

2010-01-01

# Evaluating the Role of Evolutionarily Conserved Regions of LEDGF/p75 in HIV-1 Infection

Jose Adrian Garcia

*University of Texas at El Paso*, jose.525@gmail.com

Follow this and additional works at: [https://digitalcommons.utep.edu/open\\_etd](https://digitalcommons.utep.edu/open_etd)



Part of the [Molecular Biology Commons](#)

---

## Recommended Citation

Garcia, Jose Adrian, "Evaluating the Role of Evolutionarily Conserved Regions of LEDGF/p75 in HIV-1 Infection" (2010). *Open Access Theses & Dissertations*. 2681.

[https://digitalcommons.utep.edu/open\\_etd/2681](https://digitalcommons.utep.edu/open_etd/2681)

This is brought to you for free and open access by DigitalCommons@UTEP. It has been accepted for inclusion in Open Access Theses & Dissertations by an authorized administrator of DigitalCommons@UTEP. For more information, please contact [lweber@utep.edu](mailto:lweber@utep.edu).

**EVALUATING THE ROLE OF EVOLUTIONARILY CONSERVED REGIONS OF  
LEDGF/p75 IN HIV-1 INFECTION**

**JOSE A. GARCIA-RIVERA**  
**Department of Biological Sciences**

**APPROVED:**

---

**Manuel Llano, M.D., Ph.D., Chair**

---

**Igor C. Almeida, Ph.D.**

---

**German Rosas-Acosta, Ph.D.**

---

**Mahesh Narayan, Ph.D.**

---

**Kristin L. Gosselink, Ph.D.**

---

**Martin E. Fernandez-Zapico, M.D.**

---

**Patricia D. Witherspoon, Ph.D.**

**Dean of the Graduate School**

Copyright ©

By

Jose A. Garcia-Rivera

2010

## **Dedication**

I would like to dedicate this work to my baby brother Daniel who has been my inspiration for every single goal I have chosen to pursue in my life, you are always in my heart no matter how far we may be apart...



**EVALUATING THE ROLE OF EVOLUTIONARILY CONSERVED REGIONS OF  
LEDGF/p75 IN HIV-1 INFECTION**

**by**

**JOSE A. GARCIA-RIVERA, B.S.**

**DISSERTATION**

**Presented to the Faculty of the Graduate School of**

**The University of Texas at El Paso**

**in Partial Fulfillment**

**of the Requirements**

**for the Degree of**

**DOCTOR OF PHILOSOPHY**

**Department of Biological Sciences**

**THE UNIVERSITY OF TEXAS AT EL PASO**

**August 2010**

### **Acknowledgements**

I would like to thank first and foremost my family, who has stood behind me 100% in all the decisions I have made throughout my whole life. I sincerely believe that I would not have been able to complete my graduate studies without their support. I love you all!

I would like to extend my sincerest thanks to Dr. Manuel Llano who has not only been a mentor to me but also a friend. This work wouldn't have been possible without his relentless concern and support throughout my project. I would also like to thank my fellow lab members who thanks to their friendship and collaborative support makes going to work every day a pleasure. Last but not least I would like to thank all the members of my dissertation committee who have provided valuable insights on how to improve my experimental procedures.

## Abstract

LEDGF/p75 is an important cellular co-factor for lentiviral integration. LEDGF/p75-deficient cells are markedly resistant to HIV-1 infection and re-expression of the wild type protein rescues infectivity. Although the molecular mechanism of LEDGF/p75 in HIV-1 integration is not yet known, this co-factor activity requires the interaction of LEDGF/p75 with both the host chromatin and viral integrase. In order to evaluate the involvement of other LEDGF/p75 regions in HIV-1 infection we constructed a panel of deletion mutants targeting clusters of charged residues that are evolutionarily conserved and predicted to be post-translationally modified. These mutants were evaluated for their ability to rescue HIV-1 infection in a LEDGF/p75-deficient human T cell line (T<sub>L3</sub> cells) and to interact with chromatin and integrase. Our results indicate that serine residues 271, 273, and 275 are involved in the HIV-1 cofactor role of LEDGF/p75. HIV-1 infection is impaired in LEDGF/p75-deficient CD4<sup>+</sup> T cells expressing the LEDGF/p75 mutant S271A/S273A/S275A to levels observed with LEDGF/p75 mutants where chromatin binding is deficient. However, this mutation did not affect chromatin or integrase binding. More importantly, this mutant tethers EGFP-tagged integrase to the host chromatin in cells. According to five different global phosphoproteomic studies done by other laboratories, these residues are targeted for phosphorylation in both human and mouse cells. In correlation with this, *in silico* analysis predicted that these serine residues are a substrate for Protein Kinase Casein Kinase II. This data strongly suggests that the role of LEDGF/p75 in HIV-1 integration may also involve other protein functions in addition to integrase and chromatin binding.

## Table of Contents

Acknowledgements.....	v
Abstract.....	vi
Table of Contents.....	vii
List of Tables.....	ix
List of Figures.....	x
Chapter 1: General Introduction.....	1
1.1.    HIV-1 Life Cycle.....	1
1.2.    Molecular Aspect of HIV-1 Viral Entry.....	3
1.3.    Molecular Aspects of HIV-1 Integration.....	5
1.4.    HIV-1 Vaccine Strategies.....	10
1.5.    LEDGF/p75's Involvement in HIV-1 Integration.....	12
1.6.    Project Hypothesis.....	17
1.7.    Significance and Aims.....	18
Chapter 2: Implication of Serine Residues 271, 273, and 275 in the Human Immunodeficiency Virus Type 1 Cofactor Activity of Lens Epithelium-Derived Growth Factor/p75.....	23
2.1.    Introduction.....	23
2.2.    Materials and Methods.....	26
2.3.    Results.....	35
2.4.    Discussion.....	62
2.5.    Acknowledgements.....	65
Chapter 3: LEDGF/p75 is a Member of the Transcriptional Elongation Complex.....	66
3.1.    Introduction.....	66

3.2.	Materials and Methods.....	69
3.3.	Results.....	72
3.4.	Discussion.....	76
3.5.	Acknowledgments.....	79
Chapter 4: Final Conclusions and Future Directions.....		80
4.1.	Overview and Final Conclusions.....	80
4.2.	Future Directions.....	84
References.....		85
Appendix.....		95
Abbreviations.....		95
List of Publications and Manuscripts.....		97
Curriculum Vitae.....		98
Supplemental Material.....		100

## **List of Tables**

<b>Table 1.</b> Potential LEDGF/p75 Interactors.....	67
<b>Table 2.</b> LEDGF/p75 mutants.....	104

## List of Figures

<b>Figure 1.</b> HIV-1 life cycle.....	2
<b>Figure 2.</b> NHEJ Pathway.....	6
<b>Figure 3.</b> Post Integration Repair.....	8
<b>Figure 4.</b> LEDGF/p75 interacts with HIV-1 integrase.....	13
<b>Figure 5.</b> LEDGF/p75 is essential for HIV-1 infection.....	14
<b>Figure 6.</b> LEDGF/p75 tethering model.....	15
<b>Figure 7.</b> Newly proposed molecular mechanism of LEDGF/p75 in HIV-1 integration.....	17
<b>Figure 8.</b> LEDGF/p75 chromatin binding and integrase binding domains.....	20
<b>Figure 9.</b> Salt extraction and infectivity data.....	38
<b>Figure 10.</b> Evaluation of the HIV-1 cofactor activities of LEDGF/p75 deletion mutants.....	43
<b>Figure 11.</b> HIV-1 cofactor activity of different LEDGF/p75 CR3 mutants.....	47
<b>Figure 12.</b> HIV-1 DNA integration in T <sub>L3</sub> cells expressing LEDGF/p75 S271A/S273A/S275A.....	51
<b>Figure 13.</b> Chromatin-binding activity of LEDGF/p75 mutants.....	54
<b>Figure 14.</b> Interaction of LEDGF/p75 mutants with HIV-1 integrase.....	57
<b>Figure 15.</b> Integrase to chromatin tethering assay.....	60
<b>Figure 16.</b> Interaction of LEDGF/p75 with components of the elongation complex.....	73
<b>Figure 17.</b> Quantitative confocal colocalization of LEDGF/p75 with proteins of the transcriptional machinery.....	75
<b>Figure 18.</b> Research Implications.....	82

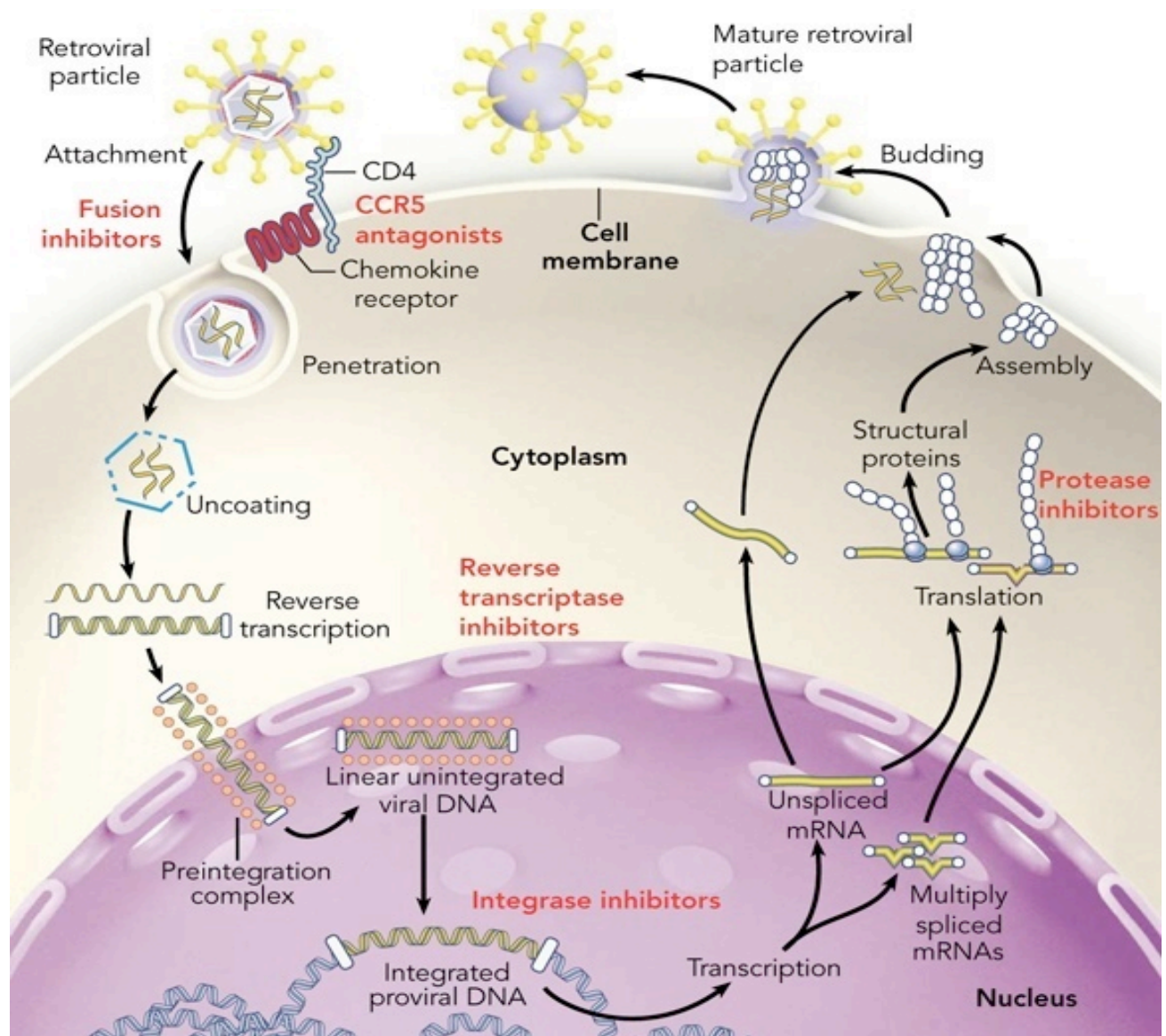
<b>Figure S1. LEDGF/p75 protein sequences alignment.....</b>	<b>100</b>
--------------------------------------------------------------	------------



## **Chapter 1: General Introduction**

### **1.1 HIV-1 Life Cycle**

HIV-1 is a retrovirus that enters target cells after fusion of viral and cellular membranes. Fusion is mediated by a specific interaction of glycoproteins on the surface of the virus with receptors on the cell's surface. Once the viral RNA genome is successfully released into the cytosol of the host cell, it is then immediately reverse transcribed into double-stranded cDNA containing a copy of the viral long terminal repeat (LTR) at each end. This viral DNA associates with viral and cellular proteins to form a nucleoprotein complex known as the preintegration complex (PIC), which is translocated into the nucleus. Integration of the viral DNA into the host chromatin is then catalyzed by the viral enzyme integrase (IN), which performs two enzymatic activities essential for viral integration; 3' processing of the viral DNA and DNA strand transfer of the viral genome into the host chromatin. 3' processing occurs soon after reverse transcription when a dinucleotide from each HIV-1 LTR is cleaved off, leaving two sticky ends on either end of the viral DNA [1, 2]. IN's second enzymatic activity, DNA strand transfer, takes place after the PIC enters the nucleus and locates a suitable integration site resulting in viral integration [3]. After strand transfer, single stranded gaps flank the ends of the viral genome. Efficient viral integration requires DNA repair of the gaps by host repair machinery. Only after integration can the viral genome be efficiently transcribed and new progeny produced, making integration a crucial step in HIV-1 infection [4] (Fig.1).



**Figure 1.** *HIV-1* life cycle. Human immunodeficiency virus (HIV) is a retrovirus that can lead to *acquired immunodeficiency syndrome* (AIDS), a condition in humans in which the immune system begins to fail, leading to life-threatening opportunistic infections (google images, 2007).

## 1.2. Molecular Aspects of HIV-1 Viral Entry

The entry of HIV-1 into target cells requires the binding of the outer envelope glycoprotein (gp120) to the CD4 receptor [5] and a coreceptor, CCR5 [6] or CXCR4 [7]. gp120 is composed of five conserved regions (C1-C5) and five hypervariable regions (V1-V5) [8]. This protein is heavily glycosylated with 50% of its molecular mass made up of carbohydrates and contains between 20-25 recognition sites for N-linked glycosylation [9]. N-glycans within gp120 play roles in both viral entry and in shielding the virus against neutralizing antibodies [10].

gp120 contains five loops (V1-V5) with V3 being the most immunodominant and most targeted by neutralizing antibodies [10]. This loop interacts with both co-receptors, CCR5 and CXCR4 [11]. N-Glycans within V1 and V2 (g2-g7) are thought to act as a barrier by preventing binding of V3-specific neutralizing antibodies to the V3 loop [12]. It has been demonstrated that N-glycans within the V1/V2 variable regions of gp120 are indispensable for both maintaining viral functionality and acting as masking epitopes against neutralization [12]. In a study done by Wolk *et al* it was shown that N-glycans g4, g5, g6 and g7 were important for viral infectivity since mutants lacking these N-glycans were highly deficient in viral entry. It was also shown that mutants lacking g2, g3, g5 and g6 were highly sensitive to neutralizing antibodies meaning that these N-glycans play a role in shielding HIV-1 from neutralizing antibodies [12].

Once gp120 binds to its receptor (CD4) and co-receptor (CCR5 or CXCR4) a conformational change is triggered that causes the insertion of the hydrophobic N-terminus of gp41 into the target cell membrane [13]. This HIV-1 transmembrane protein (gp41) typically contains three to four highly conserved sites for N-glycan attachment which are located in a short stretch of amino acids (20 to 30 residues) [14]. Experimental data has demonstrated that these

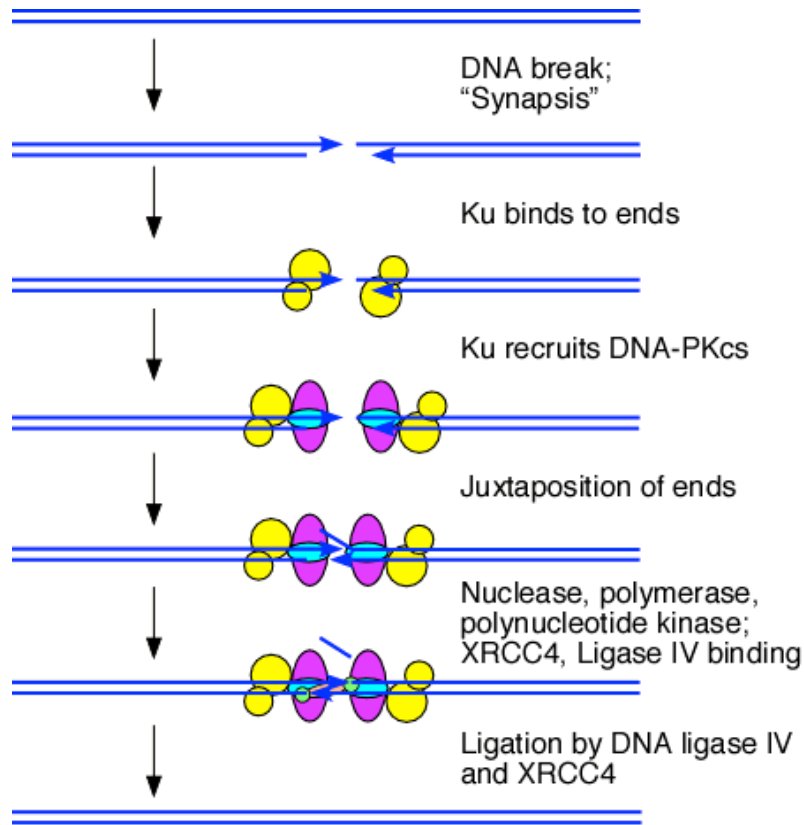
sites are largely dispensable for HIV-1 replication. This was shown in an elegant study done by Johnson *et al* in which they systematically mutated the N-glycan attachment sites of gp41 not only individually but also in all possible combinations and then measured the effect of each of these mutants on viral replication. What they reported was that only the quadruple mutant (g1234) in which all four asparagines residues which serve as N-glycan attachment sites were mutated to glutamine residues did not replicate in any cell line and were also not fusogenic [14]. The reason behind this mutation was the fact that glutamine is structurally similar to asparagines, differing by only a single methylene group.

### 1.3. Molecular Aspects of HIV-1 Integration

Integrase performs two major catalytic activities: 3'-processing, an endonucleolytic cleavage at each 3'-OH end of the viral cDNA and a strand transfer reaction, which leads to the insertion of the processed viral DNA into the target DNA. Once integrated, the provirus persists in the host cell and serves as a template for the transcription of viral genes and replication of the viral genome leads to the production of new viral progeny.

Post-integration repair (PIR) is an essential step in the retroviral lifecycle [15]. PIR occurs after 3'-processing and DNA strand transfer. An intermediate is produced during the initial integrase-mediated linkage that occurs between the viral and host DNA in which the proviral DNA is flanked by short, single-stranded gaps in the host-cell DNA. Integration is then completed by PIR which is composed of four distinct steps: (1) trimming of the 2-bp flaps from the 5'-ends of the proviral DNA, (2) filling in of the single-stranded gaps, (3) ligation of the trimmed 5'-ends to the filled in host DNA strands, and (4) reconstitution of the appropriate chromatin structure at the integration site [15, 16].

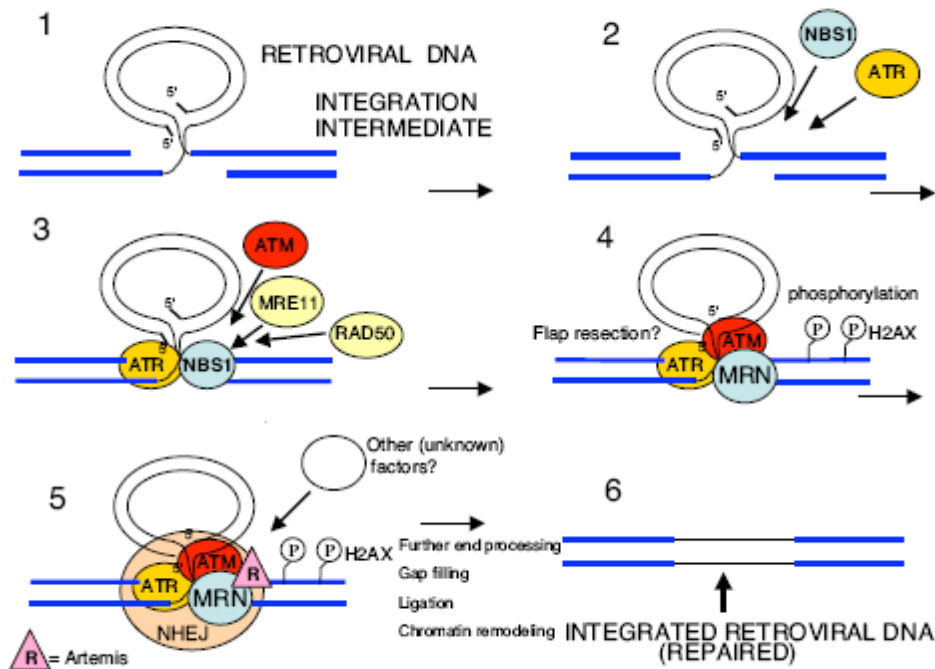
It has been proposed that Nijmegen breakage syndrome-1 protein (NBS1), which is one of the earliest known sensors of double-stranded DNA breakage (DSB) is involved in HIV-1 PIR. Smith *et al.* demonstrated that NBS1 associates with viral DNA and mediates efficient integration. They also proposed that retroviral PIR employs the *non-homologous end joining* (NHEJ) pathway (Fig. 2).



**Figure 2.** *NHEJ Pathway.* DNA double stranded breakage in which the DNA ends are non homologous are mainly repaired through the NHEJ pathway. This pathway employs a number of proteins including Ku, DNA-PK<sub>cs</sub>, and Ligase IV in order to repair DNA lesions [15].

The NHEJ pathway is a repair mechanism used by cells to repair DSB [17]. This pathway employs several proteins that function in concert with each other to repair DNA damage. Once a DSB occurs, Ku, a heterodimer that is an abundant DNA-binding protein with ATPase activity, binds strongly to the DNA ends. Ku then recruits to the DNA ends DNA-PK<sub>cs</sub> which is trimeric protein consisting of DNA-PK<sub>cs</sub>, Ku70, and Ku86. Artemis, which is a substrate of DNA-PK, is then recruited by DNA-PK<sub>cs</sub>, which then phosphorylates Artemis along with itself. XRCC4, which is another protein recruited by Ku whose recruitment is independent of DNA-PK<sub>cs</sub> and Artemis, then recruits DNA Ligase IV thus allowing the ends of the DSB to be ligated together [18, 19].

It has been proposed by a number of laboratories that PIR hijacks the host cell's DSB pathways thus allowing retroviruses such as HIV-1 to fully integrate their viral DNA within that of the host cell (Fig. 3).



**Figure 3. Post Integration Repair.** According to the model proposed by Smith *et al.* this process incurs a number of steps. (1) First, integrase catalyzes the formation of the integration intermediate within the host cells's DNA, (2) NBS1 and ataxia telangiectasia-mutated kinase (ATR) are then each independently recruited to the site of integration, (3) NBS1 then recruits other proteins such as MREI I, RAD50, and ATM to the integration site, (4) the 5'-end DNA flaps of the proviral DNA are then trimmed and ATM phosphorylates H2AX (although it was found that this step is dispensable for PIR), (5) it is then thought that the Artemis and NHEJ proteins are recruited to the site of integration and that these proteins along with other cellular factors mediate the final steps of PIR. These steps could include further end processing, gap filling, ligation and chromatin remodeling [15].



Initial studies on PIR have shown that it seems to use a somewhat different type of cellular machinery than DSB repair. For instance DSB repair requires the phosphorylation of histone H2AX which is mediated by the ATM and ATR kinases but in the case of PIR this step appears to be dispensable [20]. Another difference between DSB and PIR is that DSB requires both NBS1 and ATM to recruit ATR but in regard to PIR it is recruited independently of these two proteins, this difference in ATR recruitment could be explained by the thought that the single stranded DNA gaps, which flank the integration site, are sufficient to recruit the ATR protein [15].

#### **1.4. HIV-1 Vaccine Strategies**

So far, all attempts to develop an effective vaccine against HIV-1 have failed. This is mainly due to high mutational rate presented by the virus allowing it to escape immunological surveillance [21-23]. Glycans present on the viral envelope have a crucial role in enabling the virus to efficiently fuse to target cells in order to insert its genome within the host [24]. These glycans also function to disguise the virus from the immune system by acting as “self” antigens on the viral surface [25]. Therefore agents that interact with these glycans could not only interfere with viral entry but also allow the virus’s envelope glycan shield to be breached thus allowing previously shielded immunodominant epitopes of the viral envelope to be recognized by the immune system [26]. These carbohydrate-binding agents (CBAs) may serve as an important tool in vaccine development against HIV-1 [23].

It has been postulated that CBAs could be the first chemotherapeutics with a dual mechanism of antiviral action. First, through direct antiviral activity by binding directly to glycans on the virus’s surface thus blocking viral entry and second through indirect antiviral activity resulting from a progressive loss of protective glycans from the viral envelope allowing for an immune response to take place against previously hidden immunogenic epitopes in the viral envelope [23].

CBAs can inhibit HIV-1 infection of dendritic cells (DCs) and DC-directed HIV-1 transfer [27]. This transfer takes place during the initiation of infection when DCs present at the site of viral entry (skin dermis, placenta, and genital mucosa) capture HIV-1 virions through their DC-SIGN receptors and present them to immature T cells in the draining lymph nodes [23]. At the moment of presentation virus particles are present at the cell-cell junctions and this is thought

to act as an infectious synapse in which the passage of virions between the two cell types is facilitated, thus causing infection of T cells [28].

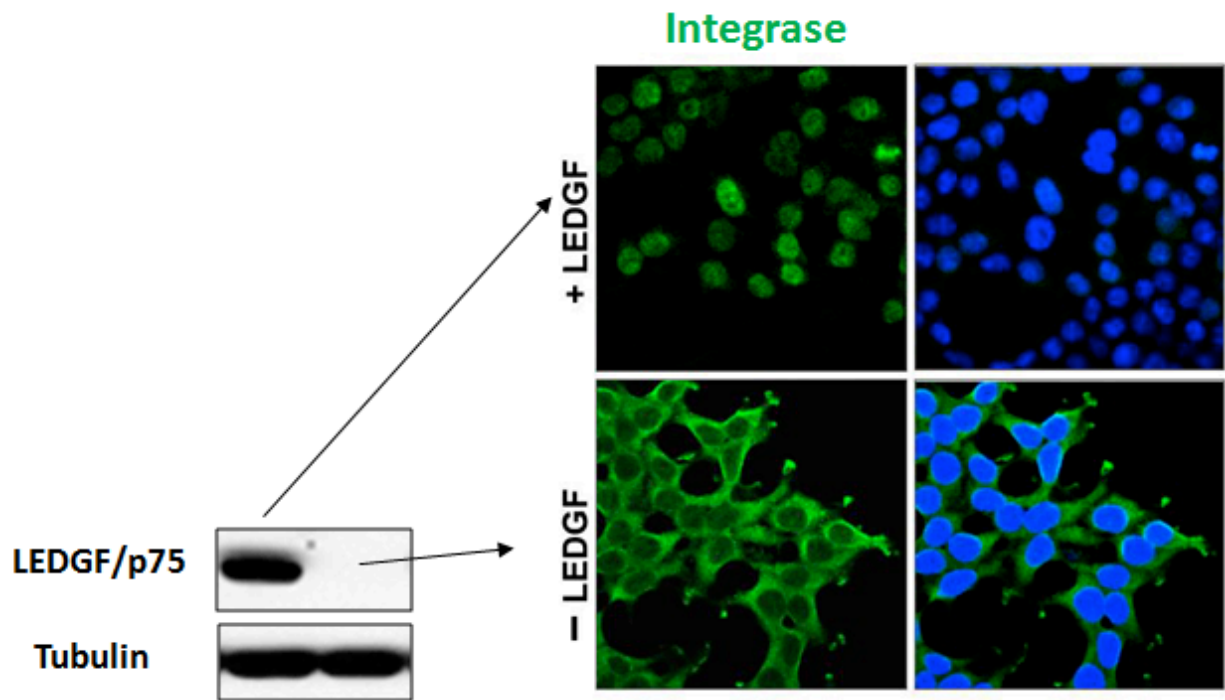
There has been much evidence indicating that the glycan shield present on the viral envelope helps the virion escape immune surveillance [23]. HIV-1 strains lacking the highly conserved N-linked glycan at position 306 within the V3 loop of gp120 are highly sensitive to neutralization [29]. Studies with SIVmac239, a strain of simian immunodeficiency virus that is highly resistant to neutralization by polyclonal antisera and monoclonal antibodies have shown that the elimination of N-glycan attachment sites within gp120 result in a dramatic increase in sensitivity to neutralization by monoclonal antibodies [30]. These results indicate that the deletion of as few as two N-glycans within the viral envelope can trigger a significant immune response [23].

CBAs directly interact with glycans present on the envelope gp120 of HIV-1. This interaction progressively forces the virus to delete N-glycans in gp120 in an attempt to escape recognition and in doing so exposes viral epitopes previously hidden, thus triggering an immune response to neutralize the invading pathogen. It is important to note that CBAs do not need to be taken up by a virus-infected cell in order to exert their antiviral activity and in turn do not disturb natural cellular glycan formation [23].

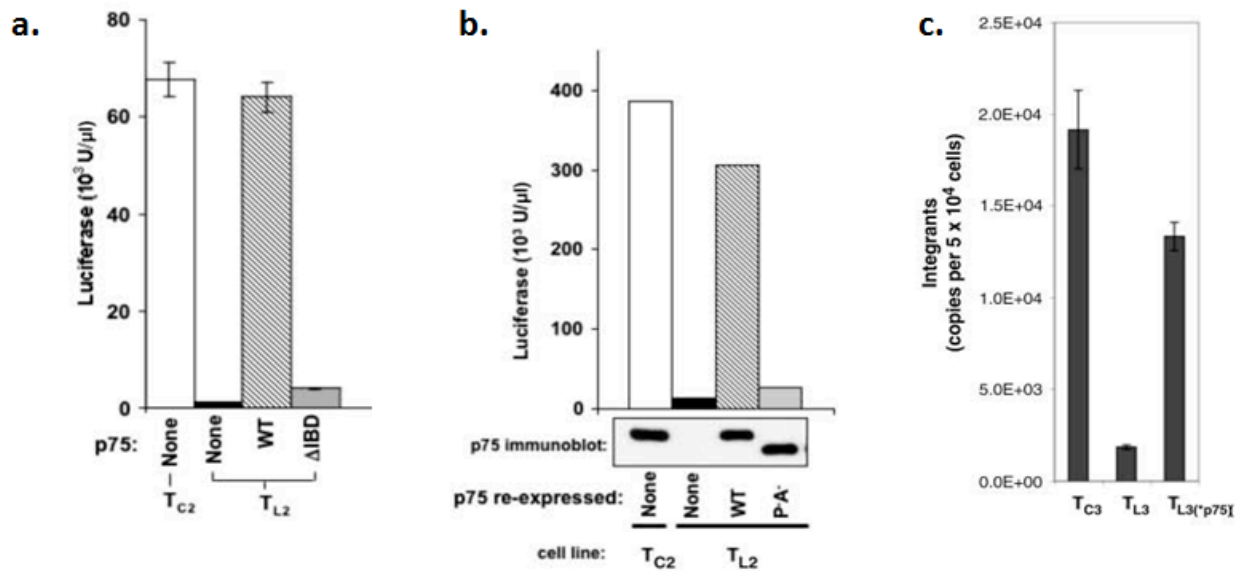
To date, no existing antiviral chemotherapeutic has demonstrated the ability to work in concert with the patient's immune system to further increase therapeutic pressure on mutated forms of the HIV-1 virus. Therefore CBAs have the potential to become the first strategy to combine drug mediated virus suppression with the induction of a specific immunological response to the pathogen [26].

### **1.5. LEDGF/p75's Involvement in HIV-1 Integration**

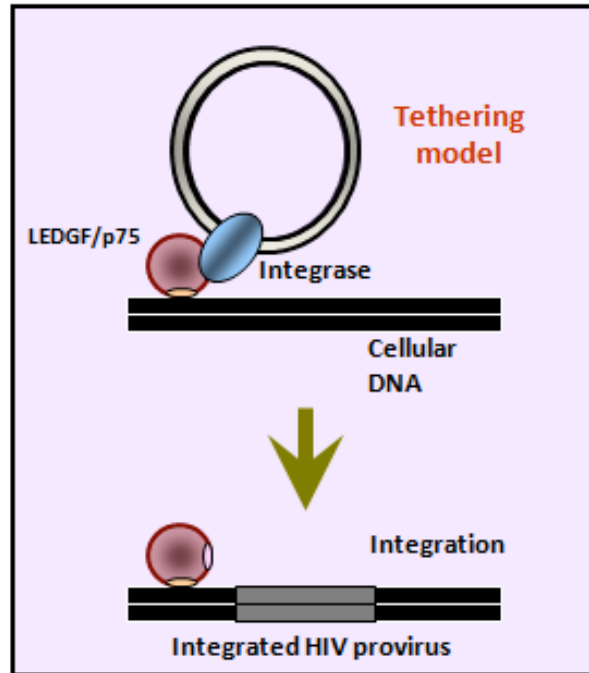
HIV-1 integration is a complex phenomenon that requires the concerted action of viral and cellular proteins. The lens epithelium-derived growth factor/p75 (LEDGF/p75) is a cofactor for lentiviral integration. LEDGF/p75 is a nuclear protein that belongs to the family of hepatoma-derived growth factors. This HIV-1 cofactor is an alternatively spliced product from the gene, PSIP1 [31, 32]. Increased interest has drawn among HIV-1 researchers since the discovery that this protein interacts with HIV-1 integrase (Fig. 4) [33]. Subsequently, it was demonstrated that LEDGF/p75 is essential for HIV-1 integration (Fig. 5) [4]. In this study, it was found that HIV-1 target cells deficient in LEDGF/p75 were resistant to HIV-1 integration and therefore to infection [4]. Infectivity was rescued in these cells by re-expression of LEDGF/p75 wild type but not of mutants impaired in their chromatin or integrase binding activity. It was then postulated that LEDGF/p75 acts as a molecular tether attracting the nuclear imported PIC to the host chromatin (Fig. 6).



**Figure 4.** *LEDGF/p75 interacts with HIV-1 integrase.* HIV-1 IN is un-tethered from chromatin by depletion of LEDGF/p75 and displays a pan-cellular distribution within the cell [34].



**Figure 5.** *LEDGF/p75 is essential for HIV-1 infection.* (a) LEDGF/p75 reexpression in LEDGF/p75 deficient cells rescues HIV infectivity while reexpression of a LEDGF/p75 mutant lacking the integrase binding domain (IBD) does not. (b) LEDGF/p75 reexpression in LEDGF/p75 deficient cells rescues HIV infectivity while reexpression of a LEDGF/p75 mutant lacking the chromatin binding domain made up by the PWWP and AT Hook motifs does not. (c) Alu-PCR results. Reexpression of LEDGF/p75 rescues the number of integrates per cell. [4]



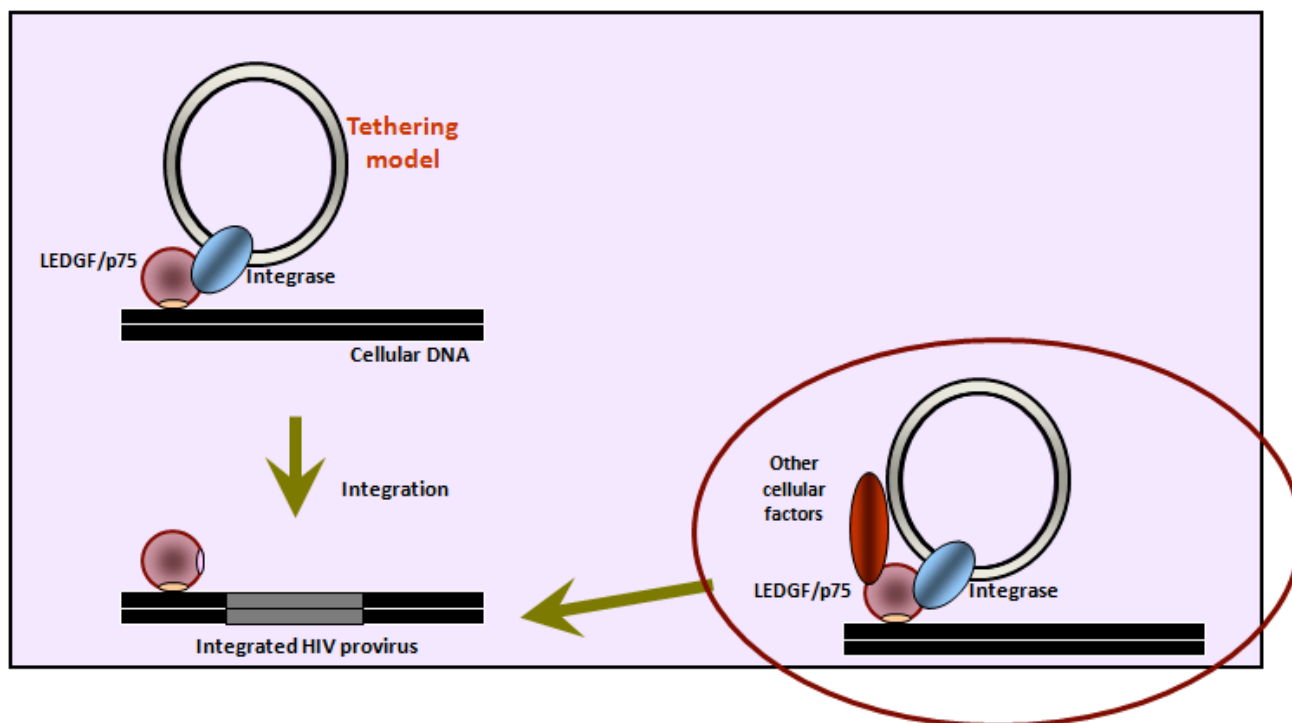
**Figure 6.** *LEDGF/p75 tethering model.* LEDGF/p75 acts as a tether between viral integrase and chromatin facilitating the integration of the viral DNA into the host genome [4].

Site selection for HIV-1 integration is not a random process; it is bias towards gene rich sites rather than non-coding regions of DNA [35]. Lentiviruses such as HIV-1 greatly prefer transcription units over gene-poor regions and integration occurs nearly equally along gene lengths, and positively correlates with transcriptional activity [36]. HIV-1 integration site distribution is altered in LEDGF/p75 deficient cells with more than half of the favoring of transcription units being eliminated with the depletion of LEDGF/p75 [35]. These results suggest that the frequency and distribution of HIV-1 integration is in large part governed via the interaction between IN and cellular LEDGF/p75 [37].



## 1.6. Project Hypothesis

Our hypothesis is that evolutionary conserved regions of LEDGF/p75, rich in charged residues, therefore solvent exposed, and predicted to harbor post-transcriptional modifications may play an important role in the HIV-1 cofactor activity of LEDGF/p75. We propose that these regions function by recruiting other cellular factors that facilitate HIV-1 integration (Fig. 7).



**Figure 7.** *Newly proposed molecular mechanism of LEDGF/p75 in HIV-1 integration.* Contrary to the tethering model, this model proposes that LEDGF/p75 recruits other cellular factors that allow integration to occur more efficiently.

## 1.7. Significance and Aims

### Significance

HIV-1 integration is an obligated step in its life cycle, viral replication can only occur after the formation of the provirus. LEDGF/p75 is a required cellular co-factor for HIV-1 integration [4]. Understanding the molecular mechanism of LEDGF/p75 in HIV-1 integration will significantly contribute to the identification of new targets for anti-HIV-1 drug development.

#### ***Specific Aim 1: Generation of stable cell lines expressing LEDGF/p75 mutants.***

Using different prediction software we identified seven main LEDGF/p75 evolutionarily conserved regions which are rich in charged residues and that are targeted by different post-translational modifications. Deletion mutants lacking these regions were generated by PCR mutagenesis and stably expressed in LEDGF/p75-deficient human CD4<sup>+</sup> T cells.

**1.1 Selection of LEDGF/p75 regions to evaluate.** Regions were chosen for evaluation based on stringent selection criteria. (1) The regions must be evolutionarily conserved from amphibians to humans. (2) They must contain clusters of charged residues. (3) These regions should serve as potential targets of post-translational modifications. (4) Their role in HIV-1 infection is unknown.

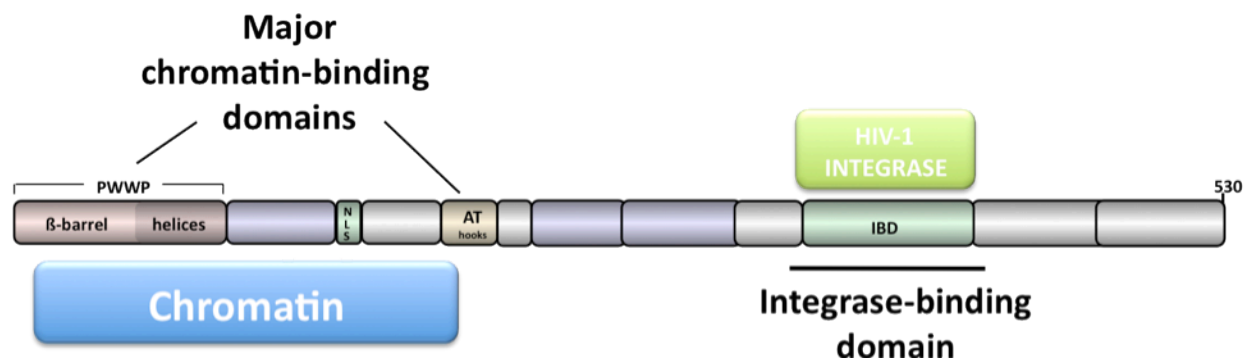
#### **1.2 Stable expression of LEDGF/p75 mutants in LEDGF/p75-deficient human CD4<sup>+</sup> T cells.**

The HIV-1 co-factor activity of LEDGF/p75 was evaluated by re-expressing different deletion mutants in LEDGF/p75-deficient CD4<sup>+</sup> T cells (T<sub>L3</sub> cells). These cells were generated by

transduction with a lentiviral vector expressing an shRNA against LEDGF/p75. Using this procedure, a 97% reduction of the LEDGF/p75 mRNA levels was permanently achieved [4]. HIV-1 infection is severely impaired in T<sub>L3</sub> cells and infectivity is rescued by re-expression of LEDGF/p75. Retroviral transduction was used to stably express the different LEDGF/p75 deletion mutants in T<sub>L3</sub> cells.

**Specific Aim 2: Evaluate the role of different LEDGF/p75 regions in HIV-1 infection.**

T<sub>L3</sub> cells stably expressing the different LEDGF/p75 deletion mutants were challenged with replication defective HIV-1 luciferase viruses. Infectivity was evaluated five days later by measuring luciferase activity in the challenged cells. As control, T<sub>L3</sub> expressing no LEDGF/p75, LEDGF/p75 wild type or mutants lacking the chromatin binding domain or the integrase binding domain were used [38, 39] (Fig. 8).



**Figure 8.** *LEDGF/p75 chromatin binding and integrase binding domains.* LEDGF/p75 interacts with chromatin and integrase through its chromatin binding domain (CBD) and integrase binding domain (IBD), respectively. The CBD is composed of the PWWP and two AT hook motifs.

***Specific Aim 3: Evaluate the influence of different LEDGF/p75 regions in the tethering activity of LEDGF/p75.***

The tethering activity of LEDGF/p75 depends on its simultaneous binding to chromatin and integrase. Chromatin binding and integrase binding activities were evaluated for each of the LEDGF/p75 deletion mutants.

***3.1 To determine the effect of different mutations on the chromatin binding activity of LEDGF/p75.*** Chromatin binding was assessed by the chromatin binding assay [39] that is based on the resistance of chromatin-bound proteins to extraction in isotonic buffers containing 1% Triton X-100. Whereas, the strength of chromatin binding was determined by the salt extraction assay which is based on the capacity of increased ionic strengths to break down charged interactions between proteins and DNA.

***3.2 To determine the effect of different mutations on the integrase binding activity of LEDGF/p75.*** Integrase binding activity was evaluated by the capacity of LEDGF/p75 mutants to protect integrase from proteasome-mediated degradation [40]. In addition, integrase interaction was further proved by co-immunoprecipitation techniques. Finally the combined capacity of LEDGF/p75 mutants of chromatin and integrase binding activities (integrase to chromatin tethering activity) was interrogated in LEDGF/p75-deficient cells expressing eGFP-tagged HIV-1 integrase.

**Specific Aim 4: Evaluation of potential LEDGF/p75 protein-protein interactors through mass spectrometry.**

Data obtained from LEDGF/p75 deletion and point mutants suggested that other cellular factors may interact with LEDGF/p75 and promote its HIV-1 co-factor activity. In order to discover possible LEDGF/p75 protein-protein interactors we will perform a mass spectrometry analysis of LEDGF/p75.

**4.1 To conduct a mass spectrometry analysis of LEDGF/p75.** Using cells deficient for LEDGF/p75 and cells reexpressing LEDGF/p75 a chromatin bound protein fraction was obtained and run out on an SDS-PAGE gel which was then silver stained and protein bands excised. These bands were then subjected to mass spectrometry analysis.

**4.2 To confirm potential protein-protein interactions.** Once potential LEDGF/p75 protein-protein interactors are elucidated, their association with LEDGF/p75 will be confirm through the use of immunoprecipitation experiments.

## **Chapter 2: Implication of Serine Residues 271, 273, and 275 in the Human Immunodeficiency Virus Type 1 Cofactor Activity of Lens Epithelium-Derived Growth Factor/p75**

### **2.1 Introduction**

The lens epithelium-derived growth factor/p75 (LEDGF/p75) is a cellular cofactor for HIV-1 DNA integration implicated in the efficiency and genomic location of the viral integration process [4, 35, 37, 41, 42]. Cells lacking LEDGF/p75 showed a severe defect in HIV-1 infection characterized by decreased levels of integrated viral DNA in the absence of other pre-integration defects in the viral life cycle [4, 37]. In addition, in the absence of LEDGF/p75, HIV-1 DNA integration occurred less frequently in active transcription units and more often in close proximity to CpG islands [35, 37, 41]. These data indicated that LEDGF/p75 is an important target for anti-HIV-1 drug development.

LEDGF/p75 is a ubiquitously expressed chromatin-bound protein evolutionarily conserved from bony fishes to humans. The chromatin-binding domain of LEDGF/p75, located in the N-terminal region of the protein, was identified by analysis of deletion mutants [39, 43]. LEDGF/p75 mutants lacking the evolutionarily conserved PWWP and two AT hook motifs failed to bind to chromatin during both interphase and mitosis. However, chromatin binding of mutants lacking only one of these protein regions was normal or minimally defective [39]. These data indicated that LEDGF/p75 chromatin binding requires the functional interaction of the PWWP domain and two AT hook motifs [39]. More recently, mutagenesis analysis of evolutionarily conserved residues within the PWWP suggested that a hydrophobic cavity in this protein domain has a fundamental role in the chromatin binding of LEDGF/p75 [44].

Co-immunoprecipitation of HIV-1 integrase and LEDGF/p75 drove the interest of retrovirologists to this cellular protein [33]. The interaction of LEDGF/p75 with HIV-1 integrase has been extensively characterized in non-infected cells where integrase is expressed as a sole protein by stable or transient plasmid transfection, and in *in vitro* systems using purified recombinant proteins [45-48]. The interaction of LEDGF/p75 with integrase occurs through a C-terminal region called the integrase binding domain (IBD) [38, 49, 50]. The IBD was identified by evaluating the interaction of HIV-1 integrase with LEDGF/p75 deletion mutants using different cellular and biochemical assays. LEDGF/p75 tethers HIV-1 integrase to the host chromatin [51, 52], protects integrase from proteasomal-mediated degradation [40], promotes integrase multimerization [53], and enhances its enzymatic activity in *in vitro* integration assays [54]. The molecular basis of this protein-protein interaction was further defined after x-ray diffraction analysis of an IBD- integrase catalytic core domain complex [55, 56].

Although the molecular mechanism of the HIV-1 cofactor activity of LEDGF/p75 has not been completely understood yet, it is well established that it requires its chromatin and integrase binding activities [4, 37]. HIV-1 infection of LEDGF/p75-deficient cells was rescued upon re-expression of LEDGF/p75 wild type (WT) but not of deletion mutants lacking the chromatin- or the integrase-binding domains [4]. These observations suggested that chromatin-bound LEDGF/p75 could tether lentiviral pre-integration complexes to the host chromatin allowing efficient viral integration [4]. This tethering model was also supported by the role of LEDGF/p75 in HIV-1 DNA integration site distribution [35, 37, 41].

In order to evaluate if LEDGF/p75 regions, not involved in the tethering mechanism, are required for its HIV-1 cofactor activity, we studied a panel of deletion mutants. These mutants lacked different protein regions that are evolutionarily conserved and predicted to be solvent



exposed as well as harboring different protein post-translational modifications. HIV-1 cofactor activity and integrase to chromatin tethering capacity were evaluated for each of these mutants. Using this systematic approach, we have identified that the LEDGF/p75 serine residues 271, 273 and 275 are involved in its HIV-1 cofactor activity without influencing its integrase to chromatin tethering activity. Our results indicate that, in addition to its tethering activity, other molecular events seem to be implicated in the HIV-1 cofactor role of LEDGF/p75.

## 2.2 Materials and Methods

**Plasmids.** *LEDGF/p75 expression plasmids.* pFLAG LEDGF/p75 was used for transient expression experiments. This plasmid has a human cytomegalovirus immediate early gene (CMV) promoter driving the transcription of a LEDGF/p75 cDNA containing seven synonymous mutations in the target site of the twenty-one shRNA 1340 (AACCGGCAACACGAAGAGGCAA, changed nucleotides are underlined) [51]. This shRNA is present in all the LEDGF/p75-deficient cell lines used in this report. The LEDGF/p75 open reading frame was PCR amplified and cloned BamHI / ApaI into pCMV-FLAG expression plasmid, upstream of the FLAG sequence. pCMV-FLAG was derived from pCMV-Myc [51] by substituting Myc by the FLAG tag epitope. Myc was removed ApaI / BglII and a DNA linker containing the FLAG sequence flanked by ApaI / BglII sticky ends was inserted.

Stable expression of LEDGF/p75 WT or mutants was achieved by retroviral transduction. LEDGF/p75 was expressed from the murine leukemia virus (MLV) expression plasmid pJZ308 [4]. LEDGF/p75 mutants were generated by PCR with the Phusion™ Site-Directed Mutagenesis Kit (Finnzymes, Inc) using specific primers (primers sequence available upon request) and following the manufacturer's instructions. All the constructs described above were verified by overlapping DNA sequencing of the complete LEDGF/p75 cDNA.

*HIV-1 integrase expression plasmids.* pHIN-eGFP-IRES-P is a CMV-driven HIV-1 integrase-eGFP expression plasmid that was used to generate stable cell lines expressing this fusion protein. This plasmid was derived from pHIN [51] by substituting the Myc tag with eGFP and introducing an internal ribosome entry site (IRES)-puromycin resistance gene (PAC) cassette downstream to the eGFP open reading frame. The IRES-PAC cassette was PCR amplified from pEFIRESP with primers 5'-

TATAAGATCTAATTCACGCGTGGTACCTCTAG-3' and 5'-  
TATAAGATCTGGTCGTGCGCTCCTTTCGGTCG-3' and then introduced into a unique BglII  
site in pHNeGFP.

*Retroviral vector plasmids.* Plasmids used for production of retroviral vectors, pHIVluc, JZluc, and pTSINcherry/p75 were previously described [4], pCMVΔR8.91 and pMD.G were a gift of D. Trono.

**Cell lines.** *LEDGF/p75-deficient cell lines.* The human CD4<sup>+</sup> T cell line -LEDGF/p75-deficient, T<sub>L3</sub> cells [4], were used for stable expression of LEDGF/p75 mutants. These cells were generated by transduction of SupT1 cells with a HIV-derived vector expressing a shRNA against LEDGF/p75. As control we used T<sub>C3</sub> cells that were also derived by transduction of SupT1 cells with a HIV-derived vector expressing a scrambled shRNA sequence [4]. T<sub>L3</sub> cells express 97% less LEDGF/p75 mRNA than T<sub>C3</sub> cells as determined by real time PCR [4]. To generate T<sub>L3</sub> cells expressing different LEDGF/p75 mutants, cells were transduced with a MLV-derived vector expressing the LEDGF/p75 mutants (pJZ-LEDGF-FLAG) and selected in G418 (600μg /ml) as described previously [4]. Robust polyclonal G418-resistant cell lines were obtained and characterized by immunoblotting with an anti-LEDGF (BD Transduction Laboratories, catalog number 611714) or anti-FLAG monoclonal antibodies (Mabs) (Clone M2, Sigma). The LEDGF/p75-deficient HEK293T-derived cell line si1340/1428 [51] was used for transient expression of LEDGF/p75 proteins.

*HIV-1 integrase cell lines.* LEDGF/p75-deficient HEK293T cells expressing Myc-tagged HIV-1 integrase (LH4 cells, [40]) were used to evaluate the capacity of the LEDGF/p75 mutants to protect HIV-1 integrase from proteosomal-mediated degradation. In addition, 2L<sub>KD</sub>-IN-eGFP

cells were used to evaluate the integrase to chromatin tethering capacity of LEDGF/p75 mutants. These cells are LEDGF/p75-deficient HEK293T cells stably expressing HIV-1 integrase C-terminally tagged with eGFP. To generate 2L<sub>KD</sub> IN-eGFP cells, HEK293T cells were plated at  $3 \times 10^6$  in a 75-cm<sup>2</sup> flasks and calcium-phosphate transfected the next day with 20µg of the expression plasmid pCMV-IN-eGFP-IRES-P linearized at the prokaryotic backbone. Stably transfected cells were obtained after selection in the presence of puromycin (3µg/ml) and integrase-eGFP expression was verified by immunoblotting with an anti-eGFP Mab and by fluorescence-activated cell sorting (FACS) analysis. Subsequently, LEDGF/p75 deficiency was achieved in these cells by transduction with an HIV-based lentiviral vector expressing a LEDGF/p75-specific shRNA 1340 [4]. This lentiviral vector integrates into the host genome a cassette containing *in cis* a U6 small nuclear RNA promoter that drives the expression of a LEDGF/p75-specific shRNA and a CMV promoter directing the expression of a mCherry fluorescent protein. This expression system allows selection of LEDGF/p75 knockdown cells based on their mCherry fluorescence levels [4]. Knockdown levels of LEDGF/p75 were verified further by immunoblotting with an anti-LEDGF Mab.

SupT1-derived cell lines were grown in RPMI1640 while HEK293T-derived cells were grown in DMEM and both culture mediums were supplemented with 10% of heat-inactivated fetal calf serum, 2mM L-glutamine and 1% penicillin/streptomycin.

**Generation of retroviral vectors.** Procedures previously described [4] were followed for the production of the different retroviruses used here. Briefly, MLV-derived vectors were produced in Phoenix A packaging cells by calcium-phosphate co-transfection of 15µg of pJZLEDGF/p75 wild type or pJZLEDGF/p75 mutants or pJZluc and 5µg of the Vesicular Stomatitis Virus

glycoprotein G (VSV-G) expression plasmid, pMD.G. Forty-eight hrs after transfection, the viral supernatants were harvested and concentrated by ultracentrifugation at 124,750g for two hrs on a 20% sucrose cushion. HIV-derived vectors expressing anti-LEDGF shRNA were produced by calcium-phosphate co-transfection of HEK293T with 15µg of pTSINcherry/p75 [4], 15µg of pCMVΔR8.91 and 5µg of pMD.G. The viral-containing supernatant was collected and concentrated by ultracentrifugation as described above. Single-round infection, luciferase-expressing HIV (HIVluc) was prepared by calcium-phosphate co-transfection of HEK293T cells with 15µg of pHIVluc [4] and 5µg of pMD.G. Viral supernatant was collected forty-eight after transfection and aliquots stored at -80°C until use.

**Single-round viral infectivity assay.** T<sub>L3</sub>, T<sub>C3</sub>, and LEDGF/p75-expressing T<sub>L3</sub> cells were plated at 1x10<sup>5</sup> cells in 500µl of RPMI1640 culture medium in 24-well plates and infected with HIVluc or MLVluc viral supernatants. Five days post-infection, cells were collected by centrifugation at 1000g for six mins and the pellet lysed in 100µl of PBS-1% Tween 20 for 15 mins on ice. Cellular lysates were centrifuged at 22,000g for 2 mins and supernatant used for quantification of luciferase activity. An aliquot of 20µl of the cellular lysate supernatant was mixed with 45µl of substrate (*Bright-Glow*<sup>TM</sup> Luciferase Assay System, Promega) and luciferase activity was quantified using a microplate luminometer.

**Immunoblotting.** Cellular lysates were resolved by SDS-PAGE and transferred overnight to PVDF membranes at 100 mAmp at 4°C. Membranes were blocked in TBS containing 10% milk for one hour and then incubated with the corresponding primary antibody diluted in TBS-5% milk-0.05% Tween 20 (antibody dilution buffer). FLAG-tagged LEDGF/p75 was detected with

anti-FLAG Mab (1/500, M2, Sigma), non-tagged LEDGF/p75 was detected with anti-LEDGF Mab (1/250), Myc-tagged HIV-1 integrase was detected with anti-Myc Mab (1/500, clone 9E10, Covance). As loading control, anti-alpha tubulin Mab (clone B-5-1-2, Sigma) was used at a 1/4000 dilution. Membranes were incubated overnight at 4°C with anti-FLAG, -LEDGF, and -Myc Mabs whereas anti-alpha tubulin Mab were incubated for two hrs at 25°C. Primary antibody-bound membranes were washed in TBS-0.1% Tween 20 and bound antibodies detected with goat anti-mouse Igs-HRP (Sigma) diluted 1/2000 in antibody dilution buffer followed by chemoluminescence detection.

**Chromatin-Binding Assay.** Previously described procedures [39] were followed with minor modifications. Figure 3 of reference [39] shows a validation experiment where multiple proteins located in different cellular compartments were evaluated. Briefly,  $18 \times 10^6$  T<sub>L3</sub>-derived cells expressing different LEDGF/p75 mutants were washed in PBS and distributed in three aliquots containing equal amount of cells. Two of the samples were lysed for 15 mins on ice in 100µl of CSK I buffer (10mM Pipes pH6.8, 100mM NaCl, 1mM EDTA, 300mM sucrose, 1mM MgCl<sub>2</sub>, 1mM DTT, 0.5% Triton X-100) containing protease inhibitors (final concentration: leupeptine 2µg/ml, aprotinin 5µg/µl, PMSF 1mM, pepstatin A 1µg/ml). The third sample was lysed for 15 mins on ice in 100µl CSK I buffer supplemented with 350mM NaCl and protease inhibitors, centrifuged at 22,000g for 3 mins at 4°C and supernatant saved for further analysis (total fraction, T). Cells lysed in CSK I buffer were centrifuged at 1000g for 6 mins at 4°C and the supernatant pooled (non-chromatin-bound fraction, S1). S1 supernatants were clarified further by centrifugation at 22,000g for 3 mins and supernatant transfer to a fresh tube while pellets were washed once in 200µl of CSK I buffer. One of these pellets was resuspended in CSK I 350mM

NaCl buffer and incubated on ice for 15 mins followed by centrifugation at 22,000g for 3 mins at 4°C and the supernatant was collected for further analysis (chromatin-bound fraction, P1). The other pellet, obtained after cell lysis in CSK I, was resuspended in 100µl of CSK II buffer supplemented with protease inhibitors, 4 units of turbo DNase (Ambion) and 11µl of 10X turbo DNase reaction buffer. DNase treatment of this pellet was conducted at 37°C for 30 mins and then followed by extraction with (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 250mM for 15 mins at 37°C. The DNase/(NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> treated sample was centrifuged at 22,000g for 3 mins and the supernatant saved for analysis (chromatin-bound fraction, S2). The resulting pellet was further extracted with CSK I 350mM NaCl for 15 mins on ice, centrifuged at 22,000g for 3 mins and the obtained supernatant (non-chromatin-bound fraction, P2) collected for analysis. A volume of 15.7µl of S1, P1, P2 and T and 20µl of S2 (amounts equivalent to 0.9x10<sup>6</sup> cells) was evaluated by immunoblotting using an anti-FLAG Mab.

**Salt extraction assay.** 36x10<sup>6</sup> T<sub>L3</sub> cells or T<sub>L3</sub> cells expressing LEDGF/p75 mutants were washed with PBS and distributed in six samples, each containing an equal amount of cells. One of the samples was resuspended in 100µl of Laemmli sample buffer, boiled for 9 mins and centrifuged at 22,000g for 3 mins and the resulting supernatant was further analyzed (total fraction, T). The other five cellular aliquots were lysed on ice for 15 mins in 100µl of CSK I containing protease inhibitors and supplemented with increasing concentrations of NaCl (final concentration: 100, 150, 200, 350 and 500mM). Lysed cells were centrifuged at 22,000g for 3 mins at 4°C and the supernatant (salt-extracted fraction) collected for analysis. The obtained pellets (salt-resistant fraction) were subsequently boiled in 100µl of Laemmli sample buffer, centrifuged at 22,000g for 3 mins at 4°C and the resulting supernatant collected. The presence of

LEDGF/p75 in the total, salt-extracted or salt-resistant fractions was analyzed by immunoblotting with an anti-FLAG Mab using 20µl of cell lysate, an equivalent amount of  $6 \times 10^4$  cells.

**Immunoprecipitation.** The interaction of the LEDGF/p75  $\Delta$ CR3 mutant with Myc-tagged HIV-1 integrase was evaluated by immunoprecipitation in the LEDGF/p75-deficient HEK293T-derived cell line si1340/1428 [40]. Cells were plated at  $0.45 \times 10^6$  cells/well in a 6-well plate and co-transfected the next day by calcium-phosphate with 2µg of pFLAG-LEDGF/p75  $\Delta$ CR3, or pFLAG-LEDGF/p75 WT or pFLAG-LEDGF/p75  $\Delta$ IBD and 2µg of the HIV-1 integrase expression plasmid pHIN. Forty-eight hrs after transfection, cells were lysed in 300µl of RIPA buffer (150mM Tris-HCl, pH 8.0, 150mM NaCl, 0.5% DOC, 0.1% SDS, 1% NP-40) supplemented with protease inhibitors. Cell lysates were clarified by centrifugation at 22,000g for 3 mins and supernatant used for immunoprecipitation using goat anti mouse Igs-coated magnetic beads (Pierce). Beads (100µl) were previously incubated for 20 mins on ice with 3µg of anti-FLAG Mab diluted in RIPA buffer. Then, beads were separated from the unbound antibodies, mixed with the cell lysate and rotated for 3hrs at 4°C. After this incubation, beads were washed three times in RIPA buffer and bound proteins eluted by boiling in 30µl of Laemmli sample buffer. Immunoprecipitated proteins were analyzed for the presence of Myc-tagged HIV-1 integrase and LEDGF/p75-FLAG by immunoblotting with anti-Myc and anti-FLAG Mabs, respectively.

**Integrase protection assay.** LEDGF/p75-deficient HEK293T cells expressing Myc-tagged HIV-1 integrase were used to evaluate the effect of different LEDGF/p75 deletion mutants on



integrase protein stability. This assay is based on the integrase binding-dependent capability of LEDGF/p75 to protect integrase from proteasome-mediated degradation [40]. Integrase protein levels in these LEDGF/p75-deficient cells are very low, as detected by sensitive immunoblottings, and re-expression of LEDGF/p75 WT rescues integrase levels ([40] and figures 1b and 6a-I in this report). Cells were plated at  $0.45 \times 10^6$  per well in a 6-well plate and transfected the next day by calcium-phosphate with 2 $\mu$ g of pFLAG-LEDGF/p75 WT or mutants. Forty-eight hrs after transfection, cells were lysed in 300 $\mu$ l of RIPA buffer supplemented with protease inhibitors. Cell lysates were clarified by centrifugation at 22,000g for 3 mins and supernatant used for immunoblotting with anti-Myc or anti-FLAG Mabs.

**Integrase to chromatin tethering assay.** This assay is based on the role of LEDGF/p75 in the chromatin tethering of HIV-1 integrase [51]. Chromatin-bound LEDGF/p75 interacts with integrase through the IBD, tethering the viral enzyme to chromatin during all the phases of the cell cycle. In LEDGF/p75-deficient cells, HIV-1 integrase lacks its exclusively nuclear localization during interphase and is not associated to chromatin during mitosis ([51] and figure 7 in this report). 2L<sub>KD</sub>-IN-eGFP cells, lacking LEDGF/p75 and stably expressing HIV-1 integrase-eGFP, were used in this assay. Cells were plated at  $2 \times 10^5$  in 2 mls of culture medium in a LabTek II chambered coverglasses and transfected the next day with 2 $\mu$ g of pFLAG-LEDGF/p75 wild type or mutants. Eighteen hrs after transfection, fresh culture medium was added and forty-eight hrs later cells were washed three times in PBS and fixed with 4% formaldehyde-PBS for 10 mins at 37°C. Then, cells were washed twice in PBS and stained with DAPI. The subcellular distribution of HIV-1 IN-eGFP-LEDGF/p75 complex was analyzed by fluorescence microscopy.

**HIV-1 DNA integration analysis.** Viral integration was quantified by detection of Alu-LTR and total HIV-1 cDNA by real-time PCR in T<sub>L3</sub> and T<sub>L3</sub> cells engineered to express LEDGF/p75 WT or S271A/S273A/S275A mutant. Two completely independent experiments using two different viral preparations were considered. In these experiments, 10<sup>5</sup> cells were infected with DNase-treated HΔEluc viral supernatant and cultured for ten days. Then DNA was extracted (High pure PCR template preparation kit, Roche) from 10<sup>6</sup> infected cells and 20ng of DNA were used for the detection of total HIV-1 cDNA, mitochondrial DNA and 2LTR circles while 0.2ng of DNA were used for the Alu-LTR junctions PCR. Alu-LTR and total HIV-1 cDNA products were normalized for mitochondrial DNA and 2LTR circles, respectively. All these real time PCRs were performed in a MiniOpticon system (Bio-rad) with primers and conditions previously described [4]. Fold change was calculated using the ΔCt method as recommended in the thermocycler manual and differences were expressed in percentage, considering the value for T<sub>L3</sub> LEDGF/p75 WT cells as 100%.

***In silico* analysis.** Mutations introduced in LEDGF/p75 were guided by a systematic bioinformatics analysis focused on evolutionarily conservation across different species, the presence of post-translational modifications as well as the prediction of solvent accessibility. Evolutionary conservation, as referred in this text, includes the presence of either identical or homologous residues in the compared protein sequences. LEDGF/p75 protein sequences were retrieved from the NCBI protein database and aligned using ClustalW2 [57], sequence comparisons were performed with BLASTP 2.2.19+ [58]. LEDGF/p75 protein motif search was performed with the ELM server [59] and the prediction of relative solvent accessibility with PaleAle [60].

## 2.3 Results

**Experimental strategy.** Two regions of LEDGF/p75, the chromatin-binding domain and the IBD, are required for its HIV-1 cofactor activity [4, 37]. However, the involvement of other LEDGF/p75 regions in HIV-1 infection has not been previously studied. In order to evaluate if LEDGF/p75 protein regions not implicated in chromatin binding or integrase interaction are also necessary for HIV-1 infection, a panel of deletion mutants was analyzed. These mutants were stably expressed in a LEDGF/p75-deficient human CD4<sup>+</sup> T cell line (T<sub>L3</sub> cells) [4] and their ability to rescue HIV-1 infection was evaluated. In addition, the chromatin- and integrase-binding activities of these mutants were assessed.

Although analysis of LEDGF/p75 deletion mutants have been used before to successfully map functional domains in LEDGF/p75 [38, 39, 43, 49, 61], this approach could lead to misinterpretations due to the global impact of mutations on the structure, hence the function, of the analyzed protein. In order to prevent data over-interpretation, only deletion mutants causing a decrease in the HIV-1 cofactor activity of LEDGF/p75 similar to that observed following deletion of the chromatin-binding or the integrase-binding domains were considered for further analysis. Finally, the phenotypes observed in deletion mutants were verified in LEDGF/p75 point mutants.

In order to facilitate the analysis of LEDGF/p75 mutants, a C-terminally FLAG tagged LEDGF/p75 was used. To exclude any artifactual observation due to the addition of this small tag to LEDGF/p75, tagged (WT-FLAG) and not tagged (WT) versions of this protein were evaluated for their interaction with chromatin and HIV-1 integrase and for their HIV-1 cofactor activity (Fig. 9).

The effect of this tag on the chromatin binding strength of LEDGF/p75 was evaluated using the chromatin salt extraction assay. This is based on the effect of increased salt concentrations on the binding of proteins to chromatin. T<sub>L3</sub> cells expressing LEDGF/p75 WT or WT-FLAG were lysed in 1% Triton X-100 CSK I buffer containing increasing concentrations of NaCl (100, 150, 200, 350 and 500mM). Cellular lysates were centrifuged and supernatant (salt-extracted fraction) and pellet (salt-resistant fraction) were evaluated by immunoblotting for the presence of LEDGF/p75. Total LEDGF/p75 extraction was achieved by lysing unfractionated cells in Laemmli sample buffer (Fig. 9a). Using this method, we found that both LEDGF/p75 forms persisted bound to chromatin (salt-resistant fraction) at 100 and 150mM NaCl and were fully extracted above 200mM NaCl (Fig. 9a), indicating that the FLAG tag did not affect its binding to chromatin.

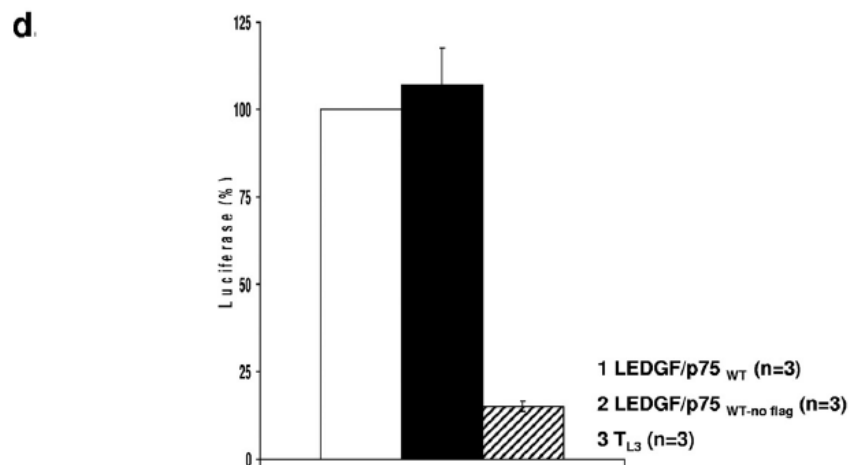
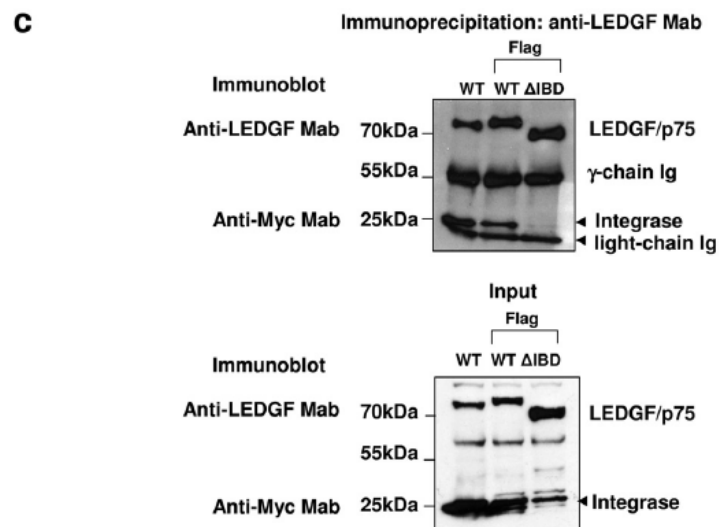
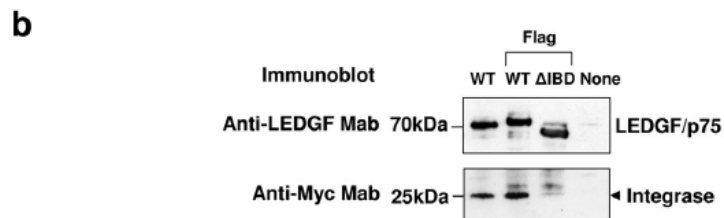
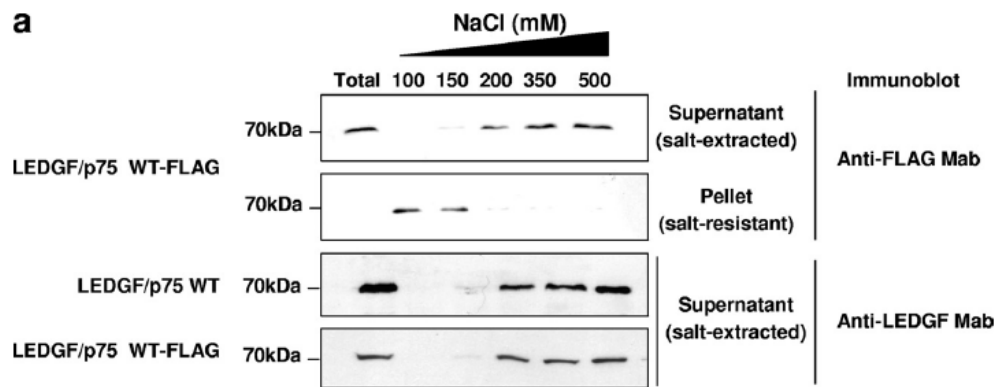
To evaluate the effect of the FLAG tag on the interaction of LEDGF/p75 with HIV-1 integrase we compared the ability of LEDGF/p75 WT and WT-FLAG to protect integrase from proteasome-mediated degradation (integrase protection assay) (Fig. 9b). LEDGF/p75-deficient HEK293T cells, stably expressing Myc-tagged HIV-1 integrase, LH4 cells, were transiently transfected with plasmids expressing LEDGF/p75 WT, WT-FLAG or  $\Delta$ IBD and integrase protein levels were evaluated by immunoblot forty-eight hrs after transfection. Steady-state HIV-1 integrase levels were very low in the absence of LEDGF/p75 and dramatically increased upon expression of LEDGF/p75 WT or WT-FLAG but not of the  $\Delta$ IBD mutant (Fig. 9b), indicating that the FLAG tag did not affect the interaction of LEDGF/p75 with HIV-1 integrase.

The influence of the FLAG tag on the direct interaction of LEDGF/p75 with HIV-1 integrase was also evaluated by co-immunoprecipitation. LEDGF/p75-deficient HEK293T cells co-expressing LEDGF/p75 WT, WT-FLAG or  $\Delta$ IBD-FLAG and Myc-tagged HIV-1 integrase

were lysed and immunoprecipitated with an anti-LEDGF/p75 Mab. Then, the presence of integrase and LEDGF/p75 was evaluated in the immunoprecipitated proteins by immunoblotting with anti-Myc or anti-LEDGF/p75 Mabs, respectively. In support of data in figure 1b, both LEDGF/p75 WT and WT-FLAG very efficiently immunoprecipitated HIV-1 integrase, while LEDGF/p75  $\Delta$ IBD did not (Fig. 9c). These data also demonstrated that the FLAG tag on LEDGF/p75 did not affect its interaction with HIV-1 integrase.

Finally, we compared the HIV-1 co-factor activity of LEDGF/p75 WT and WT-FLAG. T<sub>L3</sub> cells stably expressing or not each of these LEDGF/p75 proteins were infected with HIV luc and five days after infection luciferase levels were detected (Fig. 9d). In correlation with observations described in figure 9a to 9c, the C-terminal FLAG did not alter the LEDGF/p75 HIV cofactor activity.

In summary, data in figure 9 clearly demonstrated that LEDGF/p75 WT-FLAG is functionally equivalent to its non-modified variant and, unless it is specified, all the data described below were obtained with LEDGF/p75 FLAG tagged at the C-terminus.



**Figure 9.** (a) Salt extraction assay. TL3 cells expressing the LEDGF/p75 WT or WT-FLAG were lysed in a buffer containing increasing NaCl concentrations. Cells were separated into a soluble fraction and an insoluble fraction by centrifugation, and the presence of LEDGF/p75 was evaluated by immunoblotting with an anti-LEDGF or anti-FLAG MAb, as indicated. An unfractionated total cellular fraction (T) was included as a control. (b) Integrase protection assay. LEDGF/p75-deficient HEK293T cells stably expressing Myc-tagged HIV-1 integrase were either transiently transfected or not transfected with plasmids expressing the LEDGF/p75 WT, WT-FLAG, or  $\Delta$ IBD-FLAG. Expression levels of LEDGF/p75 proteins and HIV-1 integrase were determined by immunoblotting with anti-LEDGF and anti-Myc MAbs, respectively. (c) Coimmunoprecipitation of HIV-1 integrase with LEDGF/p75. LEDGF/p75-deficient HEK293T cells were cotransfected with plasmids expressing the LEDGF/p75 WT, WT-FLAG, or  $\Delta$ IBD-FLAG and Myc-tagged HIV-1 integrase and subjected to immunoprecipitation with an anti-LEDGF MAb. Immunoprecipitated proteins were detected by immunoblotting with anti-Myc or anti-LEDGF MAbs. Mouse antibody heavy and light chains were used as a loading control for the immunoprecipitation. LEDGF/p75 and HIV integrase proteins were detected in the samples used for immunoprecipitation (input) by immunoblotting with anti-LEDGF and anti-Myc MAbs, respectively. (d) Single-round infection of TL3 cells expressing the LEDGF/p75 WT or WT-FLAG. Cells were challenged with HIVluc, and luciferase activity was analyzed 5 days later. Luciferase levels detected in TL3 cells expressing the LEDGF/p75 WT were considered to be 100%. Error bars indicate standard deviation values calculated for a number ( $n$ ) of independent infection experiments performed on different days with different viral preparations.

**LEDGF/p75 deletion mutants.** Regions deleted from LEDGF/p75 are schematized in figure 10a. These regions were designated conserved/charged regions (CR) 1, 2, 3, 4 and 5. Region boundaries were established based on evolutionary conservation of LEDGF/p75 across different species, the presence of predicted or demonstrated post-translational modifications and the existence of clusters of polar residues that directly correlate with protein solvent accessibility.

The CR1 region contains 52 amino acids and was selected based on its evolutionary conservation in mammals and on its proposed role in the LEDGF/p75 PWWP domain chromatin binding activity [39]. This region is enriched for charged amino acids (54.8%) and 76.4% of them are identical or homologous residues (conserved) present from avian to humans (Data not shown). Interestingly, CR1 contains a cluster of eight Ser/Thr phosphosites between residues 105-135 [62-67], this cluster represents 36.3% of the total identified phosphorylated Ser/Thr residues in LEDGF/p75. In addition, Thr141 has been found phosphorylated both in human and mouse cells [62, 63, 67]. This phosphosite is contained in a predicted forkhead-associated domain (FHA) interaction motif, LEDGF/p75 residues 141-144. The FHA domain is a protein module present in prokaryotic and eukaryotic proteins involved in signaling, cell cycle control and DNA repair that binds to specific threonine phosphoporylated sequences [59].

CR2 and CR3 boundaries were previously described [39]. These regions were demonstrated to cooperate with the PWWP domain and AT hook motifs during LEDGF/p75 chromatin binding [39]. The CR2 region is 68 amino acids long, 58.8% are charged and 67% are conserved from amphibians to humans (Data not shown). Evolutionary conservation and solvent exposure is particularly interesting within residues 212-249 where 68% are present in bony fishes and 70.3% of them are charged and predicted to be exposed to the solvent [60]. In addition, although CR2 represents only 12.8% of LEDGF/p75, it contains 29.4% of its lysine



residues. Lysine is targeted by different post-translational modifications including methylation, SUMOylation, ubiquitination and glycosylation.

The CR3 region is 59 amino acids long and 64% of these residues are charged (Data not shown). The N-terminal region of CR3, residues 267-275, contains six demonstrated phosphorylated Ser/Thr sites which represent 27.2% of the LEDGF/p75 Ser/Thr identified phosphosites [64-71]. These Ser/Thr residues are predicted to be targeted by protein kinase casein kinase 2 [59]. In addition, residues 271-275, contain two phosphorylated serines that are conserved from bony fishes to humans. An evolutionarily conserved cluster of acidic residues follows these phosphosites.

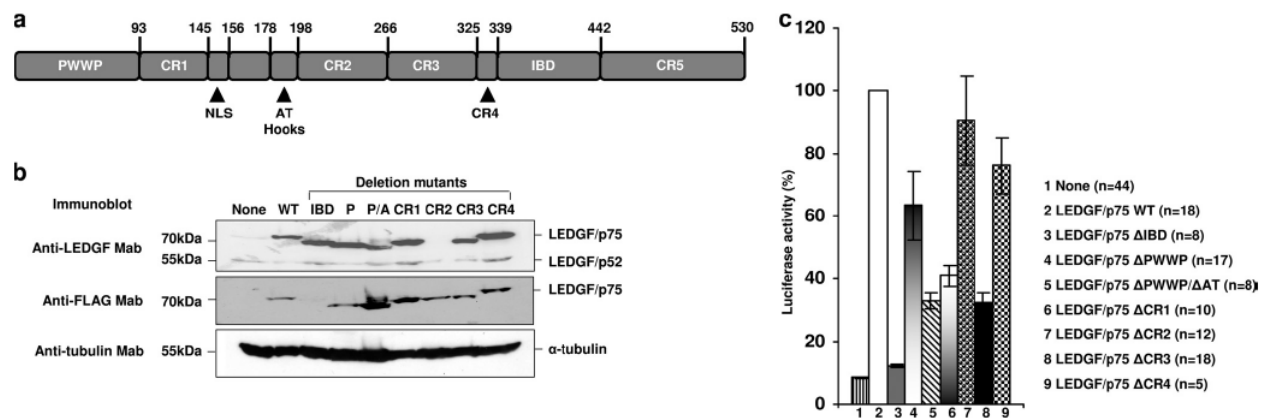
The C-terminal part of LEDGF/p75 contains the IBD. N-terminally to this domain, a region 14 amino acids long, that is rich in conserved and charged residues was deleted for analysis (CR4 region). In addition, a region 88 amino acids long, designated CR5, that is located C-terminally to the IBD was also independently deleted. Although 70% of the residues in CR5 are evolutionarily conserved, this region is not rich in charged amino acids. However, CR5 contains four demonstrated Ser/Thr phosphosites [67] and three of them are clustered in its C-terminal end.

**Role of different protein regions of LEDGF/p75 in HIV-1 infection.** T<sub>L3</sub> cells were used to evaluate the role of different LEDGF/p75 protein regions in HIV-1 infection. Deletion mutants were stably expressed by retroviral transduction of T<sub>L3</sub> cells. Similar levels of re-expressed LEDGF/p75 proteins of the expected relative molecular weight were verified by immunoblotting with anti-LEDGF/p75 and anti-FLAG Mabs (Fig. 10b). Non FLAG-tagged LEDGF/p75  $\Delta$ IBD was detected only with an anti-LEDGF/p75 Mab, while FLAG-tagged LEDGF/p75  $\Delta$ CR2,

lacking the epitope detected by the anti-LEDGF/p75 Mab, was only detected by the anti-FLAG antibody. The LEDGF/p75  $\Delta$ CR5 mutant failed to be expressed in T<sub>L3</sub> cells, although robust polyclonal G418-resistant cell lines were obtained after two independent attempts. In addition, treatment of these cell lines with the proteasome inhibitor MG132 did not rescue the expression of LEDGF/p75  $\Delta$ CR5 (Data not shown), therefore we excluded this region from further analysis.

T<sub>L3</sub>-derived cell lines expressing LEDGF/p75 deletion mutants were challenged with HIVluc [4] and five days later luciferase activity was measured (Fig. 10c). Standard deviations and the number of independent experiments performed are indicated. Luciferase levels detected in T<sub>L3</sub> cells expressing LEDGF/p75 WT were considered as 100% infectivity.

Re-expression of FLAG-tagged LEDGF/p75 WT rescued infectivity in these cells as previously reported for a non-tagged LEDGF/p75 [4]. In addition, LEDGF/p75 lacking the IBD, the PWWP domain (P), or the PWWP domain and the AT hook motifs (P/AT) exhibited 12.3%, 63.4%, and 32.9% of the HIV-1 cofactor activity of LEDGF/p75 WT, respectively. These data were also in agreement with previous observations [4, 37]. Interestingly, the deletion of CR2 or CR4 only slightly impaired the cofactor activity of LEDGF/p75 showing 90.5% and 76% of the wild type activity, respectively. On the contrary, deletion of CR1 reduced the cofactor activity to 41% while deletion of CR3 caused a larger defect exhibiting only 32.4% of the wild type activity. Importantly, the reduction in cofactor activity observed after deleting CR3 was comparable to the simultaneous deletion of PWWP and AT hooks (Fig. 10c).



**Figure 10.** Evaluation of the HIV-1 cofactor activities of LEDGF/p75 deletion mutants. (a) Schematic representation of deleted regions. (b) Different LEDGF/p75 deletion mutants were stably expressed in LEDGF/p75 knockdown TL3 cells, and their protein levels were evaluated by immunoblotting using anti-LEDGF or anti-FLAG MAb. Alpha-tubulin detection was used as a loading control. P and P/A represents PWWP and PWWP/AT hook, respectively. (c) Single-round HIV-1 infections of TL3-derived cells. Cells immunoblotted in b were challenged with HIVluc, and luciferase activity was determined 5 days later.

In order to evaluate further the role of the CR3 region of LEDGF/p75 in HIV-1 infection, we generated three smaller deletion mutants within this region (Fig. 11a). The boundaries of the new deletion mutants were established considering the sequence aspects described previously. A group of positively charged amino acids conserved from avian to humans was considered to limit CR3.1 whereas CR3.2 was defined by the presence of two clusters of charged and conserved residues. CR3.3 is a deletion within CR3.2 containing the C-terminal cluster of conserved and charged residues present in CR3.2. (Fig. 11a). LEDGF/p75 proteins lacking CR3.1, CR3.2 and CR3.3 regions were expressed in T<sub>L3</sub> cells (Fig. 11b) and their HIV-1 cofactor activity evaluated (Fig. 11d). Although, deletion of CR3.2 and CR3.3 showed 86.9% and 64.7% of the LEDGF/p75 WT cofactor activity, respectively, CR3.1 deletion decreased the activity of LEDGF/p75 to 27.3% of that observed in the wild type protein. These data suggest that LEDGF/p75 residues 267-281 have a role in HIV-1 infection.

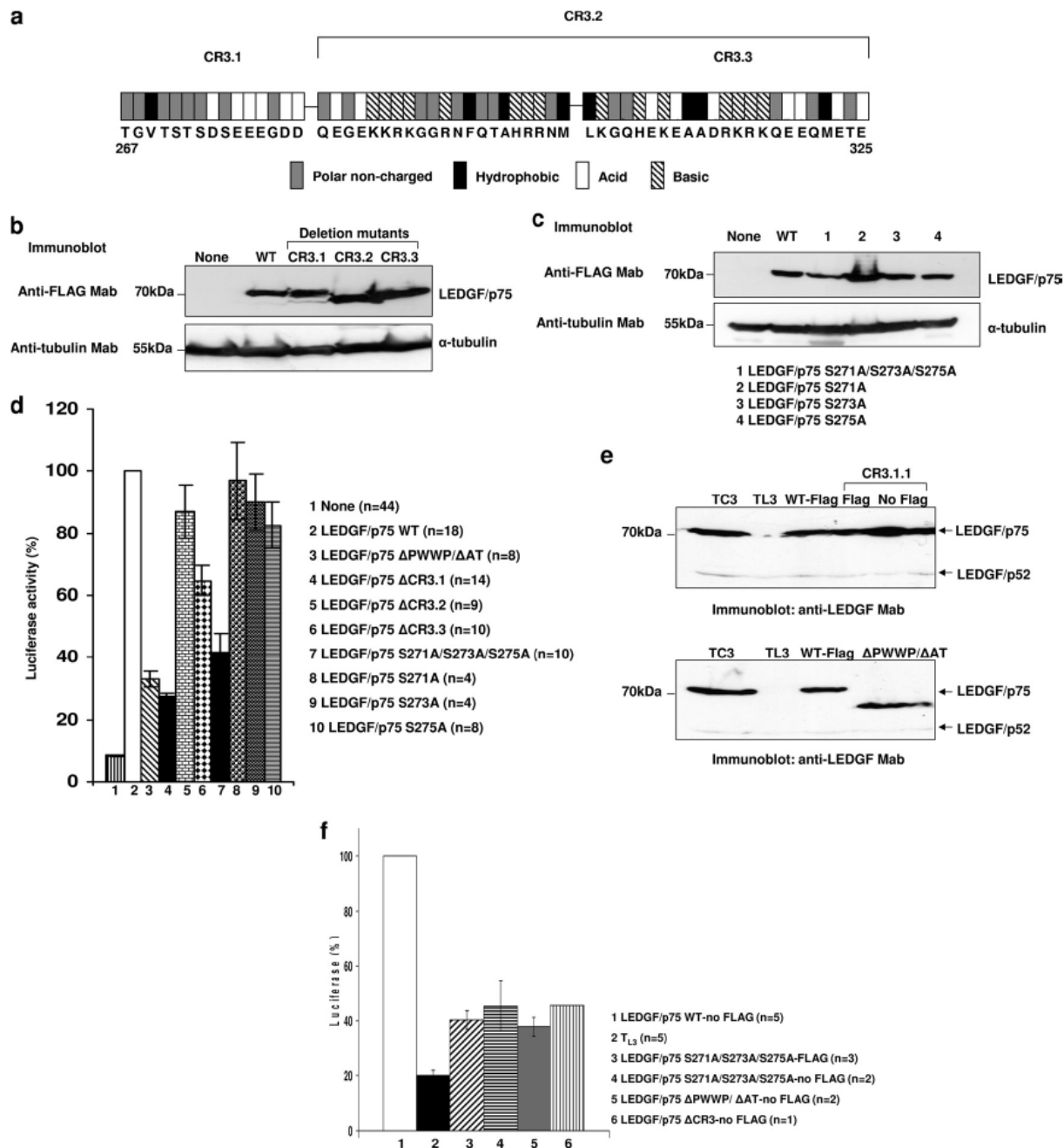
**Role of Ser 271, 273 and 275 in HIV-1 infection.** Serine residues 271, 273, 275 in LEDGF/p75 have been reported to be phosphosites in global phosphoproteomics analysis of human and mouse cells [64-71]. In addition, these residues are predicted to be targeted by protein kinase casein kinase 2 (PKCK2) [59], an enzyme that phosphorylates serine/threonine residues that are followed by an acidic residue in position +3 from the C-terminus of the phosphate acceptor [72]. In order to evaluate the role of serine residues 271, 273, 275 in HIV-1 infection, LEDGF/p75 mutants S271A/S273A/S275A, S271A, S273A and S275A were stably expressed in T<sub>L3</sub> cells (Fig. 11c) and HIV-1 infectivity evaluated (Fig. 11d). Immunoblot analysis of transduced T<sub>L3</sub> cells demonstrated similar levels of the re-expressed LEDGF/p75 mutants as compared to cells expressing the wild type protein (Fig. 11c). However, the HIV-1 cofactor activity of LEDGF/p75

S271A/S273A/S275A was only 41.5% of the activity of the wild type protein, indicating the relevance of these phosphosites. These results were confirmed by evaluating two independently derived T<sub>L3</sub> cell lines expressing LEDGF/p75 S271A/S273A/S275A in ten independent infection experiments with HIVluc. Although HIV-1 infection was reduced in T<sub>L3</sub> LEDGF/p75 S271A/S273A/S275A cells, its susceptibility to infection by a murine leukemia virus (MLV)-derived vector expressing luciferase was not altered as compared to parental cells (Data not shown). Since LEDGF/p75 is not required for MLV infection [4, 37], these data further confirmed the specificity of the observed effect, indicating that the reduced susceptibility to HIV-1 infection was not due to a global defect present in these two independently derived cell lines expressing LEDGF/p75 S271A/S273A/S275A.

Acidic residues in positions -1, +1, +2, +4 and +5, relative to the phosphate acceptor site, are present in most of the physiological substrates of PKCK2 [72]. Additional acidic residues are found at position +5 of Ser 271, +1, +4 and +5 of Ser 273 and positions -1, +1, +2, and +5 of Ser 275. This analysis predicts Ser 275 as a better substrate of PKCK2 than Ser 271 or Ser 273. To establish the implication of each of these residues in HIV-1 infection, T<sub>L3</sub> cells expressing LEDGF/p75 single point mutants S271A, S273A, or S275A were developed (Fig. 11c) and challenged with HIVluc. As shown in figure 11d, individual Ser to Ala mutations of these residues did not have any effect on the viral cofactor activity of LEDGF/p75, exhibiting from 82.6% to 96.9% of the wild type activity.

In all the HIV infection experiments described above (Figs. 10c and 11d), the LEDGF/p75  $\Delta$ CR3 and S271A/S273A/S275A mutants expressed in T<sub>L3</sub> cells were FLAG tagged at their C-terminus. Although we demonstrated that FLAG did not affect the HIV cofactor activity of LEDGF/p75 WT (Fig. 9d); we also evaluated the role of non-tagged versions of

LEDGF/p75 S271A/S273A/S275A,  $\Delta$ CR3 and  $\Delta$ PWWP/ $\Delta$ AT mutants in HIV infection. T<sub>L3</sub> cells expressing non-tagged LEDGF/p75 WT or mutants (Fig. 11e) were challenged with HIVluc and five days later luciferase was determined. As expected, a similar impairment in the HIV cofactor activity of LEDGF/p75 was observed in these mutants regardless of the presence or absence of the FLAG tag (Fig. 11f).



**Figure 11.** *HIV-1* cofactor activity of different LEDGF/p75 CR3 mutants. (a) Analysis of charged residues in CR3 used to establish the boundaries of CR3.1, CR3.2, and CR3.3. (b and c) Expression of LEDGF/p75 mutants in TL3 cells evaluated by immunoblotting with an anti-FLAG MAb. Alpha-tubulin detection was used as a loading control. (d) Single-round infection

of TL3 cells expressing LEDGF/p75 mutants. Cells evaluated in b and c were challenged with HIV-1 luciferase reporter viruses, and luciferase activity was analyzed 5 days later. Infectivity and error bars were calculated as described in the legend of Fig. 1d. (e) Immunoblotting detection of the reexpressed LEDGF/p75 WT and mutants in TL3 cells using an anti-LEDGF MAb. Detection of endogenous LEDGF/p52 was used as a loading control. (f) Single-round infection of TL3 cells expressing the nontagged LEDGF/p75 WT or mutants. Cells immunoblotted in e were challenged with HIVluc, and luciferase activity was analyzed 5 days later. Infectivity and error bars were calculated as described in the legend of Fig. 9d.



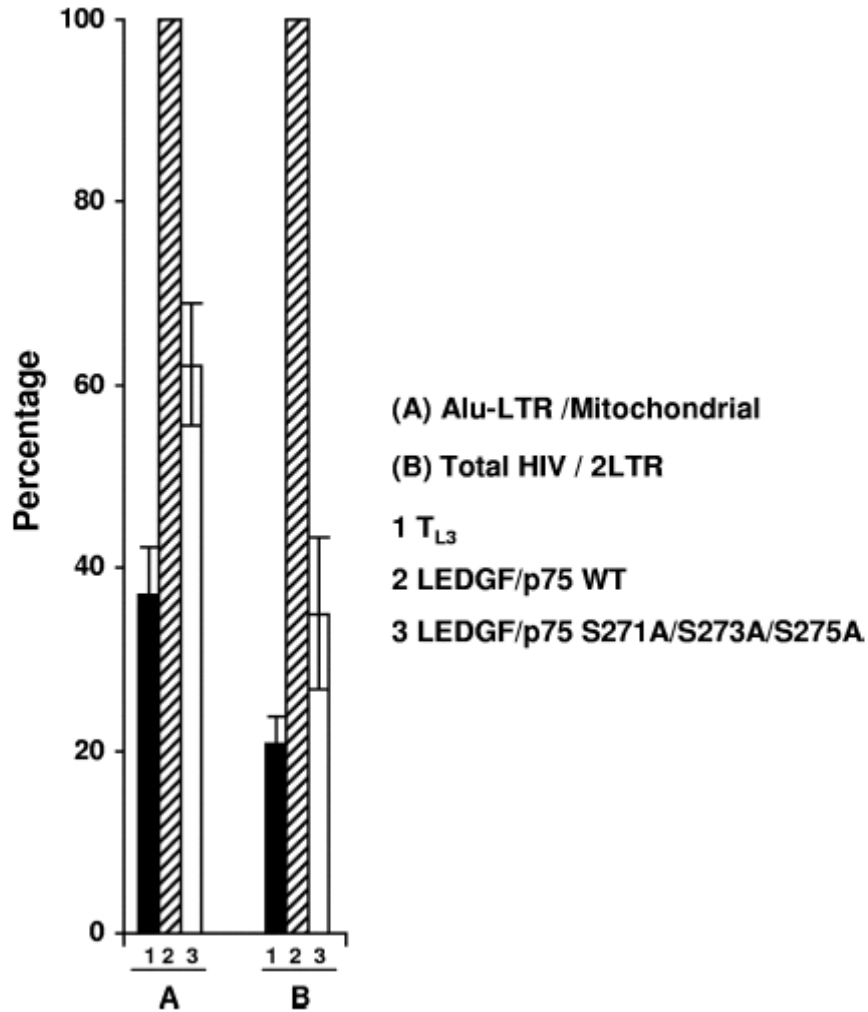
**HIV-1 integration in T<sub>L3</sub> cells expressing LEDGF/p75 S271A/S273A/S275A.** A role of LEDGF/p75 in HIV-1 DNA integration has been clearly established [4]. In order to evaluate if the LEDGF/p75 motif S271/S273/S275 was involved in viral integration, T<sub>L3</sub> cells expressing or not LEDGF/p75 WT or the S271A/S273A/S275A mutant were infected with HIVluc and ten days later analyzed. Two completely independent experiments using two different viral preparations were analyzed. DNA was isolated from infected cells and Alu-LTR junctions, mitochondrial DNA, HIV gag DNA (total HIV DNA) and HIV 2LTR circles were quantified by real time PCR. Levels of Alu-LTR junctions were normalized to mitochondrial DNA to ensure that equal cell numbers were analyzed.

As previously reported, Alu-LTR levels were decreased in T<sub>L3</sub> cells as compared to T<sub>L3</sub> cells expressing LEDGF/p75 WT, indicating that HIV-1 DNA integration is defective in the absence of LEDGF/p75 (Fig. 12). Importantly, TL3 cells expressing LEDGF/p75 S271A/S273A/S275A were also defective for viral integration and the integration levels correlated in some extension with the degree of impairment in its co-factor activity shown in figures 11d and 11f.

In order to estimate HIV-1 integration by an alternative approach, total HIV DNA (gag DNA) was determined in these cells ten days post-infection. Since infected cells passed several generations before analysis, it is expected that only HIV DNA integrated forms will persist and gag DNA will be a good indicator of viral integration. Nevertheless, to exclude the possibility of detecting also remnants of non-integrated forms of the viral genome, the levels of total HIV DNA were normalized to levels of HIV 2LTR circles. Importantly, similar 2LTR levels were detected in the different infected cell lines at the time of analysis (Data not shown). Results in figure 12 shown that total HIV DNA was also reduced in infected T<sub>L3</sub> LEDGF/p75

S271A/S273A/S275A cells as compared to cells expressing the wild type protein, indicating that HIV DNA integration is defective in T<sub>L3</sub> cells expressing this LEDGF/p75 mutant.

Importantly, the total HIV DNA levels in the analyzed cells correlated better with their susceptibility to infection than the levels of Alu-LTR junctions; compare data in figures 11d and 12. These results suggested that HIV-1 DNA integration in genomic regions distant from Alu sequences was also impaired in T<sub>L3</sub> and T<sub>L3</sub> LEDGF/p75 S271A/S273A/S275A cells.



**Figure 12.** *HIV-1 DNA integration in  $T_{L3}$  cells expressing LEDGF/p75 S271A/S273A/S275A.* Results from two independent experiments using different viral preparations are represented. DNA products were quantified by real time PCR ten days after infection. Bars represent (a) Alu-LTR DNA levels after normalization to mitochondrial DNA and in (b) total HIV DNA after normalization to 2LTR circles DNA. Normalized levels of Alu-LTR and total HIV DNA detected in  $T_{L3}$  LEDGF/p75 WT cells were considered as 100%.

**Chromatin-binding activity of LEDGF/p75 mutants.** It has been proposed that LEDGF/p75 acts as a molecular tether that links the pre-integration complex-associated integrase to the host chromatin [41, 51]. According to the tethering model, two functional LEDGF/p75 regions, the chromatin-binding domain and the IBD, are required for HIV-1 integration [4]. In order to correlate the HIV-1 cofactor activity of LEDGF/p75 CR mutants with their tethering capacity, we evaluated their chromatin binding and integrase interaction activities.

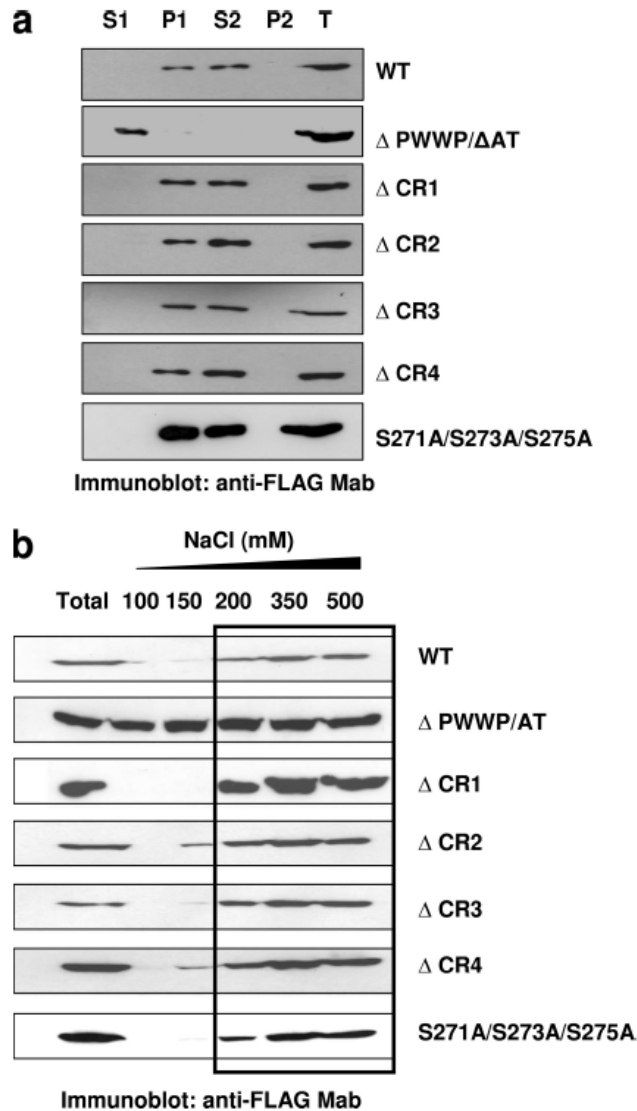
Chromatin-binding was evaluated using the previously reported chromatin-binding assay with minor modifications [39] (Fig. 13a). This procedure is based on the resistance of chromatin-bound proteins to be extracted with isotonic buffers (i.e. CSK I buffer) containing 1% Triton-X100. In agreement with previous observations [39, 44], LEDGF/p75 persisted bound to the chromatin fraction (P1 fraction) after cellular lysis in 1% Triton-X100 CSK I buffer and it was only released from this compartment by DNase and 250mM  $(\text{NH}_4)_2\text{SO}_4$  treatment (S2 fraction) (Fig. 13a). However, a LEDGF/p75 mutant lacking the chromatin-binding domains,  $\Delta\text{PWWP}/\Delta\text{AT}$ , was fully extracted by 1% Triton-X100 CSK I buffer (S1 fraction) (Fig. 13a). This assay revealed that all four LEDGF/p75 deletion mutants and LEDGF/p75 S271A/S273A/S275A were exclusively distributed to the chromatin-bound fractions (P1 and S2 fractions) as the wild type LEDGF/p75 protein (Fig. 13a). Importantly, although the HIV-1 cofactor activity of LEDGF/p75  $\Delta\text{CR3}$  or LEDGF/p75 S271A/S273A/S275A was impaired (Fig. 10c and 11d), the chromatin-binding activity of these mutants was unaffected. These results excluded defective chromatin binding as the cause for a deficient cofactor activity in these LEDGF/p75 mutants.

Although the CR-deleted LEDGF/p75 mutants are chromatin-bound, we evaluate the strength of their binding by the salt extraction method described in figure 9a. Using this method,

we found that LEDGF/p75 WT persisted bound to chromatin (salt-resistant fraction) at 100 and 150mM NaCl and was fully extracted above 200mM NaCl (Figs. 9a and 13b). Contrary to this, a mutant lacking the PWWP domain was partially extracted at 100mM and fully extracted above 150mM NaCl (Data not shown) whereas LEDGF/p75  $\Delta$ PWWP/ $\Delta$ AT, that failed to bind to chromatin, was fully extracted at 100mM NaCl (Fig. 13b). Therefore, the strength of chromatin binding determined for LEDGF/p75 WT and  $\Delta$ PWWP/ $\Delta$ AT by the salt extraction method perfectly correlated with their chromatin-binding activity observed with the previously described chromatin-binding assay (Fig. 13a).

The strength of chromatin binding of the different LEDGF/p75 CR deletion mutants was then evaluated by the salt extraction method. As shown in figure 13b, deletion of CR1 or CR3 did not affect the LEDGF/p75 chromatin binding strength since these mutant proteins were only fully extracted above 200mM NaCl. In contrast to these results, deletion of CR2 and CR4 slightly decreased LEDGF/p75 resistance to salt extraction evidenced by partial extraction of these mutants at 150mM NaCl (Fig. 13b). These data suggested that the CR2 and CR4 regions contribute to the number of charged interactions between LEDGF/p75 and chromatin. In addition, these results correlated with a previous report indicating a contribution of CR2 and the C-terminal region of LEDGF/p75 to chromatin binding [39].

Similarly to LEDGF/p75  $\Delta$ CR3, the strength of chromatin binding of LEDGF/p75 S271A/S273A/S275A was indistinguishable from the wild type protein (Fig. 13b). These results, together with the wild type phenotype observed for these mutants in the chromatin-binding assay (Fig. 13a), strongly indicated that their defective HIV-1 cofactor activity cannot be ascribed to an impaired chromatin-binding capacity.



**Figure 13.** *Chromatin-binding activity of LEDGF/p75 mutants.* **(a)** Chromatin-binding assay.  $T_{L3}$  cells expressing different LEDGF/p75 mutants were fractionated into non-chromatin-bound fractions (S1 and P2) and chromatin-bound fractions (P1 and S2) and the presence of LEDGF/p75 in these fractions was evaluated by immunoblot with an anti-FLAG Mab. An unfractionated total cellular fraction (T) was included as control. **(b)** Salt extraction analysis of LEDGF/p75 mutants stably expressed in  $T_{L3}$  cells. The salt concentration that extracted LEDGF/p75 WT from chromatin is marked with a rectangle.

**Interaction of LEDGF/p75 mutants with HIV-1 integrase.** In order to analyze the functional interaction of LEDGF/p75 mutants with HIV-1 integrase, we evaluated their capacity to shield integrase from degradation (integrase protection assay, fig. 14). LEDGF/p75-deficient HEK293T cells, stably expressing Myc-tagged HIV-1 integrase, were transiently transfected with plasmids expressing LEDGF/p75 WT or mutants and integrase protein levels were evaluated by immunoblot forty-eight hrs later.

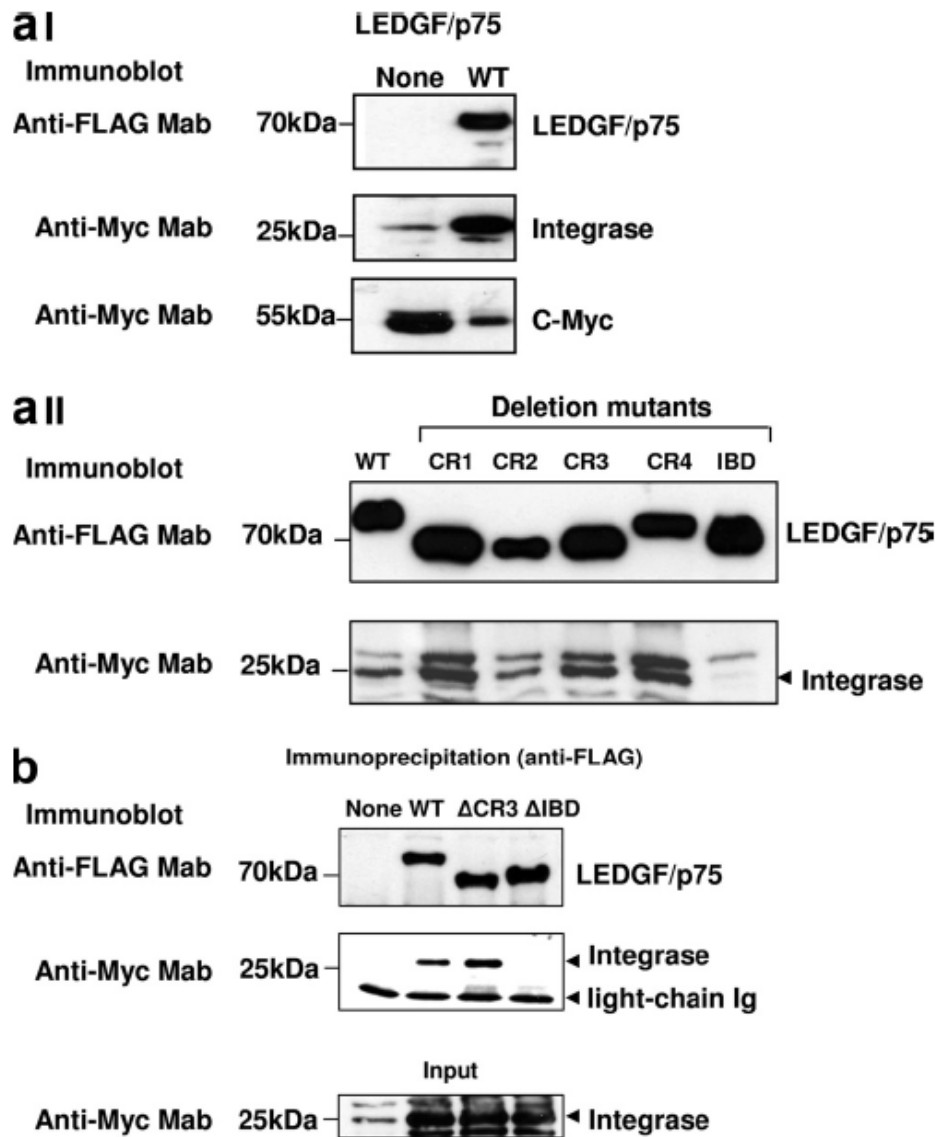
Steady-state HIV-1 integrase levels were very low in the absence of LEDGF/p75 and dramatically increased upon expression of LEDGF/p75 WT but not of LEDGF/p75  $\Delta$ IBD (Figs. 14a-I and 9b), as previously reported [40]. Using this reporter system, we evaluated the interaction of the different LEDGF/p75 deletion mutants with HIV-1 integrase. Significantly, expression of the different LEDGF/p75 CR deletion mutants (Fig. 14a-II) and of the LEDGF/p75 S271A/S273A/S275A mutant (Data not shown) prevented integrase degradation in LH4 cells. The levels of rescued integrase mirrored those of re-expressed LEDGF/p75 proteins; indicating that these mutations did not affect the capacity of LEDGF/p75 to interact with the viral protein. In addition, these data confirmed previous observations showing that the IBD was necessary and sufficient to protect integrase from degradation [40]. Furthermore, these findings clearly indicated that the reduced cofactor activity observed in LEDGF/p75  $\Delta$ CR3 or the triple point mutant LEDGF/p75 S271A/S273A/S275A was not due to a defective interaction with HIV-1 integrase.

In order to evaluate directly the interaction of LEDGF/p75  $\Delta$ CR3 with HIV-1 integrase, we performed a co-immunoprecipitation experiment (Fig. 14b). LEDGF/p75 WT, or  $\Delta$ CR3 or  $\Delta$ IBD were transiently co-expressed with Myc-tagged HIV-1 integrase in LEDGF/p75-knockdown HEK293T cells, si1340/1428 cells [51]. Forty-eight hrs later, cells were lysed in

RIPA buffer and incubated with anti-mouse Igs-coated magnetic beads loaded with anti-FLAG Mab. Immunoprecipitated proteins were later evaluated for the presence of HIV-1 integrase or LEDGF/p75 by immunoblotting with anti-Myc or anti-FLAG Mabs, respectively. Data shown in figure 14b, clearly demonstrated that deletion of CR3 did not alter the capacity of LEDGF/p75 to interact with HIV-1 integrase. These observations correlated with the data obtained in the integrase protection assay.

In summary, the results shown above indicated that the defective HIV-1 cofactor activity detected in LEDGF/p75  $\Delta$ CR3 or LEDGF/p75 S271A/S273A/S275A is not due to an impaired interaction of these mutants with HIV-1 integrase.



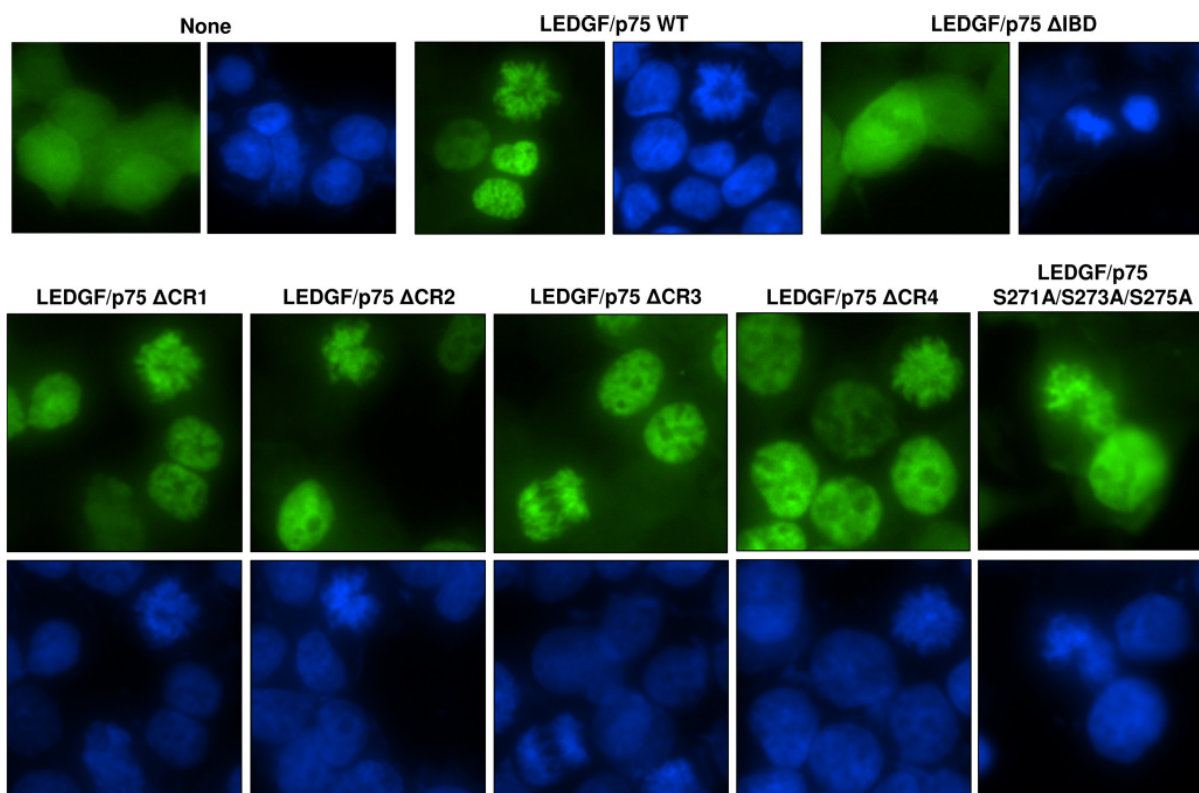


**Figure 14.** *Interaction of LEDGF/p75 mutants with HIV-1 integrase.* **(a)** Integrase protection assay. LEDGF/p75-deficient HEK293T cells stably expressing myc-tagged HIV-1 integrase were either transiently transfected or not with plasmids expressing LEDGF/p75 WT or mutants. Expression levels of LEDGF/p75 proteins and HIV-1 integrase were determined by immunoblotting with an anti-FLAG and an anti-Myc Mab, respectively. Detection of endogenous C-Myc was used as a loading control in these experiments. **(b)** Co-

immunoprecipitation of HIV-1 integrase with LEDGF/p75  $\Delta$ CR3. LEDGF/p75-deficient HEK293T cells were co-transfected with different FLAG-tagged LEDGF/p75 WT or mutants and Myc-tagged HIV-1 and subject to immunoprecipitation with an anti-FLAG Mab. Immunoprecipitated proteins were detected by immunoblotting with anti-Myc or anti-FLAG Mabs. Mouse antibodies light chains were used as a loading control for the immunoprecipitation.

### **Integrase to chromatin tethering activity of LEDGF/p75 mutants.**

We next evaluated the capacity of LEDGF/p75 mutants to tether HIV-1 integrase to chromatin (integrase to chromatin tethering assay). In this assay, we used LEDGF/p75-deficient cells expressing HIV-1 integrase-eGFP (2L<sub>KD</sub>-IN-eGFP cells). HIV-1 integrase-eGFP was distributed pan-cellularly in 2L<sub>KD</sub>-IN-eGFP cells during interphase (Fig. 15) and it was not associated with condensed chromatin during mitosis (Data not shown). However, upon re-expression of LEDGF/p75 WT, this fluorescent fusion protein was dramatically localized to the cell nuclei in interphase or to condensed chromatin during cell division (Fig. 15). However, and as expected, this striking sub-cellular redistribution was not observed after expression of LEDGF/p75  $\Delta$ IBD in 2L<sub>KD</sub>-IN-eGFP cells (Fig. 15).



**Figure 15.** *Integrase to chromatin tethering assay.* LEDGF/p75-deficient HEK293T cells stably expressing eGFP-tagged HIV-1 integrase were either transiently transfected or not with plasmids expressing LEDGF/p75 WT or mutants and analyzed by fluorescence microscopy. Chromatin was stained with DAPI.

Using the integrase to chromatin tethering assay we observed that transient expression of the four LEDGF/p75 CR mutants or LEDGF/p75 S271A/S273A/S275A consistently tethered integrase-eGFP to the nuclear compartment and to mitotic chromosomes (Fig. 15). This effect on integrase-eGFP localization was identical to the observed with the LEDGF/p75 WT, indicating that mutations of these regions did not alter the tethering capacity of LEDGF/p75. These results were also in agreement with data obtained in the chromatin-binding and integrase protection assays.

Although a reduction in the strength of the LEDGF/p75 chromatin binding was observed after deletion of CR2 and CR4 (Fig. 13b) this defect did not impair the capacity of these mutants to tether integrase to chromatin *in cells*. Even more important, the defect of LEDGF/p75  $\Delta$ CR3 or LEDGF/p75 S271A/S273A/S275A in rescuing HIV-1 infection in T<sub>L3</sub> cells cannot be attributed to an impairment of the tethering capacity of these mutants.

In sum, these results clearly demonstrated that the reduced HIV-1 cofactor activity observed in LEDGF/p75  $\Delta$ CR3 and in the LEDGF/p75 S271A/S273A/S275A is not due to impairment of their chromatin-binding activity, HIV-1 integrase binding activity, or integrase to chromatin tethering capability. These data strongly emphasized the relevance of LEDGF/p75 serine residues 271, 273 and 275 in the HIV-1 cofactor activity of a tethering competent LEDGF/p75 molecule.

## 2.4. Discussion

LEDGF/p75 is a cellular cofactor of HIV-1 DNA integration [4]. However, the molecular mechanism of this activity awaits further clarification. Analysis of LEDGF/p75 mutants indicated that chromatin- and integrase-binding properties of this cellular protein are required for its cofactor activity [4, 37]. These observations have led to propose a tethering mechanism where chromatin-bound LEDGF/p75 interacts with HIV-1 integrase tethering the PIC to the host chromatin [4]. The role of LEDGF/p75 in the HIV-1 integration site distribution further supported this tethering model [35, 37, 41]. In addition to tethering integrase to the chromatin, LEDGF/p75 protects the viral enzyme from proteasome-mediated degradation and increases its enzymatic activity *in vitro* [40, 54]. However, it is unknown if these LEDGF/p75 functions are also required for its HIV-1 cofactor activity.

Here we report molecular evidences of a new requirement for the HIV-1 cofactor activity of LEDGF/p75. We have found that serine residues 271, 273 and 275 of LEDGF/p75 importantly influence its HIV-1 cofactor activity. Mutation of these amino acids to alanine impairs the HIV-1 cofactor activity of LEDGF/p75 without altering its chromatin- or integrase-binding activities or its integrase to chromatin tethering capacity as measured by different cellular and biochemical approaches. The level of impairment in the cofactor activity of LEDGF/p75 caused by these point mutations is similar to levels observed in LEDGF/p75 mutants that are defective for chromatin binding. In addition, a pre-integration defect in the HIV-1 viral life cycle was observed in T<sub>L3</sub> cells expressing this mutant. These results suggest that serine 271, 273 and 275 mediate an unknown LEDGF/p75 function that is involved in HIV-1 DNA integration.

Beyond the identification of a role of serine residues 271, 273 and 275 in the HIV-1 cofactor activity of LEDGF/p75, our data formally demonstrate that the CR regions of the protein are not required for chromatin or integrase interaction. Noticeably, chromatin-binding, integrase-binding and integrase to chromatin tethering functions of LEDGF/p75 tolerate significant structural changes since none of the CR deletions importantly impaired them. In contrast, the HIV-1 cofactor activity of LEDGF/p75 was more sensitive to these structural changes. Although mutations affecting serines 271, 273 and 275 ( $\Delta$ CR3,  $\Delta$ CR3.1 and S271A/S273A/S275A) markedly reduced the cofactor activity of LEDGF/p75, CR1 deletion also decreased its activity. These results indicated that tethering HIV-1 integrase to chromatin, while central in the molecular mechanism of LEDGF/p75 in HIV-1 infection, is not sufficient for the full HIV-1 cofactor activity of this cellular protein.

The molecular mechanism of the serine residues 271, 273 and 275 in the HIV-1 cofactor activity of LEDGF/p75 is not known yet. This motif is evolutionarily conserved from avian to humans and is located in a region of LEDGF/p75 predicted to be highly exposed to the solvent [60]. Serine residues are targeted by different post-translational modifications including: phosphorylation, O-acetylation and O-glycosylation. Although these serines are not predicted to be O-acetylated or O-glycosylated, they have been found phosphorylated in different global phosphoproteomic analysis of human and mouse cells [64-71, 73]. Serine 271, 273 and 275 are predicted to be a substrate for phosphorylation by PKCK2. This kinase phosphorylates Ser or Thr residues in the motif S/T-X-X-E/D, where X are non-basic amino acids [72]. PKCK2-dependent phosphorylation facilitates protein-protein interactions in different cellular processes [72, 74-77]. Based on all these reported evidences we speculate that serine phosphorylation at

this motif could influence the cofactor activity of LEDGF/p75 by facilitating its interaction with other cellular proteins implicated in HIV-1 integration.



## **2.5. Acknowledgements**

This work was supported by the Public Health Service grant 1SC2GM082301-01 from the NIH and also by the University of Texas at El Paso (UTEP)-start up funds. J.A.G. was supported by a fellowship from the UTEP RISE program, NIH grant 2R25GM069621-05. J.R.K. was supported by HHMI grant 52005908. UTEP core facilities are funded by the BBRC grant 5G12RR008124.

We thank: Eric Poeschla, Mayo Clinic Rochester MN, for providing us with important reagents. Janeth Cortez and Damaris Rosado for technical assistance. Igor Almeida and Luciane Ganiko, UTEP, for technical support with the microscopic analysis. Jeremy Ross, UTEP, for important suggestions.

## Chapter 3: LEDGF/p75 is a Member of the Transcriptional Elongation Complex

### 3.1 Introduction

Since the inception of LEDGF/p75 characterization, a number of protein binding partners have been described with these including cellular proteins such as JPO2, pogZ, menin/MLL and Cdc7-ASK [78-81]. The interesting fact about all of these proteins is that they all interact with the IBD of LEDGF/p75. In an attempt to further characterize LEDGF/p75 interactors we conducted a mass spectrometry analysis of LEDGF/p75. From this analysis a number of potential LEDGF/p75 protein-protein interactors was revealed (Table 1). As part of this study it was decided to move forward with SSRP1 (Structure Specific Recognition Protein 1) and SPT16 (Suppressor of Ty 16 homolog) since they form a heterodimeric protein complex known as FACT (**F**Acilitates **C**hromatin **T**ranscription) which affects eukaryotic RNA polymerase II (Pol II) transcription elongation both *in vitro* and *in vivo* [82].

The FACT complex represents a class of chromatin structure modulator known to reorganize nucleosomal structure by the removal and/or reassembly of the histone H2A-H2B dimers. Co-immunoprecipitation experiments done with tagged recombinant proteins have shown that the Spt16 subunit interacts with H2A/H2B dimers and mononucleosomes, but not H3/H4 tetramers, whereas the SSRP1 subunit interacts only with H3/H4 tetramers and not mononucleosomes [83].

We demonstrated by co-immunoprecipitation the interaction of LEDGF/p75 with SSRP1 and Spt16, and these findings were corroborated by quantitative confocal co-localization analysis thereby demonstrating that LEDGF/p75 is a member of the transcriptional elongation complex (Fig 16 and 17).

<b><u>Protein</u></b>	<b><u>Gene</u></b>	<b><u>Gene ID</u></b>	<b><u>Species</u></b>	<b><u>Weight</u></b>	<b><u>Hits</u></b>
CSE1L	XPO2	1434	Human	110 kDa	7
Dynamin-2	P50570	47117856	Human	98 kDa	9
NOP14	NOP14	8602	Human	98 kDa	33
NAT10	NAT10	55226	Human	116 kDa	21
PWP2	PWP2	5822	Human	103 kDa	27
NCBP1	NCBP1	4686	Human	92 kDa	37
P5CS	P54886	6226882	Human	87 kDa	42
FUBP2	FUBP2	8570	Human	73 kDa	23
NOL11	NOL11	25926	Human	81 kDa	26
XRCC5	KU86	7520	Human	83 kDa	42
DDX50	DDX50	79009	Human	82.5 kDa	32
SSRP1	SSRP1	6749	Human	81 kDa	9
NOL6	NOL6	65083	Human	128 kDa	3

**Table 1.** *Potential LEDGF/p75 Interactors.* Proteins that potentially interact with LEDGF/p75 were elucidated through mass spectrometry analysis. LEDGF/p75 deficient HEK 293T were cotransfected or not (negative control) with FLAG-tagged LEDGF/p75. Cellular extracts containing the chromatin bound proteins were obtained and immunoprecipitated with anti-FLAG Mab. Proteins were separated on an SDS-PAGE gel and silver stained. Bands appearing in the sample obtained from cells expressing FLAG-tagged LEDGF/p75 but not in the non-transfected cells were excised along with the corresponding area in the negative control. These bands were then subjected to mass spectrometry analysis and protein hits appearing in the experimental control and not within the negative control were designated as potential LEDGF/p75 interactors.

### 3.2 Materials and Methods

**Cell lines.** Cell lines were derived from the human CD4<sup>+</sup> T cell line SupT1 by stably expression of an shRNA sequence targeting LEDGF/p75 (T<sub>L3</sub> cells) or a control scrambled shRNA (T<sub>C3</sub> cells). T<sub>L3</sub> cells express 97% less LEDGF/p75 mRNA than T<sub>C3</sub> cells as determined by real time PCR. LEDGF/p75-deficient T<sub>L3</sub> cells were later engineered to re-express a C-terminally FLAG-tagged LEDGF/p75. HeLa cells were used for confocal colocalization studies.

**Mass Spectrometry.** Proteins that potentially interact with LEDGF/p75 were elucidated through mass spectrometry analysis. LEDGF/p75 deficient HEK 293T were cotransfected or not (negative control) with FLAG-tagged LEDGF/p75. Cellular extracts containing the chromatin bound proteins were obtained and immunoprecipitated with anti-FLAG Mab. Proteins were separated on an SDS-PAGE gel and silver stained. Bands appearing in the sample obtained from cells expressing FLAG-tagged LEDGF/p75 but not in the non-transfected cells were excised along with the corresponding area in the negative control. These bands were then subjected to mass spectrometry analysis and protein hits appearing in the experimental control and not within the negative control were designated as potential LEDGF/p75 interactors.

**Immunoprecipitation.** The interaction of the LEDGF/p75 with the FACT complex was evaluated by immunoprecipitation using the LEDGF/p75-deficient T<sub>L3</sub> cells engineered to express FLAG-tagged LEDGF/p75 wild type (T<sub>L3</sub> LEDGF/p75 WT cells). Briefly, 18x10<sup>6</sup> T<sub>L3</sub> LEDGF/p75 WT cells were washed in PBS and lysed for 15 mins on ice in 100µl of CSK I buffer (10mM Pipes pH6.8, 100mM NaCl, 1mM EDTA, 300mM sucrose, 1mM MgCl<sub>2</sub>, 1mM

DTT, 0.5% Triton X-100) containing protease inhibitors (final concentration: leupeptine 2µg/ml, aprotinin 5µg/µl, PMSF 1mM, pepstatin A 1µg/ml). Cells lysed in CSK I buffer were centrifuged at 1000g for 6 mins at 4°C and the pellet was resuspended in 100µl of CSK II buffer supplemented with protease inhibitors, 4 units of turbo DNase (Ambion) and 11µl of 10X turbo DNase reaction buffer. DNase treatment of this pellet was conducted at 37°C for 30 mins and then centrifuged at 22,000g for 3 mins and the supernatant used for immunoprecipitation using goat anti-mouse IgG-coated magnetic beads (Pierce). Beads (100µl) were previously loaded for 20 mins on ice with 3µg of anti-FLAG Mab diluted in CSKI buffer. Then, beads were separated from the unbound antibodies, mixed with the DNase-extracted cell lysate and rotated for 2hrs at 4°C. After this incubation, beads were washed three times in CSKI buffer and bound proteins eluted by boiling in 30µl of Laemmli sample buffer. Immunoprecipitated proteins were analyzed by immunoblotting for the presence of the components of the FACT complex: SSRP1 (Santa Cruz, sc-56782) and Spt16 (Santa Cruz, sc-28734). LEDGF/p75 was detected with anti-FLAG Mab (Sigma, F1804).

**Quantitative co-localization assay.** Co-localization was evaluated with a confocal microscope using the Zeiss Zen software. In order to set up this method we evaluated co-localization of LEDGF/p75 with HIV-1 integrase. These proteins interact during all the phases of the cell cycle. As a negative control a LEDGF/p75 mutant lacking the integrase binding domain (IBD) was used. LEDGF/p75-deficient HEK 293T cells stably expressing eGFP-tagged HIV-1 integrase (2LKD-IN-eGFP cells) were plated at  $2 \times 10^5$  cells in LabTek II chambered coverglasses and transfected the next day with 2ug of pFLAG-LEDGF/p75 WT or IBD deletion mutant. Eighteen hrs after transfection fresh culture medium was added and forty-eight hours later cells were

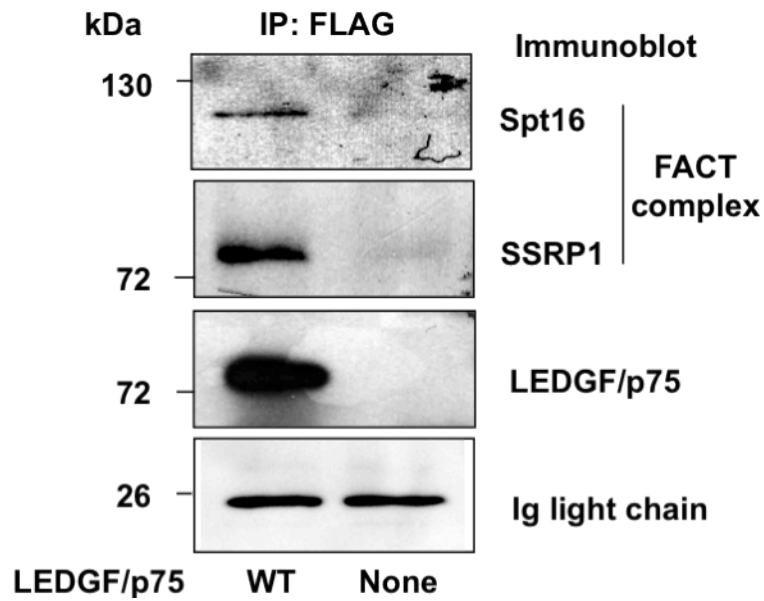
washed three times in PBS and fixed with 4% formaldehyde-PBS for 10 min at 37°C. Then, cells were washed twice in PBS and immunostained for 2 hrs at 37°C. LEDGF/p75 was detected with an anti-mouse LEDGF Mab diluted 1/100 (clone 26, BD Transduction Laboratories) or an anti-rabbit LEDGF polyclonal antibody (Pab) diluted 1/100 (Bethyl Laboratories, A300-848A) followed by incubation with anti-mouse Ig coupled to Alexa Fluor 594 (10 µg / ml, Invitrogen A21203) or with an anti-rabbit Ig coupled to Alexa Fluor 488 (10 µg/ml, Invitrogen A21206). Immunostained cells were washed and stained with DAPI. Colocalization of LEDGF/p75 with HIV-1 eGFP-integrase was analyzed with a confocal microscope and the Zeiss Zen software. In order to measure co-localization of LEDGF/p75 with proteins of the transcriptional apparatus  $2 \times 10^5$  HeLa cells were plated in LabTek II chambered coverglasses and immunostained as described above. Antibodies against SSRP1 (Santa Cruz, sc-56782) and Spt16 (Santa Cruz, sc-28734) were used. Secondary antibodies, DAPI staining and colocalization analysis was performed as described above.

### 3.3 Results

**Interaction of LEDGF/p75 with proteins involved in transcription.** LEDGF/p75-deficient SupT1 cells (T<sub>L3</sub> cells) and T<sub>L3</sub> cells expressing FLAG-tagged LEDGF/p75 were subjected to immunoprecipitation with an anti-FLAG Mab. Immunoprecipitated proteins were then evaluated by immunoblotting for the presence of different components of the elongation complex, the heterodimer SSRP1/Spt16 or FACT complex. The FACT complex remodels nucleosomes in an ATP-independent fashion allowing Pol II to access the DNA during transcriptional elongation [82].

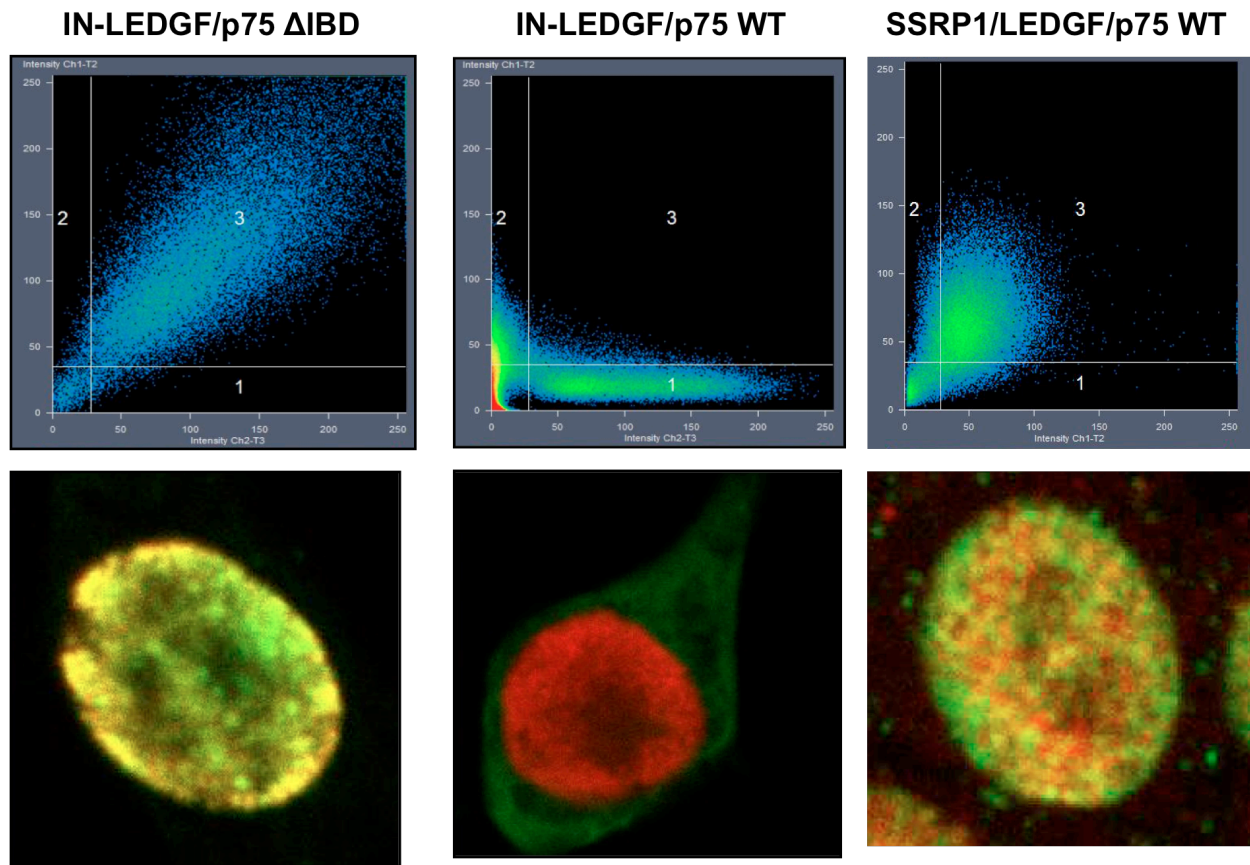
SSRP1 exists in a chromatin-bound and a non-chromatin bound fractions. Therefore, in order to evaluate the interaction of LEDGF/p75 with the FACT complex forms involved in transcriptional elongation, we lysed the cells in a buffer containing 0.1% Triton X-100 and 100 mM NaCl (CSK1 buffer). This buffer has been reported to extract only chromatin non-bound proteins. Chromatin-bound proteins were then released from the insoluble fraction of the CSKI-cellular lysate by extensive treatment with DNase. The DNase-released chromatin-bound proteins were then subjected to FLAG immunoprecipitation and coimmunoprecipitated proteins were evaluated for the presence of the FACT complex by immunoblotting with specific antibodies. Results included in figure 16 indicate that LEDGF/p75 is associated *in vivo* with both subunits, SSRP1 and Spt16, of the FACT complex, which are involved in transcriptional elongation. Since chromatin-bound proteins were extracted by DNase treatment before immunoprecipitation, the association of LEDGF/p75 with these members of the transcriptional elongation complex is independent of DNA.





**Figure 16.** *Interaction of LEDGF/p75 with components of the elongation complex.* Chromatin-bound proteins were isolated from T<sub>L3</sub> and T<sub>L3</sub> LEDGF/p75 WT cells by DNase and salt treatment and FLAG-tagged LEDGF/p75 was immunoprecipitated with an anti-FLAG Mab. The presence of the FACT complex (Spt16 and SSRP1) was evaluated in the immunoprecipitated proteins by immunoblotting with specific antibodies.

To further verify the interaction of LEDGF/p75 with SSRP1 we performed a quantitative confocal colocalization analysis (Fig. 17). As control, we calculated the co-localization of HIV-1 integrase (IN) with LEDGF/p75 WT or a mutant lacking the integrase binding domain ( $\Delta$ IBD). In correlation with previously reported data, LEDGF/p75 WT fully co-localized with IN and this co-localization was lost upon deletion of IBD (Fig. 17a-I). LEDGF/p75 WT significantly co-localized with SSRP1. Co-localization with SSRP1 validated the results of the immunoprecipitation experiment included in figure 16. Results presented in figures 16 and 17 clearly indicate that LEDGF/p75 interacts with components of the transcriptional elongation complex.



**Figure 17.** *Quantitative confocal colocalization of LEDGF/p75 with proteins of the transcriptional machinery. Colocalization controls. Colocalization of LEDGF/p75 and HIV integrase were used as controls. LEDGF/p75-deficient HEK293T cells stably expressing Myc-tagged HIV-1 integrase were transiently transfected with LEDGF/p75 WT or a mutant lacking the integrase-binding domain (IBD). LEDGF/p75 and integrase were detected with anti-LEDGF and anti-Myc antibodies, respectively. Colocalization of LEDGF/p75 with SSRP1. HeLa cells were fixed and immunostained with specific antibodies. Histograms and representative cells are shown.*

### 3.4 Discussion

A role of LEDGF/p75 in transcriptional regulation at promoters regions has been extensively described, yet the participation of LEDGF/p75 in other steps of transcriptional regulation has not been evaluated. Co-transfection of LEDGF/p75 with several promoters of stress responsive genes indicated that LEDGF/p75 regulates the transcription of these genes by binding to specific DNA sequences in their promoters [35]. However, the specific interaction of LEDGF/p75 with these sequences was not supported in an independent study. At the molecular level, it was demonstrated in *in vitro* studies that LEDGF/p75 activated transcription of reporter promoters by linking promoter-specific transcription factors with components of the basal transcription machinery. In addition, it was demonstrated LEDGF/p75 regulates the transcriptional activity of the Hox A9 gene by tethering to the promoter of this gene the histone methyl transferase menin/MLL complex [79]. Interestingly, a similar molecular tethering mechanism is involved in the role of LEDGF/p75 in HIV DNA integration.

LEDGF/p75 tethers the HIV-1 pre-integration complex to the host chromatin facilitating viral DNA integration into the host DNA. This interaction targets HIV-1 to integrate inside genes that are actively transcribed. In correlation with this, genome-wide location of LEDGF/p75 indicated that this protein is enriched inside actively transcribed genes [35]. In correlation with these experimental evidences, we found that LEDGF/p75 colocalized *in vivo* with the FACT complex, a component of the transcriptional elongation complex. Interaction with the transcriptional elongation active forms of the FACT complex was also further evidenced to occur *in vivo* and to be independent of DNA.

We do not understand yet the role of LEDGF/p75 in the transcriptional elongation complex. However, it could be involved in tethering the histone methyltransferase menin/MLL

complex to the transcription unit to catalyze trimethylation of lysine 4 on histone H3, an epigenetic mark associated with transcriptionally active chromatin and activity of P-TEFb. A similar role has been reported for the yeast COMPASS (Complex Proteins Associated with Set1) complex which is a functional analog of the human MLL [84]. The COMPASS complex is recruited by Pol II (phosphorylated at serine 5 in the C-terminal domain) to trimethylate lysine 4 on histone H3, a modification highly correlated with transcriptional elongation [85]. It is also possible that LEDGF/p75 participates in attracting to these subset of genes other components of the elongation complex.

In addition, we postulate that the presence of LEDGF/p75 at the elongation complex will facilitate HIV DNA integration by allowing the virus to exploit the cellular function of other members of this complex. For instance, the FACT complex could assist the HIV preintegration complex to gain access to the host DNA in a similar manner that FACT facilitates DNA access to the RNA polymerase during transcription. Similarly, the interaction of LEDGF/p75 with the component of the FACT complex, SSRP1, could facilitate post-integration DNA repair. SSRP1 has a role in transcription-coupled DNA repair. This is a subpathway of the nucleotide excision repair pathway that is responsible for removing DNA lesions in actively transcribed DNA strands that cause RNA polymerase to stall during transcription. HIV has a bias for integration inside actively transcribed genes [35]. Therefore, it is likely that HIV integration will cause arrest of transcription of the integration target genes. The interaction of LEDGF/p75 with SSRP1 would give an advantage to the virus to initiate efficient DNA repair following viral DNA integration. Since LEDGF/p75 is a member of the elongation complex it will be also advantageous for the virus the interaction with LEDGF/p75 to dock into chromatin areas that are

engaged in active transcription allowing efficient viral gene expression rapid after viral integration.

### **3.5 Acknowledgements**

This work was supported by grant 1SC2GM082301-01 from the National Institute of General Medical Sciences (NIGMS), National Institutes of Health (NIH), (to ML). J.R.K., P.A. and D.R. were supported by HHMI grant 52005908, J.A.G. was supported by a fellowship from the UTEP RISE program, NIH grant 2R25GM069621-05. University of Texas at El Paso (UTEP) core facilities are funded by the BBRC grant 5G12RR008124.

We thank: Eric Poeschla, Mayo Clinic, Rochester MN for providing us with LEDGF/p75-deficient cell lines and LEDGF/p75 expression plasmids. Armando Varela (UTEP) for the assistance in using the confocal microscope. Carolina Lema (UTEP) for help with the gene expression profiling analysis.

## **Chapter 4: Final Conclusions and Future Directions**

### **5.1. Overview and Final Conclusions**

Throughout the past years many significant strides have been made in regards to the role LEDGF/p75 plays in HIV-1 infection. Since the discovery that LEDGF/p75 co-immunoprecipitates with HIV-1 integrase much work has been done in an effort to further understand the molecular mechanism by which it interplays with other factors within the cell in order for integration to occur.

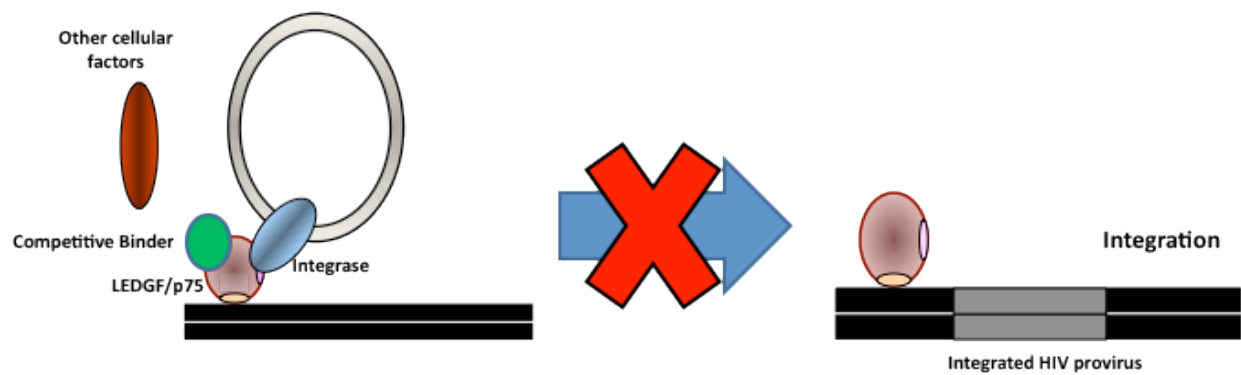
At the inception of this dissertation work, three main functional domains within LEDGF/p75 had been described, with this being: the Chromatin Binding Domain (CBD), which constitutes the functional interaction of the PWWP domain and two AT hook motifs, the Nuclear Localization Signal (NLS), and the Integrase Binding Domain (IBD). As part of my contribution to the HIV-1 field my work focused on further characterizing other potential functional regions within the LEDGF/p75 protein which met the following criteria: (1) The regions had to be evolutionarily conserved from amphibians to humans. (2) They had to contain clusters of charged residues. (3) These regions could serve as potential targets of post-translational modifications.

From the results obtained in this study it was concluded that three serine residues (S271/S273/S275) within LEDGF/p75 have a role in HIV-1 infectivity. Although the exact mechanism by which this occurs has not yet been elucidated, we hypothesize that CK2 is interacting with this region since the motif present within the region is a predicted CK2 phosphorylation site. One could speculate that this serine motif could serve as a docking site for



other cellular proteins that participate in HIV-1 DNA integration and that the prevention of this putative interaction could be an attractive target for drug development (Fig. 18)

Another region that was found to be involved in LEDGF/p75's HIV-1 cofactor activity was CR1. The deletion of this region resulted in a reduction of LEDGF/p75's cofactor activity to 41% when compared to WT. Now although these regions are implicated in HIV-1 infection they do not alter the ability of LEDGF/p75 to tether HIV-1 integrase to chromatin indicating the existence of chromatin tethering independent mechanisms in the molecular mechanism of LEDGF/p75 in HIV-1 infection.



**Figure 18.** *Research Implications.* The S271/S273/S275 motif could serve as a potential drug target site with which a competitive binder could be developed in order to block the interaction of cellular factors with this region and potentially suppress HIV-1 integration.

The discovery of the implication of the serine motif and the potential involvement of other cellular factors with LEDGF/p75 lead us to conduct a proteomic analysis of LEDGF/p75 and characterize protein-protein interactors through the use of mass spectrometry. Our hypothesis is that LEDGF/p75 may be forming a complex with other cellular factors important for HIV-1 infection as well as the cellular functions of LEDGF/p75. Through this work we elucidated a number of potential LEDGF/p75 interactors and verified by co-immunoprecipitation the interaction of LEDGF/p75 with SSRP1 and SPT16. These two proteins form a heterodimer known as the FACT (FAcilitates Chromatin Transcription) complex, which is involved in transcriptional elongation. Furthermore, we demonstrated the interaction of LEDGF/p75 with this complex by quantitative confocal co-localization studies, which evaluated the direct interaction of LEDGF/p75 with this complex. These findings correlate with other data obtained within the laboratory implicating LEDGF/p75 in the transcriptional elongation complex.

## **5.2. Future Directions**

The further characterization of the interaction of LEDGF/p75 and FACT is vital to understanding the role LEDGF/p75 plays in regard to HIV-1 integration and cellular transcriptional elongation. Work such as the mapping of which region(s) within LEDGF/p75 are implicated in this interaction would provide insight into the nature of this interaction. In addition the direct role of the FACT complex in HIV-1 DNA integration can be explored in cells transiently depleted of SSRP1 and Spt16 since stable knockdown are not viable.

In summary, our data indicates for the first time the implication of chromatin tethering-independent mechanisms in the molecular mechanism of LEDGF/p75 in HIV-1 infection.

## References

1. Pauza, C.D., *Two bases are deleted from the termini of HIV-1 linear DNA during integrative recombination*. Virology, 1990. **179**(2): p. 886-9.
2. Miller, M.D., C.M. Farnet, and F.D. Bushman, *Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition*. J Virol, 1997. **71**(7): p. 5382-90.
3. Engelman, A., K. Mizuuchi, and R. Craigie, *HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer*. Cell, 1991. **67**(6): p. 1211-21.
4. Llano, M., et al., *An essential role for LEDGF/p75 in HIV integration*. Science, 2006. **314**(5798): p. 461-4.
5. Dalgleish, A.G., et al., *The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus*. Nature, 1984. **312**(5996): p. 763-7.
6. Deng, H., et al., *Identification of a major co-receptor for primary isolates of HIV-1*. Nature, 1996. **381**(6584): p. 661-6.
7. Feng, Y., et al., *HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor*. Science, 1996. **272**(5263): p. 872-7.
8. Starcich, B.R., et al., *Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS*. Cell, 1986. **45**(5): p. 637-48.
9. Pinter, A. and W.J. Honnen, *A sensitive radioimmunoprecipitation assay for human immunodeficiency virus (HIV)*. J Immunol Methods, 1988. **112**(2): p. 235-41.

10. Parren, P.W., et al., *Antibody protects macaques against vaginal challenge with a pathogenic R5 simian/human immunodeficiency virus at serum levels giving complete neutralization in vitro*. J Virol, 2001. **75**(17): p. 8340-7.
11. Huang, C.C., et al., *Structure of a V3-containing HIV-1 gp120 core*. Science, 2005. **310**(5750): p. 1025-8.
12. Wolk, T. and M. Schreiber, *N-Glycans in the gp120 V1/V2 domain of the HIV-1 strain NL4-3 are indispensable for viral infectivity and resistance against antibody neutralization*. Med Microbiol Immunol, 2006. **195**(3): p. 165-72.
13. Chan, D.C. and P.S. Kim, *HIV entry and its inhibition*. Cell, 1998. **93**(5): p. 681-4.
14. Johnson, W.E., J.M. Sauvron, and R.C. Desrosiers, *Conserved, N-linked carbohydrates of human immunodeficiency virus type 1 gp41 are largely dispensable for viral replication*. J Virol, 2001. **75**(23): p. 11426-36.
15. Smith, J.A., et al., *Evidence that the Nijmegen breakage syndrome protein, an early sensor of double-strand DNA breaks (DSB), is involved in HIV-1 post-integration repair by recruiting the ataxia telangiectasia-mutated kinase in a process similar to, but distinct from, cellular DSB repair*. Virol J, 2008. **5**: p. 11.
16. Craigie, R., T. Fujiwara, and F. Bushman, *The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro*. Cell, 1990. **62**(4): p. 829-37.
17. Daniel, R., et al., *Evidence that stable retroviral transduction and cell survival following DNA integration depend on components of the nonhomologous end joining repair pathway*. J Virol, 2004. **78**(16): p. 8573-81.

18. Daniel, R., R.A. Katz, and A.M. Skalka, *A role for DNA-PK in retroviral DNA integration*. Science, 1999. **284**(5414): p. 644-7.
19. Downs, J.A. and S.P. Jackson, *Involvement of DNA end-binding protein Ku in Ty element retrotransposition*. Mol Cell Biol, 1999. **19**(9): p. 6260-8.
20. Daniel, R., et al., *Histone H2AX is phosphorylated at sites of retroviral DNA integration but is dispensable for postintegration repair*. J Biol Chem, 2004. **279**(44): p. 45810-4.
21. Rehmann, B. and M. Nascimbeni, *Immunology of hepatitis B virus and hepatitis C virus infection*. Nat Rev Immunol, 2005. **5**(3): p. 215-29.
22. Letvin, N.L., *Progress and obstacles in the development of an AIDS vaccine*. Nat Rev Immunol, 2006. **6**(12): p. 930-9.
23. Balzarini, J., *Targeting the glycans of glycoproteins: a novel paradigm for antiviral therapy*. Nat Rev Microbiol, 2007. **5**(8): p. 583-97.
24. Scanlan, C.N., et al., *Exploiting the defensive sugars of HIV-1 for drug and vaccine design*. Nature, 2007. **446**(7139): p. 1038-45.
25. Schilling, R., et al., *A multivalent HIV-vaccine: development of a plasmid DNA for the expression of HIV envelope glycoproteins with hypervariable V3-loop domains*. Vaccine, 2006. **24**(21): p. 4648-50.
26. Balzarini, J., *Targeting the glycans of gp120: a novel approach aimed at the Achilles heel of HIV*. Lancet Infect Dis, 2005. **5**(11): p. 726-31.
27. Turville, S.G., et al., *Sugar-binding proteins potently inhibit dendritic cell human immunodeficiency virus type 1 (HIV-1) infection and dendritic-cell-directed HIV-1 transfer*. J Virol, 2005. **79**(21): p. 13519-27.

28. Wu, L. and V.N. KewalRamani, *Dendritic-cell interactions with HIV: infection and viral dissemination*. Nat Rev Immunol, 2006. **6**(11): p. 859-68.
29. Bolmstedt, A., et al., *Enhanced immunogenicity of a human immunodeficiency virus type 1 env DNA vaccine by manipulating N-glycosylation signals. Effects of elimination of the V3 N306 glycan*. Vaccine, 2001. **20**(3-4): p. 397-405.
30. Johnson, W.E., et al., *Assorted mutations in the envelope gene of simian immunodeficiency virus lead to loss of neutralization resistance against antibodies representing a broad spectrum of specificities*. J Virol, 2003. **77**(18): p. 9993-10003.
31. Ge, H., Y. Si, and R.G. Roeder, *Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation*. EMBO J, 1998. **17**(22): p. 6723-9.
32. Ge, H., Y. Si, and A.P. Wolffe, *A novel transcriptional coactivator, p52, functionally interacts with the essential splicing factor ASF/SF2*. Mol Cell, 1998. **2**(6): p. 751-9.
33. Cherepanov, P., et al., *HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells*. J Biol Chem, 2003. **278**(1): p. 372-81.
34. Engelman, A., *AIDS/HIV. A reversal of fortune in HIV-1 integration*. Science, 2007. **316**(5833): p. 1855-7.
35. Marshall, H.M., et al., *Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting*. PLoS ONE, 2007. **2**(12): p. e1340.
36. Schroder, A.R., et al., *HIV-1 integration in the human genome favors active genes and local hotspots*. Cell, 2002. **110**(4): p. 521-9.



37. Shun, M.C., et al., *LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration*. Genes Dev, 2007. **21**(14): p. 1767-78.
38. Cherepanov, P., et al., *Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase*. J Biol Chem, 2004. **279**(47): p. 48883-92.
39. Llano, M., et al., *Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75*. J Mol Biol, 2006. **360**(4): p. 760-73.
40. Llano, M., et al., *Lens epithelium-derived growth factor/p75 prevents proteasomal degradation of HIV-1 integrase*. J Biol Chem, 2004. **279**(53): p. 55570-7.
41. Ciuffi, A., et al., *A role for LEDGF/p75 in targeting HIV DNA integration*. Nat Med, 2005. **11**(12): p. 1287-9.
42. Vandekerckhove, L., et al., *Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus*. J Virol, 2006. **80**(4): p. 1886-96.
43. Turlure, F., et al., *A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo*. Nucleic Acids Res, 2006. **34**(5): p. 1653-75.
44. Shun, M.C., et al., *Identification and characterization of PWWP domain residues critical for LEDGF/p75 chromatin binding and human immunodeficiency virus type 1 infectivity*. J Virol, 2008. **82**(23): p. 11555-67.
45. Turlure, F., et al., *Human cell proteins and human immunodeficiency virus DNA integration*. Front Biosci, 2004. **9**: p. 3187-208.

46. Van Maele, B., et al., *Cellular co-factors of HIV-1 integration*. Trends Biochem Sci, 2006. **31**(2): p. 98-105.
47. Poeschla, E.M., *Integrase, LEDGF/p75 and HIV replication*. Cell Mol Life Sci, 2008.
48. Engelman, A. and P. Cherepanov, *The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication*. PLoS Pathog, 2008. **4**(3): p. e1000046.
49. Vanegas, M., et al., *Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering*. J Cell Sci, 2005. **118**(Pt 8): p. 1733-43.
50. Busschots, K., et al., *Identification of the LEDGF/p75 binding site in HIV-1 integrase*. J Mol Biol, 2007. **365**(5): p. 1480-92.
51. Llano, M., et al., *LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes*. J Virol, 2004. **78**(17): p. 9524-37.
52. Maertens, G., et al., *LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells*. J Biol Chem, 2003. **278**(35): p. 33528-39.
53. McKee, C.J., et al., *Dynamic Modulation of HIV-1 Integrase Structure and Function by Cellular Lens Epithelium-derived Growth Factor (LEDGF) Protein*. J Biol Chem, 2008. **283**(46): p. 31802-12.
54. Cherepanov, P., *LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro*. Nucleic Acids Res, 2007. **35**(1): p. 113-24.
55. Cherepanov, P., et al., *Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75*. Proc Natl Acad Sci U S A, 2005. **102**(48): p. 17308-13.

56. Hare, S., et al., *A novel co-crystal structure affords the design of gain-of-function lentiviral integrase mutants in the presence of modified PSIP1/LEDGF/p75*. PLoS Pathog, 2009. **5**(1): p. e1000259.
57. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0*. Bioinformatics, 2007. **23**(21): p. 2947-8.
58. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs*. Nucleic Acids Res, 1997. **25**(17): p. 3389-402.
59. Puntervoll, P., et al., *ELM server: A new resource for investigating short functional sites in modular eukaryotic proteins*. Nucleic Acids Res, 2003. **31**(13): p. 3625-30.
60. Pollastri, G., et al., *Accurate prediction of protein secondary structure and solvent accessibility by consensus combiners of sequence and structure information*. BMC Bioinformatics, 2007. **8**: p. 201.
61. Maertens, G., et al., *Identification and characterization of a functional nuclear localization signal in the HIV-1 integrase interactor LEDGF/p75*. J Biol Chem, 2004. **279**(32): p. 33421-9.
62. Beausoleil, S.A., et al., *Large-scale characterization of HeLa cell nuclear phosphoproteins*. Proc Natl Acad Sci U S A, 2004. **101**(33): p. 12130-5.
63. Shu, H., et al., *Identification of phosphoproteins and their phosphorylation sites in the WEHI-231 B lymphoma cell line*. Mol Cell Proteomics, 2004. **3**(3): p. 279-86.
64. Trinidad, J.C., et al., *Comprehensive identification of phosphorylation sites in postsynaptic density preparations*. Mol Cell Proteomics, 2006. **5**(5): p. 914-22.

65. Molina, H., et al., *Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry*. Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2199-204.
66. Villen, J., et al., *Large-scale phosphorylation analysis of mouse liver*. Proc Natl Acad Sci U S A, 2007. **104**(5): p. 1488-93.
67. Dephoure, N., et al., *A quantitative atlas of mitotic phosphorylation*. Proc Natl Acad Sci U S A, 2008. **105**(31): p. 10762-7.
68. Kim, J.E., S.R. Tannenbaum, and F.M. White, *Global phosphoproteome of HT-29 human colon adenocarcinoma cells*. J Proteome Res, 2005. **4**(4): p. 1339-46.
69. Yang, F., et al., *Phosphoproteome profiling of human skin fibroblast cells in response to low- and high-dose irradiation*. J Proteome Res, 2006. **5**(5): p. 1252-60.
70. Collins, M.O., et al., *Phosphoproteomic analysis of the mouse brain cytosol reveals a predominance of protein phosphorylation in regions of intrinsic sequence disorder*. Mol Cell Proteomics, 2008. **7**(7): p. 1331-48.
71. Han, G., et al., *Large-scale phosphoproteome analysis of human liver tissue by enrichment and fractionation of phosphopeptides with strong anion exchange chromatography*. Proteomics, 2008. **8**(7): p. 1346-61.
72. Meggio, F. and L.A. Pinna, *One-thousand-and-one substrates of protein kinase CK2?* Faseb J, 2003. **17**(3): p. 349-68.
73. Trinidad, J.C., et al., *Quantitative analysis of synaptic phosphorylation and protein expression*. Mol Cell Proteomics, 2008. **7**(4): p. 684-96.

74. Semplici, F., et al., *CK2-dependent phosphorylation of the E2 ubiquitin conjugating enzyme UBC3B induces its interaction with beta-TrCP and enhances beta-catenin degradation*. *Oncogene*, 2002. **21**(25): p. 3978-87.
75. Arriagada, G., et al., *Phosphorylation at serine 208 of the 1alpha,25-dihydroxy Vitamin D3 receptor modulates the interaction with transcriptional coactivators*. *J Steroid Biochem Mol Biol*, 2007. **103**(3-5): p. 425-9.
76. Abbott, K.L., et al., *Enhanced binding of RNAP II CTD phosphatase FCP1 to RAP74 following CK2 phosphorylation*. *Biochemistry*, 2005. **44**(8): p. 2732-45.
77. Genovese, N.J., et al., *Casein kinase II motif-dependent phosphorylation of human papillomavirus E7 protein promotes p130 degradation and S-phase induction in differentiated human keratinocytes*. *J Virol*, 2008. **82**(10): p. 4862-73.
78. Bartholomeeusen, K., et al., *Lens epithelium-derived growth factor/p75 interacts with the transposase-derived DDE domain of PogZ*. *J Biol Chem*, 2009. **284**(17): p. 11467-77.
79. Yokoyama, A. and M.L. Cleary, *Menin critically links MLL proteins with LEDGF on cancer-associated target genes*. *Cancer Cell*, 2008. **14**(1): p. 36-46.
80. Bartholomeeusen, K., et al., *Differential interaction of HIV-1 integrase and JPO2 with the C terminus of LEDGF/p75*. *J Mol Biol*, 2007. **372**(2): p. 407-21.
81. Hughes, S., et al., *Transcriptional co-activator LEDGF interacts with Cdc7-activator of S-phase kinase (ASK) and stimulates its enzymatic activity*. *J Biol Chem*. **285**(1): p. 541-54.
82. Orphanides, G., et al., *FACT, a factor that facilitates transcript elongation through nucleosomes*. *Cell*, 1998. **92**(1): p. 105-16.

83. Orphanides, G., et al., *The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins*. Nature, 1999. **400**(6741): p. 284-8.
84. Koues, O.I., et al., *Roles for common MLL/COMPASS subunits and the 19S proteasome in regulating CIITA pIV and MHC class II gene expression and promoter methylation*. Epigenetics Chromatin. **3**(1): p. 5.
85. Krogan, N.J., et al., *COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression*. J Biol Chem, 2002. **277**(13): p. 10753-5.

## Appendix

### List of Abbreviations

ATR	Ataxia talangiectasia-mutated kinase
AIDS	Acquired immune deficiency syndrome
CBD	Chromatin binding domain
COMPASS	Complex Proteins Associated with Set 1
CR	Charged region
DAPI	4',6-diamidino-2-phenylindole
DC	Dendritic cells
DC-SIGN	Dendritic cell-specific ICAM-3 grabbing non-integrin
DMEM	Dulbecco's modified Eagle's medium
DSB	Double-stranded breakage
FACT	Facilitates Chromatin Transcription
gp120	glycoprotein 120
HIV-1	Human immunodeficiency virus-1
IBD	Integrase binding domain
IN	HIV-1 integrase
LEDGF	Lens-epithelium-derived growth factor
LTR	Long terminal repeat
mAb	Monoclonal antibody
NBS1	Nijmegen breakage syndrome-1 protein
NHEJ	Non homologous end joining
PBS	Phosphate-buffered saline

PIC	Preintegration complex
PIR	Post-integration repair
PWWP	Pro-Trp-Trp-Pro motif
shRNA	Short hairpin RNA
Spt16	Suppressor of Ty 16 homolog
SSRP1	Structure Specific Recognition Protein 1
TBS	Tris-buffered saline
WT	Wild type



### List of Publications and Manuscripts

1. **Garcia-Rivera JA**, Bueno M, Morales E, Kuggelman JR, Rodriguez DF, Llano M (2009) Implication of Serine Residues 271, 273 and 275 in the HIV-1 Cofactor Activity of LEDGF/p75. *J Virol*. 2010 Jan;84(2):740-52.
2. Meehan AM, Saenz DT, Morrison JH, **Garcia-Rivera JA**, Peretz M, Llano M, Poeschla EM (2009) LEDGF/p75 Proteins with Alternative Chromatin Tethers Are Functional HIV-1 Cofactors. *PLoS Pathog*. 2009 Jul;5(7).
3. Bueno M, **Garcia-Rivera JA**, Morales E, Kuggelman JR, Rosas-Acosta G, Llano M (2010) SUMOylation of the Lens Epithelium-derived Growth Factor/p75 attenuates its transcriptional activity on the Heat Shock Protein 27 promoter. *J Mol Biol*. 2010 Apr
4. Astiazaran P, Morales E, Bueno M, Kuggelman JR, **Garcia-Rivera JA**, Llano M (2010) Integrase Modulates the Interaction of LEDGF/p75 and Chromatin. *Retrovirology*. Submitted.
5. Kuggelman JR, **Garcia-Rivera JA**, Bueno M, Astiazaran P, Reyes D, Fernandez-Zapico M, Llano M (2010) The Lens Epithelium-Derived Growth Factor p75 is a Member of the Transcription Elongation Complex. In Process.

## Curriculum Vitae

Jose A. Garcia-Rivera was born on October 28, 1984 in Santurce, Puerto Rico. The oldest of two children of Jose Garcia-Ortiz and Madelene Garcia-Rivera, he received his Bachelor of Science in Industrial Chemistry from the Inter American University of Puerto Rico in 2006. While pursuing his bachelor's degree, he worked in the Applied Biotechnology Research Laboratory at the Inter American University of Puerto Rico where his research involved the characterization of the 5' upstream region of the *Plasmodium yoelii* breast cancer resistant protein gene in three different drug resistant plasmids.

In Fall 2006, he joined the doctoral program in Biological Sciences at the University of Texas at El Paso (UTEP), El Paso, Texas where he has been working under the supervision of Dr. Manuel Llano. His research focuses on the understanding of the molecular mechanism of LEDGF/p75 in regard to HIV-1 integration. He was awarded the RISE Scholars Fellowship in 2006 and has worked as a research assistant ever since.

He has participated in numerous conferences including the Annual Biomedical Research Conference for Minority Students in Anaheim, CA (2006), the SACNAS Conference in both Kansas City, MS and Dallas, TX (2007, 2009), the 3<sup>rd</sup> International Retroviral Integrase Conference in Woods Hole, MA (2008), the ASM Conference in Las Cruces, NM (2008), the 3<sup>rd</sup> Annual Research Colloquium at Texas Tech University, El Paso, TX (2009), the St. Jude National Graduate Student Symposium in Memphis, TN (2010), and the Cold Spring Harbor Laboratory-Retroviruses in Cold Spring Harbor, NY (2010).

In addition he has been the recipient of several awards such as the RISE and AGEP Scholars Fellowships (2006), the SACNAS Travel Scholarship (2007, 2009), and the UTEP Graduate School Grant (2009, 2010).

He has already published one first author peer-reviewed paper as well as two co-authored peer-reviewed papers and will soon publish at least two more articles as co-author. After graduation, he will work as a post-doctoral fellow at the Scripps Research Institute in San Diego, California. His long-term career goal is to become an independent investigator who studies protein-protein interactions between viral and cellular factors.

[illegible]

```

-
*
CR2
*
20
40
60
*
HS : NVKQPCPESDIITEEDKSKKKGQEEKQPKKQP-KKDEEGQKEBDKPRKEPDKKEGKKKEVESKRNLAKT : 69
CF : NVKQPCPSETDMVTEEDKSKKKGQEEKQPKKQI-KKDEEGQKEBDKPRKEPDKKEGKKKEVESKRNLEKT : 69
FC : NVKQPCPSESDMVTEDKSKKKGQEEKQPKKQI-KKDEEGQKEBDKPRKEPDKKEGKKKEVESKRNLSAKT : 69
BV : NVKQPCPSESDMITEEDKSKKKGQEEKQPKKQI-KKDEEGQKEBDKPRKEPDKKEGKKKEVESKRNLAKT : 69
EC : NVKQPCPSESDMITEEDKSKKKGQEEKQPKKQI-KKDEEGQKEBDKPRKEPDKKEGKKKEVESKRNLAKT : 69
MC : NVKQPCPSESDIITEEDKSKKKGQEEKQPKKQI-KKDEEGQKEBDKPRKEPDKKEGKKKEVESKRNLAKT : 69
MM : NVKQPCPBDGMVIDEDKSKKKGQEEKQPKKQI-KKDEEGQKEBDKPRKEPDKKEGKKKEVESKRNLAKT : 69
RN : NVKQPCPSESDMVIDEDKSKKKGQEEKQPKKQI-KKDEEGQKEBDKPRKEPDKKEGKKKEVESKRNLAKT : 69
MD : NVKPCPBSDNIVTEEDKSKKKGQEEKQPKKQI-KKDEEGQKEBDKPRKEPDKKEGKKKEVESKRNLAKT : 69
GG : LVRPSCISESLPVNEEAKKKGQDEKQPKKQI-KKDEEGQKEBDKPRKEPDKKEGKKKEVESKRNLAKT : 68
XL : LIRPSPBERGMTNFEEDKSKKKGQEEKQPKKQI-KKDEEGQKEBDKPRKEPDKKEGKKKEVESKRNLAKT : 70
OA : NVKPCPSPSDNIALFEEDKSKKKGQEEKQPKKQI-KKDEEGQKEBDKPRKEPDKKEGKKKEVESKRNLAKT : 68
TN : ---QSGGSDMTS-ESDRKRKRRAQEDKRYKME-----GENKKEEGTMKD---GDGR-EPFAKRRKQSR : 56
SS : ---QSGGSDMTA-ESERKRKRRTPEEDKSKKGEEFKRLDGRKRKEDNKGRE---LEGKKDPDAKRRRQTK : 63
k pc s d eed 4k4g eek p k k ee qkei kp 4e kkeg ke e k 4 a4

```

```

80
*
CR3
100
120
*
HS : GVTSTSDS-----EEEGDDQEGEKKRRKGRNFEQTAAHRRNMILKGQHEKEAADRKRKQEEQMETE : 127
CF : GVTSTSDS-----EEEGDDQSESEKKRRKGRNFEQTAAHRRNMILKGQHEKEAADRKRKQEEQMETE : 127
FC : GVTSTSDS-----EEEGDDQSESEKKRRKGRNFEQTAAHRRNMILKGQHEKEAADRKRKQEEQMETE : 127
BV : GVTSTSDS-----EEEGDDQEGEKKRRKGRNFEQTAAHRRNMILKGQHEKEAADRKRKQEEQMETE : 127
EC : GVTSTSDS-----EEEGDDQEGEKKRRKGRNFEQTAAHRRNMILKGQHEKEAADRKRKQEEQMETE : 127
MC : GVTSTSDS-----EEEGDDQEGEKKRRKGRNFEQTAAHRRNMILKGQHEKEAADRKRKQEEQMETE : 127
MM : GVTSTSDS-----EED-DQEGEKKRRKGRNFEQAARRNMILKGQHEKEAADRKRKQEEQMETE : 126
RN : GVTSTSDS-----EED-DQEGEKKRRKGRNFEQAARRNMILKGQHEKEAADRKRKQEEQMETE : 126
MD : GFASTSDS-----EEEGDDQEGEKKRRKGRNFEQAARRNMILKGQHEKEAADRKRKQEEQMETE : 127
GG : GSA$ASDS-----EDEGEHEEGDKKKKGRNFEQSTHRRNIMRGQHEKEVTERKRRKQEEQGESE : 126
XL : DSPAQSDS-----DDDEQEDDEKKS$KSGKLSQVAQKR-SLKBP$EKDNSDKRKQEEQTVDM : 125
OA : SSTSTSDSEEEEEEEEEEDEEDQDEEKKKKGRNTEQATHRRNLLKAQHEKEVAERKRRKQEEQMEAB : 134
TN : --DTSSDS-----EEDDDDD-----KRQQPTPTETGDKDDRQRRK--QDEQK : 88
SS : --ESSSDS-----EEDDDDD-----KRQQPTPTETGDKDDRQRRK--QDEQK : 107
SDS e dze eek kgy fq h rn qheke d kr qeeg e e

```



CR4 \*

HS :	CCNKDEGKKPEVKKVKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	1
CF :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	97
FC :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	97
BV :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	97
EC :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	97
MC :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	97
MM :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	97
RN :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	97
MD :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	98
GG :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	97
XL :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	97
OA :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	97
TN :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	96
SS :	CCNKDEGKKPEVKKVEKKR	RETSMDSRLQRIHAEIKNSLKTIDNDV	RCIEALDEIASLGVTNCOQCKHTEMITTLTKK1	RRFKVSCQIMEKSTMLYNK	97

QnKde kK ev4k eKK4Etmsds4LQr6haEIknsLkIdnDvn4C6ealDE6a lQv3mqaqCKH3e6ItTLK6 R4FKvsQ ImekstMLYNK

HS :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	180	*
CF :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	192	
FC :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	192	
BV :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	192	
EC :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	192	
MC :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	192	
MM :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	192	
RN :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	192	
MD :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	193	
GG :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	193	
XL :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	190	
OA :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	189	
TN :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	130	
SS :	EKNMFLVGEGETVITQVLANKSLAEQRQHEEANKTKDQGKKGPNKK1LEKX	QTSKTIINGGSDAQDNGVOPONHNGESNEDSKINHFASTKKR-PSSFEER	181	

EKNMFL6GEGDs663QVLANKS6AEQ4QhEEAnk k gK4Gpn k e e tg k nggs aq n qhng s e kd e k ee

**Figure S1.** LEDGF/p75 protein sequences alignment. Sequences from the NCBI protein database were aligned using ClustalW2. Hs, *Homo sapiens* (NM\_033222.3); Cf, *Canis familiaris* (XP\_531939); Fc, *Felis catus* (AAU10509); Bt, *Bos taurus* (AAM90841.1); Ec, *Equus caballus* (ABH11544.1); Mc, *Macaca mulatta* (ABD77428); Mm, *Mus musculus* (Q99JF8); Rn, *Rattus norvegicus* (NP\_786941); Md, *Monodelphis domestica* (XP\_001366101); Gg, *Gallus gallus* (NP\_001026781.2); Xl, *Xenopus laevis* (NP\_001089191); Oa, *Ornithorhynchus anatinus* (XP\_001507936.1); Tn, *Tetraodon nigroviridis* (CAG07956) and Ss, *Salmo salar* (ACI34089).

<b>Mutant</b>	<b>Deletion</b>	<b>Primers</b>
$\Delta$ PWWP	1-93	TTCTCGAGGGATCCCCGGGTACCGAGCTC ATCGATGAGTCAACAGGCAGCAACTAAAC
$\Delta$ ATH	178-198	ACTCACTTTTAGATTAACAGATGC ATGGTAAACAGCCCTGTCCTTCAG
$\Delta$ PWWP/ $\Delta$ ATH	See above	See primers used for single deletion mutants
$\Delta$ IBD	340-442	TTTCTTAACTTCTGGCTTCTTTCC TCTCTTGCTGAACAACGGCAACACG
$\Delta$ CR1	94-145	TGAAAATTTCACTTTTGGATTGTTATC AGAAGGGGGAGAAAGAGAAAGGCAG
$\Delta$ CR2	199-266	TTTGGGTCTGCCTCTTGGTTTTGG ACAGGGGTTACTTCAACCTCCGATTC
$\Delta$ CR3	267-325	TTTAGCTAAATTTTCCTTTTGGATTC CAGCAGAATAAAGATGAAGGAAAGAAG
$\Delta$ CR4	326-339	CTCAGTTTCCATTTGTTCTTCTTGCTTG GTGGAGAAGAAGCGAGAAACATCAATG



ΔCR5	443-530	TGTAGACTTTTCCATGATTACCTGAC GGGCCCCGACTACAAAGACGATGACG
ΔCR3.1	267-281	TTTAGCTAAATTTTTCCTTTTGTATC CAAGAA GGTGAAAAGAAGAGAAAAGG
ΔCR3.2	282-325	ATCATCTCCTTCTTCTTCAGAATCGG CAGCAGAATAAAGATGAAGGAAAGAAG
ΔCR3.3	303-325	CATATTCCTTCTGTGAGCAGTCTGAAAG CAGCAGAATAAAGATGAAGGAAAGAAG
S271A/S273A/S275A	None	TTTAGCTAAATTTTTCCTTTTGTATC ACAGGGGTTACTGCAACCGCCGATGCTGAAG
S271A	None	TTTAGCTAAATTTTTCCTTTTGTATC ACAGGGGTTACTGCAACCTCCGATTCTGAAG
S273A	None	TTTAGCTAAATTTTTCCTTTTGTATC ACAGGGGTTACTTCAACCGCCGATTCTGAAG
S275A	None	TTTAGCTAAATTTTTCCTTTTGTATC ACAGGGGTTACTTCAACCTCCGATGCTGAAG

**Table 2.** *LEDGF/p75 mutants*. Mutants were generated by PCR with the listed primers. The amino acids deleted are indicated. Primer sequences are written from the 5' end to the 3' end. For each mutant, the reverse primer is listed first followed by the forward primer.