THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 35, Issue of August 29, pp. 33528–33539, 2003
Printed in U.S.A.


# LEDGF/p75 Is Essential for Nuclear and Chromosomal Targeting of HIV-1 Integrase in Human Cells*

Received for publication, April 7, 2003, and in revised form, June 5, 2003
Published, JBC Papers in Press, June 9, 2003, DOI 10.1074/jbc.M303594200

**Goedele Maertens‡§, Peter Cherepanov¶, Wim Pluymers¶‖, Katrien Busschots¶, Erik De Clercq¶, Zeger Debyser¶‖**, and Yves Engelborghs‡ ‡‡**

*From the ‡Laboratory of Biomolecular Dynamics, Katholieke Universiteit Leuven, Celestijnenlaan 200D, B-3001 Leuven and the ¶Rega Institute for Medical Research, Katholieke Universiteit Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium*

**We have reported that human immunodeficiency virus type 1 (HIV-1) integrase (IN) forms a specific nuclear complex with human lens epithelium-derived growth factor/transcription co-activator p75 (LEDGF/p75) protein. We now studied the IN-LEDGF/p75 interaction and nuclear import of IN in living cells using fusions of IN and LEDGF/p75 with enhanced green fluorescent protein and far-red fluorescent protein HcRed1. We show that both the N-terminal zinc binding domain and the central core domains of IN are involved in the interaction with LEDGF/p75. Both domains are essential for nuclear localization of IN as well as for the association of IN with condensed chromosomes during mitosis. However, upon overexpression of LEDGF/p75, the core domain fragment of IN was recruited to the nuclei and mitotic chromosomes with a distribution pattern characteristic of the full-length protein, indicating that it harbors the main determinant for interaction with LEDGF/p75. Although the C-terminal domain of IN was dispensable for nuclear/chromosomal localization, a fusion of the C-terminal IN fragment with enhanced green fluorescent protein was found exclusively in the nucleus, with a diffuse nuclear/nucleolar distribution, suggesting that the C-terminal domain may also play a role in the nuclear import of IN. In contrast to LEDGF/p75, its alternative splice variant, p52, did not interact with HIV-1 IN *in vitro* and in living cells. Finally, RNA interference-mediated knock-down of endogenous LEDGF/p75 expression abolished nuclear/chromosomal localization of IN. We conclude, therefore, that the interaction with LEDGF/p75 accounts for the karyophilic properties and chromosomal targeting of HIV-1 IN.**

The human immunodeficiency virus (HIV),[1] the causative agent of AIDS, belongs to the *Lentiviridae* genus of retroviruses. The early steps of HIV replication include reverse transcription of the diploid viral RNA genome into a double-stranded linear DNA replica and integration into a host cell chromosome. Reverse transcription takes place in the cytoplasm of the infected cell and results in the formation of a compact and stable preintegration complex (PIC), containing the viral reverse-transcribed genome and a number of virion-derived and cellular proteins. HIV and other lentiviruses are able to productively infect non-dividing, terminally differentiated cells, a feature distinguishing them from oncoretroviruses, which require cell division for productive infection (1–4). Previous work has characterized the nuclear import of HIV-1 PICs as an active, energy-dependent process (5), yet its mechanism has remained a puzzle. The determinants of HIV nuclear import that have been suggested so far are: the nuclear localization signals (NLSs) of the viral matrix (MA), Vpr, and integrase (IN) proteins, and the central DNA flap (for reviews see Refs. 6–8). The latter has been reported to be essential for nuclear import of HIV PICs and viral replication (9). Although the effect of the central DNA flap appears to be viral strain- and host cell-dependent (10, 11), its insertion in HIV-derived lentiviral vectors clearly augments transduction efficiency (12) and nuclear import (13). The karyophilic properties of MA and its role in HIV nuclear import (14) are unclear (15–17). Vpr is also not strictly required for HIV replication and DNA integration in non-dividing cells (15, 18, 19). It seems plausible that the DNA flap, MA, and Vpr, albeit redundant, exert additive and/or inter-dependent effects on HIV nuclear import. IN, on the other hand, is an attractive candidate for the role of the PIC import factor. (i) It is essential for the viral replication and spread of infection in primary cells and most T-cell lines (20); (ii) it is present in the PIC; (iii) its karyophilic properties have been demonstrated by many groups (for references see below). Unfortunately, mutations in IN have pleiotropic effects on viral replication, including alterations in viral particle morphology, defects in reverse transcription and integration (21), confounding a detailed genetic analysis of its functions.

HIV-1 IN is a 32-kDa protein, initially produced as part of the Gag-Pol precursor polyprotein and released after cleavage by the viral protease during maturation of the virion. IN is responsible for the catalysis of the insertion of the viral DNA into the host cell chromosome (for reviews see Refs. 22–24). Like all retroviral INs, HIV-1 IN is composed of three domains. The N-terminal domain harbors an HHCC-type zinc binding

* Work at the Rega Institute and in the Laboratory of Biomolecular Dynamics was supported by Concerted Research Action Fund Grant GOA/2001/02 from the Flemish Government. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Aspirant of the Fund for Scientific Research Flanders (Fonds voor Wetenschappelijk Onderzoek Flanders).

‖ Postdoctoral fellow of the Fund for Scientific Research Flanders (Fonds voor Wetenschappelijk Onderzoek Flanders).

** To whom correspondence may be addressed. Tel.: 32-16-332160; Fax: 32-16-332131; E-mail: zeger.debyser@med.kuleuven.ac.be.

‡‡ To whom correspondence may be addressed. Tel.: 32-16-327160; Fax: 32-16-327974; E-mail: yves.engelborghs@fys.kuleuven.ac.be.

[1] The abbreviations used are: HIV, human immunodeficiency virus; BSA, bovine serum albumin; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonic acid; Ct, C-terminal; EGFP, enhanced green fluorescent protein; GFP, green fluorescent protein; IN, integrase; LEDGF, lens epithelium-derived growth factor/transcription co-activator p75; MA, matrix protein; NLS, nuclear localization signal; Nt, N-terminal; Ni-NTA, nickel-nitrilotriacetic acid; PBS, phosphate-buffered saline; PIC, preintegration complex; siRNA, short interfering RNA.

domain and has been implicated in the multimerization of the protein (25). The core domain contains the catalytic site and possesses structural elements necessary for sequence-specific recognition of the viral long terminal repeat (26). The arginine/lysine-rich C-terminal domain of IN also contributes to the multimerization of the protein (27) and is thought to be involved in DNA binding. In addition, binding to DNA has been shown to induce oligomerization of HIV-1 IN *in vitro* (28). When expressed in or microinjected into human cells, HIV-1 IN accumulates in the nuclei (17, 29–33). During mitosis, IN stably associates with condensed chromosomes (30). Recent studies with digitonin-permeabilized cells have shown that nuclear import of HIV-1 IN can occur in the absence of cytosolic extracts, requires ATP hydrolysis, and is GTPase Ran-independent (34). Similar observations have been made on the nuclear import of the NLS receptor, importin $\alpha$ (35). Therefore, the virus either does not rely on the classical importin- and Ran-dependent nuclear import mechanism or is able to take advantage of an alternative pathway.

Recently, we have shown that in human cells HIV-1 IN forms a specific nuclear complex with lens epithelium-derived growth factor/transcription co-activator p75 (LEDGF/p75) (36). Recombinant LEDGF/p75 protein effectively promoted HIV-1 IN strand transfer activity *in vitro*. Of note, LEDGF/p75 was found to be up-regulated in HIV-infected cells (37). All these observations suggested that LEDGF/p75 could play a role in retroviral DNA integration. Although its precise cellular function remains elusive, several reports have implicated LEDGF/p75 in gene expression and cellular stress response (38–40). LEDGF/p75 has been reported to be a DNA-binding protein, with specificity for stress response DNA elements (40). According to the abundant mouse and human LEDGF/p75 mRNA-derived expressed sequence tags in GenBank™, LEDGF/p75 is expressed at all stages of development in a variety of organs and tissues, including skin, bone, thymus, brain, mammary gland, testis, and embryonic and hematopoietic stem cells. A second protein product, p52, can be produced from the same gene as LEDGF/p75 as a result of alternative splicing of the pre-mRNA (38, 41). At least *in vitro*, p52 was found to be a more general and stronger transcriptional co-activator than LEDGF/p75. In addition, the *in vitro* interaction of p52 with the ASF/SF2 splicing factor suggests that p52 might play a role in both transcription and splicing (42).

In the present work, we demonstrate that HIV-1 IN and LEDGF/p75 are intimate binding partners in human cells and that LEDGF/p75 is crucial for targeting IN to the nucleus and the chromosomes. We also show that both the N-terminal zinc binding and the catalytic core domains of IN are involved in the interaction with LEDGF/p75, whereby the core domain plays the dominant role.

## EXPERIMENTAL PROCEDURES

*EGFP and HcRed1 Fusion Constructs*—The full-length IN synthetic gene ($IN^s$) (30) or its fragments were PCR-amplified using *Pfu* DNA polymerase (Stratagene) and an appropriate pair of primers (see below). Each primer contained either a recognition site for *Xho*I (sense primers) or for *Hin*dIII (antisense primers). In-frame stop codons were included in the antisense primers. The primers used were: $C_s$, 5′-GGGGGCTCGAGCAGACTGCAGAAGCAGATCACC; $C_{as}$, 5′-GGGGAAGCTTGGACTTAGTCCTC; $N_s$, 5′-GGGGGCTCGAGCAGATTCCTGGACGGCATTGAC; $N_{as}$, 5′-GGGGAAGCTTACATAGCCTCGCC; $Core_s$, 5′-GGGGGCTCGAGCAGACACGGGCAGGTTGATTGC; $Core_{as}$, 5′-GGGGAAGCTTACTCTTTGGTCTGG. The PCR fragments were subcloned between the *Xho*I and *Hin*dIII restriction sites of the pEGFP-C2 vector (Clontech). A series of EGFP fusion constructs were made: pEGFP-$IN^s$, coding full-length HIV-1 IN, using the primers $N_s$ and $C_{as}$; pEGFP-$IN^s$/ΔC, expressing IN with a deletion of the C-terminal domain, using the primers $N_s$ and $Core_{as}$; pEGFP-$IN^s$/ΔN, expressing IN without the N-terminal zinc binding domain, using the primers $Core_s$ and $C_{as}$; pEGFP-$IN^s$/Nt (using the primers $N_s$ and $N_{as}$), pEGFP-$IN^s$/core (using $Core_s$ and $Core_{as}$), pEGFP-$IN^s$/Ct (using $C_s$ and $C_{as}$) coding the N-terminal, the central, and the C-terminal IN domain, respectively. The H12N mutant of the synthetic gene was engineered by PCR with the primers H12N (5′-GACGGCATTGACAAGGCTCAGGAGGAGAACGAGAAGTACCACTC) and T3 (5′-AATTAACCCTCACTAAAGGG) using *Pwo* DNA polymerase (Roche Applied Science) and $pIN^s$ (30) as the template. The second PCR was performed on the resulting amplicon with the primers $N_s$ and T3; the final PCR product was digested with *Xho*I and *Hin*dIII and subcloned into pEGFP-C2 to obtain pEGFP-$IN^s$(H12N). To obtain pHcRed1-$IN^s$, for expression of HIV-1 IN fused to the C terminus of the far-red fluorescent protein, HcRed1 (43), the *Xcm*I/*Eco*RI restriction fragment of $pIN^s$ (30) was cloned between the *Xho*I and *Eco*RI sites of pHcRed1-C1 (Clontech) after treatment of the *Xcm*I and the *Xho*I termini of the DNA fragments with T4 DNA polymerase.

To generate HcRed1-labeled LEDGF/p75, the *Bam*HI/*Eco*RI fragment of pCP6H75 (36), spanning the LEDGF/p75 open reading frame, was subcloned between the *Bgl*II and *Eco*RI sites of pHcRed1-C1. The coding sequences of HcRed1 and LEDGF/p75 were placed into frame via *Bsp*EI restriction and mung bean nuclease (Invitrogen, Groningen, The Netherlands) digestion followed by re-ligation to obtain the plasmid pHcRed1-p75. To generate an analogous fusion between p52 and HcRed1, the *Xho*I/*Eco*RI fragment of pHcRed1-p75 was replaced by the *Xho*I/*Eco*RI fragment of pKB-Nat52 (see below), resulting in pHcRed1-p52. The plasmid pEGFP-p75 expressing EGFP-tagged human LEDGF/p75 protein was obtained by inserting the *Bam*HI/*Eco*RI fragment of pCP6H75 (36) between the *Bgl*II and *Eco*RI sites of pEGFP-C2. Prior to ligation the *Bam*HI and *Bgl*II termini of the DNA fragments were filled in, using T4 DNA polymerase. All plasmid constructs used in this work were verified via sequence analysis to confirm absence of mutations.

*Plasmids for Bacterial Expression of LEDGF/p75, p52, HIV-1 IN, and HIV-1 $IN^{H12N}$*—The plasmid pKB-IN6H was used for the expression of the C-terminally tagged form of HIV-1 IN. To obtain pKB-IN6H, the IN gene (derived from the NL4–3 HIV-1 clone) was PCR-amplified from pINSD (44) using the primers 5′-AATACGACTCACTATAGGG (T7 promoter primer) and 5′-GCGCGTCGACATCCTCATCCTGTCTAC (INSalI primer); the resulting PCR fragment was digested with *Nde*I and *Sal*I and subcloned into the pET-20b(+) vector (Novagen).

To create pGM-INH12N-6H, for bacterial expression of the C-terminally $His_6$-tagged $IN^{H12N}$ mutant, a DNA fragment containing the HIV-1 IN open reading frame with the mutation was engineered in two consecutive PCR reactions. First, pKB-IN6H was used as the template and a PCR was performed with the forward primer 5′-GATAAGGCCCAAGAAGAAAATGAGAAATATCACAG and the INSalI primer (see above). The resulting amplicon was used as template in the second reaction with the primer 5′-ATATACATATGTTTTTAGATGGAATAGATAAGGCCCAAG and the INSalI primer. The final PCR fragment was digested with *Nde*I and *Sal*I and cloned into the pET-20b(+) vector.

The constructs pCP-Nat75 and pKB-Nat52 were used for bacterial expression of non-tagged LEDGF/p75 and p52 proteins, respectively. A DNA fragment, containing the LEDGF/p75 open reading frame, was amplified using pCP-6H75 (36) as a template and 5′-TGACTCGCGATTTCAAACC and 5′-CCGCGAATTCTAGTTATCTAGTGTAGAATCCTTC as the primers. To obtain the pCP-Nat75 plasmid, this PCR fragment was digested with *Eco*RI and inserted between the *Nde*I and *Eco*RI sites of the pRSETB vector (Invitrogen). Prior to ligation, the *Nde*I terminus of the vector DNA was filled in using T4 DNA polymerase. To produce pKB-Nat52, a DNA fragment containing the p52 open reading frame was constructed by PCR using the primers 5′-TGACTCGCGATTTCAAACC and 5′-GGCGAATTCTACTGTAGATTACATGTTGTTGGTGCTCAGTTTCCATTTGTTCC. The resulting fragment was digested with *Eco*RI and cloned between the *Eco*RI and *Nde*I sites of pRSETB.

*Cell Culture and Transfections*—HeLa cells were grown in Dulbecco's modified Eagle's medium supplemented with GlutaMAX™, 10% fetal calf serum, and 0.02 mg/ml gentamicin (Invitrogen) at 37 °C and 5% $CO_2$ in a humidified atmosphere. Cells were seeded the day before transfection in 8-well LabTek chambered coverglass cuvettes (VWR International, Leuven, Belgium). Transfection of HeLa cells was performed at ~80% confluency using Lipofectamine 2000 reagent (Invitrogen) with 0.36 $\mu$g of plasmid, following the instructions from the manufacturer. Double transfections were accomplished following the same protocol except that transfections were performed with 0.18 $\mu$g of each plasmid.

For transient LEDGF/p75 knock-down experiments, HeLa cells were transfected with small interfering RNA (siRNA) synthetic duplexes using the Gene Silencer transfection reagent (Gene Therapy Systems) or co-transfected with siRNA plus the pEGFP-$IN^s$ plasmid using Lipo-

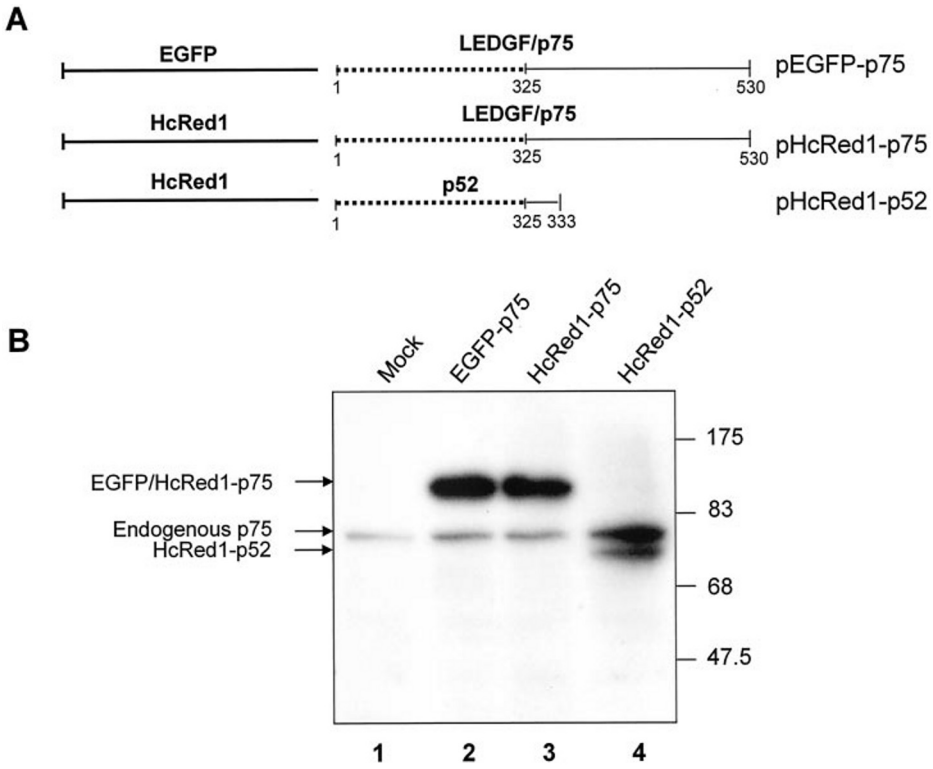


FIG. 1. **HcRed1- and EGFP-fused LEDGF/p75 and p52 proteins: constructs and expression in HeLa cells.** *A*, LEDGF/p75 and the splice variant, p52, were expressed as fusions with EGFP or HcRed1. LEDGF/p75 (530 amino acids) and p52 (333 amino acids) have 325 N-terminal residues in common (indicated by the *dotted bar*). The 8 amino acids at the C terminus of p52 are distinct from LEDGF/p75. The names of the relevant expression plasmids are indicated. *B*, transient expression of HcRed1-p75, EGFP-p75, and HcRed1-p52. HeLa cells were transfected as explained under "Experimental Procedures." Expression was analyzed at 48 h after transfection by Western blotting using monoclonal anti-LEDGF/p75-p52 antibody. *Lane 1*, mock-transfected HeLa cells; *lanes 2–4*, cells were transfected with pEGFP-p75 (*lane 2*), pHcRed1-p75 (*lane 3*), and pHcRed1-p52 (*lane 4*). Endogenous p52 protein was not detectable by Western blotting with the antibody used. The positions and sizes of the molecular size markers are shown. Positions of HcRed1-and EGFP-fused LEDGF/p75, HcRed1-p52 and the endogenous LEDGF/p75 are indicated at the *left side* of the panel.

fectamine 2000 according to established protocols (45). Fetal calf serum was added to the medium at 5 h after transfection. The effect of the RNA interference was studied 48–60 h after transfection.

*Western Blotting and Indirect Immunofluorescence*—For Western blot detection, protein extracts separated in 11% SDS-PAGE gels were electroblotted onto polyvinylidene difluoride membranes (Bio-Rad, Nazareth, Belgium). The polyclonal anti-GFP antibody was purchased from Invitrogen; the mouse anti-LEDGF p75/p52 antibody was from BD Biosciences (Erembodegem, Belgium). The polyclonal rabbit anti-HIV-1 IN antibody has been described previously (30). Secondary horseradish peroxidase- or alkaline phosphatase-conjugated goat anti-mouse and goat anti-rabbit antibodies were from Dako Diagnostics (Leuven, Belgium). Detection was carried out using ECL+ chemiluminescent horseradish peroxidase substrate (Amersham Biosciences) or with 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium chromogenic substrate for alkaline phosphatase. The broad range prestained protein marker mix (New England Biolabs, Hertfordshire, United Kingdom) was used for estimation of the molecular weights.

Indirect immunofluorescent detection of endogenous LEDGF/p75 was performed as previously described (36). Cells grown in LabTek II glass chamber slides (VWR International) were fixed by incubation with 4% formaldehyde in PBS for 10 min, washed with PBS, and permeabilized/fixed with ice-cold methanol. The cells were then blocked in 10% fetal calf serum, 20 mM ammonium chloride, and PBS for 30 min and incubated with monoclonal anti-LEDGF p75/p52 antibodies (used at a dilution of 1:300), followed by Alexa-488- or Alexa-568-conjugated goat anti-mouse antibody (Molecular Probes, Leiden, The Netherlands). The nuclear DNA was stained with 5 $\mu$M ToPro3 iodide (Molecular Probes).

*Laser Scanning Microscopy and Image Analysis*—Confocal microscopy was performed using an LSM 510 unit (Zeiss, Zaventem, Belgium). SYTO 17 (Molecular Probes) was used to stain DNA of live cells. Prior to image acquisition, cells were washed with serum-free OptiMEM (Invitrogen). All two- and three-color images with a resolution of 1024 × 1024 pixels were acquired in the multi-track mode. EGFP was excited at 488 nm (AI laser), HcRed1 at 543 nm (HeNe laser), and SYTO 17 at 633 nm (HeNe laser). After the main dichroic beam splitter (HFT 488/543 for EGFP and HcRed1, HFT 488/543/633 for EGFP and SYTO 17) the fluorescence signal was divided by a secondary dichroic beam splitter (NFT 635 VIS or NFT 543) and detected in the separate channels using the appropriate filters: BP 505–530 (for EGFP), LP560 (for HcRed1), and LP650 (for SYTO 17). In this set-up, no cross-talk between the green and red channels was observed.

The colocalization of fluorescently-tagged proteins in the nucleus was quantified using the Image-Pro Plus version 4.5 software (Media Cybernetics, Carlsbad, CA) and expressed in terms of the correlation coefficient ($r_p$) (same as Pearson's $r$ correlation) (46).

$$r_p = \frac{\sum_i ((S1_i - \langle S1\rangle)\cdot(S2_i - \langle S2\rangle))}{\sqrt{\sum_i (S1_i - \langle S1\rangle)^2 \cdot \sum_i (S2_i - \langle S2\rangle)^2}} \qquad \text{(Eq. 1)}$$

*S1* and *S2* represent the signal intensities of pixels in the first and second channel, respectively; $\langle S1\rangle$ and $\langle S2\rangle$ are the average intensity of the first channel and second channel, respectively. The correlation coefficient is a value between −1 and +1, with −1 corresponding to negative correlation between images and +1 corresponding to a total overlap of the images from the two channels. It reflects similarity of image patterns and does not depend on intensities of the images.

*Recombinant Proteins*—Non-tagged LEDGF/p75 and p52 proteins were produced from the plasmids pCP-Nat75 and pKB-Nat52, respectively, in the Endo I-free PC1 *Escherichia coli* host strain (*E. coli* B, BL21(DE3), Δ*endA*::Tc$^R$, pLysS) (47). Expression was induced in LB medium at 29 °C by addition of 0.5 mM isopropyl-1-thio-β-D-galactopyranoside. Cells harvested 3 h after induction were disrupted using a French press in 450 mM NaCl, 30 mM Tris, pH 7.0. The supernatant obtained by centrifugation of the lysate was passed through a 1-ml HiTrap heparin column (Amersham Biosciences, Uppsala, Sweden) to capture LEDGF/p75 or p52, and the protein was eluted by a linear gradient of NaCl concentration in 30 mM Tris, pH 7.0. The fractions containing LEDGF/p75 or p52 were pooled and further purified by cation exchange chromatography on a 1-ml HiTrap SP Sepharose column (Amersham Biosciences).

To produce C-terminally $His_6$-tagged wild type HIV-1 IN, PC1 *E. coli* cells harboring pKB-IN6H were grown in LB medium to an optical density of 0.8 and induced by addition of 0.5 mM isopropyl-1-thio-β-D-galactopyranoside, at 29 °C for 3 h. The protein was purified essentially as described for N-terminally tagged HIV-1 IN (48). In brief, cells were lysed using a French press in 1 M NaCl, 7.5 mM CHAPS, 30 mM Tris, pH 7.4, and the soluble $His_6$-tagged IN protein was enriched by batch adsorption to Ni-NTA-agarose (Qiagen, Hilden, Germany). Protein eluted with 200 mM imidazole, 1 M NaCl, 7.5 mM CHAPS, 30 mM Tris, pH 7.4 was further purified on a 1-ml HiTrap heparin column (Amersham Biosciences). The $His_6$-tagged H12N mutant was induced in PC1 cells from pGM-INH12N-6H and purified in a similar way. Purified recombinant LEDGF/p75, p52, IN, and IN$^{H12N}$ proteins were concen-

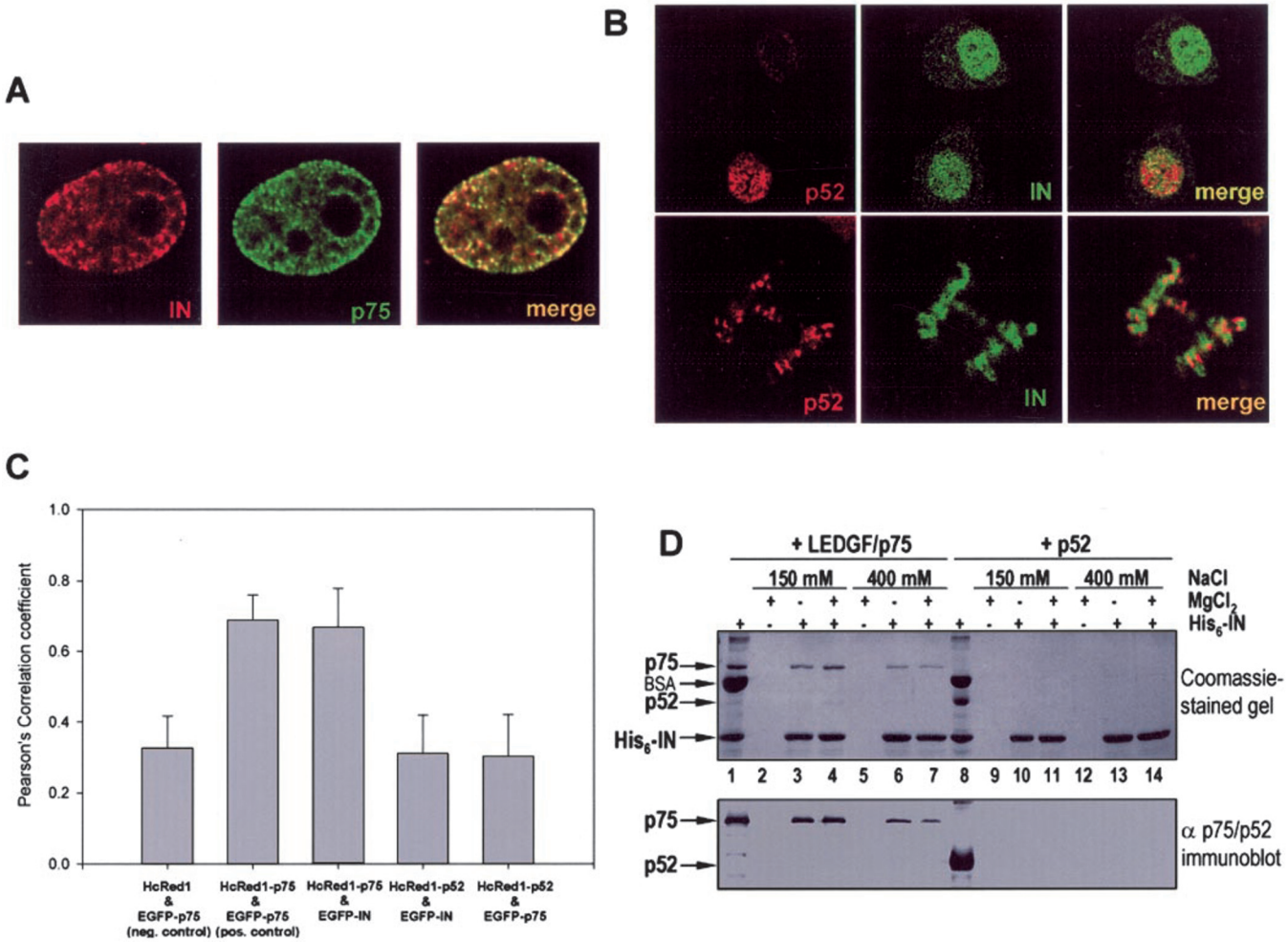


FIG. 2. **HIV-1 IN interacts with LEDGF/p75, but not with p52.** *A*, confocal images of HeLa cells co-expressing HcRed1-IN with EGFP-p75: HcRed1-IN (*left panel*), EGFP-p75 (*middle panel*), and merge (*right panel*). Both fusion proteins display a nuclear staining and colocalize with each other. *B*, confocal images of HeLa cells co-expressing HcRed1-p52 with EGFP-IN: HcRed1-p52 (*left panels*), EGFP-IN (*middle panels*), and merge (*right panels*). p52 and IN show a discrete nuclear staining during interphase (*upper three panels*); however, their nuclear patterns are distinct form each other. During mitosis (*lower three panels*), both proteins are concentrated on the condensed chromosomes, but no clear colocalization can be observed. *C*, numeric analysis of colocalization. The degree of colocalization was quantified as described under "Experimental Procedures" and expressed in terms of Pearson's correlation coefficient ($r_p$). Average values and standard deviations obtained from minimum 10 cells for each condition are shown. The correlation coefficients, determined for the HcRed1-p52/EGFP-IN and HcRed1-p52/EGFP-p75 pairs, are approximately equal to that for co-expressed HcRed1 and EGFP-p75 ($r_p = 0.32 \pm 0.09$). This can be viewed as a negative control for colocalization. HcRed1-p75 and EGFP-IN co-localize with an $r_p$ of $0.67 \pm 0.11$, which is very close to the $r_p$ value for the HcRed1-p75/EGFP-p75 pair ($0.69 \pm 0.07$) (positive control). *D*, recombinant LEDGF/p75 forms a stable complex with $His_6$-tagged HIV-1 IN *in vitro*, whereas p52 does not. The pull-down assay was performed as described under "Experimental Procedures." The specific buffer conditions, *i.e.* NaCl concentrations and presence or absence of $MgCl_2$ in the reactions are indicated *above* the gels. Recombinant LEDGF/p75 (*lanes 3–7*) or p52 (*lanes 10–14*) was incubated with $His_6$-tagged HIV-1 IN, and the complexes were recovered on a $Ni^{2+}$-chelating resin. In the control samples (*lanes 2* and *9*), recombinant IN was omitted. *Lanes 1* and *8* reflect protein input in the reactions; BSA, $His_6$-IN, LEDGF/p75, or p52 were loaded in the same amounts as were present in the binding reactions. The respective positions of LEDGF/p75, BSA, p52, and $His_6$-IN are indicated on the *left side* of the gel; the gel was stained using Coomassie R250. The same samples were analyzed by Western blotting with anti-LEDGF p75/p52 antibody, for a more sensitive detection of the p52 protein.

trated by ultrafiltration using Centricon 10 (Millipore, Brussels, Belgium), supplemented with 5 mM dithiothreitol plus 10% glycerol, and kept frozen at −80 °C.

*$His_6$ Tag Integrase Pull-down Assay*—Binding of IN to LEDGF/p75 or p52 was assayed in 25 mM Tris-HCl, pH 7.4, 0.1% Nonidet P-40, 20 mM imidazole containing 100 or 400 mM NaCl, in the presence or absence of 1 mM $MgCl_2$ (binding buffer). 1 μg of recombinant $His_6$-tagged HIV-1 IN or $His_6$-$IN^{H12N}$ was incubated with 1–3 μg of LEDGF/p75 or p52 in 200 μl of binding buffer supplemented with 2 μg of bovine serum albumin (BSA). Following a 30-min incubation at 4 °C, the mixtures were centrifuged to remove aggregated protein. The samples were supplemented with 40 μl of Ni-NTA-agarose and stirred for an additional 30 min. The agarose beads were recovered by centrifugation and washed with 500 μl of binding buffer. Bound proteins were eluted in 40 μl of binding buffer supplemented with 200 mM imidazole and 1% SDS and analyzed by SDS-PAGE electrophoresis followed by staining with Coomassie-R250 or by Western blotting with monoclonal anti-LEDGF p75/p52 antibody for more sensitive detection.

*Synthetic siRNA Duplexes*—The LEDGF/p75-specific siRNA duplexes L1 and L3 were designed according to Elbashir *et al.* (45) and were synthesized by Xeragon (Germantown, MD) and Dharmacon Research (Lafayette, CO), respectively. The GFP22 (49) siRNA (Xeragon) and the Scramble II duplex (SDII) (Dharmacon) were used as non-interfering controls. All siRNA duplexes were purchased in pre-annealed form and were deprotected and treated according to the instructions from the manufacturer. The L1, L3, and SDII duplexes contained a pair of 3′-terminal deoxythymidine (TT) overhangs each. The sequences of the siRNA duplexes and locations of their target sites within LEDGF/p75 mRNA are shown in Fig. 9*A*.

## RESULTS

*Colocalization of Fluorescently Labeled HIV-1 IN and LEDGF/p75 in Living Cells*—We have recently reported that HIV-1 IN forms a specific complex with the endogenous nuclear LEDGF/p75 protein in human cells (36). Indirect immuno-

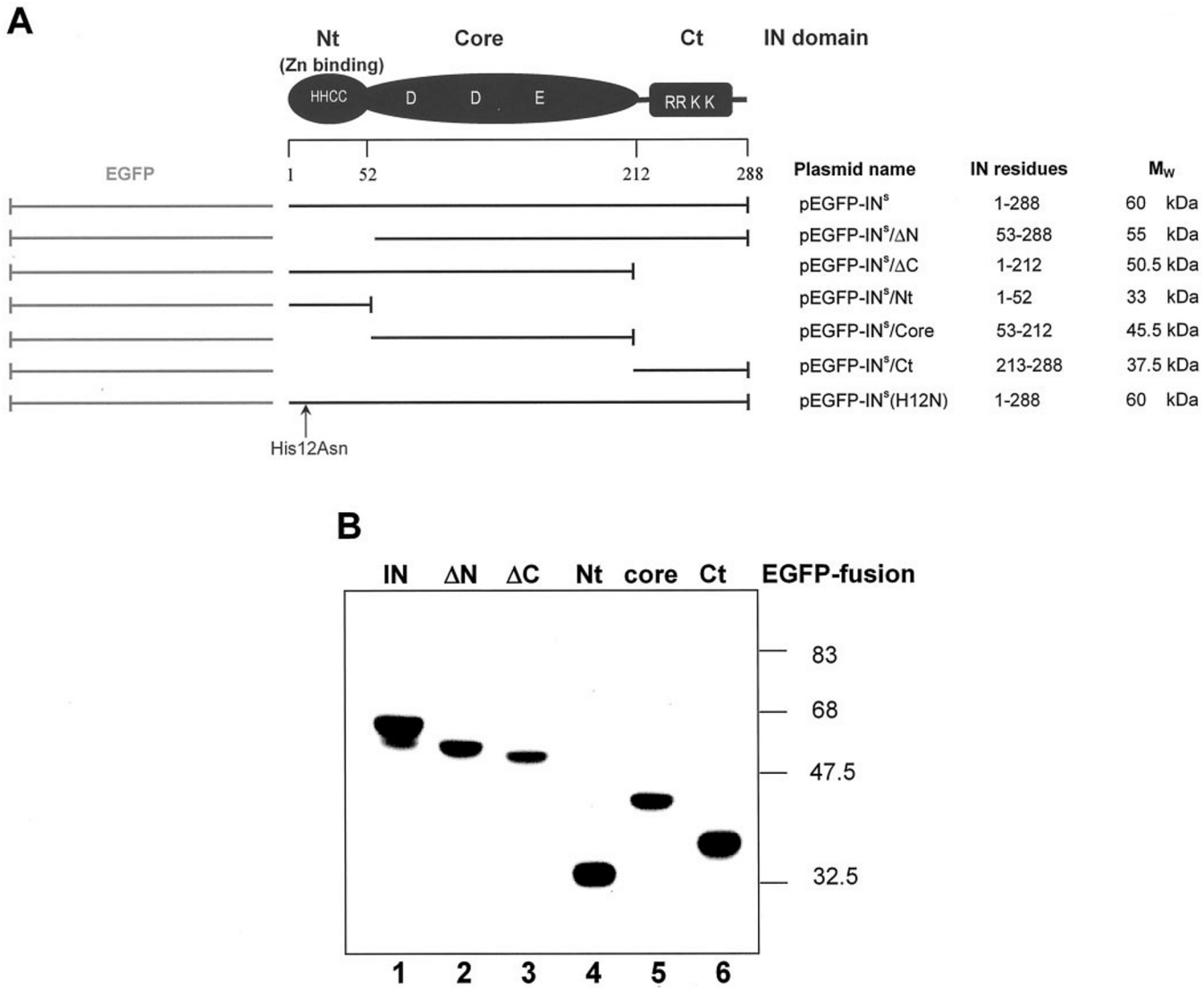


FIG. 3. **Structures of the EGFP-fused HIV-1 IN mutants used in this study.** *A*, the structural domains of HIV-1 IN and the design of the EGFP-fused HIV-1 IN mutants. The N-terminal HHCC zinc binding domain, catalytic core, containing the DD(35)E motif, and the positively charged C-terminal domain are indicated. The names of the corresponding expression plasmids and the expected molecular weights of the expressed fusion proteins are shown. The EGFP-IN$^{H12N}$ mutant contains full-length IN carrying a point mutation in the zinc binding domain. *B*, transient expression of the EGFP-fused IN mutants in HeLa cells. Expression was analyzed 48 h after transfection by Western blotting using polyclonal anti-IN antibody. Material from $8 \times 10^3$ cells was loaded in each well, except for the EGFP-IN/core sample, where twice as much was used to achieve an equal detection. Cells were transfected with pEGFP-IN$^s$ (*lane 1*), pEGFP-IN$^s$/ΔN (*lane 2*), pEGFP-IN$^s$/ΔC (*lane 3*), pEGFP-IN$^s$/Nt (*lane 4*), pEGFP-IN$^s$/core (*lane 5*), and pEGFP-IN$^s$/Ct (*lane 6*). The positions and sizes of the molecular weight markers are indicated at the *right side* of the panel.

staining of FLAG-tagged IN and LEDGF/p75 showed that the two proteins co-localize in the nucleus. For studies in intact cells, we constructed vectors for expression of EGFP- and HcRed1-tagged human LEDGF/p75 and HIV-1 IN (Figs. 1*A* and 3*A*, respectively). Western blot analysis of HeLa cells transfected with the fusion constructs using anti-LEDGF p75/p52 antibodies (Fig. 1*B*) or anti-HIV-1 IN (Fig. 3*B*) and anti-EGFP (data not shown) revealed specific bands at the expected positions. Separately expressed full-length HIV-1 IN and LEDGF/p75, fused to EGFP or HcRed1, showed nuclear localization and distribution patterns in HeLa cells, in agreement with previous reports (36, 50). When HcRed1-IN was co-expressed with EGFP-p75 (Fig. 2*A*) or EGFP-IN together with HcRed1-p75 (Fig. 4*A*), a clear nuclear colocalization was observed, suggesting that fusing both binding partners to the fluorescent proteins did not abolish their interaction. As a quantitative measure of colocalization, we determined the correlation coefficient ($r_p$) (same as Pearson's $r$ correlation; see Ref. 46) for the distributions of the fluorescently tagged IN and LEDGF/p75 proteins obtained from the green and red channels (Fig. 2*C*). The correlation coefficient for EGFP-IN and HcRed1-p75 in the nucleus was determined to be $0.67 \pm 0.11$, which is very close to the $r_p$ value for the colocalization of EGFP-p75 and HcRed1-p75 ($0.69 \pm 0.07$), which served as a positive control (Fig. 2*C*). In contrast, when EGFP-p75 fusion and free HcRed1 were co-expressed, neither colocalization nor cross-talk between the green and the red channels was observed, and a low correlation coefficient ($r_p = 0.32 \pm 0.09$) was obtained (Fig. 2*C*).

*p52 Does Not Interact with HIV-1 IN*—A second protein, p52, can be expressed from the same gene as LEDGF/p75 as a result of alternative splicing of the pre-mRNA (38, 41). p52 (333 amino acids) and LEDGF/p75 (530 amino acids) share 325 N-terminal amino acid residues (Fig. 1*A*). The endogenous levels of p52 in HeLa, HEK-293, or CEM cells were too low to be detected by Western blotting (data not shown), although, at least in the case of recombinant proteins, the sensitivity of detection of p52 was higher than that of LEDGF/p75 (compare *lanes 1* and *8* on the blot shown in Fig. 2*D*). The fact that p52

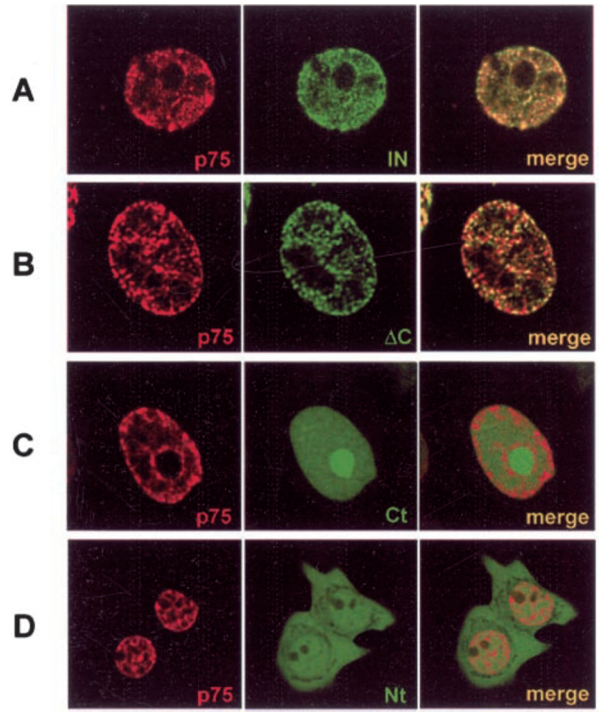


FIG. 4. **Intracellular distribution of EGFP-IN, EGFP-IN/ΔC, EGFP-IN/Ct, and EGFP-IN/Nt in the presence of HcRed1-p75.** HeLa cells were co-transfected with pHcRed1-p75 plus pEGFP-IN$^s$ (*A*), pEGFP-IN$^s$/ΔC (*B*), pEGFP-IN$^s$/Ct (*C*), and pEGFP-IN$^s$/Nt (*D*). Expression of the fluorescent fusions was studied by confocal laser scanning microscopy at 24 h after transfection. The HcRed1, EGFP signals, and their overlays are shown in the *left*, *middle*, and *right* image of each panel, respectively.

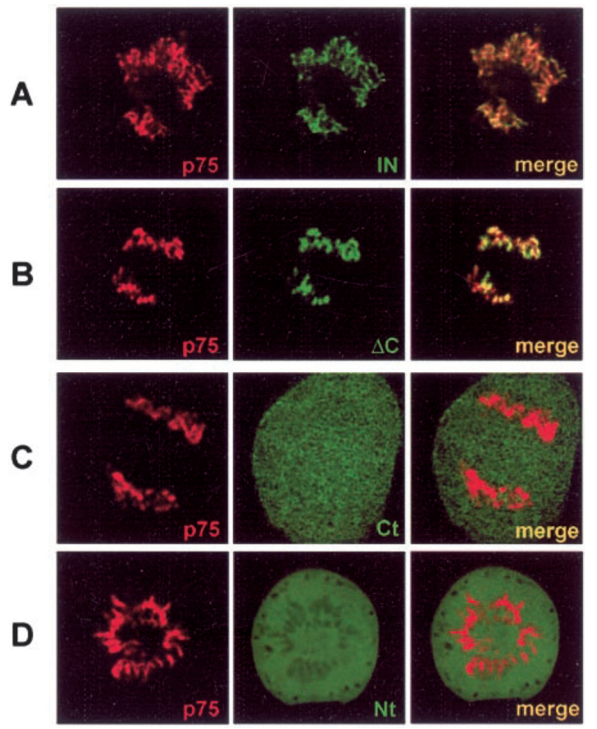


FIG. 5. **Intracellular distribution of the EGFP-IN deletion mutants during mitosis.** HcRed1-p75 is concentrated on chromosomes in all conditions. HcRed1-p75, EGFP-IN deletion, and the merged images are shown in the *left*, *middle*, and *right* images of each panel. EGFP-IN (*A*) and EGFP-IN/ΔC (*B*) colocalize with the condensed chromosomes and HcRed1-p75, whereas the EGFP-IN/Ct and /Nt (*C* and *D*, respectively) do not.

TABLE I
*Analysis of the intracellular distribution and the colocalization with LEDGF/p75 of the separate domains of HIV-1 integrase*

| EGFP fusion | Nuclear localization[a] | Colocalization with mitotic DNA | Colocalization with LEDGF/p75 |
|---|---|---|---|
| IN | +++ | Yes | Yes |
| IN/ΔC | +++ | Yes | Yes |
| IN/Core | =/++ | No/yes[b] | No/yes[b] |
| IN$^{H12N}$ | =/++ | No/yes[b] | No/yes[b] |
| IN/ΔN | ++/− | No/yes[b] | No/yes[b] |
| IN/Nt | = | No | No |
| IN/Ct | +++ | No | No |

[a] Accumulation of the EGFP-fused protein in the nucleus in comparison to the cytoplasm. +++, predominantly nuclear; =, evenly distributed throughout the whole cell; ++/−, present in both nucleus and cytoplasm, more concentrated in the nucleus; =/++, nuclear accumulation is observed upon overexpression of LEDGF/p75.

[b] Colocalization with mitotic DNA and LEDGF/p75 is observed upon overexpression of HcRed1-p75. The degree of concentration in the nucleus and the colocalization with LEDGF/p75 and condensed chromosomes are dependent on the expression levels of LEDGF/p75.

was not detected in a complex with HIV-1 IN (36) could be explained by low expression levels of p52 in the original cell line. To establish whether HIV-1 IN can associate with p52, we co-expressed HcRed1-tagged p52 and EGFP-IN in HeLa cells. The distribution of the HcRed1-p52 fusion protein was as shown previously for EGFP-tagged p52 (50) and was clearly different from the distribution of EGFP-IN (Fig. 2*B*). The mean correlation coefficient for HcRed1-p52 and EGFP-IN colocalization in the nucleus was determined to be 0.32 ± 0.09, which is close to the $r_p$ value for our negative controls (Fig. 2*C*). In addition, HcRed1-p52 did not show colocalization with EGFP-p75 ($r_p$ = 0.30 ± 0.12), when the two fusions were co-expressed in HeLa cells (Fig. 2*C*).

Direct interaction between IN and LEDGF/p75 was confirmed *in vitro* by a $His_6$ tag IN pull-down assay with $Ni^{2+}$-chelating agarose (Fig. 2*D*). After incubating $His_6$-tagged HIV-1 IN with recombinant LEDGF/p75, the samples were centrifuged to remove aggregated protein. No LEDGF/p75 could be detected in these pellets; however, a fraction of IN was observed in the aggregated form when the binding buffer with 100 mM NaCl was used (data not shown). After binding to Ni-NTA-agarose and elution, recombinant LEDGF/p75 could be readily recovered in a stable complex with $His_6$-tagged IN in a variety of conditions (*lanes 3–7*). In accord with the observed lack of intracellular colocalization between the EGFP-IN and HcRed1-p52 fusion proteins, p52 failed to form a stable complex with $His_6$-tagged IN in similar conditions and was undetectable in the pulled-down fractions (*lanes 10–14*). In addition, unlike LEDGF/p75, recombinant p52 protein failed to stimulate enzymatic activity of HIV-1 IN in the mini-HIV assay (36) (data not shown).

*Intracellular Distribution of EGFP-IN/ΔC, /Nt, and /Ct Fusions*—Next, we wanted to study the contribution of the individual HIV-1 IN domains to the nuclear localization of the protein and their interaction with LEDGF/p75 in living cells. For this purpose, we prepared a series of EGFP-fused HIV-1 IN deletion constructs (Fig. 3*A*). Transient transfection of HeLa cells resulted in efficient expression of the fusion proteins with predicted molecular masses, as observed by Western blotting with anti-IN antibodies (Fig. 3*B*) or anti-GFP antibodies (data not shown). Fusions were visualized by laser scanning microscopy (Figs. 4–7). An overview of the specific properties of the different IN deletion mutants in terms of their intracellular distribution and interaction with mitotic DNA and LEDGF/p75 is given in Table I.

Transient expression of EGFP-IN resulted in a specific irregular nuclear distribution pattern, characteristic for IN (Fig. 4*A*). Strikingly, the absence of the C-terminal DNA binding

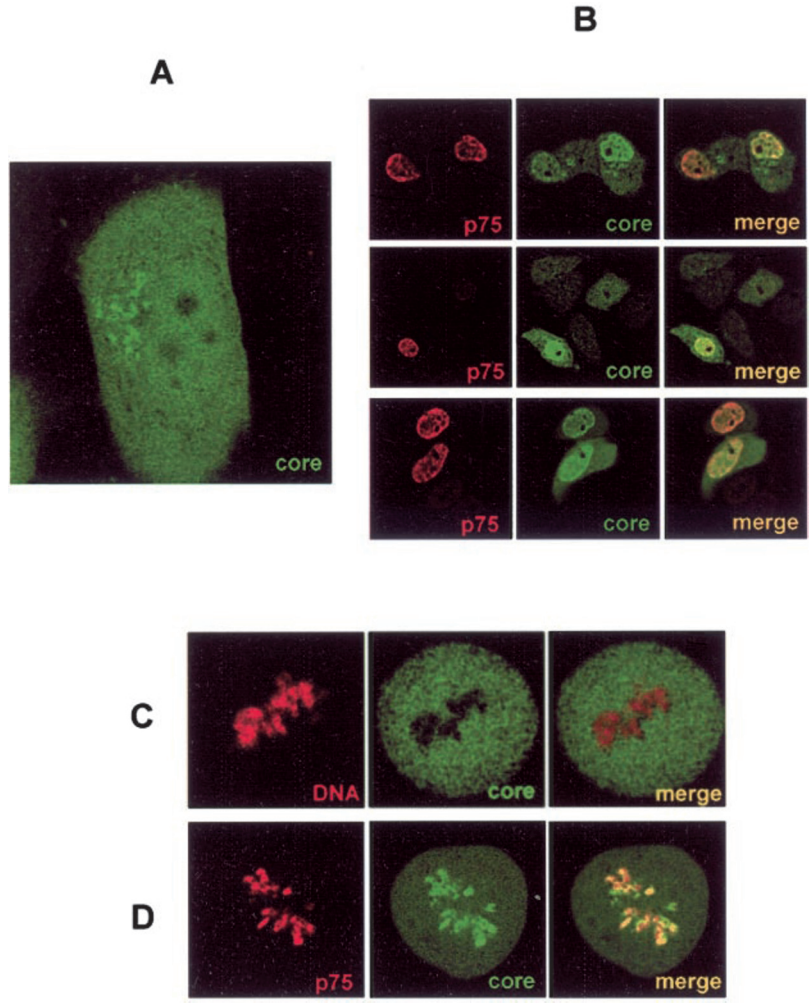


FIG. 6. **Overexpression of LEDGF/p75 recruits the core domain fragment of IN to the nucleus and chromosomes.** *A*, EGFP-IN/core distribution in HeLa cells under conditions of normal endogenous expression of LEDGF/p75. The core domain is distributed throughout the whole cell. *B*, upon overexpression of HcRed1-p75, the core domain is recruited to the nucleus. The distributions of co-expressed HcRed1-p75 and EGFP-IN/core and the overlay of both images are shown. *C*, HeLa cell expressing EGFP-IN/core during mitosis. No interaction of EGFP-IN/core with the condensed chromosomes could be detected. Chromosomal DNA was stained with the cell-permeable dye SYTO 17. *D*, upon overexpression of HcRed1-p75, the core domain fragment of IN becomes concentrated on the condensed chromosomes.

domain of IN did not have any significant effect on the intracellular distribution. The observed nuclear pattern for the EGFP-IN/ΔC deletion mutant was the same as for the full-length protein (Fig. 4*B*). The presence of the other IN domains, however, did play a crucial role in IN intracellular localization. Although EGFP-IN/Nt showed a dispersed cytoplasmic-nuclear localization pattern (Fig. 4*D*), the EGFP-IN/Ct was exclusively nuclear and concentrated in nucleoli (Fig. 4*C*), as confirmed by indirect immunostaining using anti-nucleolin antibody (data not shown). Although the predicted molecular mass of EGFP-IN/Ct protein (37.5 kDa for a monomer) is smaller than the exclusion limit of the nucleopore complex (45–60 kDa) (for review, see Ref. 51), taking its diffuse nuclear distribution into account, these data suggest the absence of a specific intranuclear binding partner and favor the idea of an active nuclear import mechanism for EGFP-IN/Ct.

Both EGFP-IN and EGFP-IN/ΔC showed clear colocalization with co-expressed HcRed1-p75 (Fig. 4, *A* and *B*). The distinct intracellular distribution patterns observed for EGFP-IN/Ct and EGFP-IN/Nt correlated with the absence of colocalization with HcRed1-LEDGF/p75 (Fig. 4, *C* and *D*, respectively). Whereas the nuclear localization of full-length EGFP-IN and EGFP-IN/ΔC fusions appeared to be more pronounced in cells expressing HcRed1-p75 (data not shown), no significant changes in the intracellular distribution pattern of EGFP-IN/Nt and EGFP-IN/Ct were observed upon co-expression of HcRed1-p75.

Both HIV-1 IN and LEDGF/p75 have been shown to be associated with condensed chromosomes during mitosis (30, 50). In agreement with these reports, fusions of full-length HIV-1 IN and LEDGF/p75 to fluorescent proteins displayed strong chromosomal staining in mitotic HeLa cells when expressed together (Fig. 5*A*) or separately (data not shown). Unexpectedly, the absence of the C-terminal DNA binding domain in the EGFP-IN/ΔC fusion did not abolish or diminish its association with chromosomes (Fig. 5*B*). In addition, the EGFP-IN/Ct protein, containing only the C-terminal domain, did not display detectable chromosomal localization, suggesting that the C terminus of IN is not involved in binding to chromosomal DNA (Fig. 5*C*). Because EGFP-IN/Nt displayed a diffuse distribution in mitotic cells (Fig. 5*D*), the N-terminal domain is essential but not sufficient for the chromosomal association.

*Intracellular Distribution of EGFP-IN/core and EGFP-IN/ΔN*—In contrast with the other IN deletion mutants, the distribution of the EGFP-IN/core and EGFP-IN/ΔN displayed a dramatic dependence on the intracellular levels of LEDGF/p75. At endogenous levels of LEDGF/p75, EGFP-IN/core was distributed diffusively throughout the whole cell, with some dot-like structures close to the nucleus (Fig. 6*A*). EGFP-IN/ΔN was mostly nuclear, although even at low expression levels, it was readily detectable in the cytoplasm (Fig. 7*A*). Strikingly, in cells co-expressing HcRed1-p75, both EGFP-IN/core and EGFP-IN/ΔN accumulated in the nucleus, with a pattern similar to that of full-length IN, colocalizing with HcRed1-p75 (Figs. 6*B* and 7*B*, respectively). Moreover, nuclear localization of EGFP-IN/core and EGFP-IN/ΔN was clearly more pronounced in the cells expressing higher levels of HcRed1-p75 (compare different cells in Fig. 6*B* for EGFP-IN/core). A similar

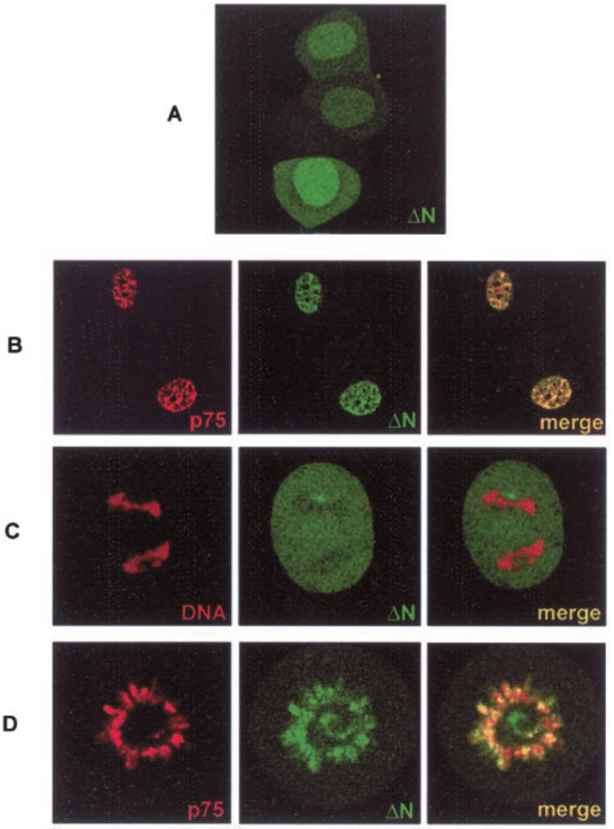


FIG. 7. **Overexpression of LEDGF/p75 recruits EGFP-IN/ΔN to the nucleus and chromosomes.** *A*, EGFP-IN/ΔN distribution in HeLa cells under conditions of normal endogenous expression of LEDGF/p75. EGFP-IN/ΔN is accumulated in the nucleus; however, a cytoplasmic pool is visible even at low expression levels of EGFP-IN/ΔN. *B*, upon overexpression of HcRed1-p75, EGFP-IN/ΔN is recruited to the nucleus. The distributions of co-expressed HcRed1-p75 and EGFP-IN/ΔN and the overlay of both images are shown. *C*, HeLa cell expressing EGFP-IN/ΔN during mitosis. No interaction of EGFP-IN/ΔN with the condensed chromosomes could be detected. Chromosomal DNA was stained with the cell-permeable dye SYTO 17. *D*, upon overexpression of HcRed1-p75, EGFP-IN/ΔN becomes concentrated on the condensed chromosomes.

dependence was observed in mitotic cells. When expressed separately from HcRed1-p75, EGFP-IN/core and EGFP-IN/ΔN showed a diffuse distribution (Figs. 6*C* and 7*C*, respectively). When co-expressed with HcRed1-p75, both EGFP-IN/core and EGFP-IN/ΔN became recruited to the chromosomes (Figs. 6*D* and 7*D*, respectively). These observations suggest that the core domain of IN is necessary and sufficient for interaction with LEDGF/p75. However, as EGFP-IN/ΔC displayed the same distribution pattern as full-length EGFP-IN at endogenous levels of LEDGF/p75, the N-terminal zinc binding domain of IN enhances affinity of the interaction. In addition, although at endogenous levels of LEDGF/p75, EGFP-IN/ΔN was concentrated in the nucleus, its intranuclear distribution appeared diffuse, more like that of the EGFP-IN/Ct fusion.

*Mutational Disruption of the Zinc Binding Domain Reduces Affinity of IN for LEDGF/p75*—The N-terminal domain of IN folds into a rigid structure stabilized by the coordination of a single $Zn^{2+}$ cation by four conserved residues (His-12, His-16, Cys-41, and Cys-44) (52, 53). Zinc remains associated with recombinant IN through the purification process (25). Mutations of the zinc coordinating residues result in reduced $Zn^{2+}$ binding by IN and destabilize the structure of the N-terminal domain of the protein (44, 54, 55). We first studied the effect of the His-12 to Asn mutation in IN ($IN^{H12N}$) on the binding to LEDGF/p75 *in vitro*. In our $His_6$ tag integrase pull-down assay, a stable complex between recombinant IN carrying the H12N mutation and LEDGF/p75 could not be detected (Fig. 8*A*), suggesting that the structural fold of the zinc binding domain is important for the interaction with LEDGF/p75. Next, we introduced the same mutation into the EGFP-IN fusion construct (Fig. 3*A*). The expression level of EGFP-$IN^{H12N}$ in HeLa cells was similar to that of the wild type EGFP-IN protein (Fig. 8*B*). Remarkably, however, the mutant lost karyophilic properties and was distributed in a diffuse manner throughout the whole cell (Fig. 8*C*). As was observed for EGFP-IN/core and EGFP-IN/ΔN, EGFP-$IN^{H12N}$ became recruited to the nucleus upon overexpression of HcRed1-p75, and colocalized with the latter (Fig. 8*D*). In addition, co-expression of HcRed1-p75 also restored the chromosomal localization of EGFP-$IN^{H12N}$ in mitotic cells (data not shown).

*Depletion of Endogenous LEDGF/p75 Affects Nuclear Accumulation of HIV-1 IN*—To prove the importance of LEDGF/p75 for nuclear import of HIV-1 IN, we used RNA interference to transiently knock-down endogenous LEDGF/p75. After transfection of HeLa or HEK-293 cells with synthetic LEDGF/p75 mRNA-specific siRNA duplexes, ~80–95% depletion in LEDGF/p75 was observed 60–72 h after transfection (Fig. 9*B*). The knock-down of LEDGF/p75 did not cause a change in cell morphology or affect the growth of the transfected cells as compared with the cells treated with non-interfering siRNA. Accordingly, a stable knock-down of LEDGF/p75 in human cell lines using hairpin siRNA has been reported to be compatible with cell survival and proliferation (56). However, knock-down of endogenous LEDGF/p75 resulted in a striking defect in nuclear accumulation of EGFP-IN in HeLa cells. EGFP-IN appeared diffusely distributed in cells co-transfected with an LEDGF/p75-specific siRNA (Fig. 9*C*). Although EGFP-IN was still detectable in the nucleus, the typical irregular distribution pattern was lost. Moreover, in mitotic cells, EGFP-IN was no longer concentrated on the chromosomes (Fig. 9*D*). Although depletion of LEDGF/p75 varied from cell to cell, indirect immunostaining using anti-LEDGF/p75 antibody confirmed the correlation between knock-down of LEDGF/p75 and the alteration in EGFP-IN distribution (data not shown). We also noticed that depletion of LEDGF/p75 was accompanied by a decrease in the EGFP-IN expression levels, as reflected by the fact that higher input laser power was necessary for efficient detection of EGFP signal after knock-down of LEDGF/p75 (data not shown). A similar loss of the nuclear/chromosomal accumulation of HIV-1 IN concomitant with a decrease in IN levels upon LEDGF/p75 knock-down was also observed in stable cell lines expressing FLAG-tagged IN (data not shown). In addition, incubation of the siRNA-transfected cells in the presence of 5 nM leptomycin B, a potent inhibitor of the CRM1-mediated nuclear export (57), for 3–6 h did not cause nuclear re-entry of EGFP-IN (data not shown). Hence, CRM1-mediated nuclear export is probably not involved in the re-distribution of IN upon knock-down of LEDGF/p75.

## DISCUSSION

*Nuclear and Chromosomal Localization of IN Deletion Mutants Correlates with Their Colocalization with LEDGF/p75*—HIV-1 IN accumulates in the nucleus in the absence of other viral proteins (17, 29–31, 33), where it forms a specific complex with the endogenous LEDGF/p75 protein (36). Because the presence of IN in the cell does not seem to affect the intranuclear distribution of LEDGF/p75 (data not shown), the latter may play the dominant role determining nuclear accumulation and possibly intracellular trafficking of IN. To study the contributions of the individual domains of HIV-1 IN to the nuclear import and intracellular interaction with LEDGF/p75, we made a series of EGFP-fused IN deletion mutants preserving

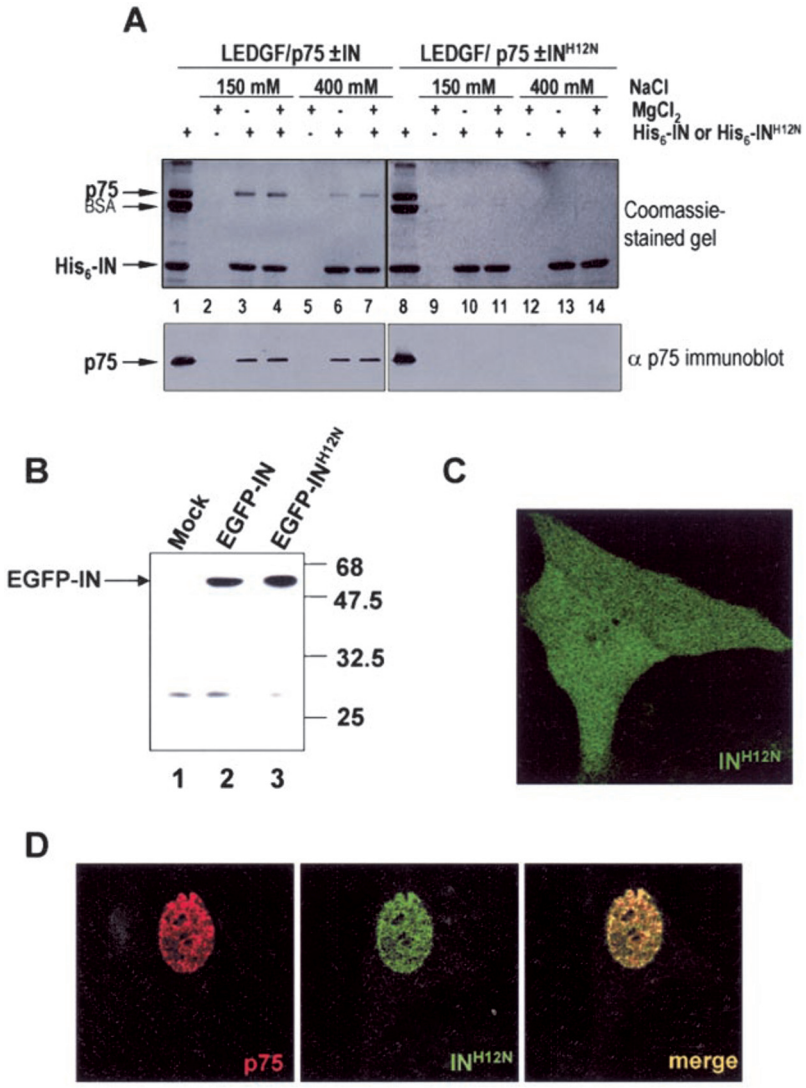


FIG. 8. **IN$^{H12N}$ has reduced affinity for LEDGF/p75.** *A*, recombinant $His_6$-tagged HIV-1 IN or similarly prepared $His_6$-IN$^{H12N}$ mutant (both 1.5 $\mu$g) were incubated with 2 $\mu$g of recombinant LEDGF/p75 and the complexes were recovered using Ni-NTA-agarose as described under "Experimental Procedures." The specific buffer conditions are indicated *above* the gels in both panels. Recombinant LEDGF/p75 was incubated with $His_6$-IN (*lanes 3–7*) or $His_6$-IN$^{H12N}$ (*lanes 9–14*). IN was omitted in *lanes 2* and *9*. *Lanes 1* and *8*, sample containing LEDGF/p75, BSA, and $His_6$-IN (*lane 1*) or $His_6$-IN$^{H12N}$ (*lane 8*) in equivalent amounts as used in the pull-down assay. Positions of LEDGF/p75, BSA, and $His_6$-IN in the gel are indicated on the *left side* of the panel. For more sensitive detection of LEDGF/p75, Western blotting was performed. *B*, transient expression levels of EGFP-IN$^{H12N}$ are similar to those of the wild-type EGFP-IN fusion. Western blot analysis was performed as in Fig. 3*B*. *Lane 1*, mock-transfected HeLa cells; *lanes 2* and *3*, cells were transfected with pEGFP-IN$^{s}$ (*lane 2*) and pEGFP-IN$^{s}$(H12N) (*lane 3*). *C*, confocal image of HeLa cells expressing EGFP-IN$^{H12N}$. The EGFP-IN$^{H12N}$ protein is homogeneously distributed throughout the whole cell. No nuclear accumulation or specific nuclear pattern could be observed in interphase cells. *D*, upon overexpression of HcRed1-p75, the EGFP-IN$^{H12N}$ mutant is recruited to the nucleus and the specific nuclear distribution pattern reminiscent of that of wild-type IN and EGFP-IN becomes apparent.

its structural domains. According to their intracellular distribution, the mutants can be subdivided in two families (Table I). The first family is composed of the EGFP fusions that displayed the irregular nuclear distribution pattern as seen with wild type IN and retained the ability to colocalize with LEDGF/p75 throughout the cell cycle. It includes the full-length IN, the ΔC and ΔN deletion mutants, the core domain fragment, and the full-length IN$^{H12N}$ mutant. The latter three mutants required overexpression of LEDGF/p75 to display nuclear/chromosomal accumulation and colocalization with LEDGF/p75. The second family consists of the mutants that showed a distinct intracellular distribution and no colocalization with LEDGF/p75. EGFP-IN/Nt and EGFP-IN/Ct fusions fall in this category. The former showed a dispersed cellular distribution, whereas the latter was concentrated in the nucleus and nucleoli; neither mutant associated with condensed chromosomes during mitosis.

The observed nuclear distribution pattern of IN mutants during interphase and their binding to condensed chromosomes in mitosis clearly correlates with their colocalization with LEDGF/p75. We can conclude that the C-terminal domain of IN is dispensable for the interaction with LEDGF/p75 and for the interaction with condensed chromosomes in living cells. On the other hand, both the N-terminal zinc binding and the core domains of IN are important for the interaction with LEDGF/p75. Indeed, deletion or mutation of the N-terminal domain abolished the specific nuclear distribution and colocalization of EGFP-IN/ΔN, EGFP-IN/core and EGFP-IN$^{H12N}$ with endogenous LEDGF/p75 in the cell. Because overexpression of LEDGF/p75 could restore nuclear/chromosomal localization of EGFP-IN/core and not of EGFP-IN/Nt, the core domain of IN harbors a signal that is both necessary and sufficient for interaction with LEDGF/p75 in cells. However, we cannot rule out that some additional cellular factors may be involved in the IN-LEDGF/p75 interaction. Because EGFP-IN/ΔC was nuclear at the endogenous LEDGF/p75 levels, we conclude that the N-terminal domain of IN enhances the affinity of the IN-LEDGF/p75 interaction. Moreover, the mutant IN$^{H12N}$ failed to interact with LEDGF/p75 in our $His_6$-tag integrase pull-down assay. Hence, the association between the core domain of IN and LEDGF/p75 might be dynamic (and thus not detected in this particular *in vitro* assay) or might require additional factors, such as chaperones, present in the cell. Whether the zinc binding domain stabilizes the complex via a direct interaction with LEDGF/p75 or via influencing the structure and/or the multimeric state of IN will be the subject of future research. Interestingly, although several described mutations in the IN zinc binding domain, including H12N, cause only a partial reduction of IN enzymatic activity *in vitro*, they do not allow replication of the mutant virus (21). We speculate that the inability of the mutant IN protein to interact with endogenously expressed cellular LEDGF/p75, might eventually ex-

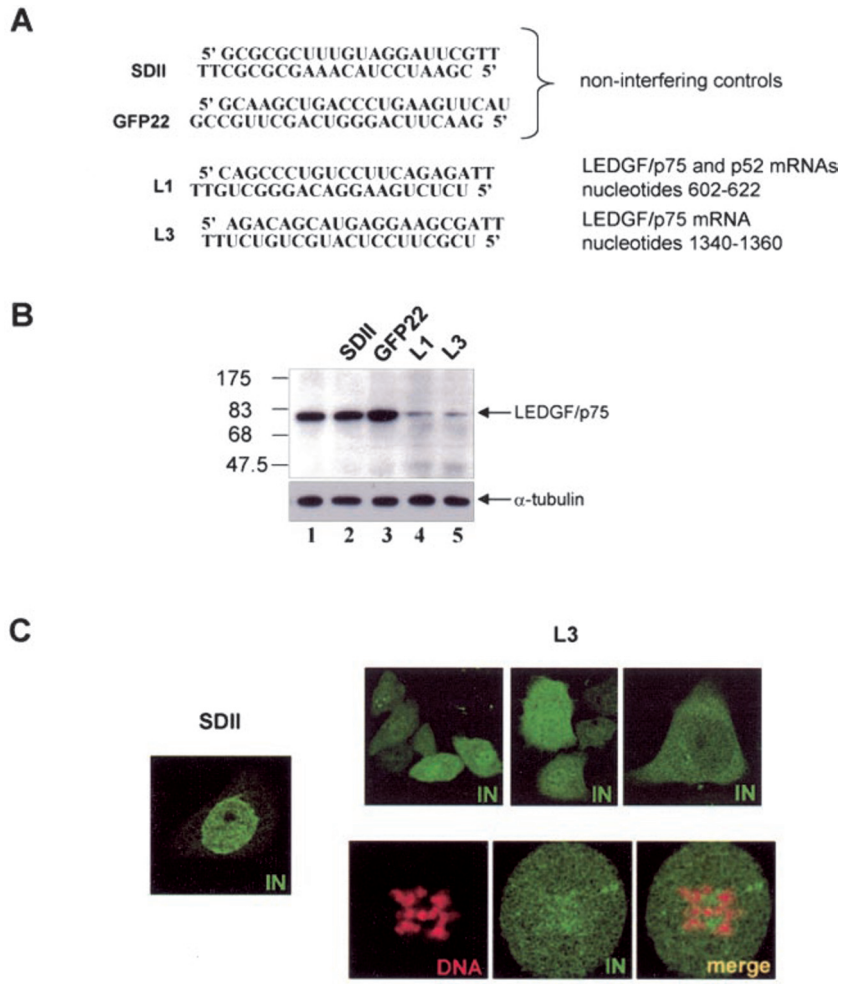


FIG. 9. **Knock-down of LEDGF/p75 by specific siRNA causes a redistribution of HIV-1 IN from the nucleus to the cytoplasm.** *A*, the sequences of the control siRNAs (SDII and GFP22) and the interfering siRNAs (L1 and L3). The locations of the target sites within LEDGF/p75 mRNA with respect to the first nucleotide of the start codon of the LEDGF/p75 open reading frame are indicated. *B*, Western blot showing decrease in LEDGF/p75 levels in HeLa cells upon transfection with LEDGF/p75-specific siRNAs. Cells were analyzed at 60 h after transfection with the siRNA duplexes, and 10 $\mu$g of total protein were loaded in each lane. *Lane 1*, non-transfected HeLa cells; *lanes 2–5*, cells transfected with SDII siRNA (*lane 2*), GFP22 (*lane 3*), L1 (*lane 4*), and L3 (*lane 5*) siRNA. As a loading control, $\alpha$-tubulin was detected in the same samples. *C*, HeLa cells were co-transfected with the full-length EGFP-IN expression construct plus the control RNA duplex (SDII) or the p75 mRNA-specific siRNA (L3). The confocal images were taken from live cells 60 h after transfection. The chromosomal DNA was stained with the cell-permeable dye SYTO 17.

plain the observed phenotype. Of note, Woodward *et al.* (58) have recently reported that the zinc-binding domain of feline immunodeficiency virus IN is essential for its nuclear localization. It remains to be seen whether feline immunodeficiency virus IN interacts with and requires LEDGF/p75 for its nuclear localization in feline and human cells.

A second protein product, p52, is expressed from the same gene as LEDGF/p75 as a result of alternative splicing of the pre-mRNA (38, 41). It appears to be much less abundant and ubiquitous than LEDGF/p75, which is indirectly reflected by the scarcity of the available human and mouse p52 mRNA-derived expressed sequence tags in GenBank™. In agreement with Nishizawa *et al.* (50), we observed a marked difference in the intranuclear distributions of the p52 and LEDGF/p75 proteins. Although p52 had a speckled distribution pattern, LEDGF/p75 appeared more heterogeneous in the nucleoplasm. The distinct nuclear distribution patterns probably reflect different binding partners and functions in the cell. At the amino acid sequence level, p52 shares 325 N-terminal residues with LEDGF/p75. Because p52 did not interact with HIV-1 IN, the C-terminal 205 residue fragment of LEDGF/p75 is likely to harbor the site of interaction with IN.

*LEDGF/p75 Is Essential for the Nuclear Accumulation of HIV-1 IN*—Transient transfection typically results in a wide range of expression levels, varying from cell to cell. Initially, we considered cells with moderate to low EGFP-IN expression levels as more representative for the *in vivo* situation, whereby only a limited number of IN molecules are present in the infected cell. In these cells, EGFP-IN was predominantly present in the nucleus. However, in cells that produced higher levels of EGFP-IN, the nuclear accumulation was less pronounced. Overexpression of HcRed1-p75 restored nuclear accumulation of EGFP-IN even at very high expression levels of the latter (data not shown). Hence, when IN overtitrates the endogenous levels of LEDGF/p75, the fraction that is free of the binding partner becomes diffusively distributed. This dependence on LEDGF/p75 expression levels was much more pronounced with the EGFP-IN/core, EGFP-IN/$\Delta$N, and EGFP-IN$^{H12N}$ mutants. Even more compelling evidence that LEDGF/p75 is essential for nuclear import of IN was provided by our siRNA experiment. The karyophilic properties of IN and its association with chromosomes were completely abolished upon knock-down of endogenous LEDGF/p75.

*The C-terminal Domain of HIV-1 IN Possesses a Cryptic NLS but Does Not Interact with Mitotic DNA*—The Lys/Arg-rich C-terminal domains of retroviral IN proteins have been implicated in nonspecific DNA binding (59). Deletion of the C-terminal domain abolishes divalent cation-independent binding of the recombinant protein to long terminal repeat as well as to unspecific DNA (60). The C-terminal domain of HIV-1 IN (residues 213–288) also harbors a part of a putative bipartite NLS ($NLS_D$: $^{211}$KELQKQITK, see Ref. 33). However, deletion of the C-terminal domain in EGFP-IN/$\Delta$C did not affect the nuclear localization of IN. This result is in agreement with Tsurutani *et al.* (61), who reported the persistence of karyophilic properties of their IN$^{\Delta 181–288}$ deletion mutant. Although the EGFP-IN/Ct fusion used in our work harbors only a part of the putative NLS, it was exclusively nuclear in HeLa cells. However, the diffuse nuclear/nucleolar distribution and the absence of colocalization with LEDGF/p75 and chromatin indicate that the

C-terminal fragment, although intrinsically karyophilic, does not associate with and is not trapped on specific nuclear structures, suggesting an active nuclear import mechanism. We speculate that the C-terminal domain may have a role in HIV-1 IN nuclear import, although its karyophilic property seems to be masked in the context of the full-length protein. Interestingly, deletion of the N-terminal domain of IN, but not the mutation of zinc-binding His-12, could expose the NLS. Our results also indicate that the C-terminal domain of HIV-1 IN is not essential for binding to chromosomal DNA and therefore, probably, is not involved in integration site selection. However, it remains possible that, within the viral PIC, the C-terminal domain has a distinct configuration and might play a role in nuclear trafficking and binding to target DNA.

*What Is the Mechanism of HIV-1 IN Nuclear Import?*—The first mechanism proposed for HIV-1 IN nuclear import implicated the classical importin $\alpha/\beta$ pathway (33). A putative bipartite NLS within the C-terminal half of the protein was put forward. However, the mutational disruption of the suggested NLS sequence did not abolish nuclear localization of EGFP-IN fusions (61). A reduction in nuclear accumulation was observed when the putative NLS was mutated in the context of FLAG-tagged HIV-1 IN protein (62). These apparently conflicting data might have resulted from the differences between the fusion peptides used. In particular, addition of the negatively charged FLAG tag sequence to the C terminus of HIV-1 IN increases the proportion of protein in the cytosolic fraction (data not shown). In addition, some mutations might have disrupted folding or multimerization of IN resulting in reduced binding to endogenous LEDGF/p75. The involvement of the classical importin $\alpha/\beta$ in the nuclear import of HIV-1 IN has been challenged recently by Depienne *et al.* (34), who showed that, in semi-permeabilized cells, the protein can be imported via a novel importin- and Ran-independent mechanism. What can we learn from the lack of nuclear accumulation of full-length HIV-1 IN in the absence of LEDGF/p75 protein? The results of our LEDGF/p75 knock-down experiments argue against the involvement of a cellular nuclear import factor other than LEDGF/p75, directing HIV-1 IN into the nucleus. However, three alternatives must be considered. First, the observed re-distribution of IN from nucleus to cytoplasm upon LEDGF/p75 knock-down could be mediated by an active nuclear export of IN molecules that are no longer associated with LEDGF/p75. Although we cannot completely refute this hypothesis, the general CRM1-dependent nuclear export mechanism is probably not involved, because incubation of cells with leptomycin B did not cause nuclear re-entry of EGFP-IN and FLAG-tagged HIV-1 IN in the absence of LEDGF/p75 (data not shown). Second, the pool of cellular IN, not associated with LEDGF/p75, may be subjected to a post-translational modification that would disrupt its nuclear import. Instability of HIV-1 IN in human cells through ubiquitination and subsequent proteasome-dependent degradation of HIV-1 IN has been previously suggested (63). Of note, knock-down of LEDGF/p75 leads to a reduction of IN expression, likely resulting from the proteasome activity, as the effect is sensitive to specific proteasome inhibitors.[2] Other types of post-translational modifications, such as sumoylation or phosphorylation, often implicated in regulation of nuclear import should also be considered (64). Finally, LEDGF/p75 might be necessary for the correct folding of HIV-1 IN. Whether LEDGF/p75 is directly involved in active nuclear import of IN or plays a more passive role, for example, by trapping the imported IN within the nuclear structure, is currently under investigation.

Although the precise role of the LEDGF/p75 protein in the viral life cycle is yet to be unraveled, its specific interaction with HIV-1 IN and its role in nuclear accumulation of IN leaves little doubt that it is a genuine cofactor for HIV replication. Experiments to determine the role of LEDGF/p75 in HIV PIC nuclear import and DNA integration are currently under way. According to one plausible model, LEDGF/p75 is an integral part of the viral PIC stabilizing the IN multimer and targeting the PIC to the nucleus/chromatin. Alternatively, LEDGF/p75 could be a chromosome-bound receptor for the PIC, helping to direct HIV DNA integration into the transcriptionally-active regions of the chromosome. A similar mechanism has been implicated in targeting of yeast retrotransposons. Thus, integration of the Ty5 element is selectively directed to heterochromatin by the specific interaction between Ty5 integrase and Sir4p (65, 66). Furthermore, tethering of retroviral INs to specific DNA-binding proteins promotes integration close to the corresponding recognition sites *in vitro* (67, 68). According to both models, LEDGF/p75 would provide the PIC with a target DNA binding/selecting capacity, and, as a result of its association with the general transcription machinery, could account for the marked selectivity of HIV-1 DNA integration into active genes (37). According to one preliminary report (69) and our observations (data not shown), HIV-1 replication is severely impaired upon knock-down of LEDGF/p75 in HeLa-CD4 cells. A novel target for antiretroviral therapy might thus be emerging.

*Acknowledgments*—The plasmid pINSD (44) was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, National Institutes of Health from Dr. A. Engelman and Dr. R. Craigie.

REFERENCES

1. Gartner, S., Markovits, P., Markovitz, D. M., Kaplan, M. H., Gallo, R. C., and Popovic, M. (1986) *Science* **233,** 215–219
2. Weinberg, J. B., Matthews, T. J., Cullen, B. R., and Malim, M. H. (1991) *J. Exp. Med.* **174,** 1477–1482
3. Roe, T., Reynolds, T. C., Yu, G., and Brown, P. O. (1993) *EMBO J.* **12,** 2099–2108
4. Lewis, P. F., and Emerman, M. (1994) *J. Virol.* **68,** 510–516
5. Bukrinsky, M. I., Sharova, N., Dempsey, M. P., Stanwick, T. L., Bukrinskaya, A. G., Haggerty, S., and Stevenson, M. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 6580–6584
6. Fouchier, R. A., and Malim, M. H. (1999) *Adv. Virus Res.* **52,** 275–299
7. Vodicka, M. A. (2001) *Somat. Cell Mol. Genet.* **26,** 35–49
8. Sherman, M. P., and Greene, W. C. (2002) *Microbes Infect.* **4,** 67–73
9. Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L., and Charneau, P. (2000) *Cell* **101,** 173–185
10. Limon, A., Nakajima, N., Lu, R., Ghory, H. Z., and Engelman, A. (2002) *J. Virol.* **76,** 12078–12086
11. Dvorin, J. D., Bell, P., Maul, G. G., Yamashita, M., Emerman, M., and Malim, M. H. (2002) *J. Virol.* **76,** 12087–12096
12. Follenzi, A., Ailles, L. E., Bakovic, S., Geuna, M., and Naldini, L. (2000) *Nat. Genet.* **25,** 217–222
13. Van Maele, B., De Rijck, J., De Clercq, E., and Debyser, Z. (2003) *J. Virol.* **77,** 4685–4694
14. Bukrinsky, M. I., Haggerty, S., Dempsey, M. P., Sharova, N., Adzhubel, A., Spitz, L., Lewis, P., Goldfarb, D., Emerman, M., and Stevenson, M. (1993) *Nature* **365,** 666–669
15. Reil, H., Bukovsky, A. A., Gelderblom, H. R., and Gottlinger, H. G. (1998) *EMBO J.* **17,** 2699–2708
16. Fouchier, R. A., Meyer, B. E., Simon, J. H., Fischer, U., and Malim, M. H. (1997) *EMBO J.* **16,** 4531–4539
17. Depienne, C., Roques, P., Creminon, C., Fritsch, L., Casseron, R., Dormont, D., Dargemont, C., and Benichou, S. (2000) *Exp. Cell Res.* **260,** 387–395
18. Bouyac-Bertoia, M., Dvorin, J. D., Fouchier, R. A., Jenkins, Y., Meyer, B. E., Wu, L. I., Emerman, M., and Malim, M. H. (2001) *Mol. Cell* **7,** 1025–1035
19. Zufferey, R., Nagy, D., Mandel, R. J., Naldini, L., and Trono, D. (1997) *Nat. Biotechnol.* **15,** 871–875
20. Nakajima, N., Lu, R., and Engelman, A. (2001) *J. Virol.* **75,** 7944–7955
21. Engelman, A. (1999) *Adv. Virus Res.* **52,** 411–426
22. Brown, P. O. (1997) in *Retroviruses* (Coffin, J. M., Huges, S. H., and Varmus, H. E., eds) pp. 161–203, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
23. Craigie, R. (2001) *J. Biol. Chem.* **276,** 23213–23216
24. Asante-Appiah, E., and Skalka, A. M. (1997) *Antiviral Res.* **36,** 139–156
25. Lee, S. P., Xiao, J., Knutson, J. R., Lewis, M. S., and Han, M. K. (1997) *Biochemistry* **36,** 173–180
26. Esposito, D., and Craigie, R. (1998) *EMBO J.* **17,** 5832–5843

[2] P. Cherepanov, G. Maertens, E. De Clercq, Z. Debyser, unpublished data.

27. Jenkins, T. M., Engelman, A., Ghirlando, R., and Craigie, R. (1996) *J. Biol. Chem.* **271,** 7712–7718
28. Vercammen, J., Maertens, G., Gerard, M., De Clercq, E., Debyser, Z., and Engelborghs, Y. (2002) *J. Biol. Chem.* **277,** 38045–38052
29. Pluymers, W., Cherepanov, P., Schols, D., De Clercq, E., and Debyser, Z. (1999) *Virology* **258,** 327–332
30. Cherepanov, P., Pluymers, W., Claeys, A., Proost, P., De Clercq, E., and Debyser, Z. (2000) *FASEB J.* **14,** 1389–1399
31. Petit, C., Schwartz, O., and Mammano, F. (1999) *J. Virol.* **73,** 5079–5088
32. Limon, A., Devroe, E., Lu, R., Ghory, H. Z., Silver, P. A., and Engelman, A. (2002) *J. Virol.* **76,** 10598–10607
33. Gallay, P., Hope, T., Chin, D., and Trono, D. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 9825–9830
34. Depienne, C., Mousnier, A., Leh, H., Le Rouzic, E., Dormont, D., Benichou, S., and Dargemont, C. (2001) *J. Biol. Chem.* **276,** 18102–18107
35. Miyamoto, Y., Hieda, M., Harreman, M. T., Fukumoto, M., Saiwaki, T., Hodel, A. E., Corbett, A. H., and Yoneda, Y. (2002) *EMBO J.* **21,** 5833–5842
36. Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E., and Debyser, Z. (2003) *J. Biol. Chem.* **278,** 372–381
37. Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., and Bushman, F. (2002) *Cell* **110,** 521–529
38. Ge, H., Si, Y., and Roeder, R. G. (1998) *EMBO J.* **17,** 6723–6729
39. Fatma, N., Singh, D. P., Shinohara, T., and Chylack, L. T., Jr. (2001) *J. Biol. Chem.* **276,** 48899–48907
40. Singh, D. P., Fatma, N., Kimura, A., Chylack, L. T., Jr., and Shinohara, T. (2001) *Biochem. Biophys. Res. Commun.* **283,** 943–955
41. Singh, D. P., Kimura, A., Chylack, L. T., Jr., and Shinohara, T. (2000) *Gene* **242,** 265–273
42. Ge, H., Si, Y., and Wolffe, A. P. (1998) *Mol. Cell* **2,** 751–759
43. Fradkov, A. F., Verkhusha, V. V., Staroverov, D. B., Bulina, M. E., Yanushevich, Y. G., Martynov, V. I., Lukyanov, S., and Lukyanov, K. A. (2002) *Biochem. J.* **368,** 17–21
44. Engelman, A., and Craigie, R. (1992) *J. Virol.* **66,** 6361–6369
45. Elbashir, S. M., Harborth, J., Weber, K., and Tuschl, T. (2002) *Methods* **26,** 199–213
46. Gonzalez, R. C., and Woods, R. E. (2002) *Digital Image Processing*, 2nd Ed., pp. 693–753, Prentice Hall, Upper Saddle River, NJ
47. Cherepanov, P., Surratt, D., Toelen, J., Pluymers, W., Griffith, J., De Clercq, E., and Debyser, Z. (1999) *Nucleic Acids Res.* **27,** 2202–2210
48. Cherepanov, P., Este, J. A., Rando, R. F., Ojwang, J. O., Reekmans, G., Steinfeld, R., David, G., De Clercq, E., and Debyser, Z. (1997) *Mol. Pharmacol.* **52,** 771–780
49. Caplen, N. J., Parrish, S., Imani, F., Fire, A., and Morgan, R. A. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 9742–9747
50. Nishizawa, Y., Usukura, J., Singh, D. P., Chylack, L. T., Jr., and Shinohara, T. (2001) *Cell Tissue Res.* **305,** 107–114
51. Mattaj, I. W., and Englmeier, L. (1998) *Annu. Rev. Biochem.* **67,** 265–306
52. Cai, M., Zheng, R., Caffrey, M., Craigie, R., Clore, G. M., and Gronenborn, A. M. (1997) *Nat. Struct. Biol.* **4,** 567–577
53. Eijkelenboom, A. P., van den Ent, F. M., Vos, A., Doreleijers, J. F., Hard, K., Tullius, T. D., Plasterk, R. H., Kaptein, R., and Boelens, R. (1997) *Curr. Biol.* **7,** 739–746
54. van den Ent, F. M., Vos, A., and Plasterk, R. H. (1998) *J. Virol.* **72,** 3916–3924
55. Burke, C. J., Sanyal, G., Bruner, M. W., Ryan, J. A., LaFemina, R. L., Robbins, H. L., Zeft, A. S., Middaugh, C. R., and Cordingley, M. G. (1992) *J. Biol. Chem.* **267,** 9639–9644
56. Devroe, E., and Silver, P. A. (2002) *BMC Biotechnol.* **2,** 15
57. Kudo, N., Wolff, B., Sekimoto, T., Schreiner, E. P., Yoneda, Y., Yanagida, M., Horinouchi, S., and Yoshida, M. (1998) *Exp. Cell Res.* **242,** 540–547
58. Woodward, C. L., Wang, Y., Dixon, W. J., Htun, H., and Chow, S. A. (2003) *J. Virol.* **77,** 4516–4527
59. Lutzke, R. A., Vink, C., and Plasterk, R. H. (1994) *Nucleic Acids Res.* **22,** 4125–4131
60. Engelman, A., Hickman, A. B., and Craigie, R. (1994) *J. Virol.* **68,** 5911–5917
61. Tsurutani, N., Kubo, M., Maeda, Y., Ohashi, T., Yamamoto, N., Kannagi, M., and Masuda, T. (2000) *J. Virol.* **74,** 4795–4806
62. Petit, C., Schwartz, O., and Mammano, F. (2000) *J. Virol.* **74,** 7119–7126
63. Mulder, L. C., and Muesing, M. A. (2000) *J. Biol. Chem.* **275,** 29749–29753
64. Wood, L. D., Irvin, B. J., Nucifora, G., Luce, K. S., and Hiebert, S. W. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 3257–3262
65. Xie, W., Gai, X., Zhu, Y., Zappulla, D. C., Sternglanz, R., and Voytas, D. F. (2001) *Mol. Cell. Biol.* **21,** 6606–6614
66. Zhu, Y., Dai, J., Fuerst, P. G., and Voytas, D. F. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 5891–5895
67. Bushman, F. D., and Miller, M. D. (1997) *J. Virol.* **71,** 458–464
68. Katz, R. A., Merkel, G., and Skalka, A. M. (1996) *Virology* **217,** 178–190
69. Emiliani, S., Rain, J. C., Maroun, M., Moisant, F., Segeral, E., Selig, L., Legrain, P., and Benarous, R. (2003) *10th Conference on Retroviruses and Opportunistic Infections, Boston, MA, February 10–14, 2003, Abstr. 74lb*