Identification and Characterization of a Functional Nuclear Localization Signal in the HIV-1 Integrase Interactor LEDGF/p75*

Received for publication, April 28, 2004
Published, JBC Papers in Press, May 25, 2004, DOI 10.1074/jbc.M404700200

Goedeke Maertens‡‡, Peter Cherepanov**, Zeger Debyser†, Yves Engelborghs‡‡, and Alan Engelman‡‡‡

From the Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, §Laboratory of Biomolecular Dynamics, Katholieke Universiteit Leuven, Heverlee B-3001, Belgium, the ‡‡Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115, and *Rega Institute for Medical Research and KULAK, Katholieke Universiteit Leuven, Leuven B-3000, Belgium

Human lens epithelium-derived growth factor (LEDGF)/p75 protein forms a specific nuclear complex with human immunodeficiency virus type 1 (HIV-1) integrase and is essential for nuclear localization and chromosomal association of the viral protein. We now studied nuclear import of LEDGF/p75 in live and semi-permeabilized cells. We showed that nuclear import of LEDGF/p75 is GTP--dependent, and that the protein competes with the classical SV40 large T antigen nuclear localization signal (NLS) for nuclear import receptors. We identified the NLS of LEDGF/p75 through deletion analysis and site-directed mutagenesis. The LEDGF/p75 NLS, **GRKR-KAEKG** belongs to the canonical SV40-like family. Fusion of this short peptide to the amino terminus of Escherichia coli β-galactosidase rendered the fusion protein nuclear, confirming that the LEDGF/p75 NLS is transferable. Moreover, a single amino acid change in the NLS was sufficient to exclude the mutant LEDGF/p75 protein from the nucleus and abolish nuclear import of HIV-1 integrase.

The tight regulation of macromolecular transport through the nucleopore complexes (NPCs) of the nuclear envelope is vital in eukaryotic cells. The NPC has a 9-nm diffusion channel, which sets an upper limit for free diffusion to ~45–60 kDa, allowing metabolites, ions, and small macromolecules to pass through. However, NPCs allow active nuclear transport of particles larger than 25 nm in diameter (for a review, see Ref. 1). Proteins are directed in or out of the nucleus via nuclear localization signals (NLSs) or nuclear export signals, respectively.

© 2004 by The American Society for Biochemistry and Molecular Biology, Inc.

*This work was funded by National Institutes of Health Grant AI52014 (to A. E.). Work at the KULeuven was financially supported by the SBO program from the IWT Flanders (Belgium). 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.

‡‡ An Aspirant of the Fund for Scientific Research (Fonds voor Wetenschappelijk Onderzoek) Flanders.
‡‡‡ To whom correspondence should be addressed. Tel.: 617-632-4361; Fax: 617-632-3113; E-mail: Alan_Engelman@dfci.harvard.edu.

The abbreviations used are: NPC, nucleopore complex; AMP-PNP, 5′-adenylimidodiphosphate; NLS, nuclear localization signal; cNLS, classical NLS; EGFP, enhanced green fluorescent protein; GG, EGFP-GST fusion; GST, glutathione S-transferase; GTP-S, guanosine 5′-O-(3-thiotriphosphate); HcdR1-IN, HcdR1 taggged HIV-1 IN; HIV-1, human immunodeficiency virus type 1; imp-α, -β, and -7, importin-α, -β, and -7; IN, integrase; LEDGF, lens epithelium-derived growth factor; PIC, preintegration complex; WGA, wheat germ agglutinin; X-gal, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; RanGTP, Ras-related GTPase Ran protein; TB, transport buffer.

NLSSs and nuclear export signals require binding of specific import/export-receptors (importins and exportins) that are necessary for the translocation of the cargo through NPCs. The NLSs are grouped into several categories. The most commonly known are the classical NLSs (cNLSs), which are composed of a basic amino acid stretch (K/R)x-4 preceding by a Gly, Pro, or an acidic amino acid residue (2, 3) similar to the NLS of SV40 large T antigen. Bipartite NLSs as in Xenopus nucleoplasmin (4) are composed of two basic amino acid stretches interspersed by a nonconserved 10–12-amino acid spacer: (K/R)x yzK(R)x. In addition, several nonclassical NLSs have been described, examples being the M9 fragment in the heterogeneous nuclear ribonucleoprotein A1 and A2 containing a 38-amino acid stretch enriched in aromatic residues and glycine (5–7) or the nuclear targeting signals of ribosomal proteins, such as L23a (8). Furthermore, in addition to linear NLSs, discontinuous epitopes that come together upon folding into tertiary structure contribute to the nuclear import of histone proteins (9, 10) (for a review, see Refs. 11 and 12).

Proteins containing a NLS are imported into the nucleus by the importin-α/importin-β (imp-α/β) import receptors (13–15). imp-α functions as the adaptor molecule between the NLS-containing protein and imp-β. The ternary complex is directed by imp-β to the NPC, which passes through the nuclear pore by a so-called facilitated diffusion process (16). Although the exact mechanism of nuclear import per se has not been solved, it is thought that the direction of nuclear–cytoplasmic transport is dependent on the gradient of the small Ras-related GTPase Ran protein (RanGTP) across the nuclear envelope (17–20). Alternative nuclear import pathways have been described in which the NLS-containing proteins directly bind one of the nuclear import receptors of the importin-β superfamily. Heterogeneous nuclear ribonucleoproteins A1 and A2, for example, are imported by transportin (5), and histone H1 is imported by the imp-β7 heterodimer (21). The core histones and ribosomal proteins can be imported by either imp-β, imp-7, imp-5 or transportin (8, 10). Recently, examples of Ran and energy- or imp-β-independent nuclear transport mechanisms have been reported (22, 23).

Integrase (IN) is the retroviral protein responsible for integration of the DNA replica of the viral genome into a cell chromosome. When expressed on their own in the absence of other viral proteins, retroviral INs locate to cell nuclei and are therefore karyophilic (24–27). It has been suggested that human immunodeficiency virus type 1 (HIV-1) IN, as an essential component of the viral preintegration complex (PIC), plays a role in its nuclear import (24, 27–29). The mechanism of HIV-1 IN nuclear import, however, has not been fully elucidated. Recently, lens epithelium-derived growth factor (LEDGF/p75)
has been shown to associate with HIV-1 IN in human cells (30). Using transient knock-down of endogenous LEDGF/p75 via small interfering RNA, LEDGF/p75 was shown to be both necessary and sufficient for accumulation of HIV-1 IN into the nucleus (31). Although the alternative splice variant of LEDGF/p75, p52, is also karyophilic, its nuclear distribution is different from LEDGF/p75 (31, 32). p52 did not display affinity for HIV-1 IN in vitro and did not co-localize with fluorescently tagged IN in live cells (31).

The exact cellular function of LEDGF/p75 is unknown. LEDGF/p75 was originally isolated from HeLa cell extracts as an interactor of transcriptional co-activator PC4 and was therefore suggested to play a role in transcriptional regulation (33). Singh and co-workers (34–36) subsequently reported a role for LEDGF/p75 in cellular stress response and survival. LEDGF/p75 binds to heat shock and stress-related regulatory DNA elements, and cellular expression of LEDGF/p75 was increased upon heat, oxidative, or chemical stress.

Since LEDGF/p75 was shown to be essential for the nuclear accumulation of HIV-1 IN, we characterized the nuclear import pathway of LEDGF/p75 and identified its NLS. We show that a single point mutation is sufficient to confine HIV-1 IN and relocate HIV-1 IN to the cytoplasm.

**EXPERIMENTAL PROCEDURES**

**Enhanced Green Fluorescent Protein (EGFP) and EGFP-Glutathione S-Transferase (GST) (GO) Constructs—**To create pEGFP-P75/ΔC and pEGFP-P75/Ct, the fragments were PCR-amplified from pCMH75 (30) using Pfu Ultra (Stratagene, La Jolla, CA) and an appropriate set of primers. The primers used were as follows: p75/ΔC, 5′-CCGAAGATCT- TCAACGCGGATTCATCAGCTGTTCTATTACGTCGTCGTTTTACAAC; p75/ΔC2, 5′-GGGGATCCATGGGATCCCCTAAGAAAAAGAGGAAGGTCCTGCTCCAGCACCCGCGCCATCCG, GST6, 5′-CCGAAGATCTCACAGCAGAATAGCTC; p75/Ct, 5′-GGGAATTCCTACTCAAAC; p75/Ct2, 5′-CCGAAGATCTCACAGCAGAATAGCTC; p75/Ct3, 5′-GGGAATTCCTACTCAAAC. The PCR fragments were digested with BglII and EcoRI and subcloned into pEGFP-C2 (Clontech, Palo Alto, CA). The construction of pEGFP-P75 was described previously (31). The design of the EGFP-GST fusion vector pGO was based on the work of Woodward et al. (32) with some modifications. A linker was added upstream and downstream of the GST reading frame, coding for a flexible peptide of alternating Gly-Ala dimers. The following primers were used: GST1, 5′-GCAAAGATCTGGGATCCCTAAGAAAAAGAGGAAGGTCCTGCTCCAGCACCCGCGCCATCCG, GST2, 5′-CCGAAGATCTCACAGCAGAATAGCTC; GST3, 5′-GCAAGATCTTGGGCGCGGGTGCT; GST4, 5′-GGGAATTCCTACTCAAAC; GST5, 5′-GGGAATTCCTACTCAAAC; GST6, 5′-GGGAATTCCTACTCAAAC. The first PCR product was created with GST1 and GST2 primers, using PfuUltra (Stratagene, La Jolla, CA) in TB. Cells were incubated for 30 min with WGA at 37 °C. The lysate was extensively dialyzed against transport buffer (TB, 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM MgCl2, 1 mg/ml pepstatin, and 1 mM phenylmethylsulfonyl fluoride) in a Pierce dialysis bag. Both proteins were labeled using Alexa Fluor 594- and Alexa Fluor 647-conjugated probes (Molecular Probes, Eugene, OR) and incubated at 37 °C for 2 h. Immediately prior to use, the samples were loaded onto glycerol gradients and centrifuged at 200,000 × g for 4 h. Alexa Fluor 594-labeled LEDGF/p75 and Alexa Fluor 647-labeled IN were collected and diluted with 20% glycerol in TB and stored at −80 °C.

**Semipermeabilized Cells—**Rabbit reticulocyte lysate was purchased from Promega (Madison, WI). After ultracentrifugation at 100,000 × g, the lysate was extensively dialyzed against transport buffer (TB, 20 mM Hepes, pH 7.3, 110 mM potassium acetate, 2 mM magnesium acetate, 1 mM EGTA, 2 mM dithiotreitol, 1 μg/ml each aprotinin, leupeptin, and pepstatin, and 1 mM phenylmethylsulfonyl fluoride) in a Pierce dialysis cassette with a 100-kDa molecular weight cut-off. The in vitro import assay in semipermeabilized cells was based on methods described in Ref. 38. HeLa cells were seeded on LabTek coverglass coverslips and used at ~50–60% confluence. Cells were washed twice in ice-cold TB and permeabilized in 40 μM digitonin in TB on ice for 5 min. The endogenous IN loading was depleted in ice-cold TB by incubation at 4 °C for 10 min. Semipermeabilized cells were incubated at 37 °C for 30 min and contained 50% rabbit reticulocyte lysate (10%) in a reaction mixture containing 1 mM ATP, 1 mM GTP, 0.4 mg/ml X-Gal in phosphate-buffered saline containing 200 μM sodium fluoride, 4 mM potassium ferricyanide, and 2 mM MgCl2.

**Western Blot—**Cells were lysed 24 h post-transfection in 20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 0.1% SDS, and the total protein concentration was determined using the BCA protein assay (Pierce). Samples containing 15 μg of total protein were separated by 11% SDS-PAGE and electrophoresed onto polyvinylidene difluoride membranes. Monoclonal mouse anti-LEDGF/p75 5′-p52 antibody was purchased from BD Biosciences (San Jose, CA). Horseradish peroxidase-conjugated secondary goat anti-mouse antibody was from Dako (DakoCytomation California Inc., Carpenteria, CA). Detection was carried out using ECL+ chemiluminescent horseradish peroxidase substrate (Amersham Biosciences).
Nuclear Import of LEDGF/p75

reconstitution experiments, a Ran mixture containing 3 μM Ran (Jena Bioscience, Jena, Germany) and 0.5 μM nuclear transport factor 2 (Sigma) in TB was used, with or without 1 μM import receptors imp-α (human) and imp-β (Drosophila melanogaster) (Jena Biosciences). The import was terminated by washing cells with ice-cold TB followed by fixation in 2% formaldehyde (in phosphate-buffered saline). Slides were imported was terminated by washing cells with ice-cold TB followed by

**RESULTS**

**Computer Analysis**—LEDGF/p75 and its smaller splice variant p52 are karyophilic (30–32). LEDGF/p75 contains about 21% basic residues, which makes it difficult to reliably predict a NLS from the primary sequence. Thus, we turned to computer prediction programs, first using NUCDISC from the Bioinformatics Center (available on the World Wide Web at cubic.bioc.columbia.edu/cgi/var/nair/resonline.pl) (42), which looks for canonical NLSs using the k-nearest neighbor algorithm to score putative NLSs (41). Six putative NLSs within LEDGF/p75 were predicted (Table I). We also used the PredictNLS program provided by the Columbia University Bioinformatics Center (available on the World Wide Web at cubic.bioc.columbia.edu/cgi/var/nair/resonline.pl) (42), which searches for possible NLSs by comparing the protein sequence with a data base composed of experimentally determined NLSs expanded with “in silico mutagenesis”-created NLS motifs. Two overlapping NLSs were predicted by PredictNLS, both classified as SV40-like canonical NLSs (Table I).

The **NLS of LEDGF/p75** contains about 90 amino acids. The N-terminal 89 amino acids of LEDGF/p75 (Table I) were nuclear localized. Therefore, a NLS resides within the amino-terminal 325 residues of LEDGF/p75. The NLS predicted by PSORT II at amino acid position 509 (KKKP; see Table I) therefore does not play a role in the nuclear localization of LEDGF/p75.

### Table I

| Pattern | Consensus |
|---------|-----------|
| 4       | (K/R)₆    |
| 7       | (H/P/K/R)₆ |
| Bipartite| (K/R)₆ (K/R)₆,₅ |

**Site-directed Mutagenesis of the Putative LEDGF/p75 NLS**—On the basis of our deletion analysis and the computer-predicted NLSs, six putative sequences remained to be analyzed. Since one putative NLS occurred in both prediction programs, we first analyzed site-directed mutants centered around Arg¹⁴⁹ in the 146–156 amino acid stretch of LEDGF/p75 (Table I), in which Lys and/or Arg residues were changed to Ala (Figs. 2 and 3). The mutations were introduced in the pGG-P75 vector. GG-p75, with a predicted molecular mass of ~115 kDa, far exceeds the exclusion limit for free diffusion through the NPC. We first constructed pGG-P75A¹⁴⁶–¹⁴⁷ in which Arg¹⁴⁶ and Arg¹⁴⁷ were both changed to Ala. Transient expression in HeLa cells showed that altering these two arginines did not affect the nuclear localization of LEDGF/p75 (Fig. 2B). We then made the following set of mutants: pGG-P75A¹⁴⁰–¹⁴⁵ (Fig. 2C), pGG-P75A¹⁴⁵–¹⁵² (Fig. 2D), pGG-P75A¹⁴⁹–¹⁵⁰ (Fig. 2E), and pGG-P75A¹⁵⁰–¹⁵⁵ (Fig. 2F). All of these LEDGF/p75-NLS mutants were excluded from the nucleus upon transient expression in HeLa cells (Fig. 2, C–F). From this, we deduced that Lys¹⁵⁰ and/or Lys¹⁵⁵ were critical for nuclear localization (Fig. 3A). We therefore introduced single point mutations K¹⁵⁰A and K¹⁵⁵A into the expression construct. Transient expression in HeLa cells of GG-p75A¹⁴⁰ and GG-p75A¹⁵⁵ revealed that Lys¹⁵⁰ was essential for LEDGF/p75 nuclear localization under conditions wherein Lys¹⁵⁵ was dispensable (Fig. 3B). All nuclear import-defective LEDGF/p75 mutants showed a diffuse cytoplasmic distribution with varying levels of aggregation (e.g. see Figs. 2F and 3B). Western blot analysis using monoclonal anti-LEDGF/p75 antibody showed that the NLS mutants were all efficiently expressed in HeLa cells and had apparent molecular masses close to the predicted 115 kDa (Fig. 3C).

**Co-expression of pGG-P75A¹⁴⁰ Renders HIV-1 IN Cytoplasmic**—As previously described (31), fusing EGFP and red fluorescent protein HcRed1 to LEDGF/p75 and HIV-1 IN, respectively, did not perturb their nuclear co-localization upon transient expression in HeLa cells. In addition, the cellular distribution of LEDGF/p75 was not changed when fused to the larger GG double fusion (Fig. 2A). As expected, upon co-expression with HcRed1-tagged HIV-1 IN protein (HcRed1-IN), GG-p75 was co-localized with HcRed1-IN in nuclei of transfected cells (Fig. 4A). Strikingly, when HcRed1-IN was co-expressed with the NLS-defective mutant GG-p75A¹⁵⁵, both proteins became confined to the cytoplasm, forming apparent co-aggregates (Fig. 4B). Similar patterns were observed when
other GG-p75 NLS-defective mutants were co-expressed with HcRed1-IN (data not shown).

The LEDGF/p75 NLS Is Transferable—We fused the LEDGF/p75 NLS, 148GRKRKAEKQ156, to the N terminus of the E. coli β-galactosidase protein to test whether this sequence is sufficient to target a heterologous protein into the nucleus. Since the NLS of SV40 large T antigen is a canonical example of a transferable NLS (43), we also prepared a control construct that expressed the SV40 NLS fused to β-galactosidase. Upon transient transfection of HeLa cells with both constructs and staining with X-gal, β-galactosidase activity was clearly detected in the nuclei of transfected cells (Fig. 5, A and B). As expected, when wild type NLS-less protein was expressed from a control vector, β-galactosidase activity was detected throughout the cells (Fig. 5 C).

Nuclear Import of LEDGF/p75 Is a Saturable and Temperature-dependent Process and Requires Ran—To characterize the nuclear import pathway of LEDGF/p75 in more detail, we used the semipermeabilized cell system (38). In this assay, the cytoplasmic membrane is first permeabilized with digitonin, and the soluble endogenous cytosolic factors are depleted. Nuclear import of a fluorescently labeled protein can then be studied in the presence of exogenous nuclear import factors. We initially used a rabbit reticulocyte lysate as the complete source for the import factors (38). In the presence of the reticulocyte lysate, ATP, GTP, and an energy regenerating system, both Alexa 633-labeled recombinant LEDGF/p75 and rhodamine-labeled SV40-NLS substrates were readily imported into nuclei (Fig. 6A). Import was not observed when the reticulocyte lysate was omitted from the mixture (Fig. 6A). This result demonstrates that import of LEDGF/p75 is dependent on soluble cellular factors. The lectin WGA inhibits many nuclear transport mechanisms without affecting passive diffusion through the NPC by binding to N-acetylglucosamine-modified nucleoporins (44, 45). When we pretreated digitonin-permeabilized HeLa cells with 50 μg/ml WGA prior to the addition of LEDGF/p75 or bovine serum albumin-SV40 NLS to the import mixture, no nuclear import was detected (Fig. 6A). In addition, nuclear import was not observed at 4 °C or when the ATP-regenerating system was absent and ATP was replaced by nonhydrolyzable AMP-PNP (Fig. 6A). Importantly, an excess of unlabeled recombinant LEDGF/p75 was able to compete with the fluorescent LEDGF/p75 conjugate for import, which proves that nuclear import of LEDGF/p75 is a saturable process (Fig. 6B).

Nuclear import mediated by cNLSs is dependent on a gradient of RanGTP across the nuclear envelope. imp-β can only bind cargo in the cytoplasm if it is released from RanGTP via GTP hydrolysis. At the same time, imp-β can only release its cargo in the nucleus by binding to RanGTP. Therefore, imp-β-mediated import pathways are inhibited by RanQ69L, a Ran mutant incapable of hydrolyzing GTP (46). Similarly, nonhydrolyzable analogues of GTP, like GTPγS, also inhibit cNLS-dependent import (47, 48). Fig. 6C demonstrates that nuclear accumulation of LEDGF/p75 was dependent on RanGTP, since
import was not detected when RanQ69L was added to the import mixture or when GTP was replaced by GTP

LEDGF/p75 Is Imported into the Nucleus through the imp-α/β Pathway—Nuclear import of SV40 large T requires both imp-α and imp-β nuclear transport receptors. The results described above showed that LEDGF/p75 contains a SV40-like NLS and is imported into the nucleus in a temperature- and energy-dependent, saturable manner. Its dependence on Ran and GTP suggested that nuclear import of LEDGF/p75 most likely occurs through the imp-β transport receptor. To investigate whether nuclear import of LEDGF/p75 might occur via the same mechanism as SV40 large T, we performed the following

---

**Fig. 3. Lys^{150} is essential for nuclear import of LEDGF/p75.** A, summary of the changes (in boldface type) from Fig. 2 that inhibited nuclear localization of LEDGF/p75. Two residues, Lys^{150} and/or Lys^{155}, might be essential for LEDGF/p75 import. B, the single K150A mutation rendered GG-p75 cytoplasmic, whereas the K155A mutation did not affect the intracellular distribution of LEDGF/p75. C, Western blot analysis of transiently expressed GG-LEDGF/p75 mutants in HeLa cells. Expression was analyzed 24 h post-transfection by Western blot using monoclonal anti-LEDGF/p75 antibody. 15 μg of total protein was loaded in each well. Shown are nontransfected cells (lane 1) and cells expressing GG-p75 (lane 2) or its mutant derivatives K150A/R151A/K152A (lane 3), R149A/K150A/K155A (lane 4), K150A/R151A/K155A (lane 5), R150A (lane 6), K150A/K155A (lane 7), R146A/R147A (lane 8), and K155A (lane 9). The migration positions of mass standards are indicated on the left. The positions of GG-p75 and endogenous LEDGF/p75 are on the right. The protein that migrated slightly slower than the 62-kDa marker is likely to represent a minor degradation product of GG-p75.

---

**Fig. 4. An NLS-defective LEDGF/p75 confines HIV-1 IN to the cytoplasm.** A, GG-p75 and HcRed1-IN co-localize in the nuclei of transfected HeLa cells. Green, left panel, GG-p75; red, middle panel, HcRed1-IN; right panel, merged image. B, GG-p75^{150} co-expressed with HcRed1-IN traps IN in the cytoplasm. Co-aggregates can be observed in both the green and red detection channels. Green, left panels, GG-p75^{150}; red, middle panels, HcRed1-IN; right panels, merged image.
competition experiments (Fig. 7B). A 5-fold molar excess of unlabeled LEDGF/p75 was allowed to interact with the import receptors prior to the addition of the rhodamine-conjugated SV40-NLS substrate. Fig. 7B shows that the SV40-NLS substrate was not imported under these conditions. As a control, we analyzed purified NLS mutant LEDGF/p75A150 protein. As predicted, Alexa 633-labeled mutant protein was exclusively cytoplasmic at conditions that supported the ready transport of wild-type LEDGF/p75 (Fig. 7A). Consistent with this observation, the mutant protein was unable to compete with the SV40-NLS substrate for its nuclear import (Fig. 7B).

Proteins containing basic NLSs can be imported into the nucleus by imp-β either with (13) or without (49) the adaptor molecule imp-α. Another imp-β-like receptor, imp-7, has been implicated in the nuclear import of the highly positively charged linker histone H1. However, a heterodimer of imp-7 with imp-β was required for H1 nuclear import (21). We used a reconstitution assay to determine whether both imp-α and imp-β were required for LEDGF/p75. Ran GTPase alone or Ran and imp-β (Fig. 8) were insufficient to import LEDGF/p75 into the nucleus. It was only upon the addition of both imp-α and imp-β to the Ran mixture that nuclear import of LEDGF/
site-directed mutagenesis, we identified a single functional NLS within LEDGF/p75 (Table I). Through computer analysis of the amino acid sequence predicted a linear pattern. Thus, at least one NLS was expected to reside within the 325 N-terminal residues of LEDGF/p75. LEDGF/p75 and p52 share 325 amino acid residues that determine the specific nuclear localization and, as a consequence, redistribute the mutant protein to the cytoplasm. We therefore conclude that this single NLS of the eight predicted NLSs functions as a true NLS. We do note, however, that 24 h post-transfection, weak nuclear staining could be detected in about 2–5% of the cells expressing GG-p75\(^{\Delta 150}\) (or any other NLS mutant deficient in targeting GG-p75 to the nucleus) with a pattern similar to the nonmutated LEDGF/p75. We speculate that this low level of nuclear localization was a consequence of mitosis, wherein mutant proteins can bind to mitotic chromosomes following breakdown of the nuclear envelope and subsequently become captured in the nuclei. Compared with 100% nuclear import of wild type LEDGF/p75, this low level of mutant protein import does not affect our conclusions. LaCasse and Lefebvre (50) reported that in 67% of proteins that possess both DNA binding and NLS activities, the NLS overlaps with the DNA binding sequence. Cokol et al. (42) extended the data base search and expanded this number to 90%. Since NLS mutant LEDGF/p75 displayed the wild type distribution in the few cells that supported nuclear import, we speculate that the LEDGF/p75 NLS may not overlap with the protein’s DNA binding region(s).

Since transference of the LEDGF/p75 NLS onto the otherwise cytoplasmic \(\beta\)-galactosidase conferred nuclear localization to the E. coli protein (Fig. 5), we conclude that the LEDGF/p75 NLS is a linear epitope and is thus not composed of discontinuous epitopes coming together in the tertiary structure of the protein. Indeed, no specific secondary structure was predicted between residues 147 and 170 by various computer algorithms (data not shown).

LEDGF/p75 nuclear import was inhibited when mutant RanQ69L was added to the import mixture or when GTP was replaced by GTP\(\gamma\)S, which implies the function of an imp-\(\beta\) like transport receptor (18, 19, 51). Using a nuclear import reconstitution assay, we demonstrated the dependence of LEDGF/p75 on both imp-\(\alpha\) and imp-\(\beta\) for its nuclear localization (Fig. 8).

**DISCUSSION**

Identification and Characterization of the LEDGF/p75 NLS—LEDGF/p75 and its splice variant p52 are karyophilic proteins (30–32). Although both proteins are nuclear, they each have a markedly different spatial and temporal distribution in the nucleus (31, 32). LEDGF/p75 and p52 share 325 N-terminal residues (33). p52 has a unique 8-amino acid tail not present in LEDGF/p75, which does not contain an NLS-like pattern. Thus, at least one NLS was expected to reside within the N-terminal 325 residues of LEDGF/p75. Expression of GG-p75\(\Delta C\) and GG-p75/Ct deletion mutants in HeLa cells demonstrated that the C-terminal domain of LEDGF/p75 was not karyophilic and therefore does not contain a functional NLS (Fig. 2). Moreover, since the 4C mutant of LEDGF/p75 had a nuclear distribution pattern indistinguishable from the full-length LEDGF/p75, we speculate that it is not the lack of this C-terminal domain in p52 but rather the presence of the extra 8 amino acid residues that determines the specific nuclear distribution of p52. The function of the C-terminal domain of LEDGF/p75 remains to be determined. It has recently been suggested that it might play a role in the interaction of LEDGF/p75 with HIV-1 IN (31).

Computer analysis of the amino acid sequence predicted several putative NLSs within LEDGF/p75 (Table I). Through site-directed mutagenesis, we identified a single functional NLS in LEDGF/p75, \(146\)GHKRKAEEK\(^{152}\). Remarkably, a single amino acid change, K150A, was sufficient to block LEDGF/p75 nuclear localization and, as a consequence, redistribute the mutant protein to the cytoplasm. We therefore conclude that this single NLS of the eight predicted NLSs functions as a true NLS. We do note, however, that 24 h post-transfection, weak nuclear staining could be detected in about 2–5% of the cells expressing GG-p75\(^{\Delta 150}\) (or any other NLS mutant deficient in targeting GG-p75 to the nucleus) with a pattern similar to the nonmutated LEDGF/p75. We speculate that this low level of nuclear localization was a consequence of mitosis, wherein mutant proteins can bind to mitotic chromosomes following breakdown of the nuclear envelope and subsequently become captured in the nuclei. Compared with 100% nuclear import of wild type LEDGF/p75, this low level of mutant protein import does not affect our conclusions. LaCasse and Lefebvre (50) reported that in 67% of proteins that possess both DNA binding and NLS activities, the NLS overlaps with the DNA binding sequence. Cokol et al. (42) extended the data base search and expanded this number to 90%. Since NLS mutant LEDGF/p75 displayed the wild type distribution in the few cells that supported nuclear import, we speculate that the LEDGF/p75 NLS may not overlap with the protein’s DNA binding region(s).

Since transference of the LEDGF/p75 NLS onto the otherwise cytoplasmic \(\beta\)-galactosidase conferred nuclear localization to the E. coli protein (Fig. 5), we conclude that the LEDGF/p75 NLS is a linear epitope and is thus not composed of discontinuous epitopes coming together in the tertiary structure of the protein. Indeed, no specific secondary structure was predicted between residues 147 and 170 by various computer algorithms (data not shown).

LEDGF/p75 nuclear import was inhibited when mutant RanQ69L was added to the import mixture or when GTP was replaced by GTP\(\gamma\)S, which implies the function of an imp-\(\beta\) like transport receptor (18, 19, 51). Using a nuclear import reconstitution assay, we demonstrated the dependence of LEDGF/p75 on both imp-\(\alpha\) and imp-\(\beta\) for its nuclear localization (Fig. 8).

**Nuclear Import of HIV-1 IN—**Lentiviruses, in contrast to oncoretroviruses, can infect nondividing cells (52–55). The obvious requirement for productive infection of a nondividing cell is the presence of a nuclear targeting signal within the PIC. It was reported that the PIC is imported into the nucleus in an active, energy-dependent fashion. Indeed, nuclear accumulation of LEDGF/p75 was efficiently blocked by WGA and did not occur at 4 °C or in the absence of ATP (Fig. 6). Moreover, the NLS-defective K150A mutant failed to be imported in the \(\text{in vitro}\) system, suggesting that import in semipermeabilized cells and live cells occurred through functionally similar pathways.

Since the K150A and wild-type LEDGF/p75 proteins displayed similar purification profiles, our results exclude the possibility that a potential bacterial contaminant affected the observed \(\text{in vitro}\) nuclear localization profiles.

**LEDGF/p75** makes use of the imp-\(\alpha\)/\(\beta\) nuclear import pathway—Using digitonin-permeabilized HeLa cells, we showed that LEDGF/p75 was transported into the nucleus in an active, energy-dependent fashion. Indeed, nuclear accumulation of LEDGF/p75 was efficiently blocked by WGA and did not occur at 4 °C or in the absence of ATP (Fig. 6). Moreover, the NLS-defective K150A mutant failed to be imported in the \(\text{in vitro}\) system, suggesting that import in semipermeabilized cells and live cells occurred through functionally similar pathways.

Since the K150A and wild-type LEDGF/p75 proteins displayed similar purification profiles, our results exclude the possibility that a potential bacterial contaminant affected the observed \(\text{in vitro}\) nuclear localization profiles.

**LEDGF/p75** nuclear import was inhibited when mutant RanQ69L was added to the import mixture or when GTP was replaced by GTP\(\gamma\)S, which implies the function of an imp-\(\beta\) like transport receptor (18, 19, 51). Using a nuclear import reconstitution assay, we demonstrated the dependence of LEDGF/p75 on both imp-\(\alpha\) and imp-\(\beta\) for its nuclear localization (Fig. 8).

**Nuclear Import of HIV-1 IN—**Lentiviruses, in contrast to oncoretroviruses, can infect nondividing cells (52–55). The obvious requirement for productive infection of a nondividing cell is the presence of a nuclear targeting signal within the PIC. It was reported that the PIC is imported into the nucleus in an active, energy-dependent manner (56). The mechanism of lentiviral PIC nuclear import, however, has not been established. Viral proteins Vpr, matrix, and IN as well as the central DNA flap structure have been implicated in HIV nuclear import (24, 28, 57–62). Subsequent research revealed that matrix lacks an NLS and is not strictly required for infection of macrophages (63, 64). Moreover, Vpr was shown not to be essential for PIC import and virus replication in nondividing cells (63, 65, 67). In addition, the central DNA flap appears to be host cell- and viral strain-dependent (68, 69). IN is an obligatory component of the PIC and must be associated with it throughout the translocation event. Several groups suggested that IN might be the important nuclear targeting factor of the PIC. Gallay et al. (28)
reported that recombinant HIV-1 IN can be imported in semi-buffer, with Ran mixture (3 μM RanGDP and 0.5 μM nuclear transport factor 2) (Ran), plus the addition of 1 μM imp-β (β) or a 1 μM concentration of each imp-α and imp-β (α + β) with 100 μg/ml LEDGF/p75 (upper panels) or 120 μg/ml of the SV40-NLS substrate (lower panels) as a nuclear import substrate. Nuclear import of both LEDGF/p75 and SV40-NLS could only be detected when Ran, imp-α, and imp-β were present.

FIG. 8. LEDGF/p75 is imported into the nucleus through the imp-α/β pathway. Reconstitution experiments were performed in transport buffer (control), with Ran mixture (3 μM RanGDP and 0.5 μM nuclear transport factor 2) (Ran), plus the addition of 1 μM imp-β (β) or a 1 μM concentration of each imp-α and imp-β (α + β) with 100 μg/ml LEDGF/p75 (upper panels) or 120 μg/ml of the SV40-NLS substrate (lower panels) as a nuclear import substrate. Nuclear import of both LEDGF/p75 and SV40-NLS could only be detected when Ran, imp-α, and imp-β were present.

described a bipartite NLS in HIV-1 IN. Site-directed mutagenesis in the region of this bipartite NLS, however, did not abolish the nuclear localization of EGFP-fused HIV-1 IN (29). Depienne et al. (59) reported that HIV-1 IN was imported into the nuclei of semi-permeabilized cells using an uncharacterized ATP-dependent but GTP- and Ran-independent mechanism. Strikingly, the addition of cytosolic factors was not required, suggesting that the necessary import receptors remained in those digitonin-permeabilized cells (59).

In our hands, fluorescently labeled recombinant HIV-1 IN readily accumulated in the nuclei of semi-permeabilized cells in a temperature-, GTP-, and ATP-dependent fashion (data not shown). We wondered whether LEDGF/p75 would promote the nuclear import of IN in this in vitro system. When LEDGF/p75 was allowed to form complexes with IN prior to the addition to the import mixture, nuclear accumulation of IN was not enhanced. Although the exact stoichiometry of the IN/LEDGF/p75 complex was not reported, its size was estimated at around 300 kDa, probably containing a pair of IN tetramers (30). If a couple of imp-α/β heterodimers (~150 kDa) were bound to the IN/LEDGF/p75 complex, the size of the cargo-importin complex would exceed 500 kDa. We speculate that import of large IN/LEDGF/p75 complexes might require specific isoforms of imp-α and/or imp-β families, which were not functionally present in our in vitro assay. Alternatively, the structure of the NPC might be altered in the digitonin-permeabilized cell system, effectively precluding the efficient nuclear import of large protein complexes.

Multiple elements of the HIV-1 PIC have been proposed as essential for its nuclear localization, which could potentially work in an additive and/or interdependent fashion. It is not uncommon that proteins or complexes make use of several nuclear import pathways (8, 10). This would be beneficial for the virus, since one mutation aboliing an interaction with one nuclear import receptor would still allow the virus to gain the nuclear environment through an alternative pathway.

In conclusion, we mapped the nuclear targeting signal in the LEDGF/p75 protein to the short peptide (148GRKRKAEKQ166) and showed that it functioned similarly to the SV40 large T NLS in a RanGTP- and imp-α/β-dependent manner. We demonstrated that one single amino acid change, K150A, in LEDGF/p75 was sufficient to abolish the nuclear localization of the mutant protein and co-expressed HIV-1 IN. However, the specific importin-α/β isoforms involved in nuclear import of LEDGF/p75 and/or in its complex with HIV-1 IN are yet to be identified. Experiments are under way to establish the precise role of this cellular factor in retroviral replication.

Acknowledgments—Fluorescence microscopy was performed at the Confocal Microscopy Core Facility of the Brigham and Women’s Hospital (Boston, MA). We thank Dr. Philip G. Allen for technical assistance and helpful discussions.
Identification and Characterization of a Functional Nuclear Localization Signal in the HIV-1 Integrase Interactor LEDGF/p75
Goedele Maertens, Peter Cherepanov, Zeger Debyser, Yves Engelborghs and Alan Engelman

J. Biol. Chem. 2004, 279:33421-33429. doi: 10.1074/jbc.M404700200 originally published online May 25, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M404700200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 74 references, 40 of which can be accessed free at http://www.jbc.org/content/279/32/33421.full.html#ref-list-1