) or FB183 (top strand) in a region of the λ DNA close to the right operator ($O_R$) site, was performed as described (Bushman, 1994; Diamond and Bushman, 2005). Labeled products were resolved on a 6% denaturing polyacrylamide gel, visualized by autoradiography, and analyzed with ImageQuant 1.2 software (GE Healthcare Life Sciences).

## RESULTS AND DISCUSSION

### *Design and purification of λR-LEDGF/p75 fusions*

The LEDGF/p75 protein (Fig. 1B) contains a well-defined domain near the C terminus, the IBD, that binds to HIV IN (Cherepanov *et al.*, 2004, 2005). Fusions were constructed containing full LEDGF/p75 or the IBD (amino acids 326–530) linked to the DNA-binding domain of phage λ repressor [termed λR(ND)] (Fig. 2). Early studies of LEDGF/p75 suggested that it bound to DNA sequence specifically (Shinohara *et al.*, 2002), but more recent reports indicate that LEDGF/p75 binds to many different primary DNA sequences (Singh *et al.*, 2006; Turlure *et al.*, 2006). Consistent with this latter report, we found that purified LEDGF/p75 protein bound to DNA sequence nonspecifically under the conditions used to assay HIV integration *in vitro* (data not shown).

One advantage of the use of λR as a fusion partner is that we could monitor DNA binding early during development of our recombinant proteins, using assays of immunity to λ infection in *E. coli* (Fig. 2) (Ptashne, 1992). *Escherichia coli* cells expressing λR are immune to infection with λ*cI* (i.e., λ phages mutant in the *cI* gene, which encodes λR). *Escherichia coli* cells expressing low levels of λR(ND) are sensitive to λ*cI* infection because the λR(ND) alone does not dimerize efficiently and consequently binds only weakly to the λ operators (λR-binding sites) (Pabo *et al.*, 1979; Sauer *et al.*, 1979). We found initially that fusions of LEDGF/p75 or the IBD to λR(ND) did not confer high-level immunity, indicating that the fusions likely did not dimerize efficiently and so did not bind λ operator DNA tightly. We therefore modified the fusions by adding a GCN4 leucine zipper (GZ), which is known to allow tight dimerization and high-affinity DNA binding when linked to λR(ND) (Hu *et al.*, 1990). The fusions containing all three protein domains, λR(ND)-GZ-LEDGF and λR(ND)-GZ-IBD, did in fact confer high-level immunity to λ*cI* infection in *E. coli* cells, indicating tight binding to λ operators. These proteins were then overexpressed in *E. coli* and purified (Fig. 3).

### *λR-LEDGF/p75 fusions promote integration near repressor-binding sites*

We next asked whether the λR(ND)-GZ-LEDGF and λR(ND)-GZ-IBD fusions were capable of directing increased

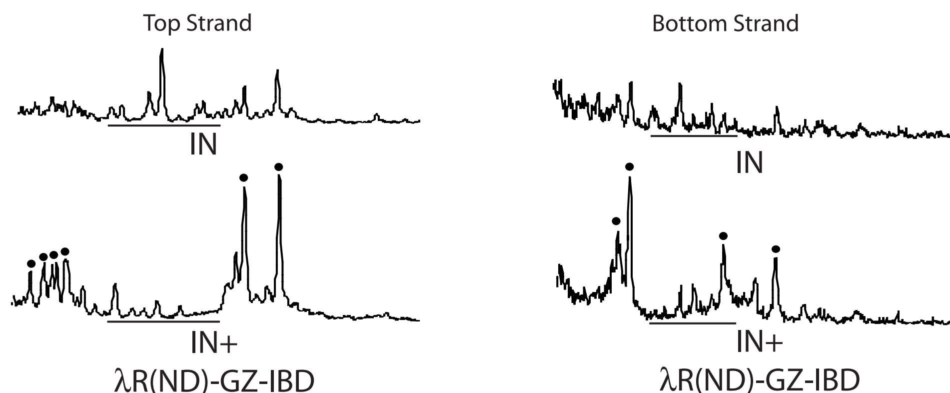


**FIG. 5.** Quantitation of integration patterns. Assays with IN only are compared with assays containing IN and λR(ND)-GZ-IBD. Dots indicate strongly affected integration sites (as in Fig. 4B and C); bars mark the location of λ operators. The results shown are representative of three replicates.

integration at target DNA sites near $\lambda$ operators. The PCR-based method used to visualize the distribution of integration sites is outlined in Fig. 4A. Integration reactions were carried out *in vitro*, using duplex oligonucleotides to mimic the viral DNA ends. Phage $\lambda$ DNA was used as the integration target, and integration was studied near the right operator ($O_R$), which contains three subsites, $O_R1$, $O_R2$, and $O_R3$, each of which binds a $\lambda$R dimer. After incubation of IN together with the LEDGF/p75 fusions and the reactant DNAs, the products were deproteinized and analyzed by PCR. One PCR primer bound to the viral DNA end, the other to a defined site in the target DNA near the $\lambda$ $O_R$. The target DNA primer was end labeled. Amplification yielded a collection of products, with the length of each amplification product reflecting the location of the initial integration event in the target DNA. The labeled integration products were then separated by electrophoresis on DNA sequencing-type gels and visualized by autoradiography. Labeled size markers were coelectrophoresed in adjacent lanes. Separate target DNA primers were used to visualize integration into the top (Fig. 4B) or bottom (Fig. 4C) DNA strand.

Reactions lacking the required metal cofactor yielded no detectable integration products (Fig. 4B and C, lanes 1). Reactions containing IN without any of the fusions showed a pattern of bands indicating that most of the target DNA phosphodiesters could be used as integration acceptor sites (Fig. 4B and C, lanes 2). Addition of $\lambda$R resulted in reduced integration in the regions of the $O_R$ because of steric interference with attack by IN complexes (Bushman, 1994), but did not affect integration elsewhere in the target DNA (Fig. 4B and C, lanes 3). For unknown reasons, one band on the top DNA strand within $O_R3$ was enhanced in reactions containing $\lambda$R, as was seen in a previous study (Bushman, 1994).

Addition of $\lambda$R(ND)-GZ-LEDGF (Fig. 4B and C, lanes 4) or $\lambda$R(ND)-GZ-IBD (Fig. 4B and C, lanes 5) also resulted in protection of the $O_R$, but in addition multiple sites in the flanking target DNA sequence were substantially increased in intensity. The increase in integration was seen on both sides of the operator on both DNA strands. Figure 5 shows a quantitative trace of integration product profiles comparing reactions containing only IN with reactions containing IN and $\lambda$R(ND)-GZ-IBD.

As a further control, reactions were also compared with those that had been carried out in the presence of purified IBD (LEDGF/p75, residues 326–530 only). These showed integration patterns identical to that of IN alone (Fig. 4B and C, lanes 6), indicating that tethering of the IBD near $O_R$ DNA via fusion to $\lambda$R(ND) is required for the local increase in integration.

### *Statistical analysis of targeting preferences*

A statistical analysis was carried out to verify that integration was increased selectively near the $O_R$. Control bands that did not change much in intensity in the various reactions were identified on either side of the $O_R$ on both DNA strands from the gels in Fig. 4B and C (bands marked $\beta$, $\delta$, $\zeta$, and $\theta$). Bands closer to the $O_R$ that were increased in intensity in the presence of the $\lambda$R-LEDGF/p75 fusions were similarly identified on the same gels (bands marked $\alpha$, $\gamma$, $\varepsilon$, and $\eta$). The intensities of these bands were quantified by PhosphorImager (GE Healthcare Life Sciences) and the ratios of affected bands to control bands were determined. For each band analyzed, two or three independent integration reactions were used, and 10–12 gel lanes were quantified. The collection of ratios ($\alpha/\beta$, $\gamma/\delta$, $\varepsilon/\zeta$, and $\eta/\theta$) was then separated into those from reactions expected to show targeting [$\lambda$R(ND)-GZ-LEDGF and $\lambda$R(ND)-GZ-IBD] and controls (IN only, or IN plus $\lambda$R, LEDGF/p75, or IBD). The mean values of the two sets of ratios were then compared by nonparametric Mann–Whitney test. This showed that differences were highly significant ($p < 0.0061$), with selectively increased integration seen near the $O_R$ on both DNA strands and on both sides of the $O_R$ in the reactions containing $\lambda$R(ND)-GZ-LEDGF and $\lambda$R(ND)-GZ-IBD. As a precaution, we repeated the statistical test, excluding the values from the gel in Fig. 4B and C, which were used in the initial selection of bands to compare, and the differences were still found to be highly significant. Taken together, these findings indicate that the LEDGF IBD, bound at the $O_R$ via fusion to $\lambda$R, can direct increased integration at nearby DNA sites *in vitro*.

### *Summary and implications*

Previously we reported that LEDGF/p75, a cellular IN-binding protein, is a mediator of HIV integration targeting *in vivo*, based on the finding that cells depleted for LEDGF/p75 showed significantly reduced integration in TUs (Ciuffi *et al.*, 2005). Here we report that tethering of LEDGF/p75-IBD to $\lambda$ operators, using $\lambda$R fusions, resulted in increased integration near these sites. These results support a model in which a protein–protein interaction between LEDGF/p75 and IN helps guide integration target site selection during HIV infection. These findings open the possibility that engineered tethering interactions might be developed to constrain integration targeting. For example, it may be possible to engineer new contact sites on retroviral INs that bind to cellular proteins that are enriched on chromosomal regions where integration is benign.

## ACKNOWLEDGMENTS

The authors thank Greg Van Duyne, Kushol Gupta, and members of the Bushman laboratory for materials and helpful discussions. This work was supported by NIH grants AI52845 and AI34786, the James B. Pendleton Charitable Trust, and Robin and Frederic Withington. A.C. was supported in part by a fellowship from the Swiss National Science Foundation.

## REFERENCES

BOESE, A., SOMMER, P., GAUSSIN, A., REIMANN, A., and NEHRBASS, U. (2004). Ini1/hSNF5 is dispensable for retrovirus-induced cytoplasmic accumulation of PML and does not interfere with integration. FEBS Lett. **578,** 291–296.

BUSHMAN, F., and MILLER, M.D. (1997). Tethering human immunodeficiency virus type 1 preintegration complexes to target DNA promotes integration at nearby sites. J. Virol. **71,** 458–464.

BUSHMAN, F., LEWINSKI, M., CIUFFI, A., BARR, S., LEIPZIG, J., HANNENHALLI, S., and HOFFMAN, C. (2005). Genome-wide analysis of retroviral DNA integration. Nat. Rev. Microbiol. **3,** 848–858.

BUSHMAN, F.D. (1994). Tethering human immunodeficiency virus 1

integrase to a DNA site directs integration to nearby sequences. Proc. Natl. Acad. Sci. U.S.A. **91,** 9233–9237.

BUSHMAN, F.D., and CRAIGIE, R. (1990). Sequence requirements for integration of Moloney murine leukemia virus DNA *in vitro*. J. Virol. **64,** 5645–5648.

BUSHMAN, F.D., and CRAIGIE, R. (1991). Activities of human immunodeficiency virus (HIV) integration protein *in vitro*: Specific cleavage and integration of HIV DNA. Proc. Natl. Acad. Sci. U.S.A. **88,** 1339–1343.

BUSHMAN, F.D., FUJIWARA, T., and CRAIGIE, R. (1990). Retroviral DNA integration directed by HIV integration protein *in vitro*. Science **249,** 1555–1558.

CARTEAU, S., HOFFMANN, C., and BUSHMAN, F.D. (1998). Chromosome structure and HIV-1 cDNA integration: Centromeric alphoid repeats are a disfavored target. J. Virol. **72,** 4005–4014.

CHEREPANOV, P., MAERTENS, G., PROOST, P., DEVREESE, B., VAN BEEUMEN, J., ENGELBORGHS, Y., DE CLERCQ, E., and DEBYSER, Z. (2003). HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. J. Biol. Chem. **278,** 372–381.

CHEREPANOV, P., DEVROE, E., SILVER, P.A., and ENGELMAN, A. (2004). Identification of an evolutionarily-conserved domain of LEDGF/p75 that binds HIV-1 integrase. J. Biol. Chem. **279,** 48883–48892.

CHEREPANOV, P., SUN, Z.Y., RAHMAN, S., MAERTENS, G., WAGNER, G., and ENGELMAN, A. (2005). Solution structure of the HIV integrase-binding domain in LEDGF/p75. Nat. Struct. Mol. Biol. **12,** 526–532.

CIUFFI, A., LLANO, M., POESCHLA, E., HOFFMANN, C., LEIPZIG, J., SHINN, P., ECKER, J.R., and BUSHMAN, F.D. (2005). A role for LEDGF/p75 in targeting HIV DNA integration. Nat. Med. **11,** 1287–1289.

CRAIGIE, R., FUJIWARA, T., and BUSHMAN, F. (1990). The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration *in vitro*. Cell **62,** 829–837.

DIAMOND, T.L., and BUSHMAN, F.D. (2005). Division of labor within human immunodeficiency virus integrase complexes: Determinants of catalysis and target DNA capture. J. Virol. **79,** 15376–15387.

EMILIANI, S., MOUSNIER, A., BUSSCHOTS, K., MAROUN, M., VAN MAELE, B., TEMPE, D., VANDERDERKHOVE, L., MOISANT, F., BEN-SLAMA, L., WITVROUW, M., CHRIST, F., RAIN, J.C., DARGEMONT, C., DEBYSER, Z., and BENAROUS, R. (2005). Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication. J. Biol. Chem. **280,** 25517–25523.

GE, H., SI, Y., and ROEDER, R.G. (1998). Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. EMBO J. **17,** 6723–6729.

GOULAOUIC, H., and CHOW, S.A. (1996). Directed integration of viral DNA mediated by fusion proteins consisting of human immunodeficiency virus type 1 integrase and *Escherichia coli* LexA protein. J. Virol. **70,** 37–46.

HACEIN-BEY-ABINA, S., VON KALLE, C., SCHMIDT, M., LE DEIST, F., WULFFRAAT, N., MACINTYRE, E., RADFORD, I., VILLEVAL, J.L., FRASER, C.C., CAVAZZANA-CALVO, M., and FISCHER, A. (2003a). A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. N. Engl. J. Med. **348,** 255–256.

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 ST. 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., WISSLER, M., PRINZ, C., RABBITTS, T.H., LE DEIST, F., FISCHER, A., and CAVAZZANA-CALVO, M. (2003b). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science **302,** 400–401.

HENDRIX, R.W., ROBERTS, J.W., STAHL, F.W., and WEISBERG, R.A. (1983). *Lambda II*. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

HOLMAN, A.G., and COFFIN, J.M. (2005). Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites. Proc. Natl. Acad. Sci. U.S.A. **102,** 6103–6107.

HU, J.C., O'SHEA, E.K., KIM, P.S., and SAUER, R.T. (1990). Sequence requirements for coiled-coils: Analysis with $\lambda$ repressor–GCN4 leucine zipper fusions. Science **250,** 1400–1403.

KATZ, R.A., MERKEL, G., KULKOSKY, J., LEIS, J., and SKALKA, A.M. (1990). The avian retroviral IN protein is both necessary and sufficient for integrative recombination *in vitro*. Cell **63,** 87–95.

KATZ, R.A., MERKEL, G., and SKALKA, A.M. (1996). Targeting of retroviral integrase by fusion to a heterologous DNA binding domain: *In vitro* activities and incorporation of a fusion protein into viral particles. Virology **217,** 178–190.

KATZMAN, M., KATZ, R.A., SKALKA, A.M., and LEIS, J. (1989). The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the *in vivo* sites of integration. J. Virol. **63,** 5319–5327.

LLANO, M., DELGADO, S., VANEGAS, M., and POESCHLA, E.M. (2004a). LEDGF/p75 prevents proteasomal degradation of HIV-1 integrase. J. Biol. Chem. **279,** 55570–55577.

LLANO, M., VANEGAS, M., FREGOSO, O., SAENZ, D., CHUNG, S., PERETZ, M., and POESCHLA, E.M. (2004b). 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.

MAERTENS, G., CHEREPANOV, P., PLUYMERS, W., BUSSCHOTS, K., DE CLERCQ, E., DEBYSER, Z., and ENGELBORGHS, Y. (2003). LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. J. Biol. Chem. **278,** 33528–33539.

MITCHELL, R., BEITZEL, B., SCHRODER, A., SHINN, P., CHEN, H., BERRY, C., ECKER, J.R., and BUSHMAN, F.D. (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol. **2,** E234.

PABO, C.O., SAUER, R.T., STURTEVANT, J.M., and PTASHNE, M. (1979). The $\lambda$ repressor contains two domains. Proc. Natl. Acad. Sci. U.S.A. **76,** 1608–1612.

PTASHNE, M. (1992). *A Genetic Switch*. (Cell Press and Blackwell Scientific Publications, Cambridge, MA).

SAUER, R.T., PABO, C.O., MEYER, B.J., PTASHNE, M., and BACKMAN, K.C. (1979). Regulatory functions of the $\lambda$ repressor reside in the amino-terminal domain. Nature **279,** 396–400.

SCHRODER, A., SHINN, P., CHEN, H., BERRY, C., ECKER, J.R., and BUSHMAN, F.D. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. Cell **110,** 521–529.

SHERMAN, P.A., and FYFE, J.A. (1990). Human immunodeficiency virus integration protein expressed in *Escherichia coli* possesses selective DNA cleaving activity. Proc. Natl. Acad. Sci. U.S.A. **87,** 5119–5123.

SHINOHARA, T., SINGH, D.P., and FATMA, N. (2002). LEDGF, a survival factor, activates stress-related genes. Prog. Retinal Eye Res. **21,** 341–358.

SINGH, D.P., KUBO, E., TAKAMURA, Y., SHINOHARA, T., KUMAR, A., CHYLACK, L.T., and FATMA, N. (2006). DNA binding domains and nuclear localization signal of LEDGF: Contribution of two helix–turn–helix (HTH)-like domains and a stretch of 58 amino

acids of the N-terminal to the trans activation potential of LEDGF. J. Mol. Biol. **355,** 379–394.

STEVENS, S.W., and GRIFFITH, J.D. (1996). Sequence analysis of the human DNA flanking sites of human immunodeficiency virus type 1 integration. J. Virol. **70,** 6459–6462.

TAN, W., ZHU, K., SEGAL, D.J., BARBAS, C.F.R., and CHOW, S.A. (2004). Fusion proteins consisting of human immunodeficiency virus type 1 integrase and the designed polydactyl zinc finger protein E2C direct integration of viral DNA into specific sites. J. Virol. **78,** 1301–1313.

TAN, W., DONG, Z., WILKINSON, T.A., BARBAS, C.F., and CHOW, S.A. (2006). Human immunodeficiency virus type 1 incorporated with fusion proteins consisting of integrase and the designed polydactyl zinc finger protein E2C can bias integration of viral DNA into a predetermined chromosomal region in human cells. J. Virol. **80,** 1939–1948.

TURLURE, F., DEVROE, E., SILVER, P.A., and ENGELMAN, A. (2004). Human cell proteins and human immunodeficiency virus DNA integration. Front. Biosci. **9,** 3187–3208.

TURLURE, F., MAERTENS, G., RAHMAN, S., CHEREPANOV, P., and ENGELMAN, A. (2006). A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. Nucleic Acids Res. **34,** 1663–1675.

VANDEGRAAFF, N., DEVROE, E., TURLURE, F., SILVER, P.A., and ENGELMAN, A. (2006). Biochemical and genetic analysis of integrase-interacting protein lens epithelium-derived growth factor (LEDGF)/p75 and hepatoma-derived growth factor related protein 2 (HRP2) in preintegration complex function and HIV-1 replication. Virology **346,** 415–426.

VANDEKERCKHOVE, L., CHRIST, F., VAN MAELE, B., DE RIJCK, J., GIJSBERS, R., VAN DEN HAUTE, C., WITVROUW, M., and DEBYSER, Z. (2006). Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. J. Virol. **80,** 1186–1896.

WU, X., LI, Y., CRISE, B., and BURGESS, S.M. (2003). Transcription start regions in the human genome are favored targets for MLV integration. Science **300,** 1749–1751.

WU, X., LI, Y., CRISE, B., BURGESS, S.M., and MUNROE, D.J. (2005). Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. J. Virol. **79,** 5211–5214.

YODER, K., and BUSHMAN, F.D. (2000). Repair of gaps in retroviral DNA integration intermediates. J. Virol. **74,** 11191–11200.

Address reprint requests to:
*Dr. Frederic D. Bushman*
*Department of Microbiology*
*University of Pennsylvania School of Medicine*
*3610 Hamilton Walk*
*Philadelphia, PA 19104-6076*

*E-mail:* bushman@mail.med.upenn.edu

Received for publication March 24, 2006; accepted after revision July 24, 2006.

Published online: August 4, 2006.