Chromatin function regulates retroviral gene expression

Ph.D. Thesis in Molecular Biology

Supervisor: Professor Mauro Giacca

Chiara Vardabasso

Scuola Normale Superiore

TABLE OF CONTENTS

SYNOPSIS

The experimental work described in this Thesis has been mainly focused on the study of the regulation of viral gene expression, in two different members of the *Retroviridae* family, namely the lentivirus HIV-1 and the gammaretrovirus MoMLV.
The results are divided into two parts, Part A and Part B.

The work reported in **Part A** gives critical clues for the comprehension of the still obscure mechanisms that regulate function of the HIV-1 transactivator Tat in the initiation and elongation of HIV-1 transcription. In particular, we exploited a proteomic screening aimed at characterizing novel Tat partners. This approach allowed us to identify the cellular histone chaperone NAP-1 (Nucleosome Assembly Protein-1) as a Tat-binding protein. We found that this interaction exerts a positive role on Tat-driven LTR transactivation and on HIV-1 infection. To our knowledge, this is the first demonstration of an interaction between Tat and a cellular histone chaperone. We propose a mechanism by which Tat benefits from this class of proteins to relieve the repression imposed by chromatin conformation on proviral transcription.

The findings described in **Part B** are part of an ongoing project, that we derived from our previous observation that HIV-1 transcription is also regulated by long-range chromatin interactions. In particular, we originally found that a gene loop structure is imposed on the provirus upon transcriptional activation. We now show that MoMLV also adopts a transcription-dependent LTR-LTR gene loop conformation. This observation leads us to hypothesize that gene looping might be a general hallmark of

retroviral transcription. Moreover, we also demonstrate that an aberrant loop might form between the retroviral MoMLV LTRs and regulatory regions of the host cell genome. We suggest that this event might be at the basis of the phenomenon of insertional mutagenesis observed in some of the gene therapy clinical trials that so far have exploited members of the *gammaretroviridae* family of viruses.

PART A

1A. INTRODUCTION

1.1 REPLICATION CYCLE OF HIV-1: AN OVERVIEW

All viruses belonging to the *Retroviridae* family are characterized by an RNA genome that is converted to a cDNA intermediate during the viral life cycle. The reverse transcription of the viral RNA is the defining hallmark of the retroviruses, the step from which these viruses owe their name. The genome contains three essential genes common to all the members of the family, *gag*, *pol* and *env*, encoding for both structural and enzymatic proteins required for the whole viral life cycle. Seven genera of *Retroviridae* family have been defined on the basis of genome structure and nucleotide sequence relationships.

The Human Immunodeficiency Virus type 1 (HIV-1) belongs to the Lentivirus genus. Since the isolation and the identification of HIV-1 as the etiologic agent of the Acquired Immuno-Deficiency Syndrome in the early '80s (Barre-Sinoussi et al., 1983; Popovic et al., 1984), a remarkable body of knowledge has accumulated about the mechanisms of retroviral infections and pathogenesis of the retrovirus-induced diseases. Nothwistanding this progress, HIV-1 still remains a major threat to public health and a challenge for drug and vaccine development, due to several reasons related to the biology of the cells that are the target of HIV-1 and to the life cycle of the virus itself.

The viral genome consists of two RNA molecules, each approximately 10 kb-long (Alizon et al., 1984), harboring, besides the *gag*, *pol* and *env* genes, a number of genes that are specific for HIV-1: the two regulatory genes *tat* and *rev*, and four other accessory genes, namely *vif*, *vpr*, *vpu* and *nef*.

The *gag*, *pol* and *env* gene products are synthesized as single polyproteins and later processed, the first two by the viral protease and the last by a cellular protease. The *gag* gene encodes for the structural core proteins (capsid, nucleocapsid) and the matrix protein of the viral particle; the *env* gene encodes for the glycoproteins gp120 and gp41, that are embedded in the phospholipidic viral envelope; the *pol* gene encodes for the enzymes Reverse Transcriptase, Integrase and Protease, crucial for viral replication. The auxiliary genes render HIV-1 a very complex and sophisticated virus, allowing it to exhibit a vast number of regulatory pathways and remarkable mechanisms for viral persistence (Freed and Martin, Fields Virology 2001).

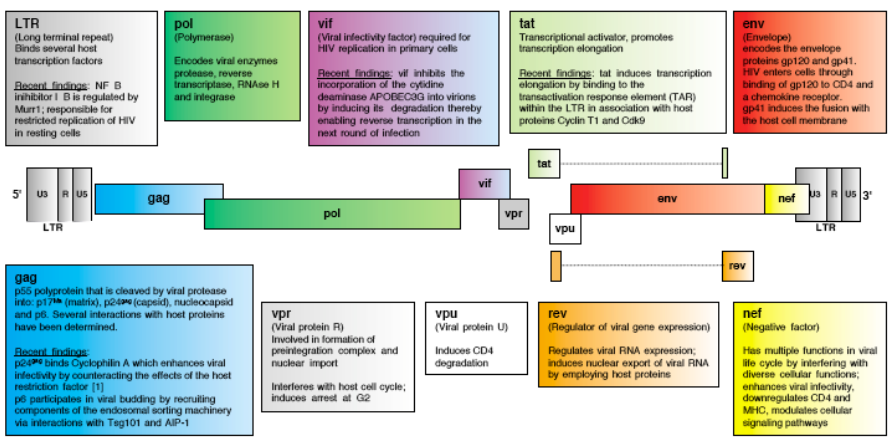


Figure 1.1 Structure of the HIV-1 genome and functions of its encoded proteins (Adapted from Trkola, 2004).

In recent years, the interplay between HIV-1 and its host cells has been extensively studied, unraveling the importance of several host factors for different steps in the viral life cycle. At the same time, strategies adopted by host cells to counteract the infection have been described, contributing to the view that the manifestation of a productive infection is the result of the balance of both pro-viral and anti-viral mechanisms.

1.1.1 A road to the nucleus: entry, uncoating and reverse transcription

HIV-1 infects primarily T helper lymphocytes, macrophages and, to some extent, microglial and dendritic cells. Entry into the host cells requires the expression of the CD4 receptor and fusion co-receptors, that are strain- and cell type specific (Kwong et al., 1998). Several chemokine receptors function as co-receptors for HIV-1, the best charachterized of which are the CC-chemokine receptor 5 (CCR5) and the CXC-chemokine receptor 4 (CXCR4), expressed by T cells, mononuclear phagocytes and dendritic cells. Based on the ability to bind these molecules, HIV-1 can be divided into different strains: R5 strain (CCR5), X4 strain (CXCR4) or R5X4 (CCR5 and CXCR4). While R5 strain can be isolated throughout the natural course of human infection, X4 (which, of note, are ony 1-2% of the dualtropic X4R5) and X4R5 emerge later during disease progression (Doms and Trono, 2000). Interestingly, and relevant for the mucosal transmission of the infection, the lectin-like receptor DC-SIGN (Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin) was shown to be involved in the viral uptake from mucosal and skin dendritic cells, that can then vehicle the virus to T cells in the lymphnodes (Geijtenbeek et al., 2000).

The conformational change induced by the first interaction between gp120 and CD4 promotes the binding to the co-receptor, eventually leading to viral entry by direct fusion with the plasma membrane (mediated by gp41) (Chan and Kim, 1998) or by endocytosis followed by glycoprotein- and dynamin-dependent fusion with intracellular compartments (Miyauchi et al., 2009).

The process leading from viral entry to uncoating is not completely understood, and apparently involves viral proteins, including p17 Matrix, Nef, Vif, as well as host proteins (Dvorin and Malim, 2003). Upon

completion of uncoating, the reverse transcription complex is formed, composed of viral RNA, Reverse Transcriptase, Integrase, p17 Matrix, P7 Nucleocapsid, Vpr and many cellular proteins. The course of reverse transcription is highly ordered and quite well-established, starts in the cytoplasm being completed in the nucleus, and involves the translocations of DNA intermediates resulting in the duplication of identical sequences (U3, R, U5) at both 5' and 3' ends of the final cDNA, the so-called "Long Terminal Repeats" (LTRs) (Gilboa et al., 1979). Products of reverse transcription are actively transported into the nucleus through the nuclear pores, consistent with the ability of the virus to also infect non-dividing cells, which is a hallmark of all lentiviruses (Bukrinsky, 2004). Double-stranded viral DNAs are complexed with both viral (Integrase, Nucleocapsid, Matrix, Reverse Transcriptase and Vpr) and cellular proteins, in the Pre-Integration Complex (PIC). The molecular processes underlining the PIC entry into the nucleus remain elusive, nevertheless it seems reasonable that a redundancy of mechanisms exists, at least in part mediated by the Nuclear Localization Signals (NLS) that have been identified in most of the viral proteins of the PIC. Cellular protein partners, namely BAF (Barrier-to-Autointegration Factor), LEDGF (Lens Epithelium-Derived Growth Factor)/p75, LAP2α (Lamin-Associated Polypeptide 2α), importins, Ku helicase, HMG (High Mobility Group) proteins, TNPO3 (Transportin-SR2) (Christ et al., 2008) are also involved in nuclear entry of the PIC and viral integration (for a comprehensive review, see (Suzuki and Craigie, 2007) and references therein).

These early steps of viral replication appear to be the main targets for those host proteins, generally termed as "restriction factors", that elicit an innate anti-viral response. Among these, TRIM5α was identified as the factor that blocks HIV-1 infection in simian cells (Stremlau et al., 2004).

Since its discovery, many studies have tried to address the molecular mechanisms by which TRIM5α restricts HIV-1. It has been suggested that the blocks are at reverse transcription and nuclear import steps, nevertheless the anti-viral activity of TRIM5α still remains poorly understood. Interestingly, TRIM5α action was found to be dependent on the peptidyl-prolyl isomerase Cyclophilin A (and even a TRIM5-Cyp A fusion gene was cloned, created by retrotransposition of a CypA cDNA into the TRIM5 locus, (Sayah et al., 2004). Cyclophilin A has previously been reported to bind and to induce a conformational change of the HIV-1 capsid protein, thus affecting its interactions with uncoating factors, with nuclear pore components as well as with restriction factors such as TRIM5α (for comprehensive reviews, see (Luban, 2007; Towers, 2007).

Another mechanism that can suppress HIV-1 replication is mediated by the apolipoprotein B mRNA-editing enzyme 3 family (APOBEC3) (Sheehy et al., 2002). These are cytidine deaminases that hypermutate the viral genomes (although a deaminase-independent inhibition by APOBEC was also reported (Nguyen et al., 2007; Sasada et al., 2005), causing replication defects at multiple steps. In particular, APOBEC3G and APOBEC3F are those responsible for the pre-integration complex restriction.. HIV-1 Vif protein counteracts incorporation of APOBEC3G into viral particles mainly by recruiting an E3 ubiquitin ligase complex that targets APOBEC3G to proteasomal degradation (recently reviewed in (Goila-Gaur and Strebel, 2008).

1.1.2 HIV-1 in the nucleus: integration and transcription

Once inside the nucleus, integration of the double-stranded viral DNA into the cellular genome is mediated by the Integrase protein, which binds the end of the viral DNA and catalyzes the subsequent joining reaction that establishes the HIV provirus within the host chromosome (Miller et al.,

1997). To accomplish its functions, both the tethering to chromatin and the completion of its enzymatic activity, Integrase requires several host factors, most of them already included in the PIC, such as LEDGF/p75 (Maertens et al., 2003), emerin (Jacque and Stevenson, 2006; Shun et al., 2007), BAF (Lin and Engelman, 2003), INtegrase Interactor 1 (INI1) (Kalpana et al., 1994), HMGA1 (Farnet and Bushman, 1997). Interestingly, Integrase was also shown to be subject to different Post-Translational Modifications (PTMs), that variously affect its properties: acetylation mediated by the cellular Histone Acetyl-Transferase (HAT) p300 increases its affinity for DNA template (Cereseto et al., 2005), while ubiquitination reduces the stability of the protein (Mousnier et al., 2007); moreover, ubiquitination is prevented by phosphorylation and subsequent isomerization (Manganaro et al., submitted paper).

Alternatively, if the viral DNA does not get integrated, it may follow three different fates: the ends of the viral DNA might be joined to form a 2-LTR circle; the viral genome might undergo homologous recombination resulting in a single LTR circle; viral DNA might auto-integrate into itself producing a rearranged circular structure. Notwithstanding that none of these variants produce infectious virus, they seem to be transcriptionally active, leading to selected transcription of *tat* and *nef* genes before integration (Wu and Marsh, 2001). Finally, it is worth mentioning that generation of these forms of circular DNA before integration, that can persist inside the cells, are related to the phenomenon referred to as pre-integration latency (Coiras et al., 2009; Lassen et al., 2004); (Zack et al., 1990).

Although HIV-1 integration is not site-specific, several lines of evidence indicated that it is not completely random. Early *in vitro* studies using either purified Integrase protein or PICs isolated from freshly infected cells, revealed that not only the primary sequence of the acceptor DNA, but also

its conformation altered by protein binding, might at least in part dictate the preference for certain sites of integration (Bor et al., 1995; Bor et al., 1996). However, major insights into the study of retroviral integration, derived from the *in vivo* studies, which took advantage of Linker-Mediated PCR to amplify and then sequence hundreds of junctions between viral and human DNA in cultured cells upon acute infection with retroviruses or retroviral vectors. First, Carteau and collegues showed that in an acutely infected T cell line (SupT1), HIV-1 strongly disfavored centromeric alphoid repeats (Carteau et al., 1998), although it has been later demonstrated that integrations into heterochromatic regions can occur, and might be responsible of post-integration latency (Lewinski et al., 2005). Several other high-throughput studies, exploiting a broad range of cellular models, such as different T- and non T-cell lines, primary lymphocytes and resting $CD4^+$ cells from infected individuals, revealed that HIV-1 preferentially integrates within active transcriptional units. This might turn to be an advantage for the virus, which can thus better exploit the cellular transcriptional machinery for the transcription of its own genes (Han et al., 2004; Huang et al., 2007; Mitchell et al., 2004; Schroder et al., 2002).

Interestingly, comparative analysis of integration sites of HIV-1 and other oncoretroviruses demonstrated that the mechanisms for integration site selection are peculiar for each retrovirus. In particular, Moloney Murine Leukemia Virus (MoMLV) integrates preferentially near the start site of transcriptionally active genes and near CpG islands, and targets Common Integration Sites (CISs) such as proto-oncogenes and growth-controlling genes, in a number of cellular models (Bushman et al., 2005; Cattoglio et al., 2007; Mitchell et al., 2004; Recchia et al., 2006; Wu et al., 2003).

Although the scenario is not still completely defined, and the viral determinants of the selection still remain elusive (even if a crucial role in tethering to chromatin has been ascribed to LEDGF/p75, (Maertens et al.,

2003)), what is clearly emerging is that the selection of sites for integration is strongly dependent on chromatin structure. Interestingly, the chromatin milieu encountered by the provirus also influences the levels of basal HIV-1 gene expression, although this effect is by-passed upon Tat transactivation (Jordan et al., 2003; Jordan et al., 2001). Thus, chromatin environment at the site of integration might be one of the possible mechanisms underlying the establishment of post-integration latency, which refers to the lack of expression after the insertion of viral DNA into the host genome. In particular, one form of latency might result from proviral integration into regions of heterochromatin, where transcription is repressed (Coiras et al., 2009; Han et al., 2004; Marcello, 2006; Williams and Greene, 2007).

Once integrated into the host genome, the provirus acts as any human RNA Pol II-transcribed gene, with transcription proceeding from the promoter and enhancer elements present in the 5' LTR, to the polyadenylation site comprised within the 3' LTR. Regulation of HIV-1 gene expression involves a complex interplay between the local chromatin environment, cellular transcription factors, and the viral-encoded trans-activating protein Tat. Early phases of transcription are mostly mediated by host transcription factors bound to the LTR, and allow the production of Tat, which in turn promotes an efficient initiation and a processive elongation of viral transcription. To fulfill its functions, Tat binds a hairpin structure present at the 5' end of the nascent viral RNA, named Trans-Activation-Responsive region (TAR), and establishes interactions with a vast amount of cellular partners, that eventually determine a stimulation of transcription.
The mechanisms that underlie the control of HIV-1 gene expression are described in details in the Section 1.2 of this Thesis.

Tat-mediated transcription leads to the generation of approximately thirty different viral transcripts from the provirus; 2 kb-long multiply spliced transcripts are predominant in the early phases of infection (encoding for Tat, Rev and Nef), while singly spliced RNAs of 4 kbs, and unspliced RNAs of 9 kbs become leading species later in infected cells. The first short transcripts are rapidly transported into the cytoplasm following the same pathway as cellular mRNAs (Cullen, 1998), while export to the cytoplasm of unspliced and partially spliced depends on the expression of the virally-encoded Rev protein. This protein shuttles between nucleus and cytoplasm and binds the viral transcripts through the interaction with an RNA stem-loop structure named RRE (Rev Responsive Element) located in the *env* gene (Malim et al., 1990). Rev binding to RRE determines the multimerization of the protein, and the following association with the cellular CRM-1 and Ran GTPase proteins. The nucleo-cytoplasmic gradient of GTP/GDP determines the translocation of the transcripts; indeed, the Rev/Crm-1/Ran-GTP complex, associated with viral RNA in the nucleus, interacts with the nuclear pore complex thus allowing nuclear export. Once in the cytoplasm, hydrolysis of GTP to GDP occurs, and causes the dissociation of the transcript, while Rev returns into the nucleus by binding importin and Ran-GDP, for subsequent round of export (Cullen, 1998; Suhasini and Reddy, 2009).

1.1.3 Late phases of HIV-1 infection: assembly, budding and maturation

The late phases of HIV-1 life cycle are characterized by the assembly of new viral particles, their release from the plasma membrane and their maturation. HIV-1 assembly and release are triggered by the viral Gag precursor p55, from which the structural components of the viral core derive. Following its synthesis, p55 traffics to the host cell plasma

membrane, and the pre-assembled virion buds from the lipid rafts, specific membrane sub-domains enriched in cholesterol. Interestingly, this is the same compartment in which the myristoylated form of the Nef protein is incorporated, being this a prerequisite for the biological activity of the protein itself (Geyer et al., 2001). Furthermore, it is worth mentioning the exception of primary macrophages, where HIV-1 buds into large intracellular vacuoles proposed to be late endosomes. In this way, HIV-1 can be retained in an infectious state for prolonged period of times, being released in a delayed manner similarly to secretion of exosomes, and thus being important for pathogenesis (Morita and Sundquist, 2004).

Several lines of evidence have suggested that HIV-1 requires different host factors to orchestrate its assembly and budding, such as components of the Endosomal Sorting Complex Required for Transport (ESCRT) (Garrus et al., 2001; Martin-Serrano et al., 2001) and the apoptosis-linked-gene 2 interacting protein (Alix) (Strack et al., 2003; von Schwedler et al., 2003). It is worth mentioning that cellular restriction factors inhibiting viral release (Tetherin and Calcium-modulating cyclophilin Ligand 1, CAML-1) have also been identified, and seem to be counteracted by the action of viral Vpu, through a still poorly understood mechanism (Neil et al., 2008; Varthakavi et al., 2008).

Finally, the budding virion is subject to cleavage of the Gag precursor by the viral protease; the proteolytic activity ends when the virion is already detached from the host plasma membrane, and results in the formation of mature infectious viruses.

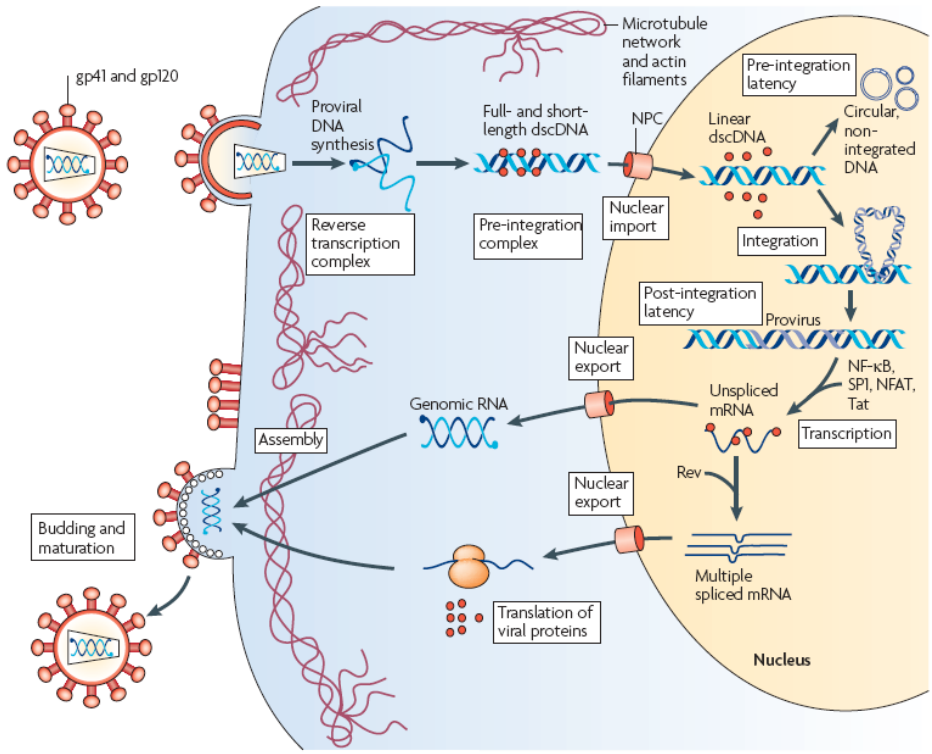


Figure 1.2 A schematic overview of HIV-1 life cycle (Adapted from Coiras et al., 2009).

1.2 CONTROL OF HIV-1 GENE EXPRESSION

The regulation of HIV-1 transcription is a complex event of significant pathological relevance, which recapitulates general concepts of cellular transcription with some peculiarities. The process of HIV-1 transcription can be subdivided into one of two distinct phases. The first phase is dependent on the HIV-1 LTR, since it contains a wide spectum of *cis*-acting consensus sites for several cellular transcription factors, that mediate early stages of viral transcription (explained in details in Section 1.2.1). The second phase immediately follows, and relies on the Tat protein, encoded by the virus itself, synthesized in the course of the first phase. By the interaction with the TAR element and through a multiplicity of different mechanisms and interactors, Tat promotes efficient initiation and elongation of viral transcription (Section 1.2.2). In addition, a further level of HIV-1 transcription regulation is related to the framework that controls general transcription of the host cells, both in terms of cellular factors distribution and availability, and in terms of nuclear positioning and/or long-range chromatin interactions established within viral sequences as well as between the provirus and host chromosomes (Section 1.2.3).

1.2.1 The HIV-1 Long Terminal Repeat (LTR) promoter

Cis-acting cellular factors binding the LTR

LTRs are generated in their symmetrical configuration during the process of reverse transcription and so they appear as “repeats” only in the viral DNA. In the context of integrated viral DNA, the major function of the LTR is the regulation of viral RNA synthesis. Different experimental approaches, such as mutagenesis of LTR-driven reporter genes in transfection experiments,

as well as DNA footprinting and Electrophoretic Mobility Shift Assays (EMSAs), have led to the identification of several binding sites for nuclear proteins within the LTR (Gaynor, 1992) and reviewed in (Pereira et al., 2000).

HIV-1 expression relies on a single transcript, which can be multiply spliced or left unspliced, that is initiated from the 5' LTR region, at a single Transcription Start Site (TSS). Indeed, the HIV-1 promoter and adjacent regulatory elements, both involved in the recruitment of RNA Pol II at the TSS, are located within the U3 region and function in the context of the 5' LTR. In addition, important regulatory motifs within the TAR (Du et al., 1993; Montano et al., 1996) as well as the U5 (el Kharroubi and Verdin, 1994; Van Lint et al., 1997) regions have also been described.

Nevertheless, it is worth mentioning that several studies have addressed the possible existence of transcripts initiating at other positions, starting from the identification of a protein, the Anti-Sense Protein (ASP) on the antisense strand (Miller, 1988). Although the question has always been controversial, recently an unbiased antisense transcription-specific RT-PCR approach detected antisense transcripts in infected cells. Furthermore, a novel poly(A) signal was identified within the *pol* gene in the antisense strand, and antisense transcription was shown to be positively modulated by Tat (Landry et al., 2007).

From a functional point of view, the LTR can be divided into one of four main regions: *(i)* the core promoter region, encompassing the TSS, which exerts a positive basal effect on transcription; *(ii)* the enhancer region, which increases the effect of the basal region; *(iii)* the modulatory region, formely called "negative regulatory element", containing several positive and negative regulatory elements, critical for modulating HIV-1 gene expression in response to various stimuli; *(iv)* the Trans-Activation-Responsive region (TAR), found within the R region in the 5' of all viral

transcripts.

A schematic representation of LTR regions is given in Figure 1.3.

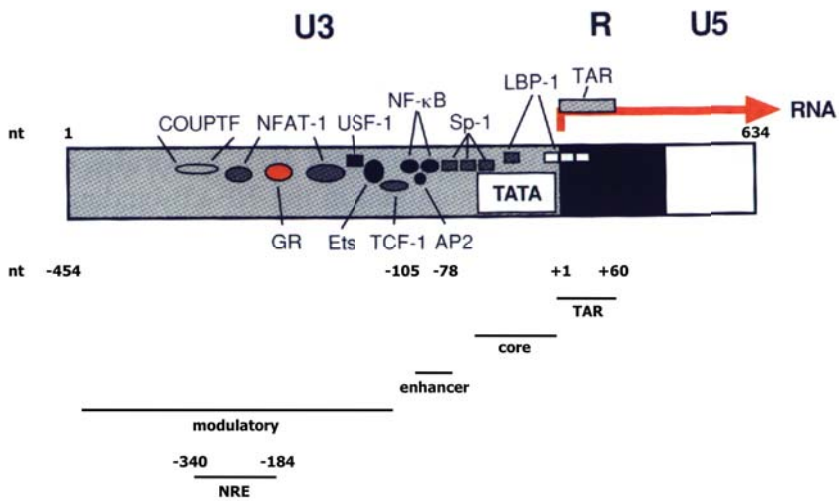


Figure 1.3 HIV-1 Long Terminal Repeat (LTR) promoter (Adapted from "Retroviruses", Coffin, Hughes and Varmus, CSHL press).

(i) The core (or basal) promoter contains the TATA box and three tandemly arranged binding sites for the constitutively expressed Sp1 transcription factors (Jones et al., 1986). Both elements are necessary for basal level of LTR-driven RNA synthesis.

As in other eukaryotic cellular promoters, the TATA box is specifically bound by the TBP (TATAA-Binding Protein) subunit of TFIID; mutations of this region resulted in a marked decrease of both basal transcription and viral replication. Once bound, TFIID constitutes a scaffold upon which other General Transcription Factors (GTFs) can assemble a functional transcription complex, the so-called pre-initiation complex (Orphanides et al., 1996), similarly to what happens for cellular genes. Indeed, TFIIB is recruited, and in turn recruits RNA Pol II to the promoter, definitively establishing the location of the TSS. To allow promoter clearance and

transcription initiation, TFIIH is also incorporated (Goodrich and Tjian, 1994), and its kinase subunit (Cyclin-Dependent Kinase, CDK7) phosphorylates the C-Terminal Domain (CTD) within the largest Pol II subunit Rpb1 (Serine 5) (Akoulitchev et al., 1995). This modification renders RNA Pol II an elongation-competent enzyme able to escape from the promoter region (Orphanides and Reinberg, 2002). Besides CDK7, also CDK8 (Maldonado et al., 1996) and the CDK9 subunit of P-TEFb (Positive Transcription Elongation Factor b) (Mancebo et al., 1997) phosphorylate the RNA Pol II CTD thus regulating transcriptional initiation and elongation (for further details, see Section 1.2.2 "Role of Tat in HIV-1 transcription").
The sites for Sp1 are located upstream of the TATA element: mutagenesis studies demonstrated that Sp1 control basal transcription, even though its effect is markedly higher on Tat-mediated transcription. Indeed, Sp1 was shown to be necessary and sufficient to recruit Cyclin T1 to the LTR, allowing P-TEFb-mediated transcription in a Tat/TAR-independent manner (Harrich et al., 1989; Yedavalli et al., 2003).
Finally, also LBP-1/YY1 (Leader Binding Protein-1/Yin Yang 1) binding sites can be found in the core promoter region: in particular, high affinity sites close to the TSS and low affinity sites close to the TATA box were described, and proved to exert a positive and a negative effect on transcription, respectively (Coull et al., 2000; Kato et al., 1991).

(ii) The enhancer region mediates the transcriptional inducibility of the provirus in response to a variety of stimuli which trigger cellular activation and proliferation (Chinnadurai, 1991; Lusic et al., 2003; Siekevitz et al., 1987). Located upstream of the core promoter, the enhancer encompasses two partially overlapping binding sites for the inducible transcription factors NF-κB (Nabel and Baltimore, 1987) and for STAT5 (Signal Tranducer and Activator of Transcription 5) (Selliah et al., 2006), respectively. Noteworthy,

also NF-AT members can bind in the same region of the NF-Bκ consensus sites, and likely play an important role particularly in T cells (Kinoshita et al., 1998). Full-length STAT5 was shown to exert a direct positive effect on activation of viral transcription (Selliah et al., 2006), while the naturally occurring C-terminally truncated STAT5 form (expressed by HIV-1 infected cells, (Bovolenta et al., 1999) acts as negative regulator of HIV-1 expression in a chronically infected cell line (U1) and in $CD4^+$ PBMCs (Peripheral Blood Mononuclear Cells) of HIV-1-positive individuals cultivated *ex vivo* (Crotti et al., 2007).

The NF-κB/Rel family regulates expression of numerous genes involved in processes such as growth, development, apoptosis, and inflammatory and immune responses. The most abundant form of NF-κB is a p50/p65 dimer, being p50 the DNA binding and p65 the transactivating domain (altough also p65 was shown to carry a DNA binding domain in its N-terminus, as reported by (Toledano et al., 1993) whose activity is tightly controlled by intracellular localization. Indeed, in the cytoplasm, NF-κB dimers are complexed with the IκB inhibitor and cannot be imported into the nucleus. Following activation of T cells by a variety of extracellular stimuli, the bound IκB is phosphorylated and subsequently degraded, while free NF-κB is translocated into the nucleus where it activates transcription of target genes (for a comprehensive review, see (Vallabhapurapu and Karin, 2009).

A further degree of complexity is added by the fact that NF-κB p65 is also regulated by interactions with cellular HATs such as CBP (CREB Binding Protein), p300 and P/CAF (p300/CBP Associated Factor), which mediate acetylation of several Lysine residues variously affecting binding to IκB, binding to DNA and transcriptional activity (Chen et al., 2002b; Kiernan et al., 2003). In particular, acetylation of Lys 310 appears to be crucial for increasing transriptional activity, and is reversed by the nicotinamide adenine dinucleotide-dependent class III HDAC (Histone De-ACetylase)

SIRT1 (Yeung et al., 2004). NF-κB has been shown to stimulate both basal and Tat-mediated expression in activated T-cells (Nabel and Baltimore, 1987; Siekevitz et al., 1987), while mutations in its binding sites affect HIV-1 infectivity to an extent that depends on the mutation type and on the levels of endogenous NF-κB.

(iii) Although the modulatory region was not shown to dramatically affect promoter activation neither *in vitro* nor in reporter assays in transfected cells, site-specific mutagenesis studies revealed that it considerably impacts viral replication in both some T cell lines and in primary lymphocytes. This region is conserved among isolated HIV strains, and it is bound by several cellular proteins such as LEF-1, Ets-1, USF, NF-AT, c-Myb and COUP-TF (reviewed in (Pereira et al., 2000)). The modulatory region has also been proposed to contain a Negative Regulatory Element (NRE), whose deletion increases LTR-driven transcription and viral replication (Rosen et al., 1985).

(iv) The HIV-1 TAR encompasses the 5'-terminal (nucleotides +1 to +59, numbering the TSS as +1) of all viral RNAs. This region functions as an RNA sequence rather than as a DNA element (Muesing et al., 1987). It folds into a highly stable, nuclease-resistant stem-bulge-loop structure which is essential for Tat-mediated LTR transactivation (Berkhout et al., 1989), as it is suggested by the fact that mutations that destabilise the stem by disrupting base-pairing abolish Tat-stimulated transcription. Furthermore, the TAR element was found to be functional only when placed in the 3' to the HIV-1 promoter and in the correct orientation and position (Selby et al., 1989).

Interestingly, footprinting analysis revealed binding sites for several transcription factors (such as AP-1, Sp1, NF-AT) also in the R/U% junction and in the U5 region.

Chromatin conformation at the HIV-1 LTR promoter

Following integration into the cell genome, and independent from the integration site, the proviral DNA, similarly to cellular genes, is organized into a chromatin structure. LTR proved to act as a very strong promoter when analyzed *in vitro* as naked DNA (Parada and Roeder, 1996), while it is almost silent when integrated into the cellular genome in the absence of stimulation (Pomerantz et al., 1990). These findings clearly indicated that chromatin conformation plays essentially a repressive role on HIV-1 transcription.

Both *in vitro* (Sheridan et al., 1997; Steger and Workman, 1997) and *in vivo* nuclease accessibility studies of the proviral chromatin structure (El Kharroubi et al., 1998; Van Lint et al., 1996; Verdin et al., 1993) showed that the 5' LTR, independent from the integration sites, is incorporated into two distinct nucleosomes, namely nuc-0 and nuc-1, precisely positioned with respect to *cis*-acting regulatory elements, and separated by a nuclease-hypersensitivity region. In the transcriptionally silent provirus, nucleosome positioning defines two large nucleosome-free areas, one spanning the core promoter/enhancer in the U3 region (nucleotides -265 to -3) and the other encompassing the primer-binding site immediately downstream to the 5' LTR (nucleotides +141 to +265). A single nucleosome called nuc-1 is located between these two open regions, and it has been shown to be specifically and rapidly destabilized concomitant with transcription activation induced by HDAC inhibitors, cytokines or by the viral transactivator Tat (Van Lint et al., 1996; Verdin et al., 1993).

Interestingly, several genome-wide studies of nucleosome positioning conducted both in yeast and in human cells revealed that most RNA Pol II-transcribed genes carry a similar chromatin conformation, with nucleosomes precisely positioned at promoters that are remodeled concomitant with transcription activation (Bernstein et al., 2004; Hartley and Madhani, 2009; Schones et al., 2008).

The position of nuc-1 in close proximity of the TSS and its displacement during transcription activation clearly suggest that chromatin exerts a crucial repressive role on proviral expression, and that disruption of nuc-1 is necessary for transcriptional activation (Figure 1.4). To further reinforce the chromatin-mediated repression of HIV-1 expression, it is worth mentioning that genomic footprinting studies performed either in activated or in silently infected cells, have indicated that most of the binding sites in the promoter are occupied by cellular transcription factors despite of the activation state (Demarchi et al., 1993). This is consistent with the view that the transcriptional activation of the integrated LTR is not mainly impaired by DNA target site accessibility, but instead occurs through the modulation of chromatin conformation.

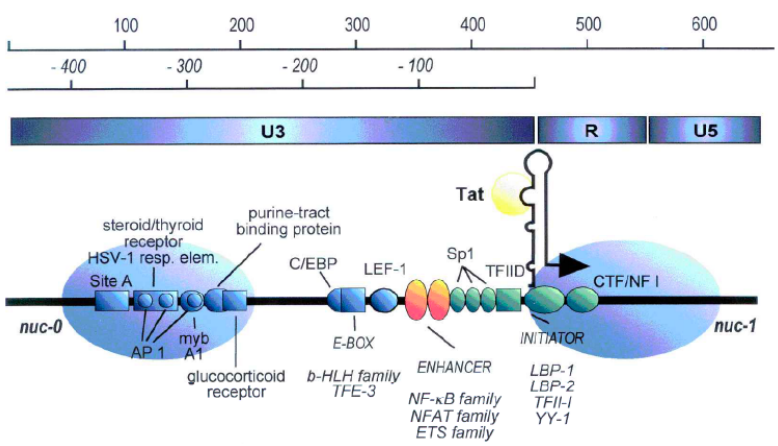


Figure 1.4 Chromatin conformation at the HIV-1 LTR promoter (Adapted from (Marcello et al., 2001b).

1.2.2 The Tat protein of HIV-1

Structure of the Tat protein

Viral replication *in vivo* requires the virally-encoded Tat protein, a 14 kDa protein conserved in the genomes of all primate lentiviruses. It is translated from multiply spliced transcripts, and its gene is located in the 3' portion of the viral genome. The *tat* gene is made up of two exons (first exon: residues 1-72, second exon: residues 73-101) and in most primary isolates encodes for a 101 amino acid-long polypeptide, although one of the most widely utilized laboratory strains (HIV-1_{HXB2}) produces a shorter albeit fully functional protein of 86 amino acids, generated from a point mutation that creates a premature stop codon (Jeang et al., 1999). No crystal structure of the protein has been obtained so far, and the best information about the three-dimensional structure of Tat, based on Nuclear Magnetic Resonance (NMR) spectroscopy, indicate that it possesses a highly flexible structure and does not exhibit obvious secondary structure elements (Bayer et al., 1995).

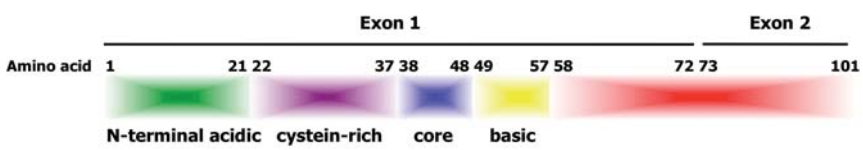


Figure 1.5 Structure of HIV-1 Tat protein.

On the basis of its amino acidic composition, the Tat sequence has been subdivided into several distinct regions, each of which is essential for Tat function: an N-terminal acidic region (aa 1-19), a cystein-rich domain (aa 20-31), a core region (aa 32-47), a basic region (aa 48-57), and a C-terminal region (aa 72-101) (Figure 1.5). The generation of a wide collection of mutants allowed the definition of a detailed structure-function

map of Tat, with different roles ascribed to each domain. The best studied region is the arginine-rich basic domain, highly conserved among HIV isolates, responsible for binding to TAR (Weeks and Crothers, 1991) and for the nuclear localization of the protein (Hauber et al., 1989). The core domain also contributes to TAR binding, while the cystein-rich region is believed to be involved in metal ion binding and, more relevant, is necessary for Cyclin T1/CDK9 binding (Garber et al., 1998). The C-terminal region appears to be dispensable for transactivation, although it retains a still unrecognized role in viral pathogenesis *in vivo*, and is less conserved than the other domains in the different isolates (Smith et al., 2003).

Post-translational modifications of Tat

Tat is also subjected to various PTMs by host cellular proteins, including acetylation *(i)*, phosphorylation *(ii)*, methylation *(iii)*, ubiquitination *(iv)*, which affect its transactivating properties, adding complexity to the transcriptional scenario at the viral promoter.

(i) The best charachterized Tat modification is acetylation, mediated by cellular HATs and occurring on lysine residues in the basic and in the cystein-rich domains of the protein. The association between Tat and HAT activity has been described by three groups almost at the same time, who showed that: Tat interacted with p300 (Hottiger and Nabel, 1998); Tat-driven LTR transactivation was defective when p300 and P/CAF expression was reduced (Benkirane et al., 1998); Tat recruited p300 and the closely related CBP to the LTR promoter as well as p300 overexpression increased LTR transactivation both in rodent and in human cells (Marzio et al., 1998). In addition, Tat was found to be itself a substrate of several HATs, namely p300, P/CAF and GCN5, and different publications identified the modified

lysines: Lys50 and Lys51 are acetylated by p300 and GCN5, while P/CAF acetylates Lys28 (Col et al., 2001; Deng et al., 2000; Deng et al., 2001; Kiernan et al., 1999; Ott et al., 1999). How acetylation at Lys28 and Lys50 might modulate Tat activity has been controversial. Despite the finding that acetylation affects Tat transactivation and viral replication (Bres et al., 2002b), the exact manner by which these PTMs impact Tat functions is still not completely understood (Bres et al., 2002a; Dorr et al., 2002; Kaehlcke et al., 2003; Mujtaba et al., 2002). Nevertheless, a model has been proposed, based on all the previous evidence, in which unmodified Tat first binds P/CAF (Kiernan et al., 1999), which in turns acetylates Lys28 thus promoting association of Tat with Cyclin T1. Then the Tat/P-TEFb complex, bound with high affinity to TAR, phosphorylates RNA pol II CTD, while p300 recruited by RNA Pol II acetylates Lys50. Lys50Ac induces the dissociation of Tat/P-TEFb from the TAR RNA (Kaehlcke et al., 2003; Kiernan et al., 1999), allowing the formation of the ternary complex Tat/P-TEFb/P/CAF that associates to the elongation complex during the whole elongation phase. Interestingly, D'Orso and collegues recently analyzed the role of Tat acetylation using Tat chimeras containing basic domains of lentiviruses in which Lys28 is not conserved (such as Bovine Immunodeficiency Viruses, BIVs) and measured affinity for TAR binding. They demonstrated that Lys28 acetylation increases the affinity and stability of the HIV-1 Tat/Cyclin T1/TAR complexes; conversely, BIVs have evolved different mechanisms to achieve high-affinity RNA binding, having a high Tat/TAR affinity that does not require neither Cyclin T1 nor Tat acetylation (D'Orso and Frankel, 2009).

The requirement for Lys28/Lys50 acetylation in ternary complex assembly/disassembly might allow for additional steps of transcriptional regulation, likely involving cycles of acetylation/deacetylation. Indeed, the nicotinamide adenine dinucleotide-dependent class III HDAC SIRT1 was

shown to specifically de-acetylate Lys50 both *in vitro* and *in vivo*. Tat and SIRT1 sinergistically activated the LTR and, conversely, Tat transactivation was defective upon SIRT1 knock-down, suggesting that this reversible acetylation allows Tat to be recycled to TAR for susequent rounds of HIV transcription (Pagans et al., 2005). Interestingly, the same authors reported that Tat blocked the ability of SIRT1 to deacetyate Lys310 in the p65 subunit of NF-κB (likely competing with p65 for SIRT1 binding), leading to a hyperactivation of expression of NF-kB-responsive genes and eventually to immune cell hyperactivation (Blazek and Peterlin, 2008).

(ii) Phosphorylation of Tat has been initially reported only *in vitro*, involving the interferon-induced, double stranded-RNA-activated protein kinase PKR (Brand et al., 1997) and the CDK2/Cyclin E complex (Deng et al., 2002). Surprisingly, in spite of the interaction of Tat with P-TEFb, no evidences of CDK9-mediated HIV-1 Tat phosphorylation has been reported so far (on the contrary, HIV-2 Tat was found to be phosphorylated by CDK9 *in vivo*) (Herrmann and Rice, 1993). More recently, Tat was shown to be phosphorylated on Serines 16 and 46 (two highly conserved residues) in the context of the integrated HIV-1 provirus. *In vivo* phosphorylation is also likely to be CDK2-mediated, and be important for HIV-1 tanscription, since mutants in Serine residues proved to be greatly impaired in viral replication (Ammosova et al., 2006). Although two serine-threonine phosphatases, namely protein phosphatase 2A (PP2A) and protein phosphates-1 (PP1) were found implicated in the regulation of HIV-1 transcription (Ammosova et al., 2003; Bharucha et al., 2002; Faulkner et al., 2003), they were so far reported to only dephosphorylate CDK9.

(iii) Both Arginine and Lysine methylation of Tat was described. Arginine methylation by the Protein Arginine Methyl-Transferase 6 (PRMT6) onTat

Arg52/Arg53 residues of Tat was found to negatively regulate HIV-1 transcription by disrupting the Tat/TAR/Cyclin T1 complex (Boulanger et al., 2005; Xie et al., 2007). More recently, methylated Tat was found to be more stable compared to the unmodified protein, due to the prevention of Tat degradation by the proteasome. Once stabilized, Tat can persist both within the cell and in the extracellular milieu, and this might be relevant for AIDS pathogenesis because Tat can thus exerts its functions beyond the infected cells, with possible involvement in the pathogenesis of AIDS-related cancer, neurodegeneration and T cell death (Sivakumaran et al., 2009).

In addition, Tat was shown to associate with and to be modified by two histone methyltransferases of the SUV39 family of SET domain-containing proteins (SETDB1 and SETDB2) *in vitro*. Reporter assays after SETDB1/2 knock-down revealed that Lys50 and Lys51 methylation exerts an inhibitory effect on viral transcription, likely competing with Lysine acetylation (Van Duyne et al., 2008).

(iv) Ubiquitination of Tat has been described by Benkirane and collegues who identified Hdm2 as the E3 ligase for ubiquitination of Lys71 both *in vitro* and *in vivo*. Interestingly, ubiquitination proved to enhance Tat transactivating properties (Bres et al., 2003). The mechanism by wich ubiquitination of Tat affects its transcriptional activity was further investigated, and insights came from the discovery that the proteasome regulates both basal and Tat-induced transcription, by a proteolytic and a non-proteolytic mechanisms, respectively. In particular, in the absence of Tat, both the 19S and 20S subunits of the proteasome are associated with the LTR and the coding regions, and negatively modulate basal transcription. This conclusion is drawn from the increase in HIV-1 expression observed after either treatment with the proteasome inhibitor

MG132 or proteasome subunit knock-down. Conversely, when ubiquitinated Tat is present, it recruits additional 19S proteasome subunit at the HIV-1 promoter, stimulating transcriptional elongation. This occurs through Tat interaction with the protein PAAF1 (Proteasomal ATPase-Associated Factor 1), which in turns favors the proteasome dissociation into 19S and 20S particles (Lassot et al., 2007). According to this model, a growing body of evidence has indicated that the 19S proteasome functions in the epigenetic regulation of gene expression, both in yeast and mammals, apparently associating with highly transcribed genes (Bhat et al., 2008; Koues et al., 2009; Lee et al., 2005). It is worth mentioning that PAAF1 seems to be required for efficient transcriptional elongation from the LTR via regulation of the histone chaperone hSpt6. Indeed, PAAF knock-down, as well as hSpt6-knock-down, led to a depletion of core histones as detected by ChIP and to an accumulation of aberrant short transcripts, suggesting that PAAF1 and hSpt6 cooperate to favor efficient transcriptional elongation by promoting the nucleosome reassembly after RNA Pol II passage (Nakamura et al., Retrovirology Volume 6 Suppl 2 "Frontiers of Retrovirology: Complex retroviruses, retroelements and their hosts", meeting abstracts).

Role of Tat in HIV-1 transcription

Transcription driven by the LTR is several hundred-fold higher in the presence of Tat than in its absence. Without Tat, the HIV-1 provirus cannot transcribe efficiently, due to a defect in elongation that leads to an accumulation of short transcripts (Kao et al., 1987). Tat is unique among eukaryotic transcriptional activators because it functions via RNA (TAR) rather than DNA promoter elements (Berkhout et al., 1989).

Interestingly, although most Tat interactions are TAR-mediated, also TAR-independent modes of Tat transactivation have also been described. Indeed, provided that the enhancer region of the LTR is intact (Berkhout et

al., 1990), Tat induces transcription from the LTR basically because it is an activator of NF-κB. Several pathways have been proposed to participate in Tat-mediated NF-κB activation. First, Tat induces nuclear translocation of NF-κB acting through the degradation of IκB, via the cellular interferon-inducible protein kinase PKR (Demarchi et al., 1996; Demarchi et al., 1999); second, Tat induces activation of NF-κB through the Tumor Necrosis Factor alpha (TNF-α)-mediated signaling pathway, since it both alters the redox state of the infected cells and induces the expression of the TNF-α gene (Kalantari et al., 2008; Schreck et al., 1991).

Through its interaction with TAR RNA and a variety of protein-protein interactions, Tat activates HIV-1 transcription by promoting the assembly of a transcriptionally competent complex at the LTR. For instance, Tat binds directly Sp1, and moreover several GTFs, including TFIID, TFIIB, TFIIH and RNA Pol II itself, recently confirmed by Raha and co-workers by Chromatin Immuno-Precipitation (ChIP) experiments (Raha et al., 2005).

The most well-charachterized Tat interaction is that with the Cyclin T1 subunit of the P-TEFb complex (Wei et al., 1998). P-TEFb is composed of the cyclin-dependent kinase CDK9 associated with Cyclin T1 or with one of the other C-type cyclins (for comprehensive reviews on P-TEFb see (Bres et al., 2008; Chen et al., 2006; Peterlin and Price, 2006). As a consequence of this binding, P-TEFb and Tat bind TAR cooperatively. The formation of the P-TEFb/Tat/TAR ternary complex is an essential step towards the assembly of the processive RNA Pol II machinery at the LTR promoter (Bieniasz et al., 1998; Garber et al., 1998; Zhou et al., 1998). Studies in murine cells, where the introduction of the human Cyclin T1 restores Tat function, gave the final proof that P-TEFb is the cellular cofactor for Tat. Indeed, mouse cells are not permissive to HIV-1 infection, due to a single Cystein-to-Tyrosine mutation at position 261 in the Cyclin T1 gene, however restoring

Tyrosine 261 renders murine cells subscetible to HIV-1 infection (Garber et al., 1998).

Studies performed using RNA interference as well as highly specific inhibitors have demonstrated that, besides being a cellular cofactor for HIV-1, P-TEFb is required for efficient expression of most of cellular genes. Notwithstanding, HIV-1 genes are the most sensitive to the availability of P-TEFb, as it is demonstrated by the fact that Tat-driven LTR transactivation is blocked by concentrations of P-TEFb inhibitors that do not exert any effect on transcription of cellular genes (Chao and Price, 2001).

It has been proposed that at the beginning of the transcription cycle, CDK7-mediated phosphorylation on Serine 5 of RNA Pol II CTD facilitates promoter clearance, but shortly after initiation, the progression of Pol II is stalled by two negative elongation factors, namely NELF (Negative Elongation Factor) and DSIF (Dichloro-1-β-D-RibofuranosylBenzimidazole riboside (DRB)-Sensitivity-Inducing Factor). To overcome this checkpoint, Tat recruits P-TEFb to the stalled Pol II, forming the stable ternary complex with Tat/TAR/Cyclin T1 (which in turn is regulated by Tat acetylations, as previously described). P-TEFb phosphorylates DSIF, NELF and Serine 2 on RNA Pol II CTD, thus causing NELF dissociation and the conversion of DSIF into a positive elongation factor, finally allowing the production of full-length transcripts (reviewed in (Barboric and Peterlin, 2005).

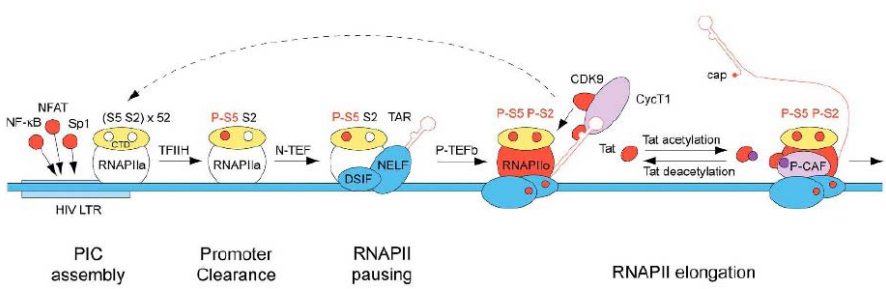


Figure 1.6 Activation of HIV-1 transcription by Tat (Adapted from Barboric et al., 2005).

It is worth mentioning that P-TEFb activity is regulated by different mechanisms. First, regulation is accomplished by controlling the expression levels of both Cyclin T1 and CDK9 (Zhou et al., 2000); second, a number of CDK9 PTMs have been described so far, namely phosphorylation (Fong and Zhou, 2000), ubiquitination (Kiernan et al., 2001) and acetylation (Sabo et al., 2008), affecting CDK9 properties; third, nucleo-cytoplasmatic shuttling of CDK9 could also be modulated finely tuning P-TEFb transactivating activity (Napolitano et al., 2002).

In addition, several lines of evidence have indicated that activity of P-TEFb is tightly controlled by its associations with both positive and negative regulators, such as Brd4 (Bromodomain-containing protein-4) and HEXIM1 (Hexamethylene bis-acetamide Inducible) proteins/7SK small nuclear RNA (snRNA), respectively.

Brd4 and HEXIM1/7SK exist in two mutually exclusive CycT1/CDK9-containing complexes: about half of cellular P-TEFb associates with Brd4, resulting in an active, low-molecular-weight complex. Generally, Brd4 recruits P-TEFb to chromatin through its interaction with acetylated histones, but it proved to be dispensable for HIV-1 transcription, since it can be functionally substituded by Tat (Yang et al., 2005). In addition, Brd4 has been recently correlated with HIV-1 transcriptional inhibition: indeed, it inhibits CDK9 inducing its phosphorylation, and as a matter of fact ChIP experiments detected phospho-CDK9 at the silent HIV-1 promoter (Bisgrove et al., 2007; Zhou et al., 2009).

The other half of cellular P-TEFb exists in an inactive high-molecular-weight complex with the 7SK snRNA and the HEXIM proteins (Nguyen et al., 2001; Yang et al., 2001; Yik et al., 2003). Interestingly, HEXIM1 binds the same region of Cyclin T1 that is involved in Tat binding; indeed, both GST pull-down experiments and size exclusion chromatography revealed a mutually

exclusive binding of the two effectors to Cyclin T1, suggesting a model where Tat and HEXIM1 compete for Cyclin T1 binding (Schulte et al., 2005). Furthermore, the 7SK-binding motif in HEXIM1 contains clusters of positively charged residues, highly homologous to the basic TAR-binding domain of Tat. A similar RNA-protein recognition mechanism might regulate the formation of both the Tat/TAR/P-TEFb and the HEXIM1/7SK/P-TEFb ternary complexes, which might help convert the inactive HEXIM1/7SK-bound P-TEFb into an active complex for Tat-activated and TAR-dependent HIV-1 transcription (Michels et al., 2004; Yik et al., 2004). Therefore it has been speculated that the TAR RNA/Tat lentiviral system has evolved to subvert the cellular 7SK RNA/HEXIM1 system. Interestingly, Ott and co-workers recently reported that acetylation of Cyclin T1 regulates the equilibrium of active and inactive P-TEFb, by triggering the dissociation of CyclinT1/CDK9 from 7SK RNA/HEXIM1; nevertheless, either an acetylation-defective mutant of Cyclin T1 or the wild type protein synergize with Tat in LTR transactivation at the same levels (Cho et al., 2009).

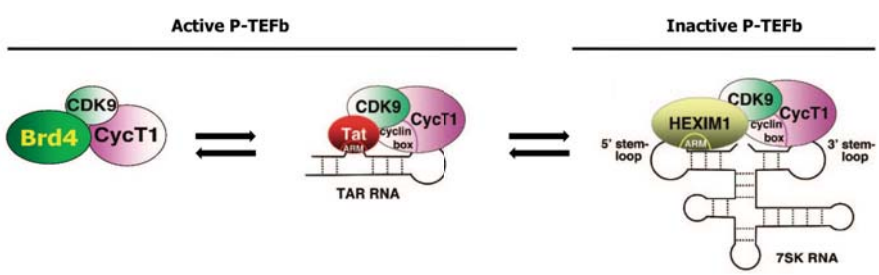


Figure 1.7 The Brd4/P-TEFb, the Tat/TAR/P-TEFb and the HEXIM1/7SK/P-TEFb complexes (Adapted from Zhou *et al.*, Microbiol. Mol. Biol. Rev. 2006)(Zhou and Yik, 2006).

Further insights into Tat transactivation come from the work of K. Jones and collaborators, which highlighted the role of the Ski-Interacting Protein SKIP as a positive regulator of Tat/P-TEFb-mediated transcription.

Depending upon the cellular context, SKIP can function either as a co-activator or as a co-repressor, and it also has a still obscure function in splicing. It has been shown to associate with the active form of P-TEFb and be required for Tat transactivation both *in vitro* and *in vivo*, as assessed by both over-expression and knock-down reporter assays. SKIP is recruited together with Tat and P-TEFb at the HIV-1 promoter and, since it also favors transcription elongation *in vitro*, it is likely that it acts through nascent RNA to overcome pausing by RNA Pol II (Bres et al., 2005). Recently, by RNAi-ChIP approaches, SKIP was shown to act downstream of Tat/P-TEFb to recruit both c-Myc and its partner TRAPP (a scaffold for HATs), and the Menin subunit of the MLL1 histone methyltransferase complex (TriMethH3K4) (Bres et al., 2009; Zhou et al., 2004).

Interestingly, also the co-repressor CTIP2 (COUP-TF Interacting Protein 2) also clearly emerged as a negative regulator of the P-TEFb complex. This protein co-purifies with the inactive P-TEFb, and inhibits the enzymatic activity of CDK9. Tat expression redirected CTIP2 occupancy from 7SK to TAR RNA, but did not impact the recruitment of the P-TEFb complex. (Rohr *et al.*, Retrovirology Volume 6 Suppl 2 "Frontiers of Retrovirology: Complex retroviruses, retroelements and their hosts", meeting abstracts 2009).

Tat and chromatin dynamics in transcription activation

Besides activating HIV-1 gene expression by increasing transcriptional processivity, Tat exerts a crucial role in relievieng the block in transcriptional activation imposed on the LTR by chromatin structure. That chromatin asociated with the HIV-1 promoter undergoes a structural rearrangement concomitantly with transcriptional activation is a well-established notion (El Kharroubi et al., 1998; Verdin et al., 1993). Numerous studies aimed at unraveling the mechanisms by which Tat

remodels chromatin structure to allow efficient transcription accumulated in recent years. Understanding how chromatin is modified and governs HIV expression is also critical for understanding proviral latency with the final goal of developing new therapeutical strategies (Richman et al., 2009)(see Discussion of this Thesis).

The epigenetic regulation of HIV-1 transcription occurs by covalent modifications of the N-termini of the core histones in nucleosomes *(i)* and by the action of ATP-dependent chromatin remodeling complexes *(ii)*.

Altough it is primarily directed by Tat in conjunction with various cellular proteins, it is worth mentioning that also other viral proteins, such as Vpr and Integrase, also participate in chromatin control of HIV-1 expression, interacting with histone modifiers or ATP-dependent chromatin remodeling complexes (reviewed in (Suzuki and Craigie, 2007).

(i) HAT-mediated histone acetylation at gene promoters has been shown to be necessary, although not sufficient, for transcriptional activation, by directly altering chromatin structure, and by acting as a molecular tag for the recruitment of chromatin-modifying complexes (for recent reviews, see (Choi and Howe, 2009; Li et al., 2007). Each HAT has its own lysine specificity within the tails of H3 and H4, according to the "histone code" hypothesis that dictates the epigenetic regulation of transcription (Jenuwein and Allis, 2001).

Tat has ben extensively shown to target several HATs, including p300/CBP (Hottiger and Nabel, 1998; Marzio et al., 1998), P/CAF (Benkirane et al., 1998), GCN5 (Col et al., 2001), Tip60 (Kamine et al., 1996) and TAFII250 (Weissman et al., 1998), which are likely responsible for nuc-1 destabilization and displacement occurring at LTR upon transactivation (Van Lint et al., 1996; Verdin et al., 1993). ChIP assay has been explored to demonstrate that, both in cells containing the sole LTR or the entire HIV-

1 provirus (latently infected cell lines), cellular HATs, namely CBP, P/CAF and GCN5, are all recruited to the LTR upon Tat activation. Once recruited, these HATs acetylate histones H3 and H4 of nuc-1, relieving chromatin repression, prior to the onset of productive viral transcription (Lusic et al., 2003). Of notice, knock-down of P/CAF and p300 dramatically reduced Tat transactivation of an integrated HIV-1 promoter (Bres et al., 2002b).

Histone acetylation might promote transcription also because acetylated histones are more easily transferred from nucleosomes to histone chaperones, thus contributing to create a histone-depleted environment, which correlates with increased gene expression (Ito et al., 2000).

In addition, Tat has been shown to modulate the enzymatic activity of HATs, as it induces a conformational change in p300/CBP that significantly increases the HAT activity of p300 on histone H4 (Deng et al., 2001).

While histone hyperacetylation results in increased chromatin accessibility and transcriptional activation, conversely HDAC-mediated removal of acetyl groups leads to transcriptional repression. Several lines of evidence have demonstrated that distinct LTR-bound complexes recruit HDACs at the HIV-1 promoter where they inhibit viral expression. TFs YY1 and LSF (Late SV40 transcription Factor) cooperatively recruit HDAC1, and this is paralleled by a reduced nuc-1 acetylation and by LTR down-regulation (Coull et al., 2000). Conversely, Tat counteracts chromatin repression by displacing HDAC1 thus increasing histone acetylation in nuc-1 (He and Margolis, 2002).

Furthermore, in latently infected cells NF-κB p50/HDAC1 complexes have been shown to constitutively bind the silent LTR, where histones proved to be hypoacetylated, while synthesis of full-length transcripts could be rescued by Tat expression (Williams et al., 2006). Recently, also CBF-1 (C-promoter Binding Factor-1) has also been described to potently inhibit LTR transcription via histone deacetylation; since this protein is over-expressed

in quiescent or unstimulated cells, its involvement in the establishment of latency was postulated (Tyagi and Karn, 2007). Consistent with the crucial role of chromatin structure and dynamics for the control of proviral expression, multiple studies have indicated that HDAC inhibitors, including TSA (Thricostatin A), trapoxin, valproic acid and sodium butyrate, lead to transcriptional activation in latently infected cells, holding an important role as potential therapeutics (Savarino et al., 2009; Van Lint et al., 1996)}. However, taking into account that transcription factors involved in HIV1 activation, namely NF-κB and Sp1, are also acetylated, whether the effect of inhibitors is due to the inhibition of HDACs or of factor deacetylases is still not clear.

Although histone acetylation has been the prominent covalent modification of histones described so far, other modifications, such as phosphorylation or methylation on specific residues, have also been identified at the induced LTR to be marks of active transcription.
For example, phosphorylation of Serine 10 in the tail of the histone H3 could be mediated by either NF-κB or the mitogen-activated protein kinase pathways via Iκ-B kinase-α or MSK1/2, respectively.
In addition, trimethylation of lysine 4 of histone H3 (TriMethH3K4) has been detected upon TNFα-induction of latently infected cells. A drop in this modification has been appreciated upon treatment with flavopiridol, an inhibitor of the kinase activity of CDK9, and this is consistent with the observation that the P-TEFb-associated SKIP protein also associates with the Menin component of the MLL1 histone methyltransferase complex (Bres et al., 2009; Zhou et al., 2004).

Generally, besides histone deacetylation, another histone covalent modification that correlates with heterochromatin assembly and gene

silencing is trimethylation of Lys9 of histone H3 (TriMethH3K9) (Grewal and Moazed, 2003). It is mediated by the methyltransferase Suv39H1 and it recruits HP1 proteins, proving to be crucial for heterochromatin establishment (Cheutin et al., 2003). Once TriMethH3K9-bound, HP1s in turn recruit more Suv39H1, ensuring heterochromatin maintenance and spread (Maison and Almouzni, 2004).

Elegant Suv39H1 and HP1γ knock-down experiments performed in several systems, including cell lines carrying silent proviruses, infected T-cell lines and PBMCs from infected individuals, showed that Tat-mediated transcription was greatly increased in cells in which the proteins had been silenced compared to control cells. Accordingly, ChIP experiments revealed that both SUV39H1 and HP1γ occupancy of the LTR was significantly reduced upon Tat activation, along with a decrease in TriMethH3K9. Collectively, these findings clearly suggested that Suv39H1, HP1γ, and TriMethH3K9 exert a crucial role in chromatin-mediated repression of HIV-1 gene expression and, moreover, that Tat-mediated transactivation of the HIV-1 promoter, besides co-activator recruitment, involves chromatin derepression (du Chene et al., 2007). However, it is worth mentioning that the role of HP1 proteins in heterohromatin establishment at the HIV-1 genome has been controversial. Indeed, Marban and collegues reported that all the three isoforms (α, β and γ) were detected in the silent HIV-1 genome and released concomitant with transcriptional activation (Marban et al., 2007). Moreover, an experimental approach similar to that used by Benkirane and collegues, consisting in a combination of siRNA-mediated knock-down of HP1s and ChIP analysis of proviral occupancy of HP1s and RNA Pol II, recently showed that the β isoform, but not the γ isoform, of HP1, and the non-processive RNA Pol II are present at the silent LTR where they function as negative regulators. Upon activation, the β isoform is released, concomitant with histone H3 phosphorylation and acetylation,

and replaced by HP1γ, both at the promoter and at the coding regions, together with the processive RNA Pol II. These results are suggestive of a positive role of the γ isoform (Mateescu et al., 2008).

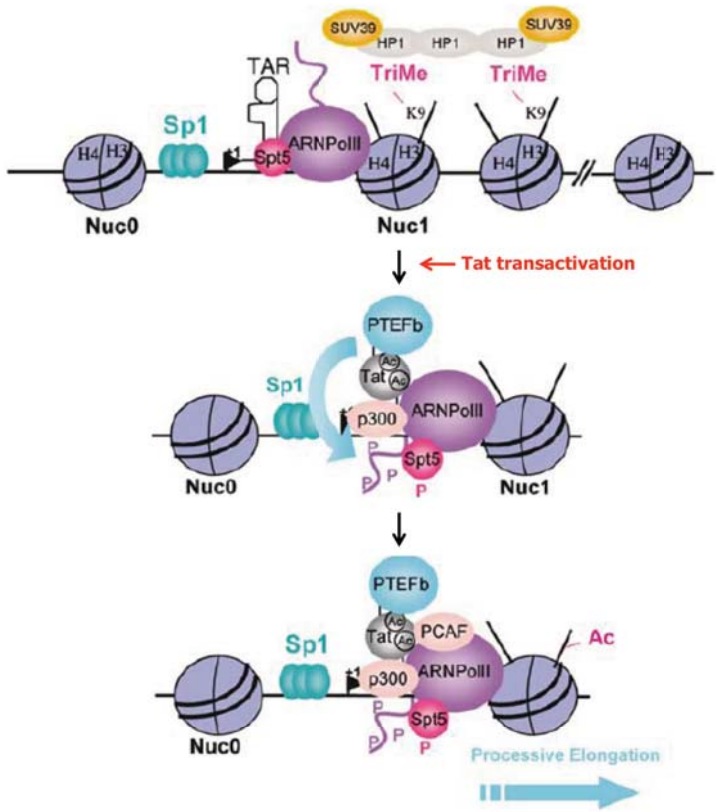


Figure 1.8 A model for chromatin dynamics in the regulation of Tat-mediated HIV-1 expression (adapted from du Chéné et al., EMBO J 2007).

Interestingly, a role in the repressive chromatin establishment at the LTR promoter has also been ascribed to the co-repressor CTIP2: in microglial cells, Sp1- and COUP-TF-bound CTIP2 recruit both HDACs (1 and 2) and SUV39H1, thus contributing to heterochromatin establishment and proviral silencing (Marban et al., 2005; Marban et al., 2007).

However, the mechanisms by which Tat relieves chromatin repression still

remain elusive: it has been hypothesized that either it associates with an enzymatic activity that directly modifies TriMethH3K9, or that it decoys SIRT1 HDAC activity, leading to an accumulation of hyperacetylated histones (Pagans et al., 2005).

(ii) Several lines of evidence have recently underlined the necessity of the SWI/SNF family of ATP-dependent chromatin remodeling complexes in Tat transactivation of HIV-1 promoter.

First, Henderson and collaborators, using an immobilized-template assay with a probe spanning the 3' boundary of nuc-1, reported that, following PMA (Phorbol Myristate Acetate) treatment, BRG-1 (the ATPase subunit of human SWI/SNF) is targeted to the LTR, and that this recruitment is dependent on the transcription factor ATF-3 (Henderson et al., 2004). Accordingly, BRG-1 knock-down, as well as INI-1 (INtegrase Interactor-1, another ATPase) knock-down, greatly impaired Tat-mediated transactivation (Mahmoudi et al., 2006), while both *in vitro* and *in vivo* BRG-1 was detected in the same complex with acetylated Tat (Agbottah et al., 2006).

The Brm SWI/SNF ATPase was also shown by ChIP assay to be recruited at the HIV-1 promoter in a Tat-dependent manner. Interestingly, this interaction is disrupted by Tat acetylation at Lys50, suggesting that Tat recruits Brm before the p300-mediated dissociation from TAR RNA (Treand et al., 2006).

Beyond HIV-1 transcription: additional activities of Tat

Besides its critical function in the control of HIV-1 expression, Tat also affects other steps of viral life cycle *(i)*, as well as different host cell functions *(ii)* and host gene expression *(iii)*.

(i) Over 10 years ago, it was reported that Tat might play a role in reverse transcription of viral RNA, as it was demonstrated by the fact that HIV-1 variants carrying mutated *tat* genes were unable to retrotranscribe their genome efficiently upon PBMC infection (Harrich et al., 1997). The mechanisms by which Tat regulates reverse transcription still remain elusive, since Tat has never been shown to be incorporated into virions; however, it has been suggested that a few Tat molecules might be incorporated, likely binding the two copies of genomic RNA, even if they are undetectable by the current available techniques. Hence, Tat might conceivably be associated with the reverse transcription complex and directly participate in the process of viral DNA synthesis, probably promoting loading of tRNA primer onto viral RNA or suppressing non-specific DNA elongation (Kameoka et al., 2002).

(ii) A more robust set of evidences suggest that Tat affects a remarkable number of host cell functions.

For example, Tat-mediated induction of apoptosis has been first described more than a decade ago (Li et al., 1995; Westendorp et al., 1995). Multiple evidence has demonstrated that Tat pro-apoptotic effect is accomplished through a multiplicity of mechanisms, including the up-regulation of Fas ligand, TRAIL (Tumor necrosis factor-Related Apoptosis-Induced Ligand), Bax and Caspase-8, and the down-regulation of Bcl2 (reviewed in (Romani et al., 2009). Of notice, Tat induces apoptosis by also altering the dynamics of microtubule polymerization inside the cell (Chen et al., 2002a). Besides binding $\alpha\beta$-tubulin dimers, Tat physically associates also with the polymerized microtubules, both *in vitro* and *in vivo*, resulting in microtubule stabilization and enhancement of tubulin polymerization (de Mareuil et al., 2005). Interfering with microtubule polymerization and depolymerization finally results in an alteration of the permeability of the mitochondrial

membrane and the release of cytochrome c from mitochondria, a critical event in triggering apoptosis (commented in (Giacca, 2005)). Moreover, in the course of a biochemical fractionation assay of T cell extracts searching for Tat-associated kinases that phosphorylate RNA Pol II CTD, the microtubule-associated LIS1 protein, known to regulate microtubule dynamics by interacting with dynein, was identified (Epie et al., 2005).

Interestingly, Tat was shown to affect translation: in a yeast two-hybrid screening aimed at identifying cellular partners of the second exon of Tat, the human translation elongation factor-1δ (EF-1δ) was characterized, and this interaction proved to impact translation of cellular, but not viral mRNAs (Xiao et al., 1998).

Finally, a growing body of evidence has recently suggested that a further level of complexity in the virus-host interplay is related to the RNA interference (RNAi) pathway, and in particular Tat has emerged as a suppressor of the RNAi machinery (reviewed in (Kumar and Jeang, 2008). Relationships between HIV-1 infection and microRNAs (miRNAs) are very complex and still remain elusive, notwithstanding mechanisms have begun to be unraveled: cellular miRNAs control HIV-1 replication and latency targeting viral mRNAs (Huang et al., 2007); viral genomes produce viral interference RNAs (viRNAs) that can target both viral and cellular mRNAs (Bennasser et al., 2004); HIV-1 infection can alter the expression of cellular miRNAs that target factors involved in HIV-1 replication (Triboulet et al., 2007). In this scenario, Tat physically interacts with Dicer, causing the partial repression of the ability of Dicer itself to process precursor dsRNAs into small interfering RNAs (siRNAs) (Bennasser and Jeang, 2006; Bennasser et al., 2005). Furthermore, the TAR RNA has been shown to function as an RNAi-decoy, which sequesters the Dicer-interacting protein TRBP2 (TAR RNA-Binding Protein 2), thus preventing the formation of a functional RISC (RNA-Induced Silencing Complex) (Bannwarth and

Gatignol, 2005; Christensen et al., 2007).

(iii) The activation of transcription exerted by Tat is not limited to the HIV-1 provirus, since expression of a number of genes has been found to be transactivated, even though to a less extent, by Tat itself. For example, Tat up-regulates the expression levels of several cytokines, including IL-2, IL-6, IL-8, IL-10 and TNF-α through the activation of the PKR-NF-κB pathway (Chang et al., 1995; Demarchi et al., 1999). Of interest, the IL-2 promoter is bound by the T-cell specific global gene regulator SATB1 (Special AT-rich sequence Binding protein 1), which in turn recruits HDAC1 leading to transcriptional repression. Tat was also found to bind SATB1, thus, Tat-mediated IL-2 up-regulation was apparently due to a competitive displacement of HDAC1 bound to SATB1, leading to increased acetylation of IL-2 promoter (Kumar et al., 2005).

Recently the Tat-mediated up-regulation of Matrix Metallo-Proteinase-9 (MMP-9) in astrocytes has been demonstrated to be dependent on the activation of the MAP kinase-NF-κB pathway (Ju et al., 2009).

Conversely, genome-wide expression profiling in T cell lines revealed that several cellular genes, mostly transcription factors or mediators involved in pathways that control cellular proliferation and differentiation, were down-regulated upon Tat transactivation (de la Fuente et al., 2002; Gibellini et al., 2002). Recently, De Marco and collaborators exploited single cell quantitative Fluorescent *in situ* RNA hybridization (RNA FISH) to demonstrate that, in a cell line harboring a Tat-inducible HIV-1 vector integrated within the *HMBOX1* gene, Tat induced LTR trans-activation and repressed *HMBOX1* expression independent from vector integration, in both alleles. The authors suggested that HIV-1 might target for integration genes that are repressed by Tat, taking an advantage for the virus during transcriptional activation (De Marco et al., 2008).

A different mechanism has been described for the Tat-mediated down-modulation of MHC II (Major Histo-Compatibility Complex class II) genes, since Tat has been reported to compete with the MHC II transactivator CIITA for the binding to Cyclin T1 (Kanazawa et al., 2000).

Finally, it is worth mentioning that HIV-1 transcription can be influenced by cellular transcription, and viceversa, through transcriptional interference. The preference of HIV-1 to integrate into actively transcribed regions (Lewinski et al., 2005; Schroder et al., 2002) raises the possibility that transcription of provirus might be positively or negatively affected by transcription of host neighbouring genes. Indeed, this has been extensively investigated (Greger et al., 1998). Recently, the Peterlin's group took advantage of T cell lines carrying silent HIV-1 integrations within actively transcribed genes to describe this event (Jordan et al., 2003). In particular, they showed that the elongating RNA Polymerase, originating from the host gene promoter, confers transcriptional interference on HIV-1 expression because it physically excludes pre-initiation complex formation on the LTR. Accordingly, this effect could be counteracted by inhibition of the host gene promoter or by activation of provirus with Tat (Lenasi et al., 2008). In addition, Han and co-workers developed an elegant system in a colon carcinoma cell line to monitor the effect of read-through activity on HIV-1 integrated into an actively transcribed gene, and they surprisingly found that it inhibited HIV-1 expression for convergent orientated provirus but enhanced HIV-1 expression when provirus was in the same orientation as the host gene (Han et al., 2008). How these controversial findings might be reconciled with the previous evidence is still to be understood (Perkins and Proudfoot, 2008).

Extracellular Tat

Plenty of literature has indicated that Tat might contribute to HIV disease as an extracellular protein, and several groups have shown that it possesses the unusual property to traffic between cells, being secreted by infected cells and internalized by cells when present in the extracellular milieu, depending on the integrity of its basic domain (Frankel and Pabo, 1988; Green and Loewenstein, 1988). Extracellular Tat has been reported to exert a number of pleiotropic activities when present in the extracellular environment, the most well-known of which is its pro-angiogenic activity on endothelial cells, assessed both *in vitro* and *in vivo*, likely occurring by its interaction with the Flk-1/KDR receptor for VEGF-1 (Vascular Endothelial Growth Factor-A) (Albini et al., 1998; Albini et al., 1996; Mitola et al., 1997). Furthermore, Tat has been shown to interact also with chemokyne receptors (Albini et al., 1998; Albini et al., 1996; Mitola et al., 1997).

Besides activating intracellular signal transduction pathways through its binding with surface receptors, Tat is also rapidly internalized by cells. Internalization occurs through the interaction with the cell surface Heparan Sulphate ProteoGlican (HSPG) receptors (Tyagi et al., 2001), followed by different endocytic routes, the most prominent of which is the caveolar endocytosis (Ferrari et al., 2003). In addition, Tat is also released by cells constitutively expressing the protein (Fittipaldi and Giacca, 2005; Tasciotti and Giacca, 2005), as well as a detectable amount of Tat is found attached to HSPGs in infected cells, suggestive of the importance of extracellular Tat of HIV-1 pathogenesis *in vivo*. Although it has become apparent that Tat is secreted through a non-canonical ER-Golgi pathway - export is not sensitive to drugs that disrupt these organelles – the mechanisms underlying Tat release still remain elusive.

1.2.3 Nuclear organization of HIV-1 Transcription

Besides the transcriptional control exerted by the elements that are encrypted within the viral genome, the LTR promoter and the Tat trans-activating protein, HIV-1 transcription is susceptible to a further, epigenetic level of regulation, related to the nuclear distribution and availability of the host factors required for viral transcription, and to the spatial arrangement that the provirus adopts inside the cell nucleus, establishing both intra- and inter-chromosomal interactions which variously affect its expression.

Nuclear distribution of proteins controlling HIV-1 expression

Thanks especially to the recent development of powerful *in vivo*-imaging approaches, a growing body of evidence has contributed to create the emerging view of large macromolecular complexes formation, such as the transcription machinery, as a highly dynamic process which involves transient and stochastic protein-protein and protein-DNA interactions (Dundr et al., 2002; Misteli, 2007). This model might be also applied to the regulation of HIV-1 transcription, and several publications have addressed the issue of studying the subnuclear localization of viral transcription and the dynamics of the proteins necessary for mRNA synthesis.

Fluorescent In Situ Hybridization (FISH)-based studies performed more than a decade ago, highlighted the proximity between HIV-1 transcription and nuclear speckles, being the latter sites where viral multi-spliced RNA accumulate (Boe et al., 1998). Nuclear speckles are irregularly shaped compartments enriched in splicing-related factors, such as small nuclear ribonucleoprotein complexes (snRNPs) and SR proteins (Lamond and Spector, 2003), likely involved in the co-transcriptional RNA processing of active genes (Brown et al., 2008). Accordingly, our laboratory, by FRET (Fluorescence Resonance Energy Transfer) and immunofluorescence

techniques, demonstrated that Cyclin T1, when over-expressed, assumes a dotted distribution, with foci appearing juxtaposed to nuclear speckles. Furthermore, when the direct interaction between Tat and Cyclin T1 was visualized inside the cells, it was found that Tat determines the re-localization of Cyclin T1 outside of the dots (Marcello et al., 2001a). Interestingly, Cyclin T1 foci were found to co-localize with other sub-nuclear structures, namely the ProMyelocytic Leukemia (PML) bodies, either when PML was over-expressed, or at the levels of expression of endogenous proteins (Marcello et al., 2003).

Besides the PML protein itself, several other proteins have been described to locate into the PML bodies, both when expressed at the endogenous levels (for example, SUMO and Sp100), and, especially when PML was over-expressed (for example, p300/CBP, GCN5, p53, pRb, Sp1) (Negorev and Maul, 2001). PML bodies have been implicated in a wide variety of cellular processes. It has been proposed that they are sites of post-translational modifications and degradation of proteins, and in several cases they have been shown to localize to gene-dense regions and actively transcribed chromatin regions, including the MHC I gene cluster (for a comprehensive review, see (Torok et al., 2009). New insights into the mechanisms by which PML regulates transcription, comes from the work of Kumar and collaborators, which shows that at the MHC I locus PML, together with the Matrix Attachment Region (MAR)-binding protein SATB1, mediates the formation of chromatin loop structures, that proved to be highly dynamic concomitantly with transcription activation (Kumar et al., 2007).

Several findings suggest that PML bodies might be involved in the regulation of HIV-1 expression as well. Notably, PML directly binds Cyclin T1, targeting it to the bodies, as detected by FRET experiments (Marcello et al., 2003). The GCN5-acetylated, transcriptionally-inactive form of CDK9

has also been reported to co-precipitate and to co-localize with PML, upon PML over-expression, both in biochemical fractionation assays (both proteins were found in the insoluble nuclear matrix) and in immunofluorescence stainings (Fogal et al., 2000; Sabo et al., 2008). Upon Tat trans-activation, AcCDK9 is displaced, as assessed by ChIP assay (Marcello et al., 2003; Sabo et al., 2008). Collectively, these observations support the view of nuclear bodies as depots where proteins such as transcription factors, co-activators and co-repressors are accumulated and released upon specific signals. Thus, PML bodies most likely modulate HIV-1 expression by regulating the availability of several factors that are required for efficient transcription, such as Cyclin T1, CDK9, p300, RNA Pol II, PML itself. These structures are highly dynamic, and equilibrium can be easily perturbed, either by Tat, that re-localizes factors in the periphery of the bodies, or by protein over-expression (PML, Cyclin T1, p300), that mediates formation of larger bodies that do not participate in transcription (Marcello et al., 2003; Marcello et al., 2004).

Spatial arrangement of HIV-1 within the cell nucleus

That chromatin organization inside the nucleus is not random and that nuclear architecture can influence gene regulation are widely accepted notions (Cremer et al., 2006; Fraser and Bickmore, 2007; Gondor and Ohlsson, 2009). Heterochromatin is distributed at the nuclear periphery, under the inner surface of the nuclear envelope, while the transcriptionally active euchromatin is dispersed in the nuclear interior. In this scenario, genes move to the periphery to switch off their expression, and move internally into euchromatin to switch them on. Accordingly, artificial tethering of a locus to the nuclear envelope resulted in gene silencing, both in yeast and in mammalian cells (Dillon, 2008).

Recently, Dieudonné and collaborators addressed the issue of how the

subnuclear localization of an HIV- provirus might affect its expression, and studied positioning of the provirus in several cell lines, all harboring silent but inducible copies of HIV-1, before and upon activation. Combining 4C (Lomvardas et al., 2006) and FISH techniques, they identified a region in the proximity of the centromere of the chromosome 12 that physically interacted with the silent provirus. Upon cell induction with phorbol esthers, this association was lost, although the transcribing provirus remained localized at the nuclear periphery (Dieudonne et al., 2009). Overall, these findings strongly support the idea that nuclear positioning to the periphery and moreover long-range chromatin interactions with pericentric heterochromatin, correlate with HIV-1 transcriptional repression.

Besides inter-chromosomal associations, recently intra-chromosomal interactions have also been shown to be crucial in the regulation of HIV-1 expression. In particular, elegant 3C – Chromosome Conformation Capture – experiments (Dekker et al., 2002), extensively described in Section 1.6) revealed the existence of a juxtaposition of the two LTRs in the HIV-1 provirus, reminiscent of gene-loop structures, already detected both in yeast and human genes (Perkins et al., 2008).

Chromatin loops are ubiquitous structural elements of chromatin, implicated at all levels of chromatin organization and function, from kb-sized loops involved in the interaction between promoters and upstream regulatory elements, to giant loops involved in the extrusion of a locus from the chromosome body (Cremer et al., 2006; Fraser, 2006; Misteli, 2007). Notably, gene loops have also been described, in which RNA Pol II-transcribed genes bridge together their 5' and 3' ends (Ansari and Hampsey, 2005; Martin et al., 2005; O'Sullivan et al., 2004; Singh and Hampsey, 2007; Tan-Wong et al., 2008), according to the well-established notion that the 3'-end processing and RNA-processing complexes physically

interact with the transcription machinery (Bentley, 2005).

Perkins and co-workers combined 3C and ChIP approaches, and further developed a plasmid-3C assay, to demonstrate that a specific gene loop conformation is imposed on the HIV-1 provirus upon transcriptional activation (Perkins et al., 2008). As a model, they used the promonocytic U1 cell line, carrying two copies of integrated HIV-1, almost silent due to defective production of Tat. Upon activation with TPA, the provirus adopted a circular conformation - detected by 3C - with the 5' and the 3' LTRs juxtaposed, and this was paralleled by an increase of phosporylated RNA Pol II (Serin 2 and Serin 5) as well as CDK9 at both proviral ends, as detected by ChIP analysis. In addition, another conformational change was revealed, between the 5' LTR and the MSD (Major Splice Donor) site, likely reflecting the functional interplay existing between these sequences to suppress the 5' LTR poly(A) signal (Ashe et al., 1995; Ashe et al., 1997).

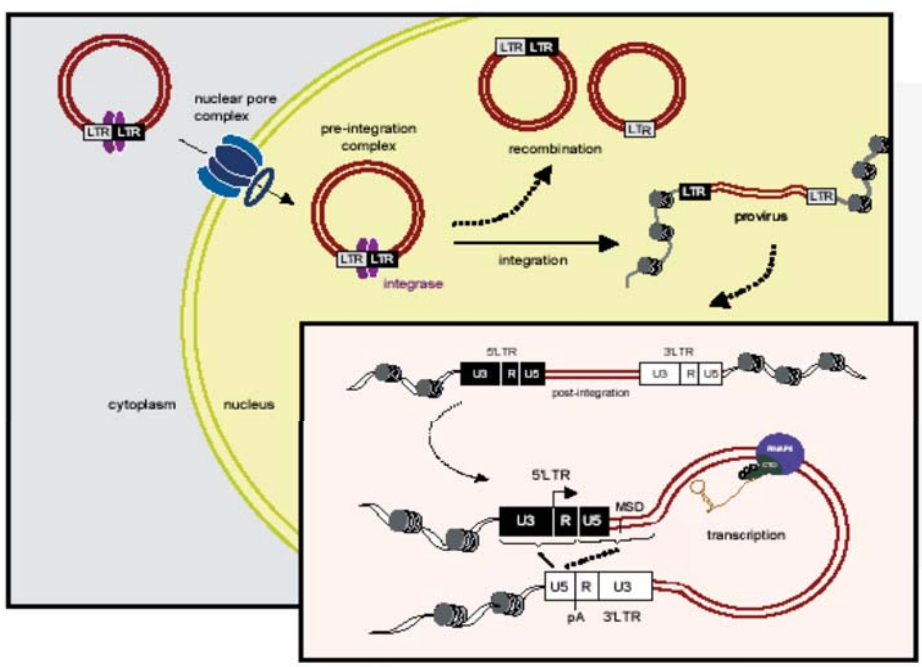


Figure 1.9 Transcription-dependent HIV-1 gene looping.

Interestingly, treatment with flavopiridol abolished loop formation, suggestive that LTR juxtaposition was dependent on ongoing transcription. Plasmid-3C experiments performed with variously modified LTRs showed that mutations of the MSD or of the 3' LTR poly(A) determined a significant drop in looping as well, while substitution with functional heterologous promoter or poly(A) elements resulted in the maintenance of HIV-1 looping properties. Taken together, these findings indicated that both ongoing transcription and pre-mRNA processing are essential for HIV-1 loop formation, and predict that these structures might be a general hallmark of active gene transcription.

1.3 THE HISTONE CHAPERONE NUCLEOSOME ASSEMBLY PROTEIN-1 (NAP-1)

1.3.1 Histone chaperones: classification and functions

Histone chaperones bind free histones and prevent their improper and premature interactions with DNA or other factors, thus controlling histone supply and incorporation into chromatin (Loyola and Almouzni, 2004; Park and Luger, 2008; Polo and Almouzni, 2006).

Althought a universally accepted classification of histone chaperones is still lacking, different criteria can be exploited to group them according to common features. Since chaperones which preferentially interact with H2A-H2B can be distinguished from those that prefer H3-H4, histone binding selectivity is considered one of those criteria. Interestingly, histone chaperones that bind linker histones as well as H2A and H3 histone variants have also been described (De Koning et al., 2007; Shintomi et al., 2005). Another useful criterion to classify histone chaperones is taking into account their way of action, so that three categories can be defined: chaperones that can bind and transport or transfer histones without necessarily involving additional partners *(i)*; multichaperone complexes that combine several histone chaperone subunits *(ii)*; chaperones that provide histone-binding capacity within large enzymatic complexes *(iii)*. An overview of the classification of histone chaperones is reported in Table 1.10.

Classification of chaperones		Histone chaperones (in species of function identification)	Main functions
 Class I: single chaperones	H3-H4	Asf1 (D.m.) Fkbp39p (S.p.) HIRA (X.l.) N1/N2 (X.l.) Spt6 (S.c.) Rtt10 (S.c.)	Histone donor for CAF and HIRA rDNA silencing Deposition factor independent of DNA synthesis H3-H4 storage in X.l. oocytes Transcription initiation and elongation Heterochromatic silencing
	H2A-H2B	Nucleoplasmin (X.l.) Nucleophosmin (NPM1) (H.s.) Chz1 (S.c.) **Nap1 (X.l.)** Nap1-related proteins: Nap1L2 (M.m.), SET/TAF1b (H.s.), CINAP (H.s., M.m.), Vps75 (S.c.) Nucleolin (H.s.)	Storage in X.l. oocytes, cytosolic-nuclear transport, replication, transcription Chz: H2AZ incorporation by SWR1 Cytosolic-nuclear transport,transcription, replicatior Transcription elongation,assist chromatin remodeling
 Class II: Multi- chaperone complex	H3-H4	CAF-1 complex (H.s.)	Deposition factor coupled to DNA synthesis, replication, repair
	H2A-H2B	FACT complex (H.s.)	Transcription elongation
 Class III: Within enzymatic complexes	H3-H4	Hif1 (S.c.) Rsf-1 (H.s.)	Assist histone acetyl-transferase (HAT) complexes Assist remodeling complexes
	ND/multiple	Arp4 (S.c.) Arp7, Arp9 (S.c.) Arp8 (S.c.) Acf1 (D.m.)	Assist remodeling complexes

Table 1.10 Classification of histone chaperones (Adapted from De Koning et al., Nat. Struct. Mol. Biol. 2007).

Histone chaperones can play a broad range of functions *in vivo* in regulating histone metabolism. Soon after their synthesis, histones interact with chaperones, either to be stored or to be translocated into the cell nucleus. Once in the nucleus, histone chaperones gain access to chromatin and play a role in all DNA transactions in which chromatin assembly is required, including *de novo* DNA replication, DNA repair and/or recombination and transcription.

DNA Replication: Three fundamental steps are necessary for duplication of nucleosome organization during replication (Polo and Almouzni, 2005, 2006): the transient disruption of parental nucleosomes *(i)*, the transfer of

parental histones onto nascent DNA *(ii)*, the deposition of newly synthesized histones *(iii)*.

Histone chaperones, together with ATP-dependent chromatin remodeling enzymes, act as histone acceptors during nucleosome disruption. Moreover, several works have described the central role of histone chaperones in depositing newly synthesyzed histones onto replicating DNA (Shibahara and Stillman, 1999; Tyler et al., 1999).

DNA Repair and/or recombination: Coordination of DNA repair and histone dynamics is required to preserve genetic and epigenetic information in response to a double strand break (DSB) within chromatin (Altaf et al., 2007; Bao and Shen, 2007; Downs et al., 2007). Phosphorylation of H2AX is induced very rapidly in quite large adjacent chromatin domains upon DSB formation. Phosphorylated H2AX is thought to provide a docking platform not only for repair factors, but also for histone modifiers and chromatin-remodeling complexes, often in concert with different histone chaperones. Furthermore, histone chaperones are involved in the subsequent steps of removal and restoration of chromatin structure; histone eviction and exchange are key mechanisms underlying these events, responsible for incorporation of newly synthesized histones and for the recycling of the evicted ones (De Koning et al., 2007; Green and Almouzni, 2002; Groth et al., 2007; Peterson and Cote, 2004).

Transcription: During transcription, histones are thought to be partially or totally evicted from the transcribed DNA (Li et al., 2007; Park and Luger, 2008). The evicted histones may then be recycled after passage of RNA Pol II or, alternatively, newly synthesized histones might be incorporated, and this cycle of eviction/deposition requires histone chaperone participation (De Koning et al., 2007; Kim et al., 2007; Workman, 2006). In order to affect histone eviction and deposition, histone chaperones need to reach transcription sites: direct interaction with elongating polymerase (Endoh et

al., 2004), binding to general transcription factors (Chimura et al., 2002) as well as interaction with the chromatin remodeling machinery (Simic et al., 2003) have all been shown to promote recruitment of histone chaperones to their site of action. Furthermore, in yeast gene expression is also modulated by incorporation of histone variants, which involves histone chaperones as well (Raisner et al., 2005; Thambirajah et al., 2009). Finally, histone chaperones can alter chromatin structure during transcription also through the links with histone modifications, in particular histone acetylation: HAT complexes might require to act in complex with a histone chaperone (Tsubota et al., 2007), or acetylation of histones might modulate interactions between histones and histone chaperones (Ito et al., 2000).

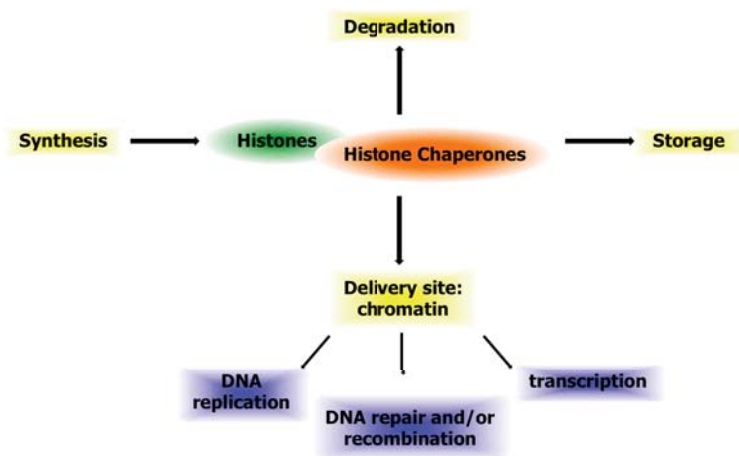


Figure 1.11 Schematic representation of histone chaperone functions.

1.3.2 Characterization of Nucleosome Assembly Protein-1

Nucleosome Assembly Protein-1 (NAP-1) was first identified in extracts from human (HeLa) and mouse cells as a protein associated with facilitated

assembly of nucleosome-like structures *in vitro* (Ishimi et al., 1984; Ishimi et al., 1983). NAP-1 is conserved among all eukaryotes (Ishimi and Kikuchi, 1991; Ito et al., 1996; Simon et al., 1994; Steer et al., 2003), and its deletion in Drosophila and in mouse has been shown to lead to embryonic lethality, suggestive of its central role inside the cells (Lankenau et al., 2003; Rogner et al., 2000).

Protein	Organism	Characteristics	References
NAP-1	Yeast, Xenopus, Drosophila	Affects gene expression (Yeast) Deletion increases embryonic lethality (Drosophila)	Ohkuni, 2003 Steer, 2003 Lankenau, 2003
NAP-1-like 1 (NAP1L1) or hNAP-1	Human	Human counterpart of yeast NAP-1	Ishimi, 1987
NAP-1-like 2 (NAP1L2)	Mouse	Expressed in brain, deletion is embryonic lethal	Rogner, 2000
NAP-1-like 3 (NAP1L3)	Human	Expressed in brain	Shen, 2001
NAP-1-like 4 (NAP1L4) or NAP-2	Human	Candidate gene for Beckwith-Wiedermann syndrome	Rodriguez, 1997
NAP-1-like 5 (NAP1L5)	Mouse		Smith, 2003
SET/TAF-Iβ	Human, Mouse	Stimulates elongation of DNA replication, inhibitor of HATs	Von Lindern, 1992 Matsumoto, 1993
CINAP	Mouse	Widely expressed in different tissues, 3-fold higher in brain	Lin, 2006
TSPY	Human	Expressed in testis, encoded by the Y chromosome	Schnieders, 1996

Table 1.12 NAP family of proteins (Adapted from Park et al., 2006).

NAP-1 belongs to the NAP family of proteins, which are characterized by sequence homology and their common ability to assemble histones into nucleosomes on DNA templates *in vitro* (Park and Luger, 2006a). While only one NAP-1 member is present in yeast, higher eukaryotes contain multiple NAP-1 homologues. The expression of many of the family members is cell- or tissue-specific, as indicated in Table 1.12.

NAP-1 is a highly acidic protein (~25%), with three clusters of acidic amino acidic residues, the longest and the most conserved of which is located in its C-terminus (Figure 1.13).

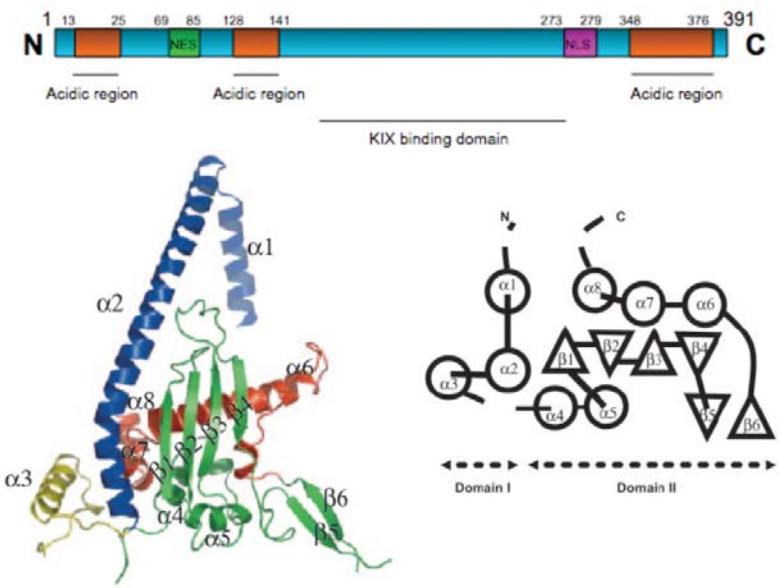


Figure 1.13 The NAP-1 protein (Adapted from Park et al., PNAS 2006).

The crystal structure of yeast NAP-1 (yNAP-1) has recently been determined (Park and Luger, 2006b) and revealed the presence of two domains, namely an N-terminal dimerization domain (domain I in Figure 1.13), and a four-stranded antiparallel β–sheet domain (domain II)

responsible for the interaction with other proteins (previously known to bind the KIX domain of p300, (Asahara et al., 2002). Althought the protein crystallized as a dimer, in physiological salt conditions it is in equilibrium between dimeric and octameric forms (Toth et al., 2005).

Both a Nuclear Localization Signal (NLS) (Mosammaparast et al., 2005; Mosammaparast et al., 2002) and a Nuclear Export Signal (NES) (Miyaji-Yamaguchi et al., 2003) can be identified in NAP-1 sequence, nevertheless the exact mechanisms by which NAP-1 is shuttled between nucleus and cytoplasm are still unclear. Intracellular localization of NAP-1 is controversial, since its predominant cytosolic distribution cannot apparently be conciled with its role in chromatin assembly (Marheineke and Krude, 1998). It might be conceived that NAP-1 is only transiently nuclear and regularly exported to the cytosol, althought the static subcellular localization of NAP-1 is cytoplasmic. Furthermore, some evidence suggest that NAP-1 localization changes during cell cycle progression, being mostly nuclear in S phase and mostly cytoplasmic in G1 and G2 phases (Ito et al., 1996; Miyaji-Yamaguchi et al., 2003; Rodriguez et al., 2000).

In particular, yNAP-1 has been shown to be phosphorylated at 11 sites *in vivo* (three serines by Caseine Kinase 2), and this modification increases nuclear import of the protein, thus affecting the progression through S phase (Calvert et al., 2008).

Besides phosphorylation, NAP-1 is also subjected to other PTMs. For example, acetylation of NAP-1, imparted by p300, occurs in a cell-cycle dependent manner: in particular, it increases during the S phase, most likely regulating cell cycle progression (Asahara et al., 2002). Two additional uncommon PTMs of NAP-1 have been described, namely polyglutamylation and polyglycylation. Polyglutamylation, originally identified on tubulin, consists in the addition of side chains composed of several glutamyl units; this modification reversibly changes the total

negative charge of the C-terminus of the protein, and is thought to play a role in the regulation of the oligomeric states and in NAP-1 nucleo-cytoplasmatic transport (Regnard et al., 2000). Polyglycylation of NAP-1 occurs on residues that are close to the polyglutamylation sites, and is supposed to block the further acidification of the C-terminus by preventing polyglutamylation itself (Ikegami et al., 2008). In this way, this PTM also participates in the balancing of nuclear import and of oligomeric state of NAP-1.

1.3.3 NAP-1 functions and interactors

NAP-1 acts as a nucleo-cytoplasmatic shuttling factor that delivers H2A/H2B histone dimers from the cytoplasm to the chromatin assembly machinery in the nucleus (Mosammaparast et al., 2005; Mosammaparast et al., 2002). In addition to its histone binding and nucleosome assembly activity, NAP-1 is implicated in transcriptional regulation (extensively described in section 1.3.4) and in the regulation of cell cycle (interacting with B-type cyclines, as described in (Altman and Kellogg, 1997; Kellogg and Murray, 1995).
Furthermore, both traditional and proteome-wide approaches combined with mass spectrometry identified NAP-1 as a protein involved in a large variety of cellular functions, such as mRNA processing and translation or cell membrane or lysine biosynthesis, with a vast number of interactors (Calvert et al., 2008; Krogan et al., 2006; Moshkin et al., 2009; Titz et al., 2006); reviewed in (Zlatanova et al., 2007). Whether all the found interactions are physiologically significant still remains to be determined. Discussing the whole spectrum of NAP-1 functions and interactors inside the cell is beyond the aim of this Thesis; therefore, in the next section I will focus on the roles exerted by NAP-1 during transcription.

1.3.4 NAP-1 and nucleosome dynamics during transcription

Chromatin structure poses significant obstacles on RNA Pol II-mediated transcription at the levels of both initiation and elongation. The participation of histone chaperones in modulating chromatin fluidity and accessibility during transcription initiation and elongation has been described in several reports, although further investigation especially concerning the roles of these proteins *in vivo*, is still needed.

Suggestive of the crucial role exerted by NAP-1 in transcriptional regulation, is a finding obtained from a microarray-based genome-wide expression analysis in Δ*nap-1* yeast cells. This analysis revealed that in the knock-out cells, about 10% of all yeast open reading frames changed their transcription levels more than 2-fold when compared to the wild type strain (Ohkuni et al., 2003). In a more recent screening performed in yeast NAP-1 was also found among the transcriptional activators (Titz et al., 2006). Besides, both *in vitro* and *in vivo* studies using reconstituted nucleosomal arrays or reporter constructs respectively, indicated that NAP-1 has a stimulatory effect on transcription, which could be further enhanced in the presence of activator or co-activator proteins (Asahara et al., 2002; Rehtanz et al., 2004; Shikama et al., 2000).

Recently, several lines of evidence have suggested a role of NAP-1 in develpmental gene regulation. Indeed, a proteomics survey of the protein interaction networks of several histone chaperones, including NAP-1, was performed in *Drosophila* embryos, and revealed the interaction of this protein with the RLAF (RPD3 LID-Associated Factors) complex. This complex possesses both histone deacetylase and TriMethH3K4-demethylase activities, and it is therefore implicated in gene silencing; in particular, a cooperation between RLAF and NAP-1 was shown to mediate repression of Notch–regulated genes (Moshkin et al., 2009). In addition, in

a study aimed at characterizing the chromatin changes associated to the reprogramming of mouse Primordial Germ Cells (PGCs) to totipotency, NAP-1 was shown to accumulate in the nuclei of reprogramming cells (while being almost completely cytoplasmatic in the neighbouring somatic cells). Collectively, these findings strongly support the view of histone chaperones, such as NAP-1, as key regulators of the global tuning of gene expression occurring during development (Hajkova et al., 2008).

Chromatin dynamics is tightly regulated by a multiplicity of mechanisms including histone modifications, chromatin remodeling, histone variant incorporation and histone eviction. The concerted action of all these mechanisms is required to displace nucleosomes at promoters concomitantly with gene activation (Bernstein et al., 2004; Cairns, 2009; Li et al., 2007; Workman, 2006), and references therein). Numerous literature data have shown that histone chaperones assist in both the deposition and the removal of promoter nucleosomes.

That transcriptionally active chromatin is depleted in H2A/H2B histones is a well-known notion (Baer and Rhodes, 1983). In this respect, NAP-1 forms a complex with H2A/H2B dimers enhancing their release, and induces a topologically-dependent H3/H4 release from *in vitro*-transcribed templates (Levchenko and Jackson, 2004). NAP-mediated removal of H2A/H2B dimers also facilitates nucleosome sliding along DNA in an ATP-independent manner (Park et al., 2005), and recruitment of transcription factors to gene promoters (Walter et al., 1995). Recently NAP-1 has also been shown to remove the histone H1 from chromatin fibers, suggesting that it might regulate transcription by modulating the local linker histone amount as well (Kepert et al., 2005).

Histone variants have also been associated with nucleosome displacement (Thambirajah et al., 2009). In particular, H2A.Z is concentrated on the

promoters of genes that are poised for activation and is more easily released from chromatin than canonical H2A. In yeast, NAP-1 is part of the SWR1 chromatin remodeling complex, which acts by either facilitating H2A-H2B dimer removal or delivering the variant histone complex (Mizuguchi et al., 2004).

Many histone chaperones also cooperate with chromatin remodeling complexes to modulate nucleosome assembly and disassembly (Angelov et al., 2006). Indeed NAP-1, together with RSC complex, disassembles nucleosomes *in vitro* (Lorch et al., 2006), while *in vivo* it cooperates with CHD (Chromo-Helicase/ATPase DNA binding) remodelers to displace nucleosomes at promoters and coding regions in an ATP-dependent manner upon gene transcription (Walfridsson et al., 2007).

Both physical and functional interactions between histone chaperones and co-repressors/co-activators (especially HATs) have also been described (Eckey et al., 2007). In particular, besides being a target of p300 itself, previous evidence has demonstrated that NAP-1 is a functional component of a complex comprising the co-activator p300, and that it increases the activity of different p300 targets, including p53 and E2F (Rehtanz et al., 2004; Sharma and Nyborg, 2008; Shikama et al., 2000). Furthermore, p300-mediated histone acetylation helps transfer H2A/H2B from nucleosomes to NAP-1 (Ito et al., 2000).

If the mechanisms described so far mainly concern the involvement of NAP-1 in transcriptional nucleosome dynamics at the level of gene promoters, it is worth mentioning that NAP-1 plays an important role during transcription elongation as well. Evidence suggests that it functions in chromatin reassembly after the passage of RNA Pol II, and that its recruitment to the coding region of actively transcribed genes is dependent on the Yra1 subunit of the TREX mRNA transcription and export complex (Del Rosario and Pemberton, 2008).

1.3.5 NAP-1 and viral infection

Increasing evidence indicate that viruses exploit human NAP-1 (hNAP-1) for several purposes. For example, Rehtanz and collegues demonstrated that the transcriptional activator E2 of Papillomavirus forms a ternary complex with hNAP-1 and p300, that is efficient in the activation of viral transcription (Rehtanz et al., 2004). Similarly, the Tax protein of the Human T cell Leukemia Virus type 1 (HTLV-1) recruits the coactivators CBP/p300 to the viral promoter, and hNAP-1 has been shown to cooperate in eviction of the acetylated histones from the chromatin template concomitant with transcription activation (Sharma and Nyborg, 2008). Interestingly, also EBNA1 protein of Epstein-Barr Virus also directly binds hNAP-1 and this interaction is crucial for viral gene expression (Holowaty et al., 2003; Wang and Frappier, 2009). Finally, in a screening aimed at identifying new interactors of HIV-1 Rev protein, hNAP-1 has been reported to bind Rev, resulting in altering Rev oligomerization state and increasing Rev availability (Cochrane et al., 2009).

AIM OF THE STUDY

The work described in PART A of this Thesis is aimed at characterizing novel interactors of the Tat transactivating protein of HIV-1. In particular, exploiting a proteomic approach, we identify the cellular histone chaperone NAP-1 (Nucleosome Assembly Protein-1) as a Tat-binding protein and we further investigate the role that this interaction exerts in the context of viral infection. We propose a mechanism by which Tat benefits from this class of proteins to relieve the repression imposed by chromatin conformation on proviral expression.

2A. RESULTS

2.1 Identification of cellular factors binding to HIV-1 Tat by proteomic analysis

In order to shed light on Tat functions, we exploited a proteomic approach directed at identifying novel cellular partners that are involved in the regulation of HIV-1 transcription by Tat. To this aim, we used an expression vector encoding the open reading frame of full length Tat (101 aa) fused with a C-terminal Flag tag (Ott et al., 1999). The Flag epitope did not interfere with the transactivating activity of Tat on the HIV promoter (data not shown). HEK 293T cells were transfected either with Tat101-Flag, or with a plasmid encoding for the Flag epitope alone as a negative control, and cell extracts were immunoprecipitated with M2 Flag antibody conjugated to agarose beads. Affinity purified Tat-Flag protein and co-purifying cellular factors were subsequently subjected to a DNase/RNase treatment, which proved to be essential to avoid the purification of a vast number of RNA-binding proteins unspecifically co-immunoprecipitating with Tat (mainly ribosomal proteins, data not shown). After elution with an excess of Flag peptide, purified proteins were separated on a 6-10% gradient SDS-PAGE gel and stained with silver stain (Figure 2.1 panel A). Individual bands that were apparent in the sample from the Tat-Flag-transfected but not in the Flag-transfected cells were excised, and the corresponding proteins were identified by mass spectrometry. ESI-MS/MS (Electrospray tandem Mass Spectrometry) analysis of peptides obtained by trypsin digestion led to the identification of four proteins, as listed in Figure 2.1, panel B. Apart from Tat-Flag itself, as expected we recognized previously described Tat partners, such as the p32 subunit of the ASF/SF2

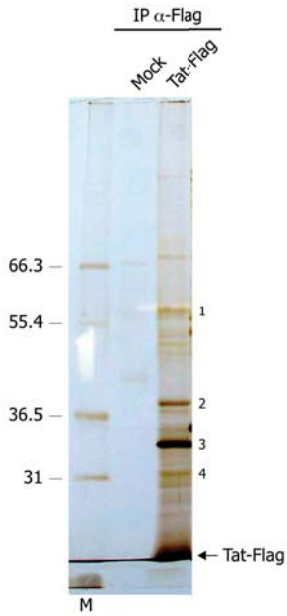
splicing factor (Petersen-Mahrt et al., 1999) and B23/nucleophosmin (Chan et al., 1997). p32 is also known as Tat-Associated Protein (TAP) due to its well-known ability to bind Tat both *in vitro* and *in vivo*; (Yu et al., 1995); acetylated Tat was shown to bind p32 thus inhibiting splicing and promoting the production of full-length transcripts (Berro et al., 2006). B23/nucleophosmin is a nucleolar protein possibly associated with ribosome assembly and/or transport, which binds Tat when both proteins are overexpressed (Li, 1997) and is required for Tat nucleolar localization but not for promoter transactivation (Stauber and Pavlakis, 1998).

Furthermore, we identified a few novel potential Tat interactors, among these, ribosomal protein S4 and hNAP-1 (human Nucleosome Assembly Protein-1), a histone chaperone that plays multiple roles during chromatin transactions (see Section 1.3 of this Thesis). The proteomic analysis was repeated and the results were also confirmed by sequencing proteins directly from the Flag beads, instead than from gel-excised bands.

Since overexpressed Tat is known to accumulate in the nucleoli due to its unspecific RNA binding capacity (e.g. (Marcello et al., 2001a)), and given the observation that the same proteomic assay resulted in the identification of a number of other ribosomal proteins when performed in the absence of RNAse, no further work was performed on the ribosomal S4 protein.

Of notice, hNAP-1 was identified by several peptides (as indicated in Figure 2.1 panel C), in three independent preparations, with $P=7.8 \times 10^{-19}$, indicating highly specific and reproducible interaction. Therefore, further investigation was conducted to better characterize the hNAP1/Tat interaction.

A



B

Sample no.	Accession no.	Molecular mass (kDa)
NAP-1	NP_631946/P55209	45.37
B23/nucleophosmin	P06748	32.57
p32 subunit of ASF/SF2	NP_001203.1	31.34
ribosomal protein S4	P62701	29.59

C

```
  1 MADIDNKEQS ELDQDLDDVE EVEEEETGEE TKLKARQLTV QMMQNPQILA
 51 ALQERLDGLV ETPTGYIESL PRVVKRRVNA LKNLQVKCAQ IEAKFYEEVH
101 DLERKYAVLY QPLFDKRFEI INAIYEPTEE ECEWKPDEED EISEELKEKA
151 KIEDEKKDEE KEDPKGIPEF WLTVFKNVDL LSDMVQEHDE PILKHLKDIK
201 VKFSDAGQPM SFVLEFHFEP NEYFTNEVLT KTYRMRSEPD DSDPFSFDGP
251 EIMGCTGCQI DWKKGKNVTL KTIKKKQKHK GRGTVRTVTK TVSNDSFFNF
301 FAPPEVPESG DLDDDAEAIL AADFEIGHFL RERIIPRSVL YFTGEAIEDD
351 DDDYDEEGEE ADEEGEEEGD EENDPDYDPK KDQNPAECKQ Q
```

Figure 2.1 Identification of Tat-interacting proteins by mass spectometry.
A. Flag-immunopecipitated material from Tat-Flag- and Flag- (indicated as mock) transfected HEK 293T cells was resolved by 6-10% gradient SDS-PAGE gel, followed by silver staining. Protein bands present exclusivelyin the sample transfected with Tat-Flag were excised from the gel and their identification attempted by ESI-MS/MS. The identified proteins, in addition to Tat-Flag, are indicated with numbers. **B.** Identified proteins are listed in the table, together with their accession numbers to NCBI or SwissProt databases, and with their molecular mass expressed in kDa. **C.** Amino acid sequence of the human NAP-1 protein. The underlined amino acid sequences correspond to peptides obtained from MS/MS analysis of three independent preparations.

2.2 hNAP-1 specifically binds HIV-1 Tat *in vivo*

To confirm the interaction between Tat and hNAP-1 found in the proteomic assay, we performed a co-immunoprecipitation analysis exploiting two

epitope-tagged versions of both proteins. When expression vectors for Tat-Flag and for an N-terminal hemagglutinin (HA)-tagged version of hNAP-1 (HA-NAP-1, (Rehtanz et al., 2004)) were transfected into HEK 293T cells, HA-NAP-1 was co-immunoprecipitated with Tat using an anti-Flag antibody (Figure 2.2, upper panels). Furthermore, the specificity of the interaction of the two proteins is underlined by the observation that no co-immunoprecipitation was observed neither when HA-NAP-1 was co-expressed with a Flag-tagged luciferase expressing plasmid, nor when Tat was co-expressed together with HA-tagged SET/TAF-I (Pegoraro et al., 2006), another member of the NAP family, despite its high sequence homology with hNAP-1 (further discussed in the Section 3.1 of this Thesis) Of note, the EBNA1 protein of EBV binds both of them, as described in Wang et al. (Wang and Frappier, 2009).

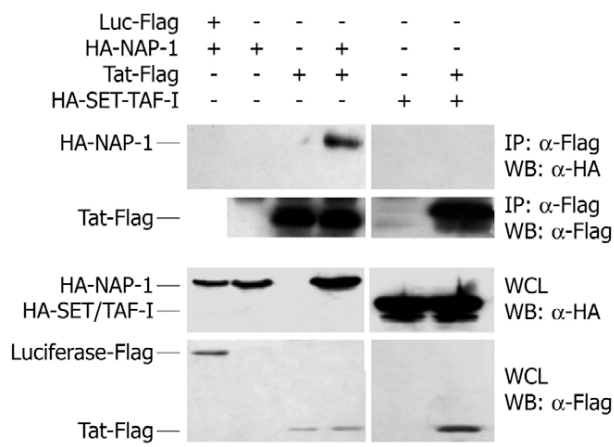


Figure 2.2 Co-immunoprecipitation of Tat with transfected hNAP-1.
The plasmids indicated on top of the figure were transfected into HEK 293T cells. The upper two panels show western blots with the indicated antibodies after immunoprecipitation using an anti-Flag antibody; the lower two panels show western blotting controls from from whole cell extracts (WCL) from transfected cells to show the levels of expression of the transfected proteins.

The same extracts were run on an SDS-PAGE gel and immunoblotted with anti-Flag and anti-HA antibodies to verify protein expression levels. Interestingly, the levels of Tat protein were significantly higher in cells also expressing hNAP-1 or hSET/TAF-I (Figure 2.2, lower panels). Whether this observation argues for a stabilizing effect that histone chaperones exert on small highly basic proteins, such as histones or Tat itself, or it is solely an artefact due to protein overexpression, is still to be determined. The same co-immunoprecipitation experiment was repeated omitting the treatment of the beads with DNase I and RNase A.

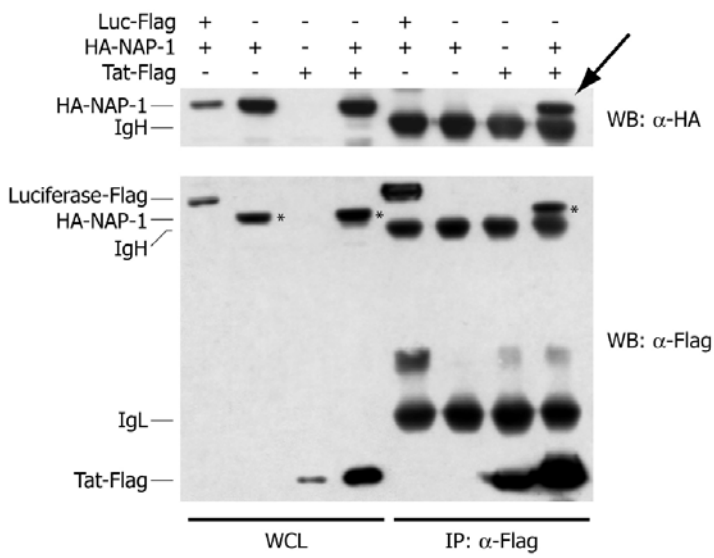


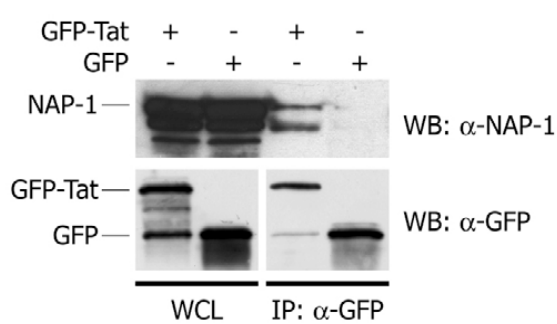
Figure 2.3 Co-immunoprecipitation of Tat with transfected hNAP-1 in the absence of DNase/RNase treatment.
The plasmids indicated on top were transfected into HEK 293T cells. Twenty-four hours after transfection, cell lysates were immunoprecipitated with an immobilized anti-Flag antibody without treatment with DNase I and RNase A. Immunoprecipitated proteins were resolved by SDS-PAGE along with input whole cell lysates (WCL; upper panel, left side) and revealed by anti-HA antibody. The arrow shows transfected hNAP-1 co-immunoprecipitated with Tat-Flag.
The same membrane as above was stripped mildly and exposed to an anti-Flag antibody to monitor the efficiency of transfection and immunoprecipitation. The bands indicated by an asterisk correspond to residual HA-hNAP-1 not removed by stripping.

As shown in Figure 2.3, the binding between Tat and hNAP-1 was confirmed, and these data are even more convincing because they exclude the possibility that treatment with nucleases might have exposed the basic domain of Tat allowing its interactions with other proteins otherwise inaccessible to them. Of note, hNAP-1 overexpression resulted in a remarkable higher expression of Tat, as in the previous experiment.

To further validate our data we performed another co-immunoprecipitation experiment, in which extracts of HEK 293T cells transfected with a GFP-tagged-Tat-encoding plasmid (Marcello et al., 2001a) and with GFP-encoding plasmid as a negative control, were immunoprecipitated with an anti-GFP antibody. The antibody (Ishimi et al., 1985) detected hNAP-1 in samples of Tat-GFP-transfected but not GFP-tranfected cells, thus indicating that HIV-1 Tat binds specifically also the endogenous hNAP-1 protein (Figure 2.4 panel A).

Finally, bacterially expressed and purified GST-Tat recombinant protein, but not recombinant GST, was also able to pull-down endogenous hNAP-1 from a HEK 293T cell extract, as shown in Figure 2.4 panel B. It is worth mentioning that both GFP- and GST-Tat retain full transcriptional and trafficking capacities as the wild-type protein, as demonstrated elsewhere (Demarchi et al., 1996; Fittipaldi et al., 2003; Marcello et al., 2001a; Marcello et al., 2004).

A



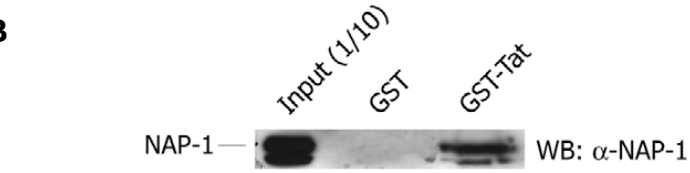


Figure 2.4 Co-immunoprecipitation of Tat with endogenous hNAP-1.
A. Co-immunoprecipitation of endogenous hNAP-1 with GFP-Tat. Extracts prepared from HEK 293T cells transfected with plasmids encoding for GFP-Tat or GFP alone were immunoprecipitated with an anti-GFP antibody and immunoblotted with anti-hNAP-1 antibody. Whole cell extracts (WCE) were run on an SDS-PAGE gel and immunoblotted with anti-GFP and anti-hNAP-1 antibodies to verify amounts of expressed proteins. **B.** GST-pulldown experiment using GST-Tat or GST alone recombinant proteins; Gst-Tat, but not control GST protein, pulled down endogenus hNAP-1 from HEK 293T whole cell extracts.

2.3 hNAP-1 binds HIV-1 Tat through its C-terminal domain

Once established that hNAP-1 and HIV-1 Tat interact *in vivo*, we sought to dissect the regions that mediate this interaction.

To map the Tat-interacting domain in the hNAP-1 protein, full-length (aa 1-391) and several deletion mutants of hNAP-1 were expressed as GST-fusion recombinant proteins (Rehtanz et al., 2004). In particular, the following deletion mutants were produced, as shown in Figure 2.5 panel A: three mutants lacking progressively longer N-terminal portions (aa 92-391; aa 162-391; aa 290-391, respectively); three mutants lacking progressively longer C-terminal portions (aa 1-290; aa 1-162; aa 1-92); three mutants lacking both N- and C-termini, and containing different parts of the central domain (aa 162-290, aa 92-162, aa 92-290). GST pull-down experiments with these fragments were performed using *in vitro* translated ^{35}S-labeled HIV-1 Tat. GST-hNAP-1 fragments immobilized on beads were incubated with ^{35}S-Tat and, after extensively washing, proteins were resolved by SDS-

PAGE. Figure 2.5 panel B shows the gel either stained with Coomassie Blue (lower panel) or after autoradiography (upper panel).

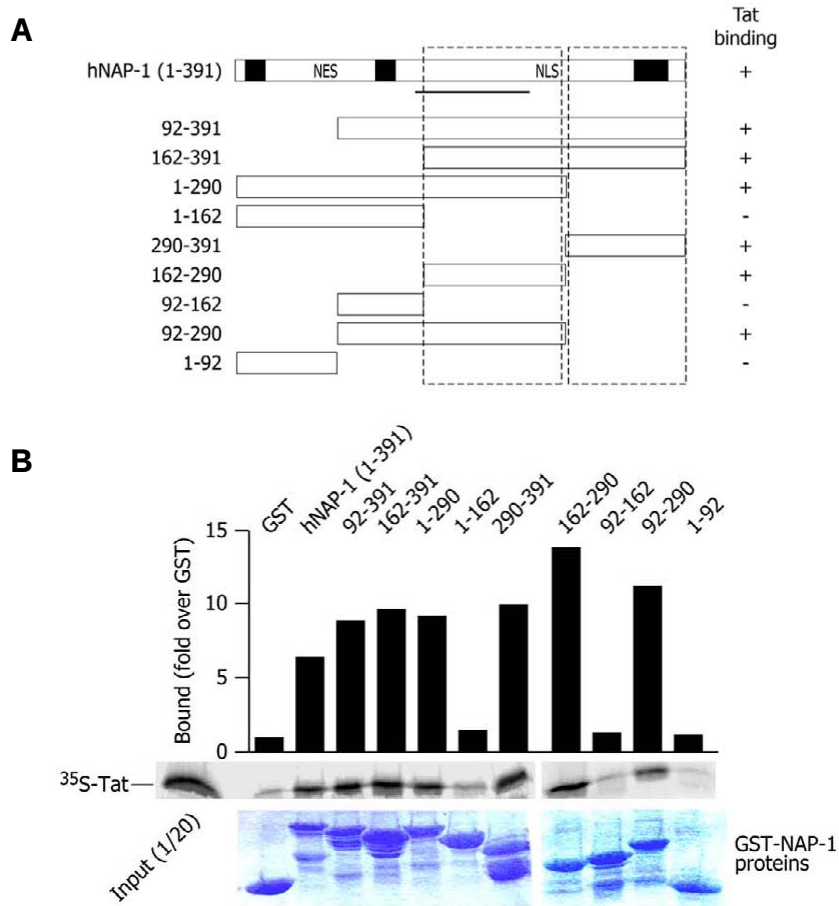


Figure 2.5 Mapping of the Tat-interacting domains in the hNAP-1 protein.

A. Schematic representation of hNAP-1 protein and of its deletion mutants obtained as GST fusion proteins. The capacity of binding to Tat (panel B) is indicated on the right side of each mutant. The two dotted boxes indicate the hNAP-1 domains interacting with Tat. **B.** Representative GST pull-down experiment using the indicated hNAP-1 mutants and radiolabeled Tat full-length protein. The autoradiography shows the amount of Tat binding to each mutant; the histogram on top shows densitometric quantification of data, expressed as fold binding with respect to background binding to GST alone (set as 1). The lower panel shows the Coomasssie stained gel at the end of the binding experiment. The experiment was repeated at least three times with similar results.

As expected, labeled HIV-Tat was found to bind GST-hNAP-1 but not GST alone. All deletants lacking the N-terminus of the protein up to aa 161 bound Tat as efficiently as the full length protein. In contrast, binding was greatly impaired when the hNAP-1 domain from residues 163 to 289 as well as the C-terminal region from residues 290 to 391 were deleted. These results indicate that Tat binds two separable domains within the hNAP-1 protein, one internal from amino acid 162 to 290 and one C-terminal from residues 290 to 391 (Figure 2.5 panel B). The former encompasses the four-stranded antiparallel β–sheet previously shown to act as the protein interaction domain (Park and Luger, 2006b) interacting with other cellular factors, without compromising histone binding (Mosammaparast et al., 2002; Shikama et al., 2000). The latter domain corresponds to the long acidic region which is also involved in the binding of other cellular proteins such as the histone acetytransferase p300 (Shikama et al., 2000).

2.4 The basic domain of HIV-1 Tat binds the C-terminal domain of hNAP-1

Conversely, next we analyzed the domains of Tat responsible for the interaction with hNAP-1. For the expression of both full-length and mutated Tat derivatives we used the 86 aa-long form of Tat, present in the widely studied HXB2 strain, which is transcriptionally active as well as Tat 101 (Jeang et al., 1999). A few deletion mutants were produced, one lacking the second exon (aa 1-72) and one lacking the N-terminal acidic domain (aa 22-86). Also point mutation mutants were employed in the assay, namely a derivative of Tat86 carrying cystein to alanine mutations at positions 22, 25 and 27 in the cystein-rich domain or arginine to alanine mutations at positions 49, 52, 53, 55, 56 and 57 in the basic domain

(Tat86 C(22-27)A and Tat86 R(49-57)A, respectively) (Marzio et al., 1998) (Figure 2.6A). These proteins, obtained as C-terminal fusions to GST, were used to pull-down ^{35}S-methionine-labelled hNAP-1 obtained by *in vitro* transcription/translation.

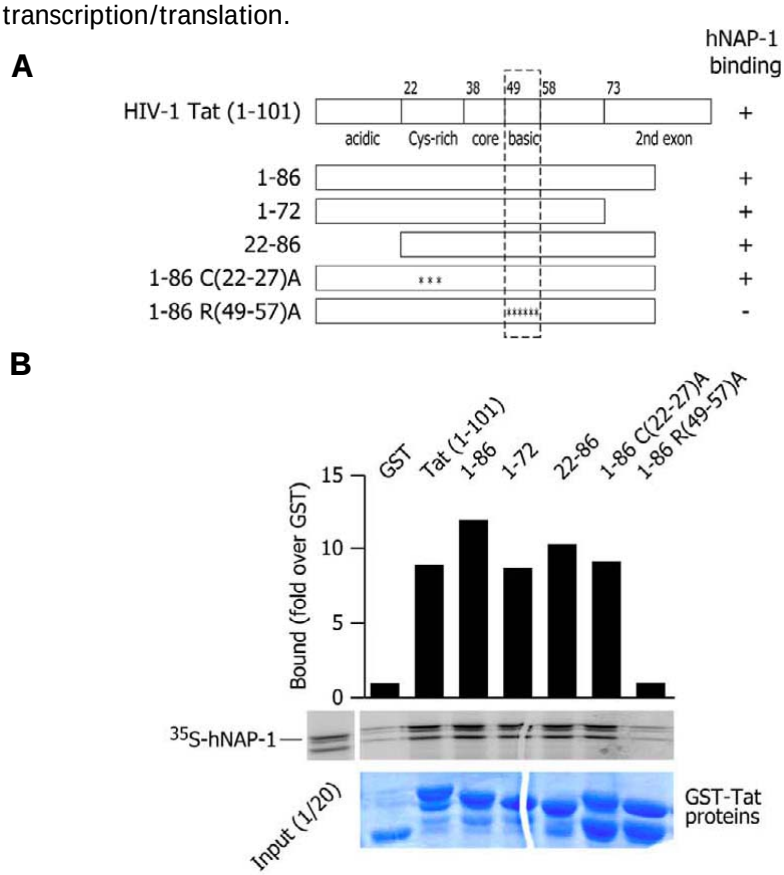


Figure 2.6 Mapping of the hNAP-1-interacting domains in the HIV-1 Tat protein.

A. Schematic representation of HIV-1 Tat protein and of its mutants obtained as GST fusion proteins. The capacity of binding to hNAP-1 (panel B) is indicated on the right side of each mutant. The dotted box correspond to the basic domain of Tat, which binds hNAP-1. **B.** Representative GST pull-down experiment using the indicated Tat derivateves and *in vitro* transcribed and translated hNAP-1 protein. The autoradiography shows the amount of hNAP-1 binding to each mutant; the histogram on top shows densitometric quantification of data, expressed as fold binding with respect to background binding to GST alone (set as 1). The lower panel shows the Coomasssie stained gel at the end of the binding experiment. The experiment was repeated at least three times with similar results.

The results obtained demonstrated that the interaction of Tat with hNAP-1 strongly depended on the integrity of the arginine-rich region of Tat, clearly indicating that Tat binds hNAP-1 through its basic domain. Similar of our observations with Tat, interactions of hNAP-1 with HIV-1 Rev and EBNA1 of EBV occur through arginine-rich regions (Cochrane et al., 2009; Holowaty et al., 2003).

2.5 hNAP-1 and Tat cooperate in the activation of HIV-1 gene expression

It is well known that one of the main biological functions of Tat is the transcriptional transactivation of LTR and, moreover, that one of the essential molecular events that parallel Tat-driven transcriptional activation is the modification of chromatin structure that essentially represses transcription from the integrated HIV-1 promoter (Easley et al., 2009; Lusic et al., 2003; Van Lint et al., 1996; Verdin et al., 1993).

hNAP-1 acts as a histone chaperone operating in both transcription and DNA replication (reviewed in (Park and Luger, 2006a; Zlatanova et al., 2007). Nucleosomes are displaced at promoters during gene activation in a process that involves histone modifications, ATP-dependent chromatin remodeling complexes, histone variants and histone eviction (Li et al., 2007). Several findings indicated that hNAP-1 participates in the regulation of dynamics of chromatin structure through a multiplicity of mechanisms (See Section 1.3.4 of this Thesis).

Based on these evidences, and in order to gain insights into the mechanism of the cooperativity between hNAP-1 and Tat proteins, we analyzed their roles in the LTR transactivation exerted by HIV-1 Tat. To this aim, a reporter construct containing the U3 and R sequences of the HIV-1 LTR

upstream of the firefly luciferase reporter gene was co-transfected into HeLa cells, together with vectors for HA-tagged hNAP-1 and HIV-1 Tat. As shown in Figure 2.7, Tat determined a 200- to 250-fold increase in the levels of gene expression, depending on the amount of transfected vector. The expression of hNAP-1 alone did not stimulate promoter activity (same activation as the basal promoter). However, the co-expression of hNAP-1 together with increasing amounts of Tat increased luciferase activity 520- and 620-fold respectively, suggesting that hNAP-1 is able to enhance Tat-mediated transactivation of LTR promoter. Part of the lysates not subjected to the luciferase assay were loaded on a SDS-PAGE gel and immunoblotted with an anti-HA antibody to verify the levels of overexpressed hNAP-1.

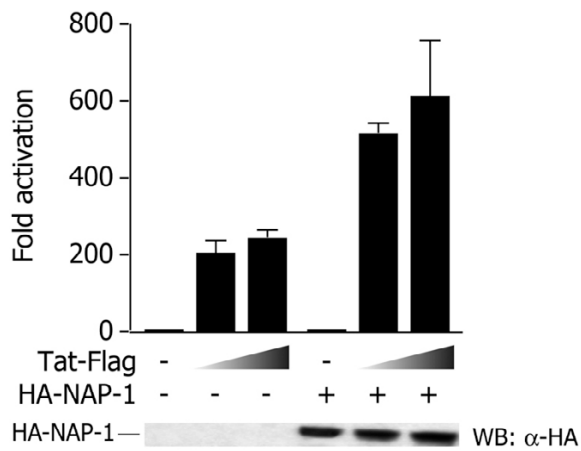


Figure 2.7 hNAP-1 cooperates with Tat in LTR transactivation.
HeLa cells were co-transfected with a reporter construct containing the HIV-1 LTR upstream of the luciferase gene, and with vectors expressing HA-tagged hNAP-1 and HIV-1 Tat, as indicated. The histogram shows mean ± standard deviation for at least three independent experiments; the results are shown as fold transactivation over LTR-luciferase reporter alone. The co-expression of hNAP-1 significantly increased Tat transactivation of the LTR promoter. The wester blot at the bottom shows the levels of transfected hNAP-1 protein in a representative experiment.

To test the requirement for endogenous hNAP-1 protein in Tat-mediated HIV-1 LTR transactivation, we performed luciferase assays with HeLa cells in which expression of hNAP-1 was down-regulated by RNAi. A specific anti-hNAP-1 siRNA oligonucleotide was designed which was able to silence up to 80% of the expression of its target from forty-eight hours after transfection onward, as assessed by Western blot analysis (Figure 2.8).

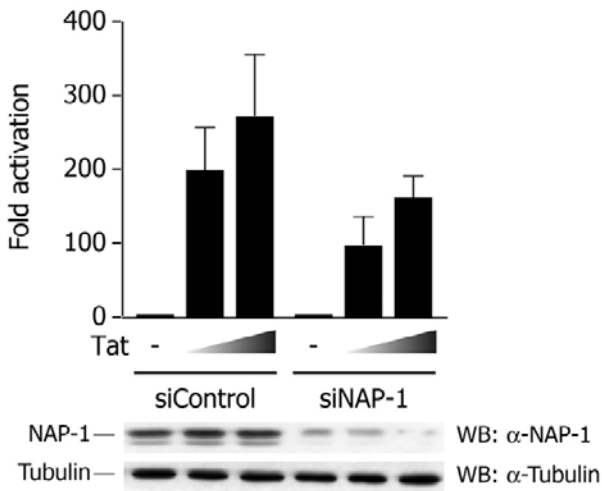


Figure 2.8 hNAP-1 knock-down greatly impaired Tat-mediated LTR transactivation.
HeLa cells were transfected with a specific siRNA against hNAP-1 or a control siRNA, and then transfected with the LTR-luciferase reporter together with Tat. The histogram shows mean ± standard deviation for at least three independent experiments; the results are shown as fold transactivation over LTR-luciferase reporter alone. The western blot at the bottom shows the levels of endogenous hNAP-1 protein and of tubulin as a control in a representative experiment.

After thirty-six hours from the beginning of siRNA treatment, cells were transfected with the reporter plasmid LTR-luciferase together with increasing amounts of HIV-1 Tat-encoding plasmid. Thirty-six hours later, luciferase assays on cell lysates were performed. As expected, in control siRNA-transfected cells we observed a dose-dependent increase in LTR

transactivation in response to Tat, while luciferase activity was greatly impaired in cells in which the expression of hNAP-1 had been silenced, compared to control cells.

Collectively, the results of these experiments strongly indicate that hNAP-1 participates in Tat-mediated control of HIV-1 gene expression.

2.6 Effect of hNAP-1 on HIV-1 infection

To examine the effect of hNAP-1 on HIV-1 expression, we used an HIV-1 vector in which a portion of *nef* gene had been replaced by the firefly luciferase gene; two frameshifts inactivate *vpr* and *env* genes, thus blocking subsequent rounds of viral replication (pNL4.3 R^- E^-). Infectious virus, pseudotyped with the Vesicular Stomatitis Virus pantropic envelope Glycoprotein (VSV-G), was produced by transfection of HEK 293T cells.

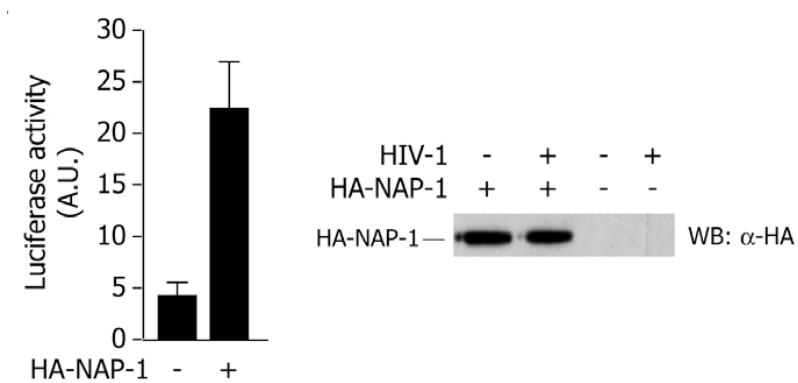


Figure 2.9 Effect of hNAP-1 overexpression on HIV-1 infection.
Overexpression of hNAP-1 enhances LTR transcription upon HIV-1 infection. HeLa cells were transfected with an expression vector for HA-hNAP-1 or with a control vector, and then infected with VSVG-luciferase HIV-1 vector. Luciferase activity was measured 24 hours after infection. The mean ± standard deviation of at least three different experiments is shown. The panel on the right side shows anti-HA immunoblotting to assess HA-hNAP-1 expression in a representative experiment.

Pseudotyped virus was used to infect HeLa cells in which hNAP-1 had been either overexpressed or knocked-down by RNAi. As shown in Figure 2.9, the overexpression of hNAP-1 (as assessed by western blot analysis) resulted in a 5-fold increase of luciferase activity in HA-hNAP-1-transfected cells compared to mock-transfected cells.

Conversely, we tested what was the effect of knocking-down the expression of endogenous hNAP-1. HeLa cells were treated with a control siRNA and with a siRNA directed against hNAP-1, and infected with the luciferase reporter virus. Twenty-four hours later we performed luciferase assay on cell lysates: consistent with the previous overexpression experiment, luciferase activity was significantly decreased in cells in which the expression of hNAP-1 had been silenced, compared to control cells (Figure 2.10).

Taken together, these results strongly support the conclusion that hNAP-1 also plays an important activating role in the context of HIV-1 expression.

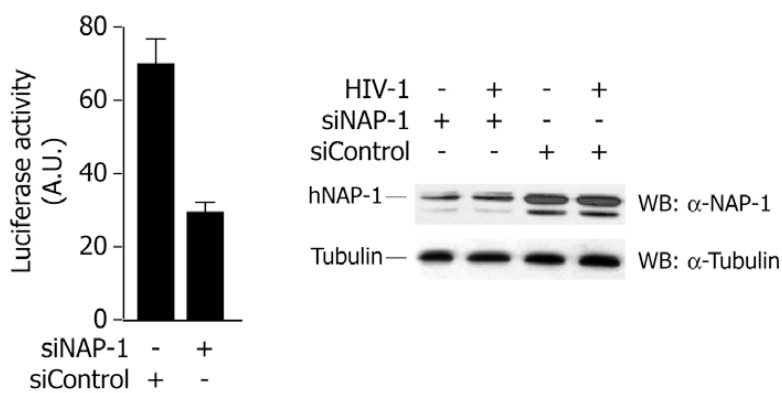


Figure 2.10 Effect of hNAP-1 knocking-down on HIV-1 infection.
Silencing of hNAP-1 impairs LTR transcription upon HIV-1 infection. HeLa cells were treated with an siRNA directed against hNAP-1 or a control siRNA. 48 hours after the beginning of siRNA treatment, cells were infected with the luciferase reporter virus, and luciferase assays were performed on cell lysates 24 hours later. The mean ± standard deviation of at least three different experiments is shown. The panel on the right side shows anti-hNAP-1 immunoblotting to assess the levels of endogenous hNAP-1 and tubulin expression in a representative experiment.

Taken together, these results strongly support the conclusion that hNAP-1 also plays an important activating role in the context of HIV-1 infection.

2.7 p300, hNAP-1 and Tat synergistically activate HIV-1 transcription

Previous work has indicated that hNAP-1 interact with the cellular transcriptional co-activator and histone acetyltransferase p300 . Similar to Tat, the binding of p300 to hNAP-1 involves two domains, one from aminoacids 123 to 230 and a second one C-terminal to aminoacid 290 (Asahara et al., 2002; Ito et al., 2000; Shikama et al., 2000). Since it has been demonstrated that Tat protein directly interacts with p300, and that p300 ia also an essential co-factor for Tat-driven HIV-1 expression (Marzio et al., 1998), we investigated the effects of hNAP-1 and p300 on Tat-mediated transactivation.

For this purpose, HeLa cells were transfected with an LTR-luciferase reporter plasmid and expression vectors for p300 and hNAP-1 together with Tat. According to our previous observations (Marzio et al., 1998), p300 enhanced Tat-driven transcriptional activation (~2 folds over Tat alone), as well as hNAP-1 enhanced Tat-mediated transactivation ~ 2.5 folds over Tat alone (see Figure 2.7). Neither p300 nor hNAP-1 alone, nor hNAP-1 plus p300 had any significant effect on transcription. However, when all the three proteins were overexpressed, transcription was further increased (~3.5 folds Tat plus hNAP-1 plus p300 over Tat alone, Figure 2.11).

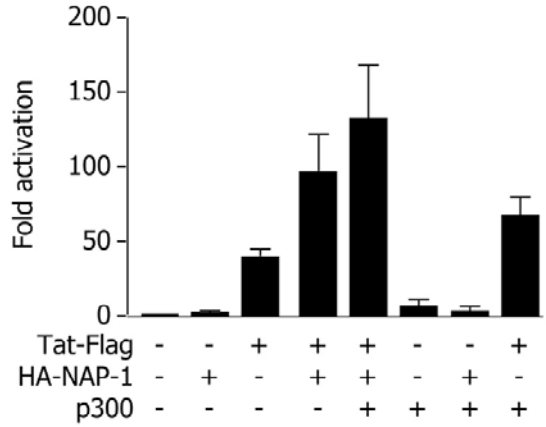


Figure 2.11 hNAP-1, Tat and the acetyltransferase p300 sinergistically activate viral transcription.
HeLa cells were co-transfected with LTR-luciferase reporter plasmid and with vectors for HIV-1 Tat, HA-hNAP-1 and p300, as indicated. The histogram shows the mean ± standard deviation of at least three independent experiments; results are expressed as fold transactivation over LTR-luciferase reporter alone.

To sought out which could be the mechanism underlying the cooperativity between Tat, hNAP-1 and p300 in LTR transactivation, we performed a co-immunoprecipitation assay in the same experimental conditions as reported for the luciferase assay described in Figure 3.11. In cells overexpressing p300, the amount of hNAP-1 protein co-immunoprecipitating with Tat was markedly increased, as shown in the upper panel of Figure 3.12. Since overexpression of p300 did not affect the levels of expression of hNAP-1 or Tat proteins (immunoblots on whole cell lysates, Figure 2.11), our findings are consistent with the possibility that p300 might stsbilize the formation of the Tat-hNAP-1 complex *in vivo*.

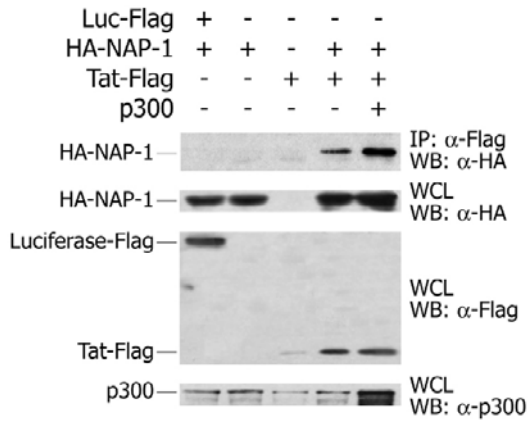


Figure 2.12 p300 enhances Tat.hNAP-1 interaction *in vivo*.
The plasmids indicated on top of the figures were transfected into HEK 293T cells. The upper panel shows immunoblot with anti-HA antibody after immunoprecipitation using an anti-Flag antibody. The lower three panels show western blotting controls from whole cell lysates (WCL) from transfected cells to show the levels of expression of the transfected proteins.

3A. DISCUSSION

Activation of the HIV-1 LTR is a complex event involving the coordinated function of several cellular proteins acting by both releasing the negative inhibition that chromatin imposes on the promoter and inducing the recruitment of elongation-competent RNA Pol II-containing complexes. Tat appears to exert an essential activatory role for both these processes. Tat function is highly dependent on specific interactions with a variety of cellular proteins. In the last decade, a number of laboratories have reported the identification of various cellular factors that mediate Tat function. These factors fall in several broad categories, including members of the basal transcriptional machinery, among which RNA Pol II itself, ubiquitous transcription factors, transcriptional co-activators, histone-acetyltransferases, and others (Brady and Kashanchi, 2005; Gautier et al., 2009; Giacca, 2004; Marcello et al., 2004). Overall, these factors only partially account for the complexity of the molecular mechanisms underlying Tat-mediated transactivation, and the overall process of proviral gene expression into the host cells still remains only partially explained.

3.1 HIV-1 Tat and the cellular histone chaperone hNAP cooperate in the activation of HIV-1 gene expression

Our proteomic screening led to the identification of yet another cellular partner, the histone chaperone hNAP-1, that appears to be involved in mediating Tat function. Full length hNAP-1 protein is 391 amino acid long, contains three acidic domains and has a long KIX-binding domain (Park and

Luger, 2006b) (Figure 3.1 panel A). NAP-1 belongs to the NAP family of proteins, whose members are characterized by an elevated sequence homology (for a recent review, see (Park and Luger, 2006a)). In particular, the KIX-binding domain (Asahara et al., 2002) and the long C-terminal acidic domain are the most conserved among NAP histone chaperones, as it is demonstrated by 47% and 68% amino acid homology in the two regions respectively between hNAP-1 and hSET-TAF-I, another member of the NAP family (Kawase et al., 1996; von Lindern et al., 1992)(Figure 3.1 panel B). Crystal structures of both yNAP1 and hSET/TAF-I were determined and their comparison revealed that the two proteins were folded similarly (Muto et al., 2007; Park and Luger, 2006b).

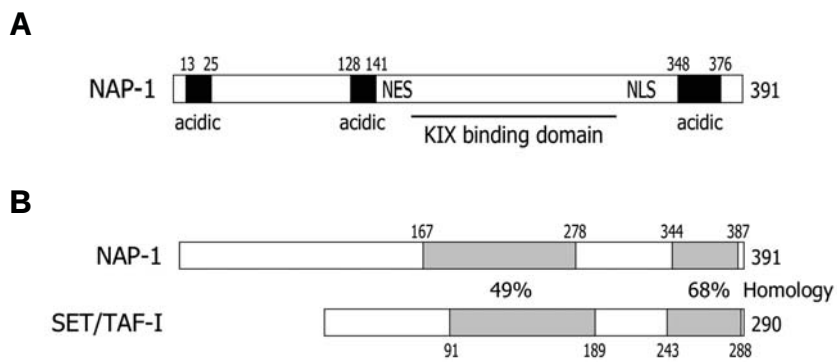


Figure 3.1 Schematic representation of hNAP-1 and sequence homology between hNAP-1 and hSET/TAF-I.
A. The acidic domains of hNAP-1 are rapresented as black boxes, with the indication of their boundary amino acids. The NLS, NES and KIX-binding domains are marked. **B.** The most conserved regions between hNAP-1 and hSET/TAF-I are shown by grey boxes, and the percentage of homology is indicated.

Interaction between Tat and hNAP-1 was confirmed both *in vitro* and inside the cell, and the domains that are involved in the binding were mapped, in both hNAP-1 and in Tat proteins. The basic region of Tat was found to bind two separable domains within hNAP-1, one internal from aminoacids 162 to

290 and one C-terminal from residues 291 to 392. These domains correspond to a series of alternate α- helix/β-sheet regions that are known to be involved in the interaction with histones and other cellular proteins (see (Park and Luger, 2006a, b) and references therein). Of notice, the observation that Tat does not bind the highly homologue C-terminal acidic domain of hSET/TAF I, argues in favor of a specific interaction between Tat and hNAP-1 which is not merely based on electrostatic interactions.

The relevance of the detected interaction between Tat and hNAP-1 was further reinforced by the observations that the over-expression of hNAP-1 stimulated Tat-mediated transactivation of the LTR in a dose-dependent manner, as well as HIV-1-expression in the context of viral infection. Conversely, the down-regulation of the protein by RNAi greatly impaired viral expression.

To our knowledge, this is the first demonstration of an interaction between Tat and a histone chaperone and a first proof of the involvement of this class of proteins in the regulation of proviral transcription. Of interest, the recent finding that Rev also binds hNAP-1, resulting in an alteration of the oligomerization state of the protein and in a consequent increase in Rev aviability, suggests that hNAP-1 can modulate HIV-1 infection exerting its positive role at several steps of the viral life cycle (Cochrane et al., 2009).

Furthermore, it is worth mentioning that interactions between hNAP-1 and other viral transactivators, such as the E2 protein of HPV (Rehtanz et al., 2004), the Tax protein of HTLV-1 (Sharma and Nyborg, 2008) and the EBNA1 (Holowaty et al., 2003; Wang and Frappier, 2009) protein of EBV, have also been described.

Of notice, and in contrast to our expectations, the proteins we identified in our proteomic screening did not detect several of the cellular proteins

previously reported to associate with Tat and to mediate some of its functions. Several possible reasons might explain this outcome.

Our proteomic screening was conducted by immunoprecipitating a Flag-epitope tagged version of Tat (which was fully active transcriptionally) followed by RNase/DNase treatment, elution with a Flag peptide and resolution of Tat-associated proteins by gradient gel electrophoresis. In particular, we found that RNase treatment was essential to avoid the purification of a vast number of RNA-binding proteins unspecifically co-immunoprecipitating with Tat (data not shown). It might well be envisaged, however, that this clearing step might also affect the binding of Tat to some of its known partners, the interaction of which is strengthened by RNA bridging. In addition, RNA removal also frees the basic domain of Tat, thus rendering this region available for the interaction with hNAP-1 (although we also demonstrated that the binding of the two proteins occurred even in the absence of the DNase/RNase treatment).

An additional explanation for the lack of other known Tat partners in our screening relates to the relative abundance of hNAP-1 in the cells, compared to other proteins such as p300 or P/CAF HATs, or Cyclin T1. Since our method relied on the identification of protein bands in silver-stained gels, a likely possibility is that we missed the detection of lower abundance proteins.

Finally, it is worth however noting that other proteomic screening aimed at the identification of cellular partners to other proteins also failed in identifying obvious candidates, while successfully discovered new factors essential for the function of the investigated proteins (Berro et al., 2006; Cochrane et al., 2009; Janardhan et al., 2004).

3.2 p300, hNAP-1 and Tat synergistically activate HIV-1 transcription

There is growing evidence that hNAP-1 plays important roles during transcriptional activation (Ohkuni et al., 2003; Rogner et al., 2000; Steer et al., 2003; Walfridsson et al., 2007). In particular, hNAP-1 and other histone chaperones both cooperate with ATP-dependent chromatin remodeling complexes (Asahara et al., 2002; Nakagawa et al., 2001) and participate in the formation of protein complexes also containing p300/CBP (Asahara et al., 2002; Ito et al., 2000; Rehtanz et al., 2004; Shikama et al., 2000). Taken together, these observations clearly suggest that hNAP-1 may serve as an interaction hub between transcriptional co-activators and chromatin. As far as p300/CBP is specifically concerned, p300 has been shown to directly bind the C-terminus of hNAP-1, namely the same region that is also involved in binding to Tat. Since the basic domain of Tat is also involved in binding to p300 (Marzio et al., 1998), we cannot rule out the possibility that p300 might act as a scaffold for the simultaneous interaction with the two proteins. While further biochemical studies are clearly needed to ascertain this possibility, it is of interest to observe that the over-expression of all the three proteins together determined an increase in the levels of LTR transcription that is higher than those obtained by over-expression of either p300 or hNAP-1 alone together with Tat. In addition, expression of p300 did not affect the levels of hNAP-1 or Tat proteins, but markedly increased their binding *in vivo*. This observation is again in favour of the view that p300 might exert a stabilizing role on the Tat-hNAP-1 interaction. This possibility would be consistent with the proposed function for hNAP-1 in regulating transcription in all p300-dependent promoters (Rehtanz et al., 2004; Sharma and Nyborg, 2008; Shikama et al., 2000).

3.3 Models for hNAP-1-mediated chromatin dynamics and transcriptional activation at HIV-1 promoter

What might be the actual mechanism by which hNAP-1 might facilitate Tat transactivation?

(i) First, over-expression of hNAP-1 significantly increases the overall levels of Tat inside the cells. This result is consistent with the possibility that the positive effect exerted by hNAP-1 on HIV-transcription is due to a hNAP-1-mediated stabilization of the Tat protein itself.

(ii) Second, and more relevant to a specific and direct role of hNAP-1 on the LTR promoter, we envision that hNAP-1 might play a positive role on HIV-1 transcription because it relieves repression exerted by chromatin at the HIV-1 promoter. Earlier results have indicated that the acetylation of histones by p300 helps transfer histones H2A and H2B from nucleosomes to hNAP-1 (Ito et al., 2000) and that, at least *in vitro*, the absence of histones H2A and H2B correlates with increased gene activity, probably by decreasing the level of chromatin folding (Baer and Rhodes, 1983; Chan and La Thangue, 2001). On the basis of these findings and of previous observations made at the level of the HTLV-1 promoter (Sharma and Nyborg, 2008), we can speculate that hNAP-1 and p300, brought to the LTR promoter through their interaction with Tat, might cooperate in the creation of an open-chromatin environment, favorable for gene expression. In addition, it is likely that histone loss results from a process of nucleosome eviction rather than nucleosome sliding, as oserved for the H3-H4 histone chaperone Asf1 (Anti-Silencing Factor 1) in yeast (Adkins et al., 2004).

Supporting this model, a recent genome-wide analysis in fission yeast has revealed that chromatin remodeling factors and NAP-1 co-localize within

promoter regions, and cooperate to disassemble nucleosomes near the TSS, an event that is linked to changes in the levels of histone acetylation (Walfridsson et al., 2007). SWI/SNF ATP-dependent chromatin remodeling complexes and acetylated histones have been shown at the LTR promoter upon transcriptional activation as well (Agbottah et al., 2006; Henderson et al., 2004; Lusic et al., 2003; Treand et al., 2006).

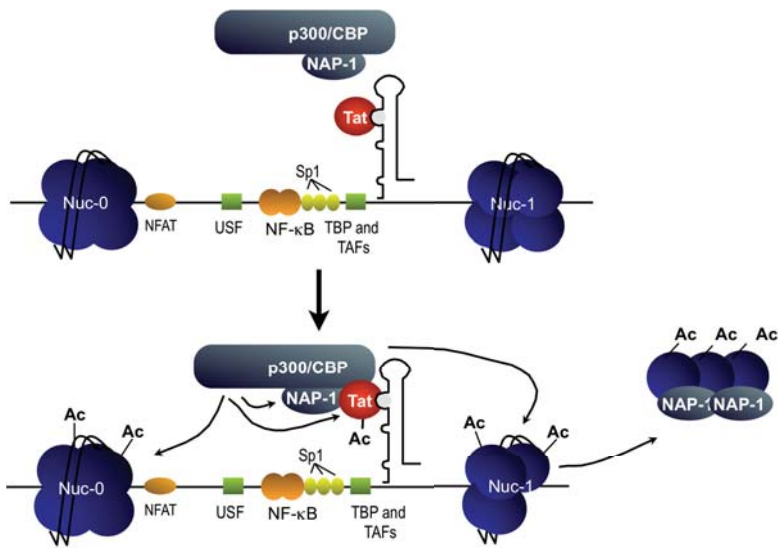


Figure 3.2 A proposed model for Tat and hNAP-1 synergy in the LTR promoter transactivation (see text for details).

Further understanding into this model might arise from ChIP experiments, which would allow one to assess the recruitment of hNAP-1 at the HIV-1 promoter upon transcriptional activation, as well as from biochemical approaches, such as micrococcal nuclease assays, that would address nucleosome occupancy on *in vitro*-chromatinized LTRs.

3.4 HIV-1 latency

Our study unraveled a novel mechanism exploited by HIV-1 Tat to overcome the repression that the chromatin environment exerts on viral gene expression. In this context, it is worth mentioning that highlightening the epigenetic mechanisms through which chromatin governs HIV-1 expression is critical for understanding the phenomenon referred to as proviral latency, which still represents one of the major challenges to find a cure for HIV-1 infection (Richman et al., 2009).

Latent HIV-1 reservoirs are established early during primary infection and constitute a major barrier to eradication, even in the presence of HAART (Highly Active Anti-Retroviral Therapy). The main reservoir is composed of latently infected resting memory $CD4^+$ T lymphocytes that harbour a transcriptionally inactive integrated provirus (Chun et al., 1997; Chun et al., 1995). The extremely long half-life of this cell population, together with a tight control of HIV-1 gene expression allows the virus to maintain hidden copies of its genome, which are in turn able to trigger a novel systemic infection upon interruption of therapy.

Given the importance of the charachterization of this reservoir in order to achieve viral eradication, significant effort has been made to define the mechanisms that establish and maintain HIV-1 latency, most of which operate at the transcriptional level, thus being influenced both by the chromatin environment at the site of viral integration and by the availability of viral and host factors.

Latency was shown to result from transcriptional interference, which depends on the integration site and on the orientation of the provirus in the cell chromosome (Han et al., 2008; Lenasi et al., 2008), or from modification of chromatin conformation and remodeling, especially as far as modifications of histone tails are concerned – for example, latently infected

Jurkat cells show highly deacetylated and trimethylated histones (Pearson et al., 2008). Finally, latency might ensue as a consequence of the absence of nuclear host transcription activators or presence of cellular transcriptional repressors (Coiras et al., 2007). Comprehensive reviewes on HIV-1 latency can be found in the recent literature (Coiras et al., 2009; Lassen et al., 2004; Marcello, 2006; Williams and Greene, 2007).

It will be interesting to understand at what extent the interaction between Tat and hNAP-1 might provide new information on the molecular mechanisms involved in the establishment and maintenance of latency.

4A. MATERIALS AND METHODS

4.1 Protein purification and identification

Twenty-four hours after transfection, ~2 X 10^8 HEK 293T cells were washed once in phosphate-buffered saline (PBS) and lysed on ice in lysis buffer (150 mM NaCl/ 20 mM HEPES pH 7.9/ 0.5% NP-40/ 1 mM EDTA/ 1mM DTT/ protease inhibitor cocktail-Roche). The cell extract was sonicated once and then centrifuged for 15' at 14000 rpm at 4°C. An aliquot of the cleared extract was kept as input, while the rest was incubated with 100 µl of packed and pre-equilibrated Flag M2 agarose beads overnight at 4°C. Beads were rinsed twice in lysis buffer, then treated with DNAse I (Invitrogen, according to manufacturer's instructions) and RNAse A (150 mM NaCl/ 10 mMTris HCl pH 7.5/ 5 mM EDTA/ 10 units RNAse A, for 30' at 37°C) and then washed in the same buffer three times. Immunocomplexes were eluted by adding 500 µg/ml Flag peptide (Sigma, St. Louis, Mo.) in lysis buffer. The eluate was concentrated by standard trichloroacetic acid precipitation and resuspended in 1X sodium dodecyl sulfate- polyacrylamide gel electrophoresis (SDS-PAGE) protein loading buffer. Proteins were then subjected to 6-12% gradient SDS-PAGE and then stained with silver stain. Mass-spectrometry-based protein identification was performed as previously described (Wysocka et al., 2003).

4.2 Cell cultures, plasmids and siRNAs

HeLa and HEK 293T cells were cultured in Dulbecco's modified Eagle's medium with Glutamax (Life Technologies, Inc.) supplemented with 10%

fetal bovine serum (Life Technologies, Inc.) and gentamicin (100 µg/ ml) at 37°C in a humidified 95% air-5% CO_2 incubator.

All hNAP-1 encoding plasmids (wild type and mutants) were a kind gift from Prof. G. Steger (Rehtanz et al., 2004). All other plasmids used have already been described elsewhere (Dorr et al., 2002; Emiliani et al., 2002; Marzio et al., 1998; Pegoraro et al., 2006).

RNA interference (RNAi) with hNAP-1 was performed against the target sequence 5' AAGGAACACGAUGAACCUAUU 3'. For the control we used against the target sequence 5' GGCTACGTCCAGGAGCGCACC 3' of the GFP protein. Synthetic double-stranded RNA oligonucleotides were purchased by Dharmacon (Chicago, IL).

4.3 Co-immunoprecipitation

For co-immunoprecipitation analyses, HEK 293T cells where transfected with the indicated plasmids using the standard calcium phosphate co-precipitation method. Twenty-four hours after transfections cells were washed once in PBS and lysed on ice in 1 ml/dish lysis buffer (150 mM NaCl/ 20 mM HEPES pH 7.9/ 0.5% NP-40/ 1 mM EDTA/ 1mM DTT/ protease inhibitor cocktail-Roche). After sonication, cleared cell extracts were incubated with pre-equilibrated Flag M2 agarose beads on a rotating wheel for four hours at 4°C. Beads were washed twice with 1 ml of lysis buffer, then they are treated with DNAse I (Invitrogen, according to manufacturer's instructions) and RNAse A (150 mM NaCl/ 10 mMTris HCl pH 7.5/ 5 mM EDTA/ 10 units RNAse A, for 30' at 37°C) and then washed in the same buffer three times.

4.4 Antibodies

Anti-hNAP-1 mouse monoclonal antiserum was a kind gift from Y. Ishimi (Ishimi et al., 1985). Mouse monoclonal anti-Flag M2 antibody, mouse monoclonal anti-tubulin, and mouse monoclonal anti-Flag M2 agarose-conjugated beads were purchased from Sigma (St. Louis, Mo.). Rat monoclonal anti-HA high affinity (3F10) antibody was purchased from Roche diagnostics.

4.5 Recombinant proteins

Glutathione S-transferase (GST), GST-Tat, GST-hNAP-1, GST-Tat mutants and GST-hNAP-1 mutants were prepared as already described (Demarchi et al., 1996). Plasmids pcDNA3-Tat101 and pcDNA3-HA-NAP-1 were used as templates to produce the in vitro ^{35}S-labeled Tat and hNAP-1 proteins, respectively, by using the TNT Reticulocyte Lysate System (Promega) according to the manufacturer's protocol.

4.6 GST pull-down assay

GST and GST-Tat recombinant proteins immobilized on agarose beads were pre-treated with nucleases (see "in vitro binding assay"). HEK 293T cells were lysed in 150 mM NaCl/ 20 mM HEPES pH 7.9/ 0.5% NP-40/ 1 mM EDTA/ 1mM DTT/ protease inhibitors (Roche). Recombinant proteins and HEK 293T extracts were incubated one hour and 30 min at 4°C, and washed four times in lysis buffer.

4.7 *In vitro* binding assay

To remove contaminant bacterial nucleic acids, recombinant proteins were pretreated with nucleases (0.25unit/µl Dnase I and 0.2 µg/µl RNase) for 1 hour at 25°C in 50 mM Tris HCl, pH 8.0/ 5 mM $MgCl_2$/ 2.5 mM $CaCl_2$/ 100 mM NaCl/ 5% glycerol/ 1 mM DTT. Subsequently, GST fusion proteins immobilized on agarose beads were washed and resuspended in NETN buffer (20 mM Tris HCl, pH7.5/ 100mM NaCl/ 1 mM EDTA/ 0.5% NP-40/ 1 mM DTT/ 1 mM phenylmethylsulfonylfluoride) supplemented with 0.2 mg/ml ethidium bromide to impede the possible formation of nonspecific interactions between residual DNA and proteins. 400 cpm of ^{35}S-labeled hNAP-1 or Tat101 proteins were added and incubated at 4°C on a rotating wheel. After one hour, bound proteins were washed twice with 0.3 ml of NETN ethidium bromide added, three times with 0.3 ml of NETN without ethidium bromide and once with 0.3 ml of 10 mM Tris HCl pH 8.0/100 mM NaCl. Finally bound proteins were separated by electrophoresis in an SDS 12% polyacrylamide gel. Gels were stained and fixed for one hour with 10% acetic acid/40% methanol/0.1% Coomassie Brilliant blue G250, and destained with 10% acetic acid/ 40% methanol. Dried gels were quantitated by Instant Imager (Packard).

4.8 Luciferase assay

Reporter gene assays were performed using pLTR-luciferase plasmid as a reporter and pcDNA3-Tat101 as an effector in the presence or absence of plasmids pcDNA3-hNAP-1 and pCMV-p300. HeLa cells were transfected using Effectene Reagent (Quiagen, according to manufacturer's protocol), with 100 ng of pLTR-luciferase, 50 ng of pcDNA3-hNAP-1 5 or 25 ng of pcDNA3-Tat101. A Renilla expression plasmid driven by the CMV promoter was cotransfected to standardize each experiment. Cells were harvested fourty-eight hours post transfections, and luciferase activity was measured

with Luciferase assay kit (Promega). The measured activities were standardized by the activities of Renilla, and transactivation was expressed as fold activation compared with the basal activity of LTR-luciferase without effectors. Representetive results of duplicate experiments that were repeated at least three times are shown in the figures with the mean and standard deviation.

For the luciferase assays performed after hNAP-1 knock down by RNAi, siRNAs were transfected using Oligofectamin Reagent (Invitrogen, according to manufacturer's protocol). After 36 hours from the beginning of siRNA treatment, cells were transfected with LTR-luciferase and CMV-Renilla plasmids and increasing amounts of pcDNA3-Tat101. Thirty-six hours later luciferase assay on cell lysates were performed.

When cells were infected with VSV-G-luciferase vectors, luciferase assays were performed twenty-four hours after the beginning of infection. For the gene-silencing experiments, cells were infected fourty-eight hours after siRNA transfection. To normalize luciferase measures, lysate concentrations were determined with Bradford reagent (Biorad, according to manufacturer's protocol).

4.9 Virus production and infections

To produce VSV-G-luciferase vectors, HEK 293T cells were transfected with pNL4.3-luciferase plasmid (Connor et al., 1995; He et al., 1995) and VSV-G encoding plasmid at a ratio 3:1, with the standard calcium phosphate co-precipitation method. Supernatant was collected 48 hours after the beginning of transfections, it was centrifuged and filtered with a 45 µm siringe. Infections with viral supernatant lasted for six hours in the

presence of polybrene (purchased from Sigma Aldrich) at a final concentation of 5 µg/ml.

PART B

1B. INTRODUCTION

1.4 INTEGRATION OF RETROVIRAL DNA INTO THE HOST CELL GENOME AND IMPLICATIONS FOR GENE THERAPY

Although not site-specific, retroviral integration into the host cell genome is not a random event. Indeed, a growing body of evidence has accumulated in the recent years showing how different members of the *Retroviridae* family exhibit distinct target site preferences (Bushman et al., 2005). In this Section, I will focus on the Moloney Murine Leukemia Virus (MoMLV), in particular I will summarize the large-scale studies that gave insights into genomic site selection for MoMLV integration, and I will describe the implications that these studies might have for the use of retroviral vectors for gene therapy.

In this respect, it is important to notice that a mounting interest of the scientific community for the issue of retroviral integration is connected to the occurrence of serious side effects in patients enrolled in gene therapy clinical trials entailing the use of MoMLV-based retroviral vectors.

1.4.1 Retroviral vectors and gene thearapy

Since the first therapeutic human gene therapy clinical trial in 1990, for the treatment of ADA-SCID (Adenosine DeAminase-Severe Combined Immuno-Deficiency) (Blaese et al., 1995) over 1300 clinical trials have been completed, are ongoing or have been approved, involving 28 different countries.

About one-fourth of these gene therapy experimentations have taken advantage of the possibility of transferring genes at high efficiency in replicating cells *ex vivo* using retroviral vectors. Indeed, these vectors stably integrate their genome into the host cell chromosomes, thus allowing the modification of the target cells. This property is particularly desirable when the permanent correction of a genetic defect is sought, as in the case of inherited disorders, or whenever gene transfer is required in permanently dividing cells, such as stem cells. Over one hundred clinical trials have therefore exploited gammaretroviruses, such as the MoMLV, for *ex vivo* transduction of bone-marrow or mobilized hematopoietic precursors followed by the reinfusion of the genetically modified cells back into the patients.

SCIDs are a heterogeneous group of inherited disorders that affect the development or the activity of the immune system, causing a profound reduction or absence of T lymphocyte function. They include ADA-SCID, the X-linked Chronic Granulomatous Disease (X-CGD) and the SCID X-linked 1 (SCID-X1). They all offer reliable models for gene therapy because they are lethal conditions that are otherwise curable only by allogenic bone marrow transplantation if HLA-matched family donors are available. In the past decade, great progress has been achieved in the treatment of SCIDs, especially in the gene therapy of ADA-SCID: treated patients exhibited transduced hematopoietic stem cells stably engrafted and differentiated into myeloid cells containing ADA after a follow-up of four years (Aiuti et al., 2009).

Unfortunately clinical benefit of gene therapy has been obscured by the occurrence of severe adverse effects during clinical trials of SCID-X1.

SCID-X1 is the most common form of SCID (40-50% of all cases), and consists in a profound immunodeficiency caused by the mutation of the

IL2RG gene that encodes for the common cytokine receptor γ chain (γc), component of the high-affinity IL-2 receptor.

After successful *in vitro* experimentation (Candotti et al., 1996; Cavazzana-Calvo et al., 1996; Hacein-Bey et al., 1998; Taylor et al., 1996), rapidly two clinical trials were performed (Cavazzana-Calvo et al., 2000; Gaspar et al., 2004), entailing the use of a MoMLV vector encoding a IL2RG cDNA to transduce autologous $CD34^+$ cells *ex vivo* which were later re-infused into the patients. The number and distribution of T cells increased rapidly, achieving normal numbers compared to age-matched control values in most of the cases.

Unfortunately, leukemiac disorders occurred in five (out of twenty) SCID-X1 treated patients: in four out of the five patients the leukemogenic clone carried the vector integrated close to or within the LIM domain-only 2 (LMO2) proto-oncogene (Nam and Rabbitts, 2006; Royer-Pokora et al., 1991; Royer-Pokora et al., 1995) and activated its transcription (Check, 2005; Hacein-Bey-Abina et al., 2008; Hacein-Bey-Abina et al., 2003; Howe et al., 2008). Activation of LMO2 promoted clonal T cell proliferation in these patients.

The phenomenon by which expression levels of cellular genes is altered upon retroviral integration (as in the case of LMO2) is referred to as "insertional mutagenesis". After the occurrence of these events, several studies tried to address the issue of insertional mutagenesis, and revealed that ~20% of genes are up-regulated following retroviral integration in T cells (Cattoglio et al., 2007; Maruggi et al., 2009). Of note, retroviral integration might also result in the down-regulation of the flanking gene, as it has been described in the case of transduction of T-cells with lentiviral vectors (Maruggi et al., 2009), and this is likely relied to the frequency of intragenic insertion and the strenght of the splice/polyA signal carried by

the vector, both higher in the case of lentiviral vectors (see Section 1.5 for further details).

However, the outcome of insertional mutagenesis induced by gene vectors is still poorly undersood, and might depend on a multiplicity of factors, including the architecture of the used vector, intrinsic developmental potential of the target cell, as well as on the extrinsic factors that influence clonal competition. In this respect it is worth mentioning that in the follow-ups of both X-CGD and ADA patients treated with *ex vivo* gene therapy, alterations in gene expression of certain genes close to the integration sites were detected, but this resulted either in non-malignant clonal expansion or in no clonal expansion at all, respectively (Cassani et al., 2009; Kang et al., 2009; Ott et al., 2006; Seger, 2008). Collectively, these findings suggest that insertion into certain genes may cause clonal amplification of transduced progenitors *in vivo*. Nevertheless, it is not clear whether clonal dominance is entirely the result of *in vivo* selection, or is favoured by the existence of highly preferred regions of retroviral integration that make clonal amplification more likely to occur (Cattoglio et al., 2007).

1.4.2 MoMLV integration into the host cell genome

The genomic features guiding retrovirus integration site selection have now been charachterized in some detail. First, two parallel studies consisting in large-scale sequencing of integration sites of MoMLV-based retroviral vectors in HeLa cells, revealed that MoMLV integration was favoured near the TSSs of genes. In particular, the authors showed that ~20% of the integrations were in regions proximal to the TSSs of genes (±5 kb). A preference in integration towards the CpG islands was also described (Mitchell et al., 2004; Wu et al., 2003).

An interesting but still controversial issue is that aimed at understanding the possible involvement of cellular proteins in determining the specificity of the integration site. Several cellular proteins have been implicated in anchoring PICs to chromatin, and the fact that some of them are different in the MoMLV or HIV-1 PICs, could be suggestive that they might at least in part account for the specificity of retroviral and lentiviral integration. In particular, several inner nuclear proteins (namely emerin, LAP2 and BAF) as well as LEDGF/p75 have been proposed to be involved in the association between viral cDNA and the host chromosome, but the evidences reported so far are still contradictory (Jacque and Stevenson, 2006; Maertens et al., 2003; Mulky et al., 2008; Shun et al., 2007).

Furthermore, it has been hypothesized that the expression program of a cell type might be instrumental in directing the pattern of integration. Therefore Recchia and co-workers used similar approaches to address retroviral integration in T cells. Indeed, a bias towards the promoters of actively transcribed genes has also been described in T cells (Recchia et al., 2006). Cattoglio and collaborators, beside further confirming these data in $CD34^+$ HSCs (Hematopoietic Stem Cells), also identified CISs (Common Integration Sites) or hot spots of retroviral integration such as proto-oncogenes, growth controlling genes and hematopoietic and immune system developmental genes (Cattoglio et al., 2007) and personal communication by Cattoglio and Mavilio). Interestingly, the same conclusions were drawn in a few recent studies, in which integrations of retroviral vectors in transduced cells from patients treated with gene therapy were analyzed. In particular, gene-therapy-treated SCID-X1 patients carried two-thirds of the insertions in or very close to genes (of which more than a half were highly expressed in $CD34^+$ progenitor cells), and one-fourth af all integrants were clustered as CISs, suggestive that insertion in many gene loci has an influence on cell engrafment, survival

and proliferation (Deichmann et al., 2007; Schwarzwaelder et al., 2007). Similarly, ADA-SCID (a form of SCID resulting from Adenosine DeAminase deficiency) patients subjected to gene therapy, also exhibited preferential vector integrations into the promoters of transcriptionally active genes, although no proto-oncogenes were targeted (Aiuti et al., 2007).

1.5 CHROMATIN ORGANIZATION AND RETROVIRAL INTEGRATION

Multiple evidence suggests that chromatin organization inside the cell nucleus is not random but rather organized into higher order structures that can impact gene regulation (Cremer et al., 2006; Fraser and Bickmore, 2007; Sexton et al., 2007). The HIV-1 provirus behaves as any RNA Pol II-transcribed gene, thus it is likely that it is affected by subnuclear localization and chromatin interactions, similar to cellular genes.

As mentioned before, it was demonstrated that positioning to the nuclear periphery and long-range chromatin interactions established with pericentric heterochromatin, correlate with HIV-1 transcriptional repression and latency (Dieudonne et al., 2009). Furthermore, previous work performed in collaboration with the Proudfoot's group in Oxford allowed our laboratory to demonstrate that HIV-1 forms a transcription-dependent gene loop structure between the 5'LTR promoter and the 3'LTR poly(A) signal. This particular conformational structure is also formed between the LTRs and the Major Splice Donor (MSD) region adjacent to the 5'LTR in HIV-1 provirus (Perkins et al., 2008). The observation that HIV-1 chromatin loops during transcriptional activation raised the question whether this is a charateristics peculiar to HIV-1 or, conversely, is a more generalized feature of all the members of the *Retroviridae* family.

Both gammaretroviral (MoMLV) and lentiviral (HIV-1) LTRs contain a poly(A) signal entirely located within the R region of the 5' and 3' LTRs. Potentially, the presence of polyadenylation signals at both ends of the viral genomes could pose an obstacle to gene expression, however both viruses have evolved a strategy to suppress the promoter-proximal (5'LTR) poly(A) signal.

HIV-1 resolves this “mechanistic dilemma” in several ways: the U3 sequences that are uniquely transcribed from the 3’ LTR have been shown to enhance polyadenylation (Ashe et al., 1995), while suppression of the 5’ poly(A) site has been found to depend on the presence of the downstream MSD and its interaction with U1 snRNP; in addition, the inactivation of the MSD has been shown to result in efficient promoter proximal polyadenylation (Ashe et al., 1997). Hence, the capacity of the HIV-1 genome to form a gene loop structure probably represents a further improvement of the virus to effectively inactivate the promoter proximal poly(A) site.

In this respect it is worth mentioning that MoMLV has a much weaker poly(A) signal, so that MoMLV did not evolve a mechanism to actively suppress the 5’ poly(A), as demonstrated in two previous studies. In one of these studies a mutational inactivation of the MSD in the MoMLV-based minigene system was shown to have little if any effect on promoter-proximal poly(A) site usage (Furger et al., 2001). In the second study, MoMLV vectors (but not HIV vectors) displayed high frequencies of read-through of the 3’poly(A) signal (Zaiss et al., 2002). A leaky 3’ poly(A) signal has the potential to mobilize cellular sequences due to the frequent 3’ poly(A) read-through. This could confer increased probability of activating or capturing cellular genes both 5’ and 3’ to the viral integration site by the means of long-range chromosomal interactions.

Collectively, these observations prompted us to investigate the looping properties of MoMLV, especially in view of the possibility that formation of aberrant chromatin loops between the LTRs and the regulatory elements of neighboring genes might be at the basis of the mutagenic potential of MoMLV vectors.

1.6 CHROMOSOME CONFORMATION CAPTURE (3C)

A powerful manner to detect long-range chromatin interactions is offered by a series of methods allowing one to probe physical proximity between chromatin elements, albeit without specifically identifying the protein components that mediate such interactions.

Chromosome Conformation Capture (3C) technique, the first technique to be developed, allows the analysis of chromatin folding in the native cellular environment, at a resolution that is even higher compared to that of the current microscopy techniques. It was initially developed by Dekker and co-workers, who studied the conformation of a complete chromosome in yeast (Dekker et al., 2002), but it was afterward modified to analyze the folding of complex gene loci in mammalian cells (Tolhuis et al., 2002).

An outline of the 3C technology is provided in Figure 1.14. Cells are treated with formaldehyde to induce both protein-protein and protein-DNA crosslinks, resulting in covalent links between interacting chromatin regions. Cross-linked chromatin is then solubilized and digested using an appropriate restriction enzyme, and digested chromatin is subsequently ligated under diluted DNA concentrations, which strongly promote intra-molecular over random inter-molecular ligation of cross-linked fragments. The cross-links are then reversed, the DNA is purified, and ligation products are detected and quantified by PCR using primers across the newly ligated ends of fragments. The relative abundance of a particular ligation product reflects the frequency with which the two restriction fragments interact inside the cell nucleus.

Plasmid 3C, recently developed by Perkins and collaborators, is based on the same principle as chromosomal 3C, but allows the analysis of transfected DNA, which is purified using a sucrose gradient; hence, a

manipulation of sequences required for gene formation is strongly facilitated (Perkins et al., 2008).

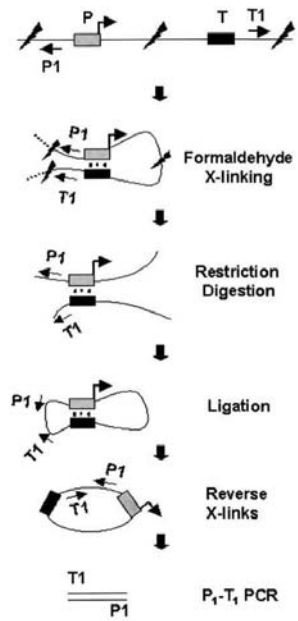


Figure 1.14 Outline of Cromosome Conformation Capture (3C) technology (Adapted from Ansari et al., 2005).

After 3C establishment, other technologies based on the 3C principle have been developed, aimed at broadening the 3C potential, and at increasing throughput, named 4C *(i),* 5C *(ii),* ChIP-loop *(iii),* 6C *(iv)* and Hi-C *(v).*

(i) While 3C permits one to only detect an interaction between known regions, the 4C approaches allow the screening of the entire genome in an unbiased manner for DNA segments that physically interact with a DNA fragment of choice (referred to as the "bait"), without any prior knowledge of their identities. All the DNA fragments captured by the bait are simultaneously amplified via inverse PCR, using two bait-specific primers that amplify from circularized ligation products (Lomvardas et al., 2006; Simonis et al., 2006; Wurtele and Chartrand, 2006; Zhao et al., 2006).

(ii) 3C-carbon copy (5C), compared to standard 3C, gives the possibility to map several hundreds of chromatin interactions for a given locus, thus allowing the generation of a complex matrix of interaction frequencies for that locus, even though the size of the studied region is quite limited. It exploits a multiplex ligation-mediated amplification step to amplify selected ligation junctions, which are subsequently analyzed by microarray detection or by high-throughput sequencing (Dostie et al., 2006).

(iii) While neither 3C nor 3C-derived approaches allow the identification of proteins, which likely mediate the different DNA interactions, a combination of 3C and ChIP was developed (ChIP-loop) to fulfill this requirement. Following formaldehyde cross-linking and enzyme digestion, chromatin is purified on an urea gradient and then immunoprecipitated using an antibody against the protein of interest. Subsequently, the precipitated DNA fragments are ligated and junctions are analyzed by PCR, as in a standard 3C assay (Kumar et al., 2007).

(iv) A further evolution of the ChIP-loop is the Combined 3C-ChIP-Cloning (6C) technology, which detects all the chromatin regions in the nucleus that interact with a given locus of interest, in a specific protein-dependent manner. Such information allows the complete mapping of all the chromatin interactions mediated by a candidate protein, thus providing an important tool to examine the role of specific proteins in nuclear organization (Tiwari et al., 2008).

(v) Recently, an exciting Hi-C approach was developed, a method that identifies in an unbiased manner the chromatin interactions across an entire genome, thus allowing the construction of spatial proximity maps of human genome at a very high resolution (1 megabase). Cells are cross-linked with formaldehyde, and DNA is digested with an enzyme that leaves 5' overhang, subsequently filled with biotinylated residues. After ligation, the resulting DNA sample consists of fragments that were originally in close

spatial proximity, marked with biotin at the junction. A Hi-C library is created by shearing the DNA and precipitating with streptavidin beads to select biotin-containing fragments; finally, the library is analyzed by massive parallel DNA sequencing, producing a catalog of all the interacting fragments (Lieberman-Aiden et al., 2009).

AIM OF THE STUDY

The work described in PART B of this Thesis is part of an ongoing project and it is aimed at assessing the capacity of gammaretroviruses (namely, the Moloney Murine Leukemia Virus, MoMLV) to form a transcription-dependent gene loop stucture as well as lentiviruses (Human Immunodeficiency Virus-1, HIV-1). Next we investigate the possibility that these higher order chromatin structures, if aberrantly generated between viral and host cell regulatory elements, might be at the basis of the mutagenic potential of MoMLV-based retroviral vectors.

2B. RESULTS

2.8 Characterization of a cellular model of retroviral integration

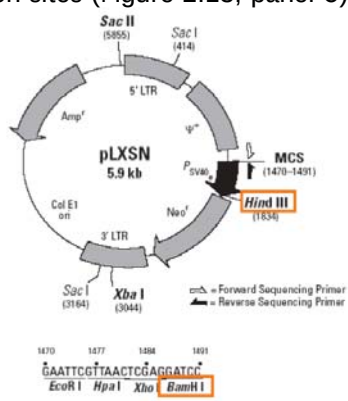
With the aim to develop a cellular system appropriate to study gammaretroviral looping, we first produced a MoMLV-derived retroviral vector by transfecting an amphotropic HEK 293-based packaging cell line with the pLXN plasmid. This plasmid was generated by digestion of pLXSN (Clontech) with BamH I and Hind III, in order to excide SV40 promoter, thus avoiding the bias related to the presence of an additional promoter, and finally obtaining a vector carrying the neomycin resistance gene (neo^R) under the direct control of the MoMLV LTR (Figure 2.13 panel A). After HeLa cell transduction, twenty neomycin-resistent clones were selected and assessed for the presence of the MoMLV provirus by real-time PCR (data not shown).

Mapping of integration sites for seven out of twenty clones was obtained by inverse PCR, performed according to Chun and co-workers (Chun et al., 1997). This approach is essentially based on a nested PCR amplification of the genomic DNA that was first digested with a restriction enzyme and then re-ligated, followed by DNA sequencing. The genomic location of the integrated MoMLV vectors in HeLa clones is indicated in Figure 2.13, panel B, together with the distance from neighboring genes, both 5' and 3' with respect to the integration sites. Interestingly, most of these clones showed integrations in the close proximity of genes, in accordance with previously published literature (Bukrinsky, 2006; Bushman et al., 2005; Cattoglio et al., 2007; Mitchell et al., 2004; Recchia et al., 2006; Wu et al., 2003). Furthermore, two out of seven clones carried the vectors integrated within the IGF1R (Insuline-like Growth Factor 1 Receptor) gene (in the fourth and

in the second intron, in clones 7 and 11, respectively). In this respect, it is worth mentioning that IGF1R gene is one of the so-called “hot spots” or CISs (Common Integration Sites) of retroviral integration, as revealed in a large-scale mapping of thousands of MoMLV integrations in $CD34^+$ hematopoietic stem cells (Cattoglio et al., 2007).

Finally, we verified the copy number of the integrated vectors by Southern Blotting, digesting DNA with EcoR I and using a probe that encompasses the neo^R gene as indicated in Figure 2.13, panel C. Only one major restriction fragment was apparent in each clone, indicating that a single integration event had occurred; the lenghts of the fragments were consistent with the size that had been predicted following the sequence analysis of the integration sites (Figure 2.13, panel C).

A



B

Clone 1: chr 15q14	19 kb at 5′ side: cardiac muscle alpha actin proprotein 43 kb at 3′ side: aquarius
Clone 5: chr 16q24.2	17kb at 5′ side: hypothetical protein 36 kb at 3′ side: soluble carrier family 7 (SLC7A5)
Clone 6-7: chr 15q26.3	Insuline-like Growth Factor 1 receptor (IGF1R) intron 4-5
Clone 8-10: chr 5q33	138 kb at 5′ side: beta2 adrenergic receptor 38 kb at 3′ side: SH3 domain and tetratricopeptide repeats 2
Clone 11: chr 15q26.3	Insuline-like Growth Factor 1 receptor (IGF1R) intron 2-3
Clone 17: chr 6p12	76 kb at 5′ side: similar to eukaryotic translation elongation factor 1β2 15 kb at 3′ side: hypothetical protein
Clone 19: chr 20q13.1	45 kb at 5′ side: adenosine deaminase 18 kb at 3′ side: WNT1 inducible signaling pathway protein 2 precursor

C

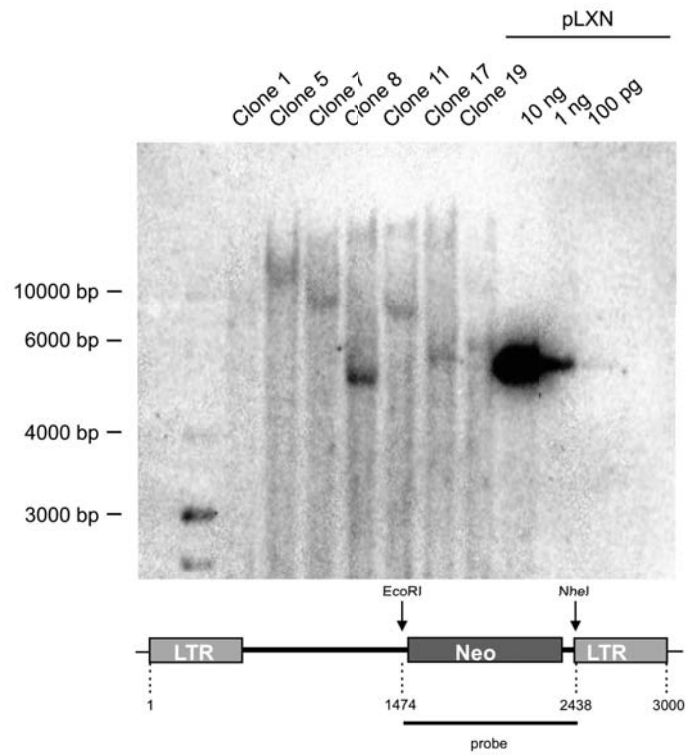


Figure 2.13 Characterization of HeLa stable clones harbouring retroviral integrations.

A. Schematic representation of the retroviral vector used to transduce HeLa cells. The plasmid pLXSN was digested with Hind III and BamH I in order to excise SV40 promoter. The resulting construct was transfected into an amphotrofic HEK 293-based packaging cell line, thus obtaining a vector which carries the neo^R gene under the control of the viral LTR. **B.** List of HeLa clones stably transduced with retroviral vectors and selected by Neomicine resistence. The integration sites were determined for 7 clones by means of inverse PCR. For each clone, the genomic location and the positions of the neighbouring genes are indicated. **C.** Approximately 20 µg of genomic DNA extracted from individual HeLa clones were digested with EcoR I and analyzed by Southern Blotting using the neo^R probe (indicated on the lower part ot the panel). Only one major restriction fragment of varying size is observed in each clone, indicating single integration event had occurred. In the last three lanes on the right, decreasing amounts of the linearized pLXN plasmid were used as positive controls.

2.9 LTR-LTR gene looping: 3C analysis of HeLa clone 11

Next we adapted the 3C methodology in order to render it suitable for the study of our model of retroviral integration. In particular, both major and minor adjustments were introduced into the previously described protocol (Perkins et al., 2008), aimed at adapting the procedure to the analysis of relatively short regions, ~3 kbs, given the size of the vector used (see the Materials and Methods section for details about the 3C protocol).

We started our analysis with clone 11, harboring the retroviral integration within the second intron of the IGF1R gene (see Figure 2.13 panel B). The viral vector sequence was found to be integrated in opposite transcriptional orientation with respect to IGF1R, in agreement with observations made by Recchia and co-workers who detected a significative preference of inverted integrations in retrovirally-transduced T cell lines (Recchia et al., 2006). For the purposes of 3C, we analyzed the flanking sequences of the integrated vector in chromosome 15, and designed PCR primers accordingly. Nuclei from HeLa clone 11 were cross-linked, and the chromatin was prepared and further digested with the restriction enzymes EcoR I and Dpn II. These enzymes met all the requirements of the technique and were tested, being the former a 6 bp-cutter with a single site inside the provirus and the latter a 4 bp-cutter that cleaves the provirus 14 times, thus providing fine mapping of the analyzed interactions. Both EcoR I and Dpn II digestions were controlled by quantitative real-time PCR, with primer pairs designed across the respective restriction sites; in both cases, the extent of digestion was higher than 98% (data not shown). Digested chromatin was then ligated under extremely diluted conditions, thus allowing the ligation of only those sequences that are physically interacting and disfavoring random inter-molecular interactions. Ligated DNA was further purified after

reversal of cross-linking, and ligation products were detected by PCR using appropriate primer sets as indicated in Figure 2.14, panels A and B.

Critical controls are essential for correct interpretation of 3C data; first, PCR amplifications are performed on chromatin that was either cross-linked or not cross-linked, and ligated or not ligated. Second, all the 3C primers were assessed for amplification of *in vitro*-generated 3C products (Figure 2.14 panels A and B, labeled as "controls"). Control template was generated by PCR amplification of the sequences surrounding individual EcoR I or Dpn II sites, and the PCR products were mixed in equimolar amounts, followed by restriction digestion and ligation to include all possible ligation products. Third, to verify the amounts of DNA that was used in 3C reactions all the chromatin preparations were amplified with the unrelated B13 primers, as previously described (Lusic et al., 2003), mapping in the lamin B2 region on chromosome 19.

Representative gels of 3C PCR products are shown in Figure 2.14. Among all the tested primer pairs, a specific PCR product was apparent with primer pairs 11E3-11E2 and 11D1-11D4, in the cases of EcoR I and Dpn II digestions, respectively, and only when chromatin was cross-linked and ligated. No evidence for interactions between the MoMLV promoter and sequences within the retroviral vector was detected, neither for EcoR I nor for Dpn II. The specificity of all the 3C products were then defined by DNA sequencing.

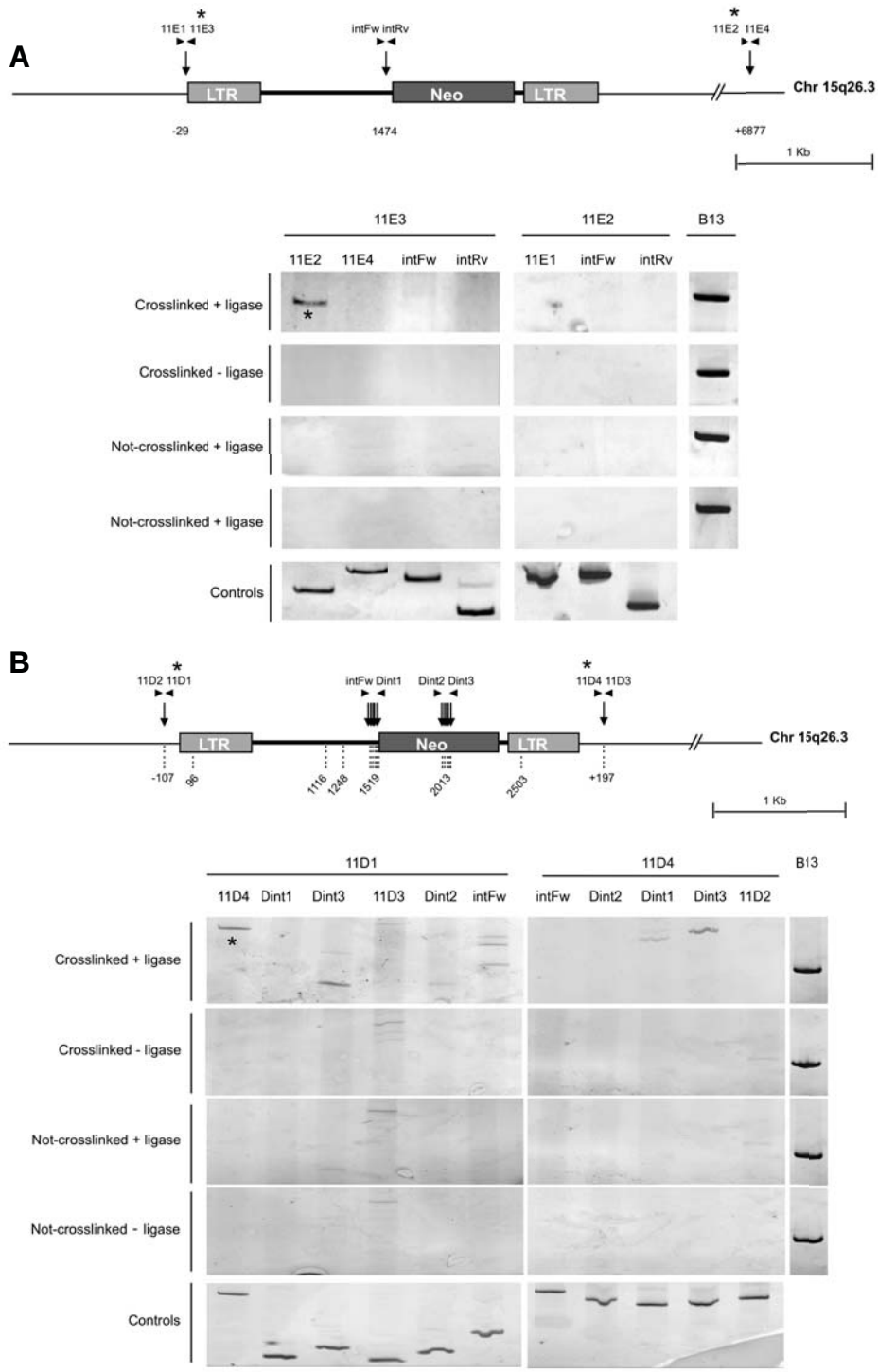

A
11E1 11E3
intFw intRv
11E2 I1E4
LTR
Neo
LTR
Chr 15q26.3
-29
1474
+6877
1 Kb
11E3
11E2
B13
11E2 11E4 intFw intRv
11E1 intFw intRv
Crosslinked + ligase
Crosslinked - ligase
Not-crosslinked + ligase
Not-crosslinked + ligase
Controls
B
11D2 11D1
intFw Dint1
Dint2 Dint3
11D4 11D3
LTR
Neo
LTR
Chr 15q26.3
-107
96
1116
1248
1519
2013
2503
+197
1 Kb
11D1
11D4
B13
11D4 Dint1 Dint3 11D3 Dint2 intFw
intFw Dint2 Dint1 Dint3 11D2
Crosslinked + ligase
Crosslinked - ligase
Not-crosslinked + ligase
Not-crosslinked - ligase
Controls

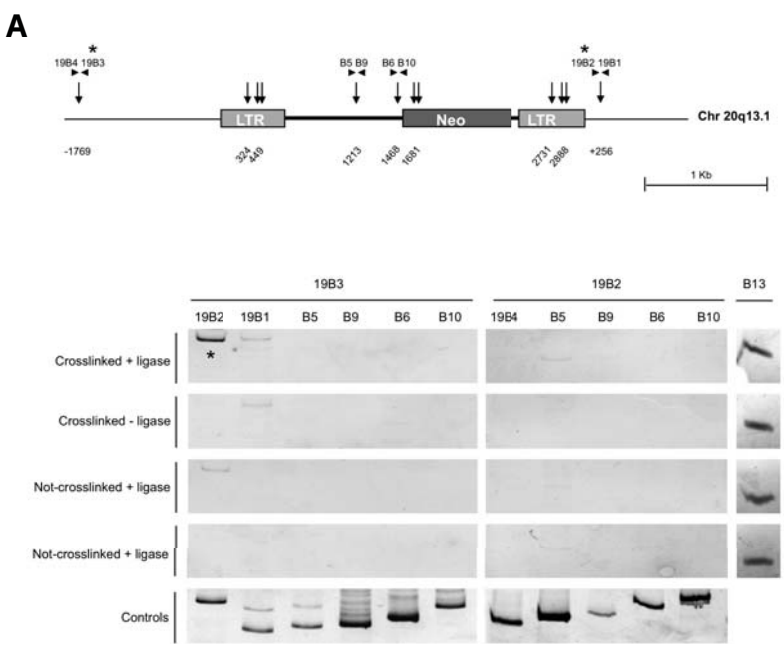
Figure 2.14 EcoR I and Dpn II 3C analysis on HeLa clone 11 cells. **A.** In the upper part of the panel the integrated provirus and flanking chromosomal sequence with restriction sites and primers for EcoR I are represented. Numbers denote distance from 5' (-) or 3' (+) proviral ends. Horizontal arrows indicate primer direction and name; vertical arrows indicate restriction sites. In the lower part, EcoR I 3C analysis in clone 11 cells using either 11E3 or 11E2 primers in combination with one of the other EcoR I primers is indicated. Common PCR primers are shown above the figure, with the second primer shown above each lane. 3C analysis was conducted both on cross-linked and not cross-linked chromatin, before and after ligation. *In vitro*-generated 3C products are labeled as controls. Primers B13 (located in the lamin B2 region in chromosome 19) were used to assess the amount of starting material. Significant interactions, present only in cross-linked and ligated chromatin, are marked by a star. **B.** In the upper part of the panel the integrated provirus and flanking chromosomal sequence with restriction sites and primers for Dpn II are represented. Primers, resctriction sites and distances from the integration site are indicated as in panel A. In the lower part, Dpn II 3C analysis in clone 11 cells using either 11D1 or 11D4 primers in combination with one of the other Dpn II primers is indicated. Primers are indicated as in panel A. 3C analysis was conducted both on cross-linked and not cross-linked chromatin, before and after ligation. *In vitro*-generated 3C products are labeled as controls; primers B13 were used to assess the amount of starting material. Significant interactions, present only in cross-linked and ligated chromatin, are marked by a star.

Primer 11E3 maps in the U3 region of the 5'LTR of MoMLV and amplifies the boundary between MoMLV and the upstream sequences; primer 11E2 amplifies the sequences surrounding the following Eco RI site that is positioned 6877 bases downstream from the MoMLV integration, whereas both 11D1 and 11D4 are in close proximity to the insertion site, 107 bp upstream and 197 bp downstream, respectively. The detection of a specific ligation product using such primers is an indication that the two LTRs are physically juxtaposed inside the cell nucleus. This is the first demonstration that a looping structure involving MoMLV retrovirus actually exists.

2.10 LTR-LTR gene looping: 3C analysis of HeLa clone 19

In order to confirm data obtained from the analysis of clone 11, we extended our studies to another HeLa clone, namely clone 19. As indicated in the panel B of Figure 2.13, this clone carries a retroviral integration in chromosome 20q13.1, in an intergenic region, with the WISP2 (WNT1 Inducible Signallilng Protein 2 precursor) gene 18 kb upstream in the same transcriptional orientation, and the ADA (Adenosine DeAminase) gene 45 kb downstream, in an opposite transcriptional orientation.

To perform 3C analysis on clone 19, we used the same approach as for clone 11, starting from the selection of the suitable restriction enzymes, followed by the design of primers spanning the restriction sites. Ban I (a 6 bp-cutter with 10 sites within the MoMLV provirus) and Dpn II proved to be the enzymes of choice; positions of restriction sites and primers for these two enzymes, both within the provirus and in the flanking regions are indicated in Figure 2.15.



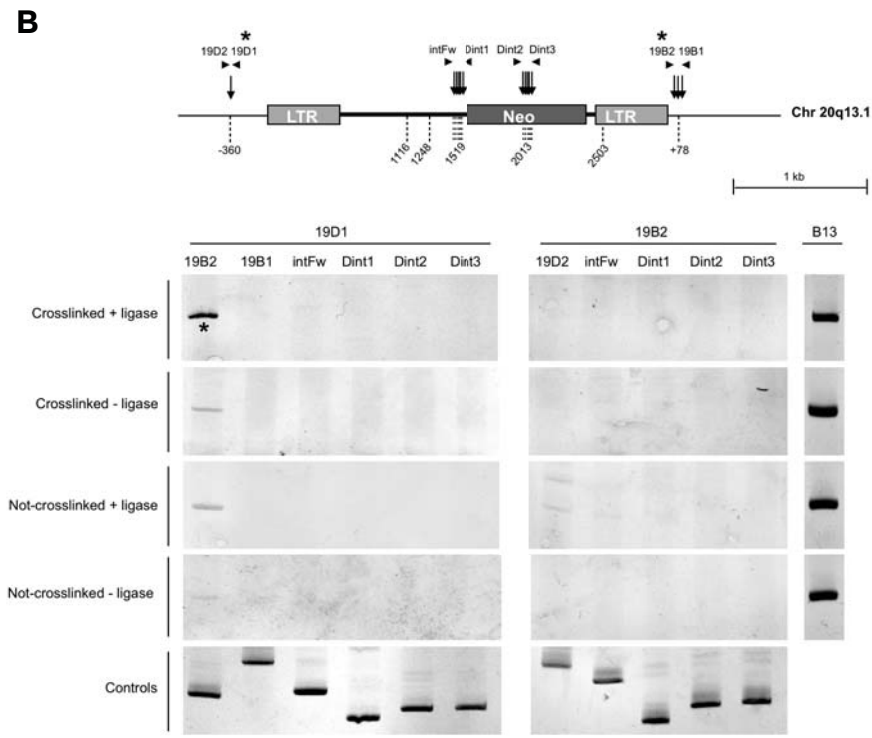


Figure 2.15 Ban I and Dpn II 3C analysis on HeLa clone 19 cells.
A. The integrated provirus and flanking chromosomal sequence with restriction sites and primers for Ban I are represented. Numbers denote distance from 5' (-) or 3' (+) proviral ends. Horizontal arrows indicate primer direction and name; vertical arrows indicate restriction sites. In the lower part of the panel, Ban I 3C analysis in clone 19 cells using either 19B3 or 19B2 primers in combination with one of the other Ban I primers is indicated. Common PCR primers are shown above the figure, with the second primer shown above each lane. Significant interactions, present only in cross-linked and ligated chromatin, are marked by a star. **B.** Restriction sites, primers for Dpn II and distances from the integration site are indicated as in panel A. In the lower part, Dpn II 3C analysis in clone 19 cells using either 19D1 or 19B2 primers in combination with one of the other Dpn II primers is indicated. Primers are indicated as in panel A. Significant interactions, present only in cross-linked and ligated chromatin, are marked

Upon completion of the 3C assay, juxtaposition of the two LTRs was also observed in clone 19, as revealed by the specific PCR bands that were

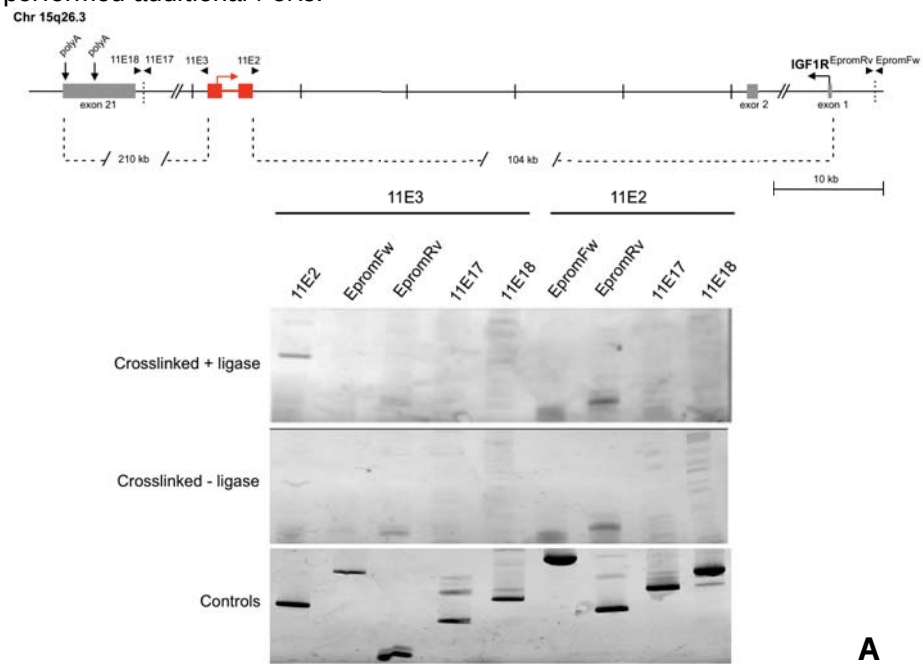
apparent using the primer sets 19B2-19B3 and 19D1-19B2 (Figure 2.15, panels A and B).

2.11 Aberrant looping involving cellular genes: 3C analysis of HeLa clones 11 and 19

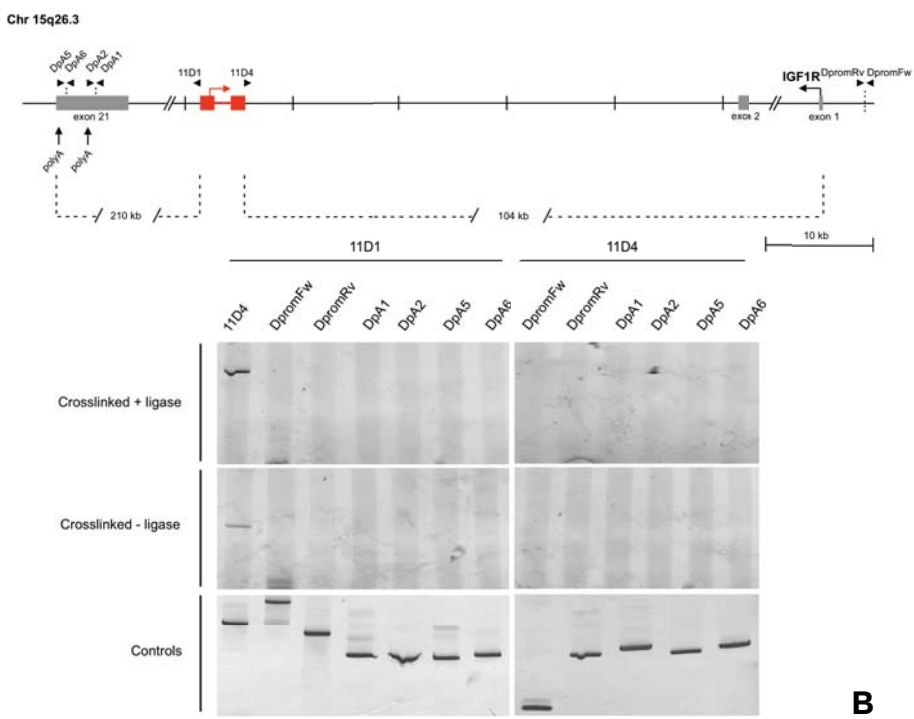
The two sets of data obtained on HeLa clone 11 and on HeLa clone 19 clearly indicate that MoMLV, similar to HIV-1, is able to form a gene-loop structure by juxtaposition of its LTRs. Since for HIV-1 looping capacity is strictly dependent on the presence of the poly(A) sequence (Perkins et al., 2008), and since the poly(A) signal of MoMLV is much weaker compared to that of HIV-1 (Furger et al., 2001; Zaiss et al., 2002), we envisioned that gammaretroviruses might mobilize cellular sequences due to the frequent 3' poly(A) read-through, resulting in the formation of aberrant loop structures and eventually in the deregulation of cellular gene expression. To test our hypothesis and to evaluate the interactions of retroviral sequences with neighboring cellular sequences positioned on the same chromosome, we further explored both clones 11 and 19, and we performed 3C analysis on larger regions of the human chromosomes harboring the proviruses.

As already mentioned, in clone 11 the vector is integrated in opposite transcriptional direction inside the second intron of the IGF1R gene. IGF1R is a long gene composed of 21 exons; the TSS and two different poly(A) sites are positioned at a distance of 104 kbs and 210 kbs with respect to the integration site. To test the possibility of the involvement of such cellular regulatory regions in the formation of atypical loop structures, we designed new primers spanning either EcoR I or Dpn II restriction sites

positioned in close proximity to promoter or poly(A) sequences, and we performed additional PCRs.



A



B

Figure 2.16 EcoR I and Dpn II 3C analysis on HeLa clone 11 cells.
A. The integrated provirus (represented in red) and the IGF1R gene with restriction sites, primers for EcoR I and distances from the integration site are represented in the upper part of the panel. Horizontal arrows indicate primer direction and name, while vertical dashed lines indicate restriction sites. In the lower part of the panel, EcoR I 3C analysis in clone 11 cells using either 11E3 or 11E2 primers in combination with one of the other EcoR I primers is indicated. Common PCR primers are shown above the figure, with the second primer shown above each lane. **B.** Restriction sites and primers for Dpn II and distances from the integration site are indicated as in panel A. In the lower part, Dpn II 3C analysis in clone 11 cells using either 11D1 or 11D4 primers in combination with one of the other Dpn II primers is indicated. Primers are indicated as in panel A.

As indicated in Figure 2.16, several primer sets were assessed, on cross-linked and not cross-linked (data not shown) chromatin. None of these primers detected participation of neighboring cellular sequences in the loop formation, neither using EcoR I nor using Dpn II.

Consistent with the absence of any interaction between retroviral LTRs and regulatory elements of IGF1R, no alterations in IGF1R expression were observed, as revealed by the quantitative Reverse Trancription-PCR (RT-PCR) showed in Figure 2.17.

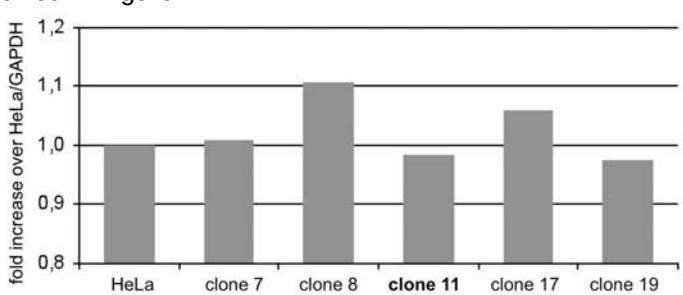
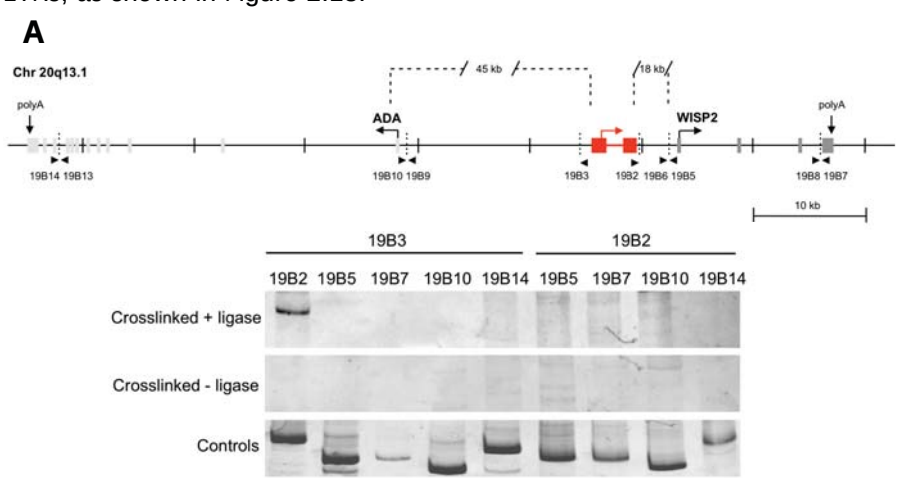


Figure 2.17 Expression analysis of IGF1R gene in clone 11.
Expression analysis of IGF1R in clone 11 cells was compared to HeLa cells and to other clones carrying different retroviral integrations. Total RNA was extracted, reverse transcribed and finally amplified by Real Time PCR. Expression in HeLa cells was set to 1, while expression in other clones was indicated as fold increase over HeLa cells. IGF1R levels were normalized to GAPDH used as a housekeeping gene.

Total RNA was extracted from HeLa and from clones 7, 8, 11, 17, 19, reverse-transcribed using random primers, and cDNA was amplified by Real Time PCR with primers designed for this purpose. Data were normalized for the expression of the housekeeping gene GAPDH (GlycerAldehyde 3-Phosphate DeHydrogenase). Interestingly, no significant differences in IGF1R expression levels were detected between the two clones carrying integrations into the IGF1R gene (namely clones 11 and 7) compared to untransduced HeLa or clones carrying different integrations (clones 8, 17, 19).

Similarly, we analyzed clone 19 for the occurrence of interactions between the LTRs and regulatory elements of flanking genes. As described above, clone 19 carries a retroviral integration in an intergenic region, between WISP2 (located 18 kb upstream in the same transcriptional orientation) and ADA (45 kb downstream, in an opposite transcriptional orientation). Primers were designed spanning the Ban I and Dpn II restriction sites in the proximity of promoters and poly(A) signals of both genes, and used in different combinations on cross-linked, digested and ligated chromatin (data on not cross-linked chromatin are not shown). As for clone 11, none of the tested cellular sequences were found to interact with the MoMLV LTRs, as shown in Figure 2.18.



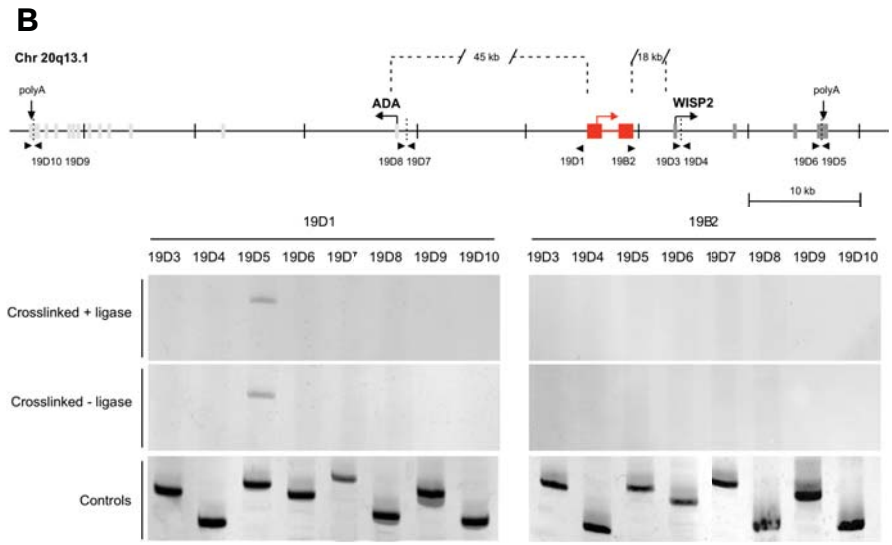


Figure 2.18 Ban I and Dpn II 3C analysis on HeLa clone 19 cells. **A.** The integrated provirus (represented in red) and ADA and WISP2 genes, with restriction sites, primers for Ban I and distances from the integration site are represented in the upper part of the panel. Horizontal arrows indicate primer direction and name, while vertical dashed lines indicate restriction sites. In the lower part of the panel, Ban I 3C analysis in clone 19 cells using either 19B3 or 19B2 primers in combination with one of the other Ban I primers is indicated. Common PCR primers are shown above the figure, with the second primer shown above each lane. **B.** Restriction sites and primers for Dpn II and distances from the integration site are indicated as in panel A. In the lower part, Dpn II 3C analysis in clone 19 cells using either 19D1 or 19B2 primers in combination with one of the other Dpn II primers is indicated. Primers are indicated as in panel A.

Next, we analyzed the expression levels of both ADA an WISP2 genes in clone 19 cells, and compared them to those obtained either from untransduced HeLa cells or from other HeLa clones harboring different integrations (clones 7, 8, 11, 17). Primers were designed across two flanking exons, thus allowing the specific amplification of retrotranscribed RNA. As expected, the absence of long-range chromatin interactions between neighbouring genes and the provirus was in correlation with the

observation that neither expression of ADA nor tf WISP2 was significantly affeected by retroviral integration (Figure 2.19)

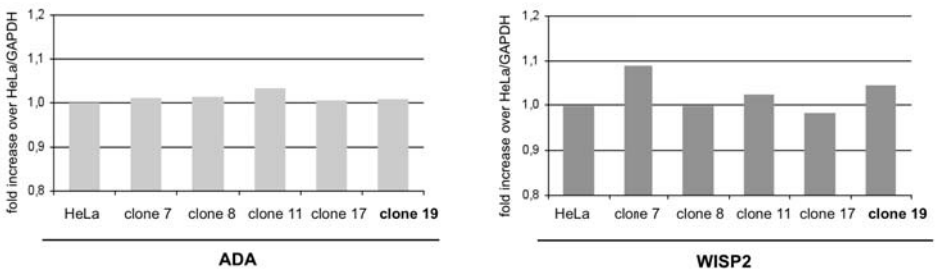


Figure 2.19 Expression analysis of ADA and WISP2 genes in clone 19.
Expression analysis of IGF1R in clone 19 cells was compared to HeLa cells and to other clones carrying different retroviral integrations. Total RNA was extracted, reverse transcribed and finally amplified by Real Time PCR.
Expression in HeLa cells was set to 1, while expression in other clones was indicated as fold increase over HeLa cells. ADA and WISP2 levels were normalized to GAPDH used as a housekeeping gene.

2.12 LTR-LTR gene looping and aberrant looping involving cellular elements: 3C analysis of HaCaT clones 27 and 28

Retroviral integration induces deregulation of nearby genes in ~20% of cases, as assessed by large-scale mapping of integration sites performed in T cells (Maruggi et al., 2009; Recchia et al., 2006). This observation clearly indicates that a higher number of clones should be analyzed in order to define in an unbiased manner the involvement of cellular elements in atypical MoMLV-mediated chromatin loops. Furthermore, the looping properties of MoMLV might be cell type-dependent, in the sense that they might be affected by the transcriptional program of the host cell.
For these reasons, in addition to the analysis on HeLa clones, we further applied the 3C technique to another cell line, a keratinocyte-derived HaCaT

cell line obtained from F. Mavilio (Department of Biomedical Sciences, University of Modena and Reggio Emilia, Modena, Italy). Mavilio and collaborators determined the integration sites for 52 different HaCaT clones, and analyzed expression profiles of genes that lie in a region spanning 50 kbs upstream and 50 kbs downstream of the integration sites (unpublished data, personal communication from Maruggi and Mavilio). Out of 83 genes analyzed, only one was observed to be significantly upregulated, namely SERPINB5 (SERin Proteinase Inhibitor clade B5) gene, in clone 28. HaCaT clone 28 harbors three different integrations, at positions 3p24.2, 10q26.13 and 18q21.33. Among the six genes analyzed in the 100 kbs across the three integrations in clone 28, three were found not to be expressed, two were expressed but their expression was not affected by MoMLV integration, and one (SERPINB5) was up-regulated. This cellular setting represented an excellent milieu for studying the physical interactions between retroviral and cellular regulatory elements.

Thus, we extended 3C analysis to clone 28 (mentioned above), as well as to another HaCaT clone, namely clone 27, carrying two proviruses integated in chromosomes 2q37.1 and 15q26.3, in which a slight up-regulation of the LRRC28 (Leucine Rich Repeat Containing 28) gene flanking the integration site has been observed.

For clone 27, we focused our attention on the retrovirus integrated in chromosome 15, and Ban I was the enzyme that met all the requirements of the 3C method (with 12 restriction sites within the provirus). Chromatin was cross-linked, digested, ligated under proper conditions and purified DNA was analyzed by PCR. As shown in Figure 2.20, the only specific band that was detected was the one amplified by the 15B2-15B3 primer set on cross-liked and ligated chromatin, indicating that LTR-LTR juxtaposition ideed occurred also in HaCaT cells.

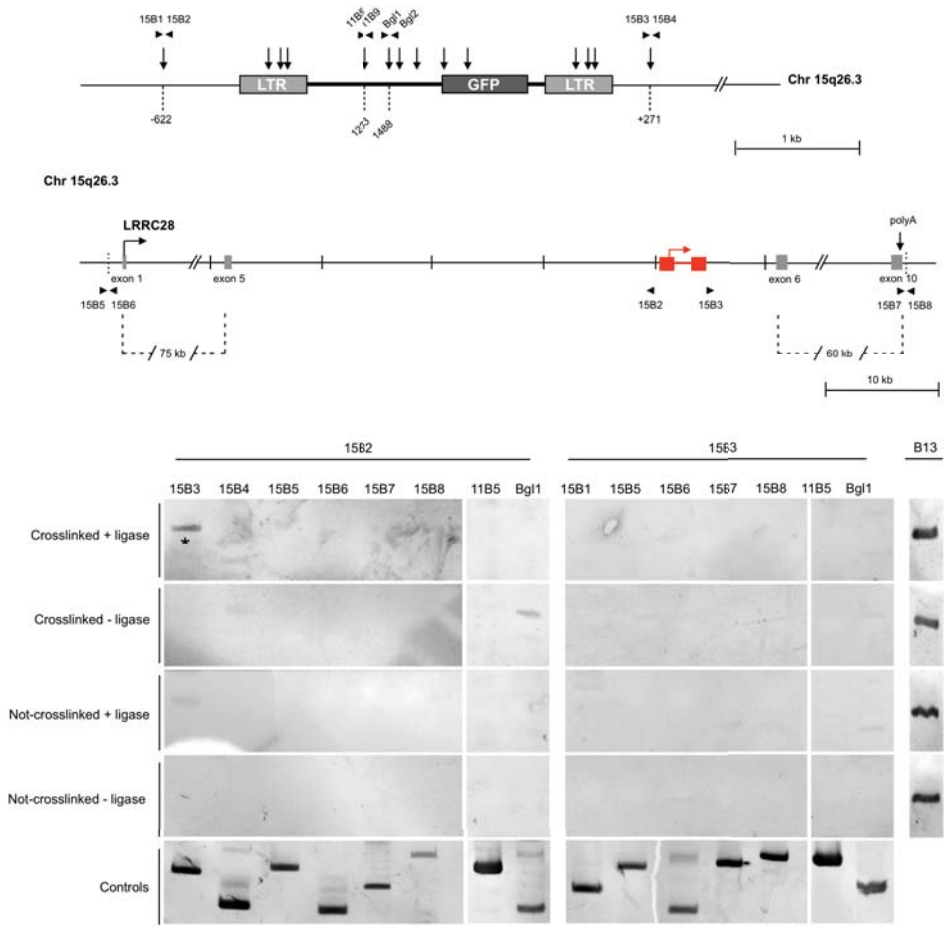


Figure 2.20 Ban I 3C analysis on HaCaT clone 27 cells.
In the upper part of the figure the integrated provirus and flanking chromosomal sequences, encompassing LRRC28 gene, with restriction sites and primers for Ban I are represented. Numbers denote distance from 5' (-) or 3' (+) proviral ends. Horizontal arrows indicate primer direction and name; vertical arrows or dashed lines indicate restriction sites. In the lower part of the figure, Ban I 3C analysis in HaCaT clone 27 cells using either 15B2 and 15B3 primers in combination with one of the other Ban I primers is indicated. Common PCR primers are shown above the figure, with the second primer shown above each lane. 3C analysis was conducted both on cross-linked and not cross-linked chromatin, before and after ligation. *In vitro*-generated 3C products are labeled as controls; primers B13 were used to assess the amount of starting material. Significant interactions, present only in cross-linked and ligated chromatin, are marked by a star.

The provirus is here integrated within the fifth long intron of the LRRC28 gene, which is transcribed in the same direction of the retroviral vector. We also addressed eventual interactions with flanking cellular elements, and for this purpose additional primers were designed along the promoter and the poly(A) signal of LRRC28.

No interactions were detected neither with the LRRC28 promoter nor with its poly(A) sequence (Figure 2.20) and, accordingly, expression of LRRC28 seemed not to be altered by MoMLV integration. An over-expression of 1.5 fold was measured over the mean expression level measured in untransduced HaCaT and in the other 51 analyzed HaCaT clones; however, this was not considered significant (unpublished data, personal communication from Maruggi and Mavilio, Figure 2.21).

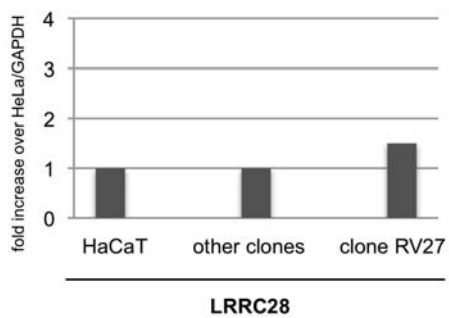


Figure 2.21 Expression analysis of LRRC28 gene in HaCaT clone 27.

Expression analysis of LRRC28 gene in clone 27 cells was compared to untransduced HaCaT cells and to other clones (51) carrying different retroviral integrations. Total RNA was extracted, reverse transcribed and finally amplified by Real Time PCR. Expression in HaCaT cells was set to 1, while expression in other clones was indicated as fold increase over HeLa cells. Expression levels were normalized to GAPDH used as a housekeeping gene.

For clone 28, we analyzed in detail only the provirus integrated in chromosome 18q21.33 (close to SERPINB5). Bgl II and Dpn II with respectively a single and 13 cleavage sites inside the provirus were used. 3C was performed according to our previously adapted protocol, with

primers designed across the restriction sites, and used in various combinations on chromatin purified at the final step of the 3C assay. As shown in Figures 2.22 and 2.23, amplification with primers 18Bg2-18Bg3 (Bgl II digestion) and 18D2-18D3 (Dpn II digestion) resulted in a specific band (confirmed by sequencing) on chromatin that was cross-linked and ligated, indicative that a LTR-LTR loop has occurred. As a negative control, all the combinations between 18Bg2 or 18D2 and internal primers (Bgl1 and Bgl2 for Bgl II digestion; D3, D4, D5, D6 for Dpn II digestion) did not generate any amplification product.

Next, we investigated the possibility that MoMLV establishes interactions with regulatory elements of nearby genes. As shown in Figures 2.22 and 2.23, the retroviral integration in chromosome 18 is in an intergenic region, with the VPS4B (Vacuolar Protein Sorting factor 4B) gene positioned 35 kbs upstream, in an opposite transcriptional orientation, and the SERPINB5 gene 19 kbs downstream, in the same transcriptional orientation of the MoMLV vector. Primers were designed across either Bgl II and Dpn II restriction sites encompassing the promoters and the poly(A) signals of both genes, and tested in PCR reactions together with 18Bg2 and 18D2, respectively. Interestingly, we detected a specific PCR band when 18Bg2-18D9 and 18D2-18D9 sets of primers were used. This result clearly suggested that the MoMLV 5'LTR promoter was juxtaposed and physically interacted with the poly(A) site of the SERPINB5 gene (Figures 2.22 and 2.23).

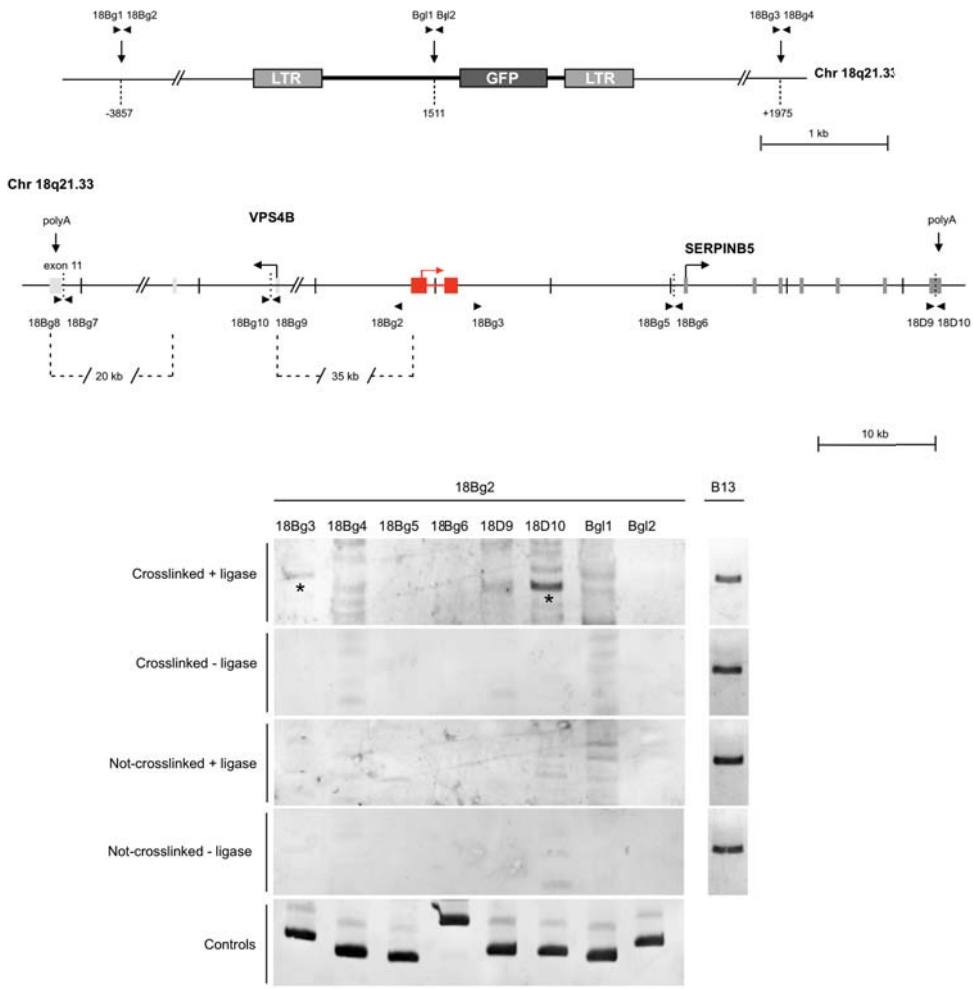


Figure 2.22 Bgl II 3C analysis on HaCaT clone 28 cells.
In the upper part of the figure the integrated provirus and flanking chromosomal sequences, encompassing VPS4B and SERPINB5 genes, with restriction sites and primers for Bgl II are represented. Numbers denote distance from 5' (-) or 3' (+) proviral ends. Horizontal arrows indicate primer direction and name; vertical arrows or dashed lines indicate restriction sites. In the lower part of the figure, Bgl II 3C analysis in HaCaT clone 28 cells using 18Bg2 primer in combination with one of the other Bgl II primers is indicated. Common PCR primers are shown above the figure, with the second primer shown above each lane. 3C analysis was conducted both on cross-linked and not cross-linked chromatin, before and after ligation. *In vitro*-generated 3C products are labeled as controls; primers B13 were used to assess the amount of starting material. Significant interactions, present only in cross-linked and ligated chromatin, are marked by a star.

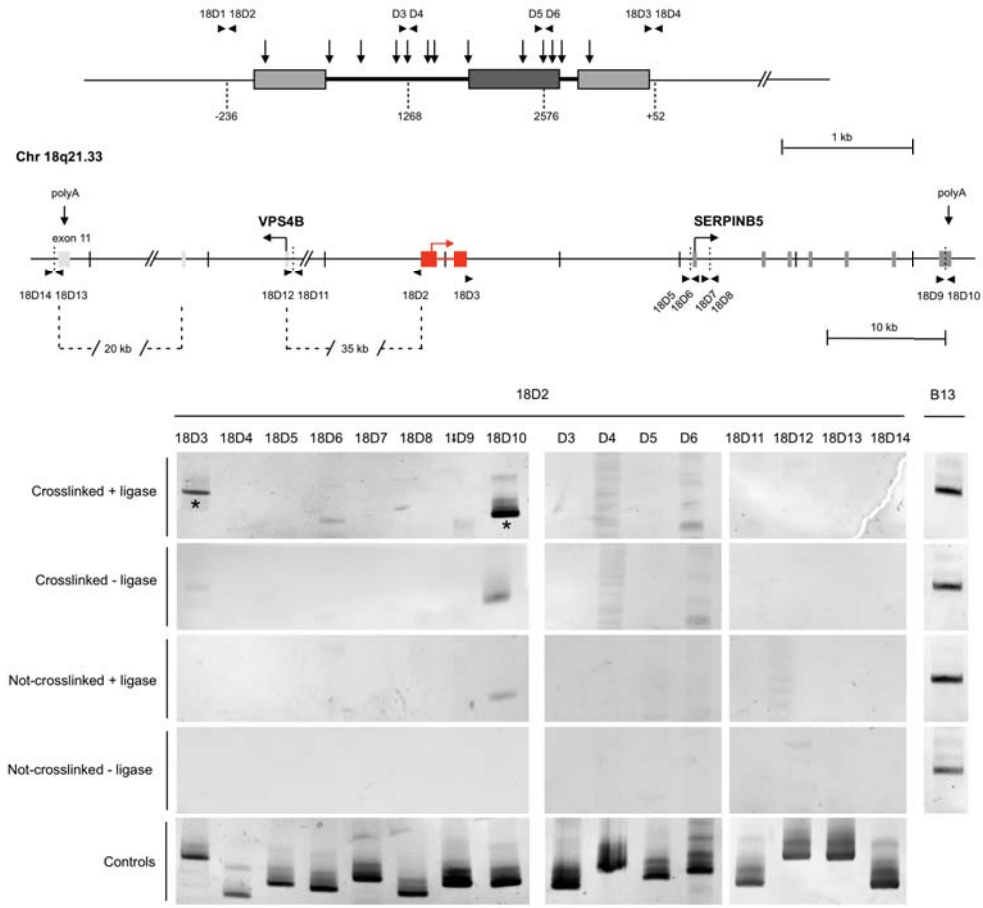


Figure 2.23 Dpn II 3C analysis on HaCaT clone 28 cells.
Same as is 3.22, with representation for Dpn II restriction sites and primers. In the lower part of the figure, Dpn II 3C analysis in HaCaT clone 28 cells using 18D2 primer in combination with one of the other Dpn II primers is indicated. Common PCR primers are shown above the figure, with the second primer shown above each lane. 3C analysis was conducted both on cross-linked and not cross-linked chromatin, before and after ligation. *In vitro*-generated 3C products are labeled as controls; primers B13 were used to assess the amount of starting material. Significant interactions, present only in cross-linked and ligated chromatin, are marked by a star.

This observation was in line with the above mentioned observation that the SERPINB5 gene was up-regulated in clone 28. In particular, the levels of

the SERPINB5 mRNA was 3.5 fold higher than the mean expression levels measured in untransduced HaCaT and in the other 51 analyzed HaCaT clones (unpublished data, personal communication by Maruggi and Mavilio, Figure 2.24).

Conversely, VPS4B was not affected by retroviral integration, and indeed no interactions were detected between this gene and the MoMLV LTRs (Figure 2.23).

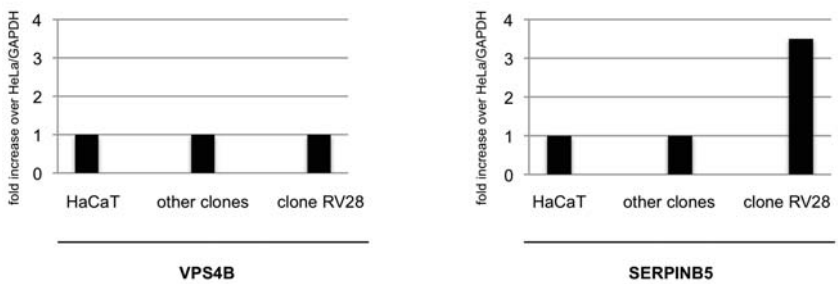


Figure 2.24 Expression analysis of VPS4B and SERPINB5 genes in HaCaT clone 28.

Expression analysis of VPS4B and SERPINB5 genes in clone 28 cells was compared to untransduced HaCaT cells and to other clones (51) carrying different retroviral integrations. Total RNA was extracted, reverse transcribed and finally amplified by Real Time PCR. Expression in HaCaT cells was set to 1, while expression in other clones was indicated as fold increase over HeLa cells. Expression levels were normalized to GAPDH used as a housekeeping gene.

Once demonstrated that the 5'LTR is juxtaposed to the poly(A) of the over-expressed SERPINB5, we next wondered which might be the mechanisms underlining this event (further discussed in Section 3.7 of this Thesis). Given that the vector and the SERPINB5 gene were really in close proximity to each other, and that they are in the same transcriptional orientation, we envisioned that a read-through transcription of the downstream gene might have occurred, consistent with the previously published literature (Zaiss et al., 2002). In order to gain a first insight into the eventual transcriptional activity in the intergenic region, we reverse-transcribed total RNA extracted

from clone 28 and from clone 27 as a control. We then amplified cDNA with primers located in that region (as indicated in Figure 2.25) and with those specific for SERPINB5 gene (between exons 2 and 3). Despite the observation that the region immediately 3' to the vector (2 kbs) was transcribed in clone 28 up to 14 folds compared to clone 27, no differences between the two clones were detected further downstream (18 kbs from the retroviral integration site).

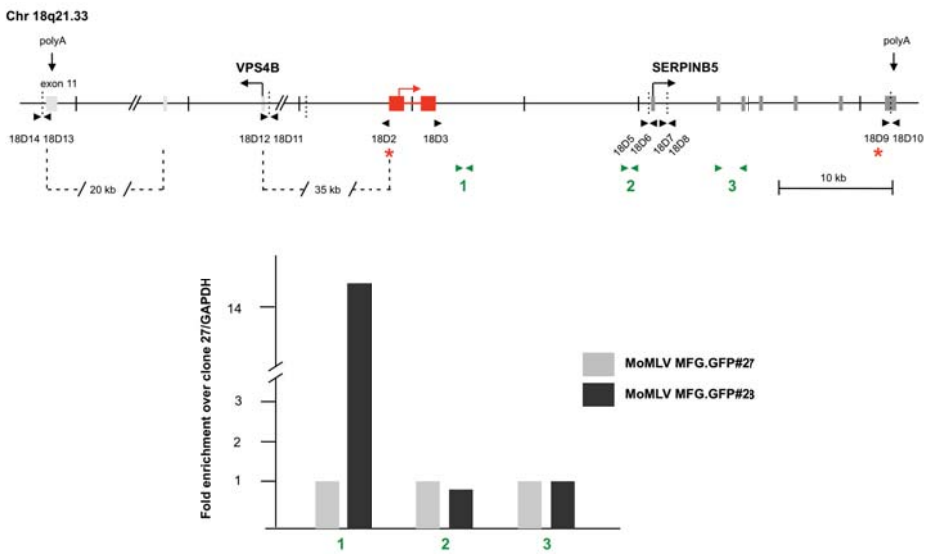


Figure 2.25 Expression analysis of intergenic region between retroviral vector and SERPINB5 gene in HaCaT clone 28.
Total RNA was extracted from clones 27 and 28, retrotranscribed and amplified by Real Time PCR with primer sets indicated as "1", "2" and "3". Expression in HaCaT clone 27 cells was set to 1, while expression in clone 28 was indicated as fold increase over clone 27 cells. Expression levels were normalized to GAPDH used as a housekeeping gene.

In this respect, it is worth mentioning that no difference in SERPINB5 expression in clone 28 was detected, indicating that clone 27 that we used to normalize our expression data is not an appropriate control (in accordance with Mavilio's data). This is not surprising, since data regarding SERPINB5 over-expression were obtained from the analysis of a number of

clones (51), and the estimated 3.5 fold up-regulation was calculated with respect to the mean value of the whole clonal population. Thus, to gain further insights into the mechanisms through which aberrant looping is associated with gene up-regulation, analysis of other HaCaT clones as well as parental HaCaT cell line are currently in progress. Moreover, we will also exploit more sensitive experimental approaches such as Northern Blotting or Run-On asssays (further discussed in Section 3.7 of this Thesis).

3B. DISCUSSION

3.5 Chromosome Conformation Capture (3C) as a powerful technique to detect chromatin loops

Inside the cell nucleus, chromatin is organized into highly dynamic and tightly regulated higher-order structures (Fraser and Bickmore, 2007; Zhao et al., 2009). The genome extensively establishes physical interactions in the form of chromatin loops, that bring distal elements of the chromosome into close physical proximity, thus contributing at various extent to the regulation of gene expression within the three-dimensional context of the nuclear architecture (Gondor and Ohlsson, 2009; Misteli, 2007). Much progress in this research area has been achieved thanks to the development of the Chromosome Conformation Capture (3C) technique (Dekker et al., 2002), which rapidly evolved into 4C-based assays (Simonis et al., 2007), and finally into the Hi-C, a striking method that permits the identification of all chromatin interactions across the entire genome at a resolution of one megabase (Lieberman-Aiden et al., 2009).

In parallel, the 3C technique itself has also greatly improved since its original development, essentially increasing its resolution, that is currently down to a few kilobases. In this way, identification of small gene loops has been recently described, initially in yeast (Ansari and Hampsey, 2005; O'Sullivan et al., 2004; Singh and Hampsey, 2007), and later on also in humans (Martin et al., 2005; Tan-Wong et al., 2008).

3.6 Gene looping between the viral LTRs is a hallmark of transcriptionally active MoMLV proviral DNA

With our work we have demonstrated for the first time that a specific gene loop structure is imposed on the integrated MoMLV *(i)*, that the regions involved in looping are the promoter/enhancer elements encompassing the 5'LTR and the poly(A) site within the 3'LTR *(ii)*, and that this LTR-LTR interaction is associated with gene activation *(iii).*

(i) Taking advantage of the 3C technique we showed here that the 5'LTR and the 3'LTR of the MoMLV are physically juxtaposed to form a gene loop structure.

A gene loop is a peculiar conformation that a gene might adopt by bringing the 3' end in spatial proximity to its 5' start site. The existence of gene loop structures is in accordance with the now widely accepted notion that the transcription machinery physically interacts with the CPF (Cleavage and Poly-adenylation Factor) 3' end-processing and RNA processing factors (Bentley et al., 2005). The first demonstration of a gene loop conformation derives from the studies on yeast genes (O'Sullivan et al., 2004), which showed an essential roles for factors such as RNA Pol II, the Ssu72 and Pta1 components of the CPF 3'-end processing complex and general transcription factor TFIIB in gene loop formation, (Ansari and Hampsey, 2005; Singh and Hampsey, 2007). The first evidence of gene loop occurrence in higher eukaryotes was described for the human mythocondrial heavy-strand rRNA, which was shown to require the human termination factor mTERF (Martin et al., 2005), followed by the demonstration that the BRCA1 (BReast CAncer susceptibility gene 1) gene also juxtaposes its 5' and 3' ends (Tan-Wong et al., 2008). Most relevant for our work, a gene loop conformation was also adopted by the HIV-1

provirus which, upon transcriptional activation, bridges together its promoter and its poly(A) signal located within the 5' and the 3' LTR, respectively (Perkins et al., 2008).

We analyzed four different cell clones, derived two from HeLa and two from HaCaT cells. These clones harboured either single (HeLa) or multiple (HaCaT) retroviral integrations in different chromosomes. In each cell line, 3C was performed using two different restriction enzymes, and reliability of our results was confirmed by a number of critical controls. In particular, the specific PCR band corresponding to the interacting fragments was only apparent on chromatin that had been cross-linked and ligated (however not in cross-linked not ligated chromatin, nor in not cross-linked chromatin); a number of primer set combinations were tested, and proved to be negative, thus confirming that the detected interaction was not merely a false positive obtained due to the length of the region analyzed (in a range of only 3-10 kbs depending on the restriction enzyme); all the primers employed were validated for their efficiency in an *in vitro*-generated template; the amount of loaded chromatin (cross-linked or not cross-linked, ligated or not ligated) was normalized.

(ii) The combination of several restriction enzymes used to digest chromatin during 3C allowed us to obtain a fine mapping of the regions involved in the interaction. Indeed, both Ban I (used for HeLa clone 19 and HaCaT clone 27) and Dpn II (used for HeLa clones 11 and 19 and for HaCaT clone 28) recognized restriction sites within the LTR (three or one, respectively), so that we could conclude that the regions responsible for the loop formation corresponded to the first 95 bp of the 5'LTR, where the Upstream Conserved Region (UCR) of the enhancer resides, and to the last 113 bp of the 3'LTR, corresponding to the R-U5 regions, including the

bipartite poly(A) signal. Interestingly, the UCR contains binding sites for several cellular factors, including YY1 (acting either as a repressor or as an activator of transcription, as in HIV-1 LTR), and NFAT (Nuclear Factor of Activated T cells, an ubiquitous transcriptional activator), and has been identified as an important activator sequence of retroviral enhancers both in murine and human cell types (Wahlers et al., 2002a). Since deletion of these sequences has been shown to attenuate the transcriptional activity of retroviral vectors in hematopoietic and fibroblastic cell lines, it is plausible that these deletions also caused a loss of the "looping" capacity of MoMLV. Moreover, this is further supporting the idea that gene loops are stuctures involved in efficient gene transcription (as it is further discussed below).

In addition, further insights into the determination of the sequences that are essential for the looping capacity might accrue from the application of the plasmid 3C assay, recently developed by Perkins and co-workers (Perkins et al., 2008). We plan to apply this technique taking advantage of MoMLV-based retroviral vectors carrying mutant UCRs or poly(A) signals of diverse origin and length, as well as mutants bearing progressive deletions of the LTRs. Whether factors such as YY1, ELP and NFAT, or even other factors are critical determinants for loop formation still remains elusive, but ChIP-loop assay will prove to be the ultimate technique to gain clues into this issue (Cai et al., 2006; Horike et al., 2005; Kumar et al., 2007).

Finally, it is worth mentioning that many other binding motifs for activating factors are located within the central region of the MoMLV enhancer, immediately downstream of the UCR, including basic helix-loop-helix, ETS (E Twenty-Six), GATA-binding factors, NF-1, and Myb (see (Wahlers et al., 2002b) and references therein). However, according to our 3C analysis with Dpn II (which cuts at position 96 of the LTR), none of the binding sites for these factors are likely to be involved in loop formation. Moreover, we can also exclude the possibility that the MoMLV loop is sustained by the

interaction between the promoter and the enhancer elements that are located within the U3 regions of both the LTRs, as previously demonstrated for the androgen receptor (Wang et al., 2005).

(iii) Gene loops have been shown to be dynamic structures that are strictly dependent on the transcriptional status of the gene. In particular, with the only exception of the BRCA1 gene, in which the 5' and the 3' ends are juxtaposed when expression is repressed, and released upon estrogen-mediated transcriptional induction (Tan-Wong et al., 2008), in general gene loops have been associated with transcriptional activation. This holds true both for several yeast genes, the "looping activity" of which dramatically diminishes upon inhibition of RNA Pol II (Ansari and Hampsey, 2005; O'Sullivan et al., 2004; Singh and Hampsey, 2007), and for the human heavy-strand mithocondrial DNA (Martin et al., 2005). Accordingly, juxtaposition of the two LTRs in the HIV-1 provirus was observed upon stimulation with Tat or phorbol esters, confirming the transcription-dependence of gene looping (Perkins et al., 2008). How exactly looping might promote transcriptional activation or viceversa, still remains to be determined. CPF complex involved in RNA Pol II recycling (Ansari and Hampsey, 2005; Singh and Hampsey, 2007), as well as ongoing transcription (Perkins et al., 2008) are essential for loop formation. Hence, a following scenario might be envisaged: loops serve to coordinate an efficient co-transcriptional RNA-processing and recycling of the polymerase, as well as to provide scaffolds for an effective reinitiation.

In this context, our study demonstrated that the interaction between the LTRs of MoMLV is linked to gene activation as well, reinforcing the view that looping might be a hallmark of transcriptionally active retroviral DNA. In fact, all the clones that we analyzed were selected for active transcription of the retroviral vector: namely, HeLa clones were selected for

neomicine resistence, directly dependent on the LTR activity. HaCaT clones were selected by FACS (Fluorescence-Activated Cell Sorter) analysis for GFP expression, which is also driven by the LTR (personal communication from Maruggi and Mavilio). In this respect, it is worth mentioning that, since HaCaT cells carry multiple integrations, we still do not know whether the GFP is expressed effectively from the integrant that we specifically considered for 3C analysis. A further confirmation of the link existing between looping and transcriptional activation might be obtained by employing RNA Pol II inhibitors, such as α-amanitin (Lusic et al., 2003), which are predicted to abolish loop formation. Indeed, a similiar approach taking advantage of flavopiridol (an inhibitor of the kinase activity of CDK9) has already been used to confirm that looping of the HIV-1 genome is a transcriptionally-dependent phenomenon (Perkins et al., 2008).

3.7 Aberrant looping as a potential mechanism for insertional mutagenesis

As we already mentioned above, both MoMLV and HIV-1 contain poly(A) signals entirely located within the R region of the 5' and 3' LTRs. While HIV-1 has evolved active mechanisms to suppress the proximal 5' poly(A) since this provides a very strong signal (Ashe et al., 1995; Ashe et al., 1997), MoMLV has a much weaker poly(A) (Furger et al., 2001; Zaiss et al., 2002). A leaky 3' poly(A) signal has the potential to mobilize cellular sequences due to the frequent 3' poly(A) read-through. This could raise increased risk of activating or capturing cellular oncogenes both upstream and downstream to the viral integration site by means of long-range chromosomal interactions, and this is exactly what we addressed in our work.

Despite the vast amount of data on the insertion sites of retroviral vectors performed in several studies and in different cell lines (Bushman et al., 2005; Cattoglio et al., 2007; Mitchell et al., 2004; Recchia et al., 2006; Wu et al., 2003), the molecular basis connecting retroviral integration to aberrant cellular gene expression - insertional mutagenesis - is still almost completely obscure. Hence, our observations suggest for the first time that a physical interaction between the retroviral promoter and cellular regulatory elements of the de-regulated gene, namely the poly(A) signal, might play a critical role.

In summary, for each of the four MoMLV integrations, we have assessed the expression levels of six genes encompassing a window of 100 kbs centered on the integration site (50 kb upstream and 50 kb downstream), namely IGF1R (HeLa clone 11), ADA (HeLa clone 19), WISP2 (HeLa clone 19), LRRC28 (HaCaT clone 27), VPS4B (HaCaT clone 28) and SERPINB5 (HaCaT clone 28). All these genes were expressed in HeLa or in HaCaT cells, but only the SERPINB5 gene was found to be significantly upregulated (3.5 folds). Our findings are consistent with several large-scale studies that extimated the de-regulation of cellular genes activation upon retroviral integration as ~20% of cases in T cells (Maruggi et al., 2009; Recchia et al., 2006). In this respect, a bioinformatic analysis recently demonstrated that retroviral vectors integrate preferentially in genomic regions enriched in cell-type specific subsets of transcription factor-binding sites, indicating that the gene expression program of the target cell is instrumental in directing retroviral integration (Felice et al., 2009). We therefore envisioned that cell-type specific transcription factors might be involved in loop formation as well, and this might contribute to explain why gene de-deregulation and aberrant looping is not observed in not all the cases and the cell types.

Notably, SERPINB5 is a tumor suppressor gene, shown to block the growth, invasion, and metastatic properties of mammary tumors (Zou et al., 1994) and this is in line with the observation that retroviral vectors have a high propensity to target genes involved in cell proliferation and signaling as well as proto-oncogenes (Baum et al., 2003; Cattoglio et al., 2007; Wu et al., 2006).

However, the molecular mechanisms connecting aberrant looping to altered gene expression is still not completely understood. We can just speculate that a read-through activity of the down-stream gene might have occurred, generating chimeric transcripts that include all or part of the coding sequences of the downstream gene (namely, SERPINB5), a possibility which would be consistent with the weak MoMLV poly(A) signal (Zaiss et al., 2002). In our work, we initially investigated this possibility, basically assessing the levels of transcription of the intergenic region between the MoMLV integration and the SERPINB5 gene by RT-PCR. Apparently, an enhanced read-through activity of the RNA Pol II in HaCaT clone 28 (in which SERPINB5 had been up-regulated), compared to clone 27, was detected in a region located immediately downstream of the retroviral integration site. However, when transcription was measured further downstream from insertion site (immediatly upstream of SERPINB5), no differences were appreciated between the two clones, indicating that chimeric transcripts are not generated. Nevertheless, these findings are preliminary, and clone 27 did not prove to be a good negative control (as previously discussed in the section 2.14 of this Thesis). Thus, we cannot rule out the possibility that the 3'LTR-read-through activity is indeed critical in the aberrant looping involving SERPINB5, and crucial information might be obtained by Northern Blotting or Run-On assays that allow a better charachterization and quantification of transcripts deriving from the

analyzed region, as well as by ChIP experiments to check the distribution of RNA Pol II on this region of chromatin.

Another possible mechanism underlining insertional mutagenesis is that a retrovirus integrated in any orientation upstream, within, or downstream of a gene might increase the level of expression by virtue of interactions between an enhancer in one of the LTRs and the cellular promoter of the gene. Should this be the case, a physical juxtaposition of the MoMLV and the promoter of the over-expressed gene would have been expected to take place which, conversely, we did not detect.

3.8 The relevance of retroviral gene loops for gene therapy

In four out of five patients in which gene therapy for SCID-X1 caused the development of leukemia, the MoMLV-based retroviral vector was inserted within (or close to) the LMO2 T cell proto-oncogene, causing its activation (Check, 2005; Hacein-Bey-Abina et al., 2008; Hacein-Bey-Abina et al., 2003; Howe et al., 2008).

Based on the findings obtained so far in our cellular models of retroviral integration, we can reasonably predict that aberrant looping between the retroviral vector and regulatory elements of LMO2 might have occurred in the leukemic patients; to demonstrate this will indeed be our ultimate goal.

Interestingly, Dean and co-workers have recently shown that the NLI/Ldb 1 complex (Nuclear LIM Interactor, acting as a protein-binding interface that facilitates multiple interactions) together with GATA-1/SCL/LMO2 binds the β-globin locus control region (LCR) as well as the promoter region after the transcriptional activation of β-globin gene. These associations coincide with loop formation, between the LCR and the gene (Song et al., 2007). GATA-1 binding motifs have also been found in the MoMLV LTRs (according to

TESS (Transcription Element Search System) *in silico* analysis, http://www.cbil.upenn.edu/cgi-bin/tess/tess). This observation intriguingly suggests that LMO2 itself might play a role as a bridging molecule in loop formation, causing its own up-regulation. A ChIP-loop assay entailing the use of an anti-LMO2 antibody will allow us to directly test this hypothesis.

Overall, a comprehensive understanding of the mechanisms involved in insertional mutagenesis is still far to be achieved. Many questions still remain unanswered, such as the role of clonal selection *in vivo*, as well as to which extent disease- , vector-, or transgene-specific factors might cooperate with insertional gene activation in inducing malignant or pre-malignant transformation (Bushman, 2007). Our analysis will provide critical insights into this field, and hopefully will give important clues for the development of safer vectors in order to further exploit the enormous potential held by gene therapy.

4B. MATERIALS AND METHODS

4.10 Cell cultures and plasmids

HeLa, HEK 293T and the amphotropic HEK 293T-derived cells were cultured in Dulbecco's modified Eagle's medium with Glutamax (Life Technologies, Inc.) supplemented with 10% fetal bovine serum (Life Technologies, Inc.) and gentamicin (100 µg/ ml) at 37°C in a humidified 95% air-5% CO_2 incubator. HeLa clones transduced with the MoMLV vectors were cultured in the same conditions described above, but medium was added with G418 (Sigma) at a concentration of 800 µg/ml. HaCaT cells were cultured in Dulbecco's modified Eagle's medium supplemented with 8% fetal bovine serum, penicillin (25 U/l), streptomicin (25 U/l), and glutamin (2mM).

pLXN plasmid was obtained upon sequential digestion with Hind III and Ban I, filling in of sticky ends with DNA Polymerase I (Klenow) (New England BioLabs), followed by ligation with T4 DNA ligase (New England BioLabs). The resulting plasmid was then checked by sequencing.

4.11 Vector production and clone selection

To produce MoMLV-based retroviral vectors, amphotropic HEK 293T packaging cells were transfected with pLXN plasmid with the standard calcium phosphate co-precipitation method. Cloroquine at a final concentration of 25 µM was added to the medium at the same time when the DNA complexes were added, to enhance transfection efficiency. Supernatant was collected 48 hours after the beginning of transfections,

centrifugated and filtered with a 45 µm siringe, and finally used to transduce HeLa cells. 36 hours after transduction, HeLa cells were splitted 1:10 and put in selection with 800 µg/ml G418. Isolated colonies were maintained in selection for 4-6 weeks.

4.12 Inverse PCR

Chromosomal regions flanking the retroviral vector in transduced HeLa clones were defined by inverse PCR as previously described by Chun and co-workers (Chun et al., 1997). DNA was extracted from HeLa clones with DNeasy Blood and Tissue kit (Qiagen), according to manufacturer's protocol. After EcoR I digestion (overnight at 37°C), DNA was phenol:chloroform extracted and ethanol precipitated, and subsequently properly diluted and ligated with 2000 U of T4 DNA Ligase (New England BioLabs), overnight at 16°C. A first PCR was performed on 20 ng of ligated DNA using the following outwards primers annealing on pLXN plasmid: inverse FW: 5' GTAGGAGACGAGAACCTAA 3'; inverse RV: 5' ATCGAGACATAGACCGCCT 3'. Herculase Enhanced DNA Polymerase (Stratagene) was used in a PCR reaction of 40 cycles of 92°C for 10 s, 52°C for 30 s, 72°C for 15 min and final extension for 5 min at 72°C. The resulting amplification products were purified using QIAquick PCR Purification kit (Qiagen, according to manufacturer's protocol), and 1 µl of the purified DNA was used as template for a second nested PCR with the following primers: 5' TTGAACCTCCTCGTTCGAC 3' and nested RV new 5' TCCTGACCTTGATCTGAAC 3'. Herculase Enhanced DNA Polymerase was used in a PCR reaction of 40 cycles of 92°C for 1 min, 56°C for 30 s, 72°C for 15 min and a final extension of 10 min at 72°C. PCR products were then run on 1% agarose gel, and the resulting band was purified with QIAquick

Gel Purification kit (Qiagen, according to manufacturer's protocol) and sequenced.

4.13 Southern Blotting

DNA was extracted as previously described in (Giacca et al., 1994) Approximately 30 µg of DNA were digested with EcoR I overnight at 37°C (New England BioLabs) and fractionated on a 0.8% agarose gel. After running, gel was subjected to a acid hydrolysis in 0.2 N HCl and, denatured in 1.5 M NaCl and 0.5 M NaOH and blotted overnight in 20X SSC transfer buffer. Membrane was neutralized in 1.5 M NaCl and 0.5 M Tris-HCl pH 7.0, air-dried and UV-crosslinked (0.5 J) before being subjected to pre-hybridation in Denhardts's buffer added with saumon sperm DNA at a final concentration of 150 µg/ml, 4 hours at 65°C. The ^{32}P-labeled probe was obtained upon EcoR I and Nhe I (New England BioLabs) digestion of pLXN plasmid, using Ready-To-Go DNA Labelling Beads (minus dCTP) (Amersham Biosciences, according to manufacturer's protocol). For the hybridation reaction, 10^6 cpm/ml were used, overnight at 65°C. After several washings in decreasing SSC concentrations, membranes were exposed to a phosphor storage screen (Cyclone, Packard).

4.14 Chromosome Conformation Capture (3C)

3C was performed as previously described by Perkins and collaborators (Perkins et al., 2008) with some modifications. Approximately $8x10^7$ cells were crosslinked with 1% formaldehyde for 10 min at room temperature. The reaction was quenched by the addition of glycine to 0.125 M. Cells

were collected and cell pellets were resuspended in lysis buffer (10mM Tris-HCl pH 8, 10 mM NaCl, 0.2% NP-40 plus protease inhibitors, 2ml/10^7 cells) for 10 min at 4°C. Cell lysis was completed with 20 strokes in Dounce homogeniser, then nuclei were washed in the appropriate restriction buffer and resuspended in a suitable volume of restriction buffer containing 0.1% SDS and incubated for 30 min at 37°C while shacking. Triton-X-100 was added to 1% and nuclei were further incubated for 15 min at 37°C to sequester SDS. The crosslinked DNA was digested overnight with the appropriate restriction enzyme (800 U). The enzyme was inactivated by the addition of SDS at a concentration of 2% and incubation at 65°C for 20 min. The extent of digestion was verified by quantitative real-time PCR using Sybr Green detection (BioRad) across the restriction sites. Digested chromatin was properly diluted (depending on the restriction enzyme used) in 800 µl ligation buffer (New England BioLabs) added with 1% Triton-X-100 and 0.1 mg/ml BSA and 4000 U of T4 DNA ligase (New England BioLabs). Ligation conditions were specific for each restriction enzyme, but generally ligation was for 4 hours at 16°C, followed by 30 min at room temperature and overnight at 4°C. Proteinase K was added, and the chromatin mixture was then incubated overnight at 65°C to reverse the crosslinks. The following day, DNA was purified by phenol:chloroform extraction and ethanol precipitation. One microliter of input was used in a PCR reaction of 30, 35 or 40 cycles (depending on the restriction enzyme) of 95°C for 30 s, 62°C for 30 s, 72°C for I min, and final extension for 5 min at 72°C. Primers used for PCR reactions were indicated in Table 4.1.

The control template was generated by PCR amplifying sequences spanning individual restriction sites to be analyzed, mixing them in equal molar amounts, followed by restriction digestion and ligation to include all possible ligation products.

Primer name	Primer sequence
11E1	ACATCTTTGTATCTCACATTTTCCAAGAT
11E2	AGAGTTTCTTTGCCAACCATAATTAGT
11E3	TTTCAATGTTATCATCTGGGAGGAGAT
11E4	CTTTGGTCTCCTTTTCCTTGTCATT
Intfw	TCAAGCCCTTTGTACACCCTAAGC
intRv	ATCCTCATCCTGTCTCTTGATCAGATC
EpromFw	CTCGAAATCTGAACGACTCCCATT
EpromRv	CCTTTGTTTTGCTTTTGTTTTAGCAAA
11E17	ATGCCAGTGCAGCTGATGTAGA
11E18	GATGTAAGAGAAGCCACTGGGC
11D1	TGGCGAATTAATGTTTCTTACAGATG
11D2	CATACAAGAGGACCTCCCAACTCA
11D3	CCCAGGCTTGTCTCGAACTC
11D4	TAAATGAAATGAGTCAAAAGCTTCTTT
Dint1	CTGCGTGCAATCCATCTTGTT
Dint2	TGCCGAGAAAGTATCCATCATG
Dint3	GGCCATTTTCCACCATGATATT
DpromFw	TATTCATTAGATGGCAGTCCTAGGGG
DpromRv	AACACGCCGCATGCACT
DpA1	AATACTGTTGGATCAGGGTTTTGTTCTT
DpA2	GCTATGATGGGCACTACACTGTTAGTG
DpA5	ACTTCCTCTAACTCCAGTGGATTGTTG
DpA6	GCTTAGAAAGCAGACAGGTGTTTTCA
19B1	GGCTCACCGCAGGCTCTA
19B2	TAAAAAATTGAAGATGGCCGGG
19B3	ACACACAGAAGTCACCATCTATAAGAAACA
19B4	CTTTCTGTGACTGTCTTTAGGGACTGTC
B5	CTCACAACCAGTCGGTAGATGTCAAGAAGAGA
B9	AAGGCTTCCCAGGTCACGATGTAGG
B6	TCCCTGGGTCAAGCCCTTTGTACAC
B10	CATCAGAGCAGCCGATTGTCTGTTGT
19B5	AAACACTAGCTTTTTCATTCCTTGCTTC
19B6	CTCTTGGGCTGGAAGGAGCT
19B7	TTATTTGGCAACCAGAGTGTGTTG
19B8	TGGAGCAGAGAAAAAGCTAGACCAT
19B9	CGCTTCTAGCTTGGGCCTG
19B10	ATCGCGCATTTCCTGGAATT
19B13	TCAGGCCAGCACTTATCCATCT
19B14	CTATCCTAGGCCCTTCTTGCTC
19D1	GTGGGGAAATAGACAATAAACACATCAGTA
19D2	CCTCTTCTTAACATCAATTGACACACC
19D3	ATCCAAATGCTGGCTCAGAGTTAG

19D4	CTCAGCATCTTCCCAGGTGAA
19D5	TGTTTGATATAGGCTGTGTGTGTAGGC
19D6	GCTGATGGAAGATGGTCCGT
19D7	GGCCTCCAGAGTCTGCTTCTG
19D8	GCTTCCATTCACTGAGCACTCAGTA
19D9	CAGAACCTCTGAAGACGCCACT
19D10	CATGCCGAGGTATACGTGTGTG
15B1	TTGTTCAGTTTTTGCCTTCCG
15B2	CTTAAATCCCCCTTGTGCCTACA
15B3	CCATTAGGGTCCAGATATCATTTTACA
15B4	CAGTAACATGCCTTAGCTCTGTGC
15B5	TGATTCTTTGTACAGATTGATAACCATGA
15B6	CTATCATTCATTCTTTTGCCTGGAG
15B7	CAGTTTTTGAGGAGTTTCAGCAAA
15B8	TTTCCAGAGGGTGTGTGCTCT
Bgl1	CCTCGATCCTCCCTTTATCCAG
Bgl2	ATGTCAGGGTCAGGGAAGTTTACA
18Bg1	GCCATTTTGCCTCTTCGTGT
18Bg2	CAGGCAATTATCTATCTATTGGGAGACT
18Bg3	GTGGCAGCAAGGAGAAGTATAAGTTAAG
18Bg4	CAGTTACCTCTATGCTGCTTTCACA
18Bg5	CTCATTTTATCGATCTTATTGACAAATGAG
18Bg6	GCAAACTAAGTCTAACCACTCATTGCT
18Bg7	ATTGTCCTTCTGATAAGTATTGGGGA
18Bg8	CGGCCCCTATTCTTTCACA
18Bg9	TAGGGCTTTGCAAAAATTACGTTC
18Bg10	TGTGAGACTCTTCACTTTTGTTGACA
18D1	TTTAGTGGAAATGGAGTCTTGCTATGT
18D2	GGACTTGAATATGGTGCAGTAGCTC
18D3	TGATTGACTACCCGTCAGCG
18D4	TCCCACAATGTCTAATATGTCAAACATT
18D5	TGGCATCTAGCAGAATGAGCTG
18D6	CTACAAATTTCTCAGACACGTTGGTG
18D7	AAGGCTGAGATTTAGGCATATGGACT
18D8	TTCCGCTTAGAAGAATGCAGTTG
18D9	AGGAATTATAGACCTCTAGTAGCTGAAATGC
18D10	AAATGACCAGATGGGACAGCTATC
18D11	TGTGACTGCGCCTTCACTC
18D12	TGACTGCTAGCGCACCGCT
18D13	GCATGTTCTATTTCTCTGGCTCTTG
18D14	CCAGGGTGTAAGTCATAGAGATGCT
15B1	TTGTTCAGTTTTTGCCTTCCG
15B2	CTTAAATCCCCCTTGTGCCTACA

15B3	CCATTAGGGTCCAGATATCATTTTACA
15B4	CAGTAACATGCCTTAGCTCTGTGC
15B5	TGATTCTTTGTACAGATTGATAACCATGA
15B6	CTATCATTCATTCTTTTGCCTGGAG
15B7	CAGTTTTTGAGGAGTTTCAGCAAA
15B8	TTTCCAGAGGGTGTGTGCTCT
B13 fw	GCCAGCTGGGTGGTGATAGA
B13 rv	CCTCAGAACCCAGCTGTGGA

Table 4.1 Primers used for 3C.

4.15 RT-PCR

Total cellular RNA was extracted from HeLa and HaCaT clones with an RNeasy Minikit (Qiagen), according to the manufacturer's protocol. RNA (3 μg) was treated with RNase-free DNase I (30 U, Roche) for one hour at room temperature, and then reverse-transcribed using 20 U MoMLV RT (Gibco-BRL), and 1 mM of random hexameric primers (Invitrogen). cDNA was amplified by quantitative real-time PCR using Sybr Green detection (BioRad). Primers were used at a final concentration of 400 nM and their sequences is indicated in Table 4.2.

Primer name	Primer sequence
IGF1R ex2 fw	TGTGAGAAGACCACCATCAACAA
IGF1R ex3 rv	CGCTTCCCACACGTGCTT
ADA ex3 fw	TCCTGGCCAAGTTTGACTACTACA
ADA ex4 rv	GCCCTCTTTGGCCTTCATCT
WISP2 ex1 fw	CATGAGAGGCACACCGAAGA
WISP2 eX2 rv	ACATGGTGTCGGGCACAG
SERP ex2 fw	ACTTGCTCAAGTGGGTGCTAAAG
SERP ex3 rv	ATGTTACTGTTTGAAATCCAAAGGGT
GAPDH fw	AAGGTGAAGGTCGGAGTCAA
GAPDH rv	AATGAAGGGGTCATTGATGG

Table 4.2 Primers used for RT-PCR

BIBLIOGRAPHY

Adkins, M.W., Howar, S.R., and Tyler, J.K. (2004). Chromatin disassembly mediated by the histone chaperone Asf1 is essential for transcriptional activation of the yeast PHO5 and PHO8 genes. Mol Cell *14*, 657-666.

Agbottah, E., Deng, L., Dannenberg, L.O., Pumfery, A., and Kashanchi, F. (2006). Effect of SWI/SNF chromatin remodeling complex on HIV-1 Tat activated transcription. Retrovirology *3*, 48.

Aiuti, A., Cassani, B., Andolfi, G., Mirolo, M., Biasco, L., Recchia, A., Urbinati, F., Valacca, C., Scaramuzza, S., Aker, M.*, et al.* (2007). Multilineage hematopoietic reconstitution without clonal selection in ADA-SCID patients treated with stem cell gene therapy. J Clin Invest *117*, 2233-2240.

Aiuti, A., Cattaneo, F., Galimberti, S., Benninghoff, U., Cassani, B., Callegaro, L., Scaramuzza, S., Andolfi, G., Mirolo, M., Brigida, I.*, et al.* (2009). Gene therapy for immunodeficiency due to adenosine deaminase deficiency. N Engl J Med *360*, 447-458.

Akoulitchev, S., Makela, T.P., Weinberg, R.A., and Reinberg, D. (1995). Requirement for TFIIH kinase activity in transcription by RNA polymerase II. Nature *377*, 557-560.

Albini, A., Ferrini, S., Benelli, R., Sforzini, S., Giunciuglio, D., Aluigi, M.G., Proudfoot, A.E., Alouani, S., Wells, T.N., Mariani, G.*, et al.* (1998). HIV-1 Tat protein mimicry of chemokines. Proc Natl Acad Sci U S A *95*, 13153-13158.

Albini, A., Soldi, R., Giunciuglio, D., Giraudo, E., Benelli, R., Primo, L., Noonan, D., Salio, M., Camussi, G., Rockl, W.*, et al.* (1996). The angiogenesis induced by HIV-1 tat protein is mediated by the Flk-1/KDR receptor on vascular endothelial cells. Nat Med *2*, 1371-1375.

Alizon, M., Sonigo, P., Barre-Sinoussi, F., Chermann, J.C., Tiollais, P., Montagnier, L., and Wain-Hobson, S. (1984). Molecular cloning of lymphadenopathy-associated virus. Nature *312*, 757-760.

Altaf, M., Saksouk, N., and Cote, J. (2007). Histone modifications in response to DNA damage. Mutat Res *618*, 81-90.

Altman, R., and Kellogg, D. (1997). Control of mitotic events by Nap1 and the Gin4 kinase. J Cell Biol *138*, 119-130.

Ammosova, T., Berro, R., Jerebtsova, M., Jackson, A., Charles, S., Klase, Z., Southerland, W., Gordeuk, V.R., Kashanchi, F., and Nekhai, S. (2006).

Phosphorylation of HIV-1 Tat by CDK2 in HIV-1 transcription. Retrovirology *3*, 78.

Ammosova, T., Jerebtsova, M., Beullens, M., Voloshin, Y., Ray, P.E., Kumar, A., Bollen, M., and Nekhai, S. (2003). Nuclear protein phosphatase-1 regulates HIV-1 transcription. J Biol Chem *278*, 32189-32194.

Angelov, D., Bondarenko, V.A., Almagro, S., Menoni, H., Mongelard, F., Hans, F., Mietton, F., Studitsky, V.M., Hamiche, A., Dimitrov, S.*, et al.* (2006). Nucleolin is a histone chaperone with FACT-like activity and assists remodeling of nucleosomes. EMBO J *25*, 1669-1679.

Ansari, A., and Hampsey, M. (2005). A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping. Genes Dev *19*, 2969-2978.

Asahara, H., Tartare-Deckert, S., Nakagawa, T., Ikehara, T., Hirose, F., Hunter, T., Ito, T., and Montminy, M. (2002). Dual roles of p300 in chromatin assembly and transcriptional activation in cooperation with nucleosome assembly protein 1 in vitro. Mol Cell Biol *22*, 2974-2983.

Ashe, M.P., Griffin, P., James, W., and Proudfoot, N.J. (1995). Poly(A) site selection in the HIV-1 provirus: inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. Genes Dev *9*, 3008-3025.

Ashe, M.P., Pearson, L.H., and Proudfoot, N.J. (1997). The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. Embo J *16*, 5752-5763.

Baer, B.W., and Rhodes, D. (1983). Eukaryotic RNA polymerase II binds to nucleosome cores from transcribed genes. Nature *301*, 482-488.

Bannwarth, S., and Gatignol, A. (2005). HIV-1 TAR RNA: the target of molecular interactions between the virus and its host. Curr HIV Res *3*, 61-71.

Bao, Y., and Shen, X. (2007). Chromatin remodeling in DNA double-strand break repair. Curr Opin Genet Dev *17*, 126-131.

Barboric, M., and Peterlin, B.M. (2005). A new paradigm in eukaryotic biology: HIV Tat and the control of transcriptional elongation. PLoS Biol *3*, e76.

Barre-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C.*, et al.* (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science *220*, 868-871.

Baum, C., Dullmann, J., Li, Z., Fehse, B., Meyer, J., Williams, D.A., and von Kalle, C. (2003). Side effects of retroviral gene transfer into hematopoietic stem cells. Blood *101*, 2099-2114.

Bayer, P., Kraft, M., Ejchart, A., Westendorp, M., Frank, R., and Rosch, P. (1995). Structural studies of HIV-1 Tat protein. J Mol Biol *247*, 529-535.

Benkirane, M., Chun, R.F., Xiao, H., Ogryzko, V.V., Howard, B.H., Nakatani, Y., and Jeang, K.T. (1998). Activation of integrated provirus requires histone acetyltransferase. p300 and P/CAF are coactivators for HIV-1 Tat. J Biol Chem *273*, 24898-24905.

Bennasser, Y., and Jeang, K.T. (2006). HIV-1 Tat interaction with Dicer: requirement for RNA. Retrovirology *3*, 95.

Bennasser, Y., Le, S.Y., Benkirane, M., and Jeang, K.T. (2005). Evidence that HIV-1 encodes an siRNA and a suppressor of RNA silencing. Immunity *22*, 607-619.

Bennasser, Y., Le, S.Y., Yeung, M.L., and Jeang, K.T. (2004). HIV-1 encoded candidate micro-RNAs and their cellular targets. Retrovirology *1*, 43.

Bentley, D.L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. Curr Opin Cell Biol *17*, 251-256.

Berkhout, B., Gatignol, A., Rabson, A.B., and Jeang, K.T. (1990). TAR-independent activation of the HIV-1 LTR: evidence that tat requires specific regions of the promoter. Cell *62*, 757-767.

Berkhout, B., Silverman, R.H., and Jeang, K.T. (1989). Tat trans-activates the human immunodeficiency virus through a nascent RNA target. Cell *59*, 273-282.

Bernstein, B.E., Liu, C.L., Humphrey, E.L., Perlstein, E.O., and Schreiber, S.L. (2004). Global nucleosome occupancy in yeast. Genome Biol *5*, R62.

Berro, R., Kehn, K., de la Fuente, C., Pumfery, A., Adair, R., Wade, J., Colberg-Poley, A.M., Hiscott, J., and Kashanchi, F. (2006). Acetylated Tat regulates human immunodeficiency virus type 1 splicing through its interaction with the splicing regulator p32. J Virol *80*, 3189-3204.

Bharucha, D.C., Zhou, M., Nekhai, S., Brady, J.N., Shukla, R.R., and Kumar, A. (2002). A protein phosphatase from human T cells augments tat transactivation of the human immunodeficiency virus type 1 long-terminal repeat. Virology *296*, 6-16.

Bhat, K.P., Turner, J.D., Myers, S.E., Cape, A.D., Ting, J.P., and Greer, S.F. (2008). The 19S proteasome ATPase Sug1 plays a critical role in regulating MHC class II transcription. Mol Immunol *45*, 2214-2224.

Bieniasz, P.D., Grdina, T.A., Bogerd, H.P., and Cullen, B.R. (1998). Recruitment of a protein complex containing Tat and cyclin T1 to TAR governs the species specificity of HIV-1 Tat. EMBO J *17*, 7056-7065.

Bisgrove, D.A., Mahmoudi, T., Henklein, P., and Verdin, E. (2007). Conserved P-TEFb-interacting domain of BRD4 inhibits HIV transcription. Proc Natl Acad Sci U S A *104*, 13690-13695.

Blaese, R.M., Culver, K.W., Miller, A.D., Carter, C.S., Fleisher, T., Clerici, M., Shearer, G., Chang, L., Chiang, Y., Tolstoshev, P*., et al.* (1995). T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. Science *270*, 475-480.

Blazek, D., and Peterlin, B.M. (2008). Tat-SIRT1 tango. Mol Cell *29*, 539-540.

Boe, S.O., Bjorndal, B., Rosok, B., Szilvay, A.M., and Kalland, K.H. (1998). Subcellular localization of human immunodeficiency virus type 1 RNAs, Rev, and the splicing factor SC-35. Virology *244*, 473-482.

Bor, Y.C., Bushman, F.D., and Orgel, L.E. (1995). In vitro integration of human immunodeficiency virus type 1 cDNA into targets containing protein-induced bends. Proc Natl Acad Sci U S A *92*, 10334-10338.

Bor, Y.C., Miller, M.D., Bushman, F.D., and Orgel, L.E. (1996). Target-sequence preferences of HIV-1 integration complexes in vitro. Virology *222*, 283-288.

Boulanger, M.C., Liang, C., Russell, R.S., Lin, R., Bedford, M.T., Wainberg, M.A., and Richard, S. (2005). Methylation of Tat by PRMT6 regulates human immunodeficiency virus type 1 gene expression. J Virol *79*, 124-131.

Bovolenta, C., Camorali, L., Lorini, A.L., Ghezzi, S., Vicenzi, E., Lazzarin, A., and Poli, G. (1999). Constitutive activation of STATs upon in vivo human immunodeficiency virus infection. Blood *94*, 4202-4209.

Brady, J., and Kashanchi, F. (2005). Tat gets the "green" light on transcription initiation. Retrovirology *2*, 69.

Brand, S.R., Kobayashi, R., and Mathews, M.B. (1997). The Tat protein of human immunodeficiency virus type 1 is a substrate and inhibitor of the interferon-induced, virally activated protein kinase, PKR. J Biol Chem *272*, 8388-8395.

Bres, V., Gomes, N., Pickle, L., and Jones, K.A. (2005). A human splicing factor, SKIP, associates with P-TEFb and enhances transcription elongation by HIV-1 Tat. Genes Dev *19*, 1211-1226.

Bres, V., Kiernan, R., Emiliani, S., and Benkirane, M. (2002a). Tat acetyl-acceptor lysines are important for human immunodeficiency virus type-1 replication. J Biol Chem *277*, 22215-22221.

Bres, V., Tagami, H., Peloponese, J.M., Loret, E., Jeang, K.T., Nakatani, Y., Emiliani, S., Benkirane, M., and Kiernan, R.E. (2002b). Differential

acetylation of Tat coordinates its interaction with the co-activators cyclin T1 and PCAF. EMBO J *21*, 6811-6819.

Bres, V., Yoh, S.M., and Jones, K.A. (2008). The multi-tasking P-TEFb complex. Curr Opin Cell Biol *20*, 334-340.

Bres, V., Yoshida, T., Pickle, L., and Jones, K.A. (2009). SKIP interacts with c-Myc and Menin to promote HIV-1 Tat transactivation. Mol Cell *36*, 75-87.

Brown, J.M., Green, J., das Neves, R.P., Wallace, H.A., Smith, A.J., Hughes, J., Gray, N., Taylor, S., Wood, W.G., Higgs, D.R.*, et al.* (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. J Cell Biol *182*, 1083-1097.

Bukrinsky, M. (2004). A hard way to the nucleus. Mol Med *10*, 1-5.

Bukrinsky, M. (2006). SNFing HIV transcription. Retrovirology *3*, 49.

Bushman, F., Lewinski, M., Ciuffi, A., Barr, S., Leipzig, J., Hannenhalli, S., and Hoffmann, C. (2005). Genome-wide analysis of retroviral DNA integration. Nat Rev Microbiol *3*, 848-858.

Bushman, F.D. (2007). Retroviral integration and human gene therapy. J Clin Invest *117*, 2083-2086.

Cai, S., Lee, C.C., and Kohwi-Shigematsu, T. (2006). SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. Nat Genet *38*, 1278-1288.

Cairns, B.R. (2009). The logic of chromatin architecture and remodelling at promoters. Nature *461*, 193-198.

Calvert, M.E., Keck, K.M., Ptak, C., Shabanowitz, J., Hunt, D.F., and Pemberton, L.F. (2008). Phosphorylation by casein kinase 2 regulates Nap1 localization and function. Mol Cell Biol *28*, 1313-1325.

Candotti, F., Johnston, J.A., Puck, J.M., Sugamura, K., O'Shea, J.J., and Blaese, R.M. (1996). Retroviral-mediated gene correction for X-linked severe combined immunodeficiency. Blood *87*, 3097-3102.

Carteau, S., Hoffmann, C., and Bushman, F. (1998). Chromosome structure and human immunodeficiency virus type 1 cDNA integration: centromeric alphoid repeats are a disfavored target. J Virol *72*, 4005-4014.

Cassani, B., Montini, E., Maruggi, G., Ambrosi, A., Mirolo, M., Selleri, S., Biral, E., Frugnoli, I., Hernandez-Trujillo, V., Di Serio, C.*, et al.* (2009). Integration of retroviral vectors induces minor changes in the transcriptional activity of T cells from ADA-SCID patients treated with gene therapy. Blood *114*, 3546-3556.

Cattoglio, C., Facchini, G., Sartori, D., Antonelli, A., Miccio, A., Cassani, B., Schmidt, M., von Kalle, C., Howe, S., Thrasher, A.J., *et al.* (2007). Hot spots of retroviral integration in human CD34+ hematopoietic cells. Blood *110*, 1770-1778.

Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., De Coene, C., Selz, F., Le Deist, F., and Fischer, A. (1996). Role of interleukin-2 (IL-2), IL-7, and IL-15 in natural killer cell differentiation from cord blood hematopoietic progenitor cells and from gamma c transduced severe combined immunodeficiency X1 bone marrow cells. Blood *88*, 3901-3909.

Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J.L., *et al.* (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. Science *288*, 669-672.

Cereseto, A., Manganaro, L., Gutierrez, M.I., Terreni, M., Fittipaldi, A., Lusic, M., Marcello, A., and Giacca, M. (2005). Acetylation of HIV-1 integrase by p300 regulates viral integration. EMBO J *24*, 3070-3081.

Chan, D.C., and Kim, P.S. (1998). HIV entry and its inhibition. Cell *93*, 681-684.

Chan, H.M., and La Thangue, N.B. (2001). p300/CBP proteins: HATs for transcriptional bridges and scaffolds. J Cell Sci *114*, 2363-2373.

Chan, P.K., Chan, F.Y., Morris, S.W., and Xie, Z. (1997). Isolation and characterization of the human nucleophosmin/B23 (NPM) gene: identification of the YY1 binding site at the 5' enhancer region. Nucleic Acids Res *25*, 1225-1232.

Chang, H.K., Gallo, R.C., and Ensoli, B. (1995). Regulation of Cellular Gene Expression and Function by the Human Immunodeficiency Virus Type 1 Tat Protein. J Biomed Sci *2*, 189-202.

Chao, S.H., and Price, D.H. (2001). Flavopiridol inactivates P-TEFb and blocks most RNA polymerase II transcription in vivo. J Biol Chem *276*, 31793-31799.

Check, E. (2005). Gene therapy put on hold as third child develops cancer. Nature *433*, 561.

Chen, D., Wang, M., Zhou, S., and Zhou, Q. (2002a). HIV-1 Tat targets microtubules to induce apoptosis, a process promoted by the pro-apoptotic Bcl-2 relative Bim. EMBO J *21*, 6801-6810.

Chen, L.F., Mu, Y., and Greene, W.C. (2002b). Acetylation of RelA at discrete sites regulates distinct nuclear functions of NF-kappaB. EMBO J *21*, 6539-6548.

Chen, Z., Zang, J., Whetstine, J., Hong, X., Davrazou, F., Kutateladze, T.G., Simpson, M., Mao, Q., Pan, C.H., Dai, S., *et al.* (2006). Structural insights into histone demethylation by JMJD2 family members. Cell *125*, 691-702.

Cheutin, T., McNairn, A.J., Jenuwein, T., Gilbert, D.M., Singh, P.B., and Misteli, T. (2003). Maintenance of stable heterochromatin domains by dynamic HP1 binding. Science *299*, 721-725.

Chimura, T., Kuzuhara, T., and Horikoshi, M. (2002). Identification and characterization of CIA/ASF1 as an interactor of bromodomains associated with TFIID. Proc Natl Acad Sci U S A *99*, 9334-9339.

Chinnadurai, G. (1991). Modulation of HIV-enhancer activity by heterologous agents: a minireview. Gene *101*, 165-170.

Cho, S., Schroeder, S., Kaehlcke, K., Kwon, H.S., Pedal, A., Herker, E., Schnoelzer, M., and Ott, M. (2009). Acetylation of cyclin T1 regulates the equilibrium between active and inactive P-TEFb in cells. EMBO J *28*, 1407-1417.

Choi, J.K., and Howe, L.J. (2009). Histone acetylation: truth of consequences? Biochem Cell Biol *87*, 139-150.

Christ, F., Thys, W., De Rijck, J., Gijsbers, R., Albanese, A., Arosio, D., Emiliani, S., Rain, J.C., Benarous, R., Cereseto, A., *et al.* (2008). Transportin-SR2 imports HIV into the nucleus. Curr Biol *18*, 1192-1202.

Christensen, H.S., Daher, A., Soye, K.J., Frankel, L.B., Alexander, M.R., Laine, S., Bannwarth, S., Ong, C.L., Chung, S.W., Campbell, S.M., *et al.* (2007). Small interfering RNAs against the TAR RNA binding protein, TRBP, a Dicer cofactor, inhibit human immunodeficiency virus type 1 long terminal repeat expression and viral production. J Virol *81*, 5121-5131.

Chun, T.W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J.A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T.C., *et al.* (1997). Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. Nature *387*, 183-188.

Chun, T.W., Finzi, D., Margolick, J., Chadwick, K., Schwartz, D., and Siliciano, R.F. (1995). In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency. Nat Med *1*, 1284-1290.

Cochrane, A., Murley, L.L., Gao, M., Wong, R., Clayton, K., Brufatto, N., Canadien, V., Mamelak, D., Chen, T., Richards, D., *et al.* (2009). Stable complex formation between HIV Rev and the nucleosome assembly protein, NAP1, affects Rev function. Virology *388*, 103-111.

Coiras, M., Lopez-Huertas, M.R., Perez-Olmeda, M., and Alcami, J. (2009). Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs. Nat Rev Microbiol *7*, 798-812.

Coiras, M., Lopez-Huertas, M.R., Rullas, J., Mittelbrunn, M., and Alcami, J. (2007). Basal shuttle of NF-kappaB/I kappaB alpha in resting T lymphocytes regulates HIV-1 LTR dependent expression. Retrovirology *4*, 56.

Col, E., Caron, C., Seigneurin-Berny, D., Gracia, J., Favier, A., and Khochbin, S. (2001). The histone acetyltransferase, hGCN5, interacts with and acetylates the HIV transactivator, Tat. J Biol Chem *276*, 28179-28184.

Connor, R.I., Chen, B.K., Choe, S., and Landau, N.R. (1995). Vpr is required for efficient replication of human immunodeficiency virus type-1 in mononuclear phagocytes. Virology *206*, 935-944.

Coull, J.J., Romerio, F., Sun, J.M., Volker, J.L., Galvin, K.M., Davie, J.R., Shi, Y., Hansen, U., and Margolis, D.M. (2000). The human factors YY1 and LSF repress the human immunodeficiency virus type 1 long terminal repeat via recruitment of histone deacetylase 1. J Virol *74*, 6790-6799.

Cremer, T., Cremer, M., Dietzel, S., Muller, S., Solovei, I., and Fakan, S. (2006). Chromosome territories--a functional nuclear landscape. Curr Opin Cell Biol *18*, 307-316.

Crotti, A., Lusic, M., Lupo, R., Lievens, P.M., Liboi, E., Della Chiara, G., Tinelli, M., Lazzarin, A., Patterson, B.K., Giacca, M.*, et al.* (2007). Naturally occurring C-terminally truncated STAT5 is a negative regulator of HIV-1 expression. Blood *109*, 5380-5389.

Cullen, B.R. (1998). Retroviruses as model systems for the study of nuclear RNA export pathways. Virology *249*, 203-210.

D'Orso, I., and Frankel, A.D. (2009). Tat acetylation modulates assembly of a viral-host RNA-protein transcription complex. Proc Natl Acad Sci U S A *106*, 3101-3106.

De Koning, L., Corpet, A., Haber, J.E., and Almouzni, G. (2007). Histone chaperones: an escort network regulating histone traffic. Nat Struct Mol Biol *14*, 997-1007.

de la Fuente, C., Santiago, F., Deng, L., Eadie, C., Zilberman, I., Kehn, K., Maddukuri, A., Baylor, S., Wu, K., Lee, C.G.*, et al.* (2002). Gene expression profile of HIV-1 Tat expressing cells: a close interplay between proliferative and differentiation signals. BMC Biochem *3*, 14.

De Marco, A., Biancotto, C., Knezevich, A., Maiuri, P., Vardabasso, C., and Marcello, A. (2008). Intragenic transcriptional cis-activation of the human

immunodeficiency virus 1 does not result in allele-specific inhibition of the endogenous gene. Retrovirology *5*, 98.

de Mareuil, J., Carre, M., Barbier, P., Campbell, G.R., Lancelot, S., Opi, S., Esquieu, D., Watkins, J.D., Prevot, C., Braguer, D.*, et al.* (2005). HIV-1 Tat protein enhances microtubule polymerization. Retrovirology *2*, 5.

Deichmann, A., Hacein-Bey-Abina, S., Schmidt, M., Garrigue, A., Brugman, M.H., Hu, J., Glimm, H., Gyapay, G., Prum, B., Fraser, C.C.*, et al.* (2007). Vector integration is nonrandom and clustered and influences the fate of lymphopoiesis in SCID-X1 gene therapy. J Clin Invest *117*, 2225-2232.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

Del Rosario, B.C., and Pemberton, L.F. (2008). Nap1 links transcription elongation, chromatin assembly, and messenger RNP complex biogenesis. Mol Cell Biol *28*, 2113-2124.

Demarchi, F., d'Adda di Fagagna, F., Falaschi, A., and Giacca, M. (1996). Activation of transcription factor NF-kappaB by the Tat protein of human immunodeficiency virus type 1. J Virol *70*, 4427-4437.

Demarchi, F., D'Agaro, P., Falaschi, A., and Giacca, M. (1993). In vivo footprinting analysis of constitutive and inducible protein-DNA interactions at the long terminal repeat of human immunodeficiency virus type 1. J Virol *67*, 7450-7460.

Demarchi, F., Gutierrez, M.I., and Giacca, M. (1999). Human immunodeficiency virus type 1 tat protein activates transcription factor NF-kappaB through the cellular interferon-inducible, double-stranded RNA-dependent protein kinase, PKR. J Virol *73*, 7080-7086.

Deng, L., Ammosova, T., Pumfery, A., Kashanchi, F., and Nekhai, S. (2002). HIV-1 Tat interaction with RNA polymerase II C-terminal domain (CTD) and a dynamic association with CDK2 induce CTD phosphorylation and transcription from HIV-1 promoter. J Biol Chem *277*, 33922-33929.

Deng, L., de la Fuente, C., Fu, P., Wang, L., Donnelly, R., Wade, J.D., Lambert, P., Li, H., Lee, C.G., and Kashanchi, F. (2000). Acetylation of HIV-1 Tat by CBP/P300 increases transcription of integrated HIV-1 genome and enhances binding to core histones. Virology *277*, 278-295.

Deng, L., Wang, D., de la Fuente, C., Wang, L., Li, H., Lee, C.G., Donnelly, R., Wade, J.D., Lambert, P., and Kashanchi, F. (2001). Enhancement of the p300 HAT activity by HIV-1 Tat on chromatin DNA. Virology *289*, 312-326.

Dieudonne, M., Maiuri, P., Biancotto, C., Knezevich, A., Kula, A., Lusic, M., and Marcello, A. (2009). Transcriptional competence of the integrated HIV-1 provirus at the nuclear periphery. EMBO J *28*, 2231-2243.

Dillon, N. (2008). The impact of gene location in the nucleus on transcriptional regulation. Dev Cell *15*, 182-186.

Doms, R.W., and Trono, D. (2000). The plasma membrane as a combat zone in the HIV battlefield. Genes Dev *14*, 2677-2688.

Dorr, A., Kiermer, V., Pedal, A., Rackwitz, H.R., Henklein, P., Schubert, U., Zhou, M.M., Verdin, E., and Ott, M. (2002). Transcriptional synergy between Tat and PCAF is dependent on the binding of acetylated Tat to the PCAF bromodomain. EMBO J *21*, 2715-2723.

Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C.*, et al.* (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. Genome Res *16*, 1299-1309.

Downs, J.A., Nussenzweig, M.C., and Nussenzweig, A. (2007). Chromatin dynamics and the preservation of genetic information. Nature *447*, 951-958.

du Chene, I., Basyuk, E., Lin, Y.L., Triboulet, R., Knezevich, A., Chable-Bessia, C., Mettling, C., Baillat, V., Reynes, J., Corbeau, P.*, et al.* (2007). Suv39H1 and HP1gamma are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. EMBO J *26*, 424-435.

Du, H., Roy, A.L., and Roeder, R.G. (1993). Human transcription factor USF stimulates transcription through the initiator elements of the HIV-1 and the Ad-ML promoters. EMBO J *12*, 501-511.

Dundr, M., Hoffmann-Rohrer, U., Hu, Q., Grummt, I., Rothblum, L.I., Phair, R.D., and Misteli, T. (2002). A kinetic framework for a mammalian RNA polymerase in vivo. Science *298*, 1623-1626.

Dvorin, J.D., and Malim, M.H. (2003). Intracellular trafficking of HIV-1 cores: journey to the center of the cell. Curr Top Microbiol Immunol *281*, 179-208.

Easley, R., Van Duyne, R., Coley, W., Guendel, I., Dadgar, S., Kehn-Hall, K., and Kashanchi, F. (2009). Chromatin dynamics associated with HIV-1 Tat-activated transcription. Biochim Biophys Acta.

Eckey, M., Hong, W., Papaioannou, M., and Baniahmad, A. (2007). The nucleosome assembly activity of NAP1 is enhanced by Alien. Mol Cell Biol *27*, 3557-3568.

El Kharroubi, A., Piras, G., Zensen, R., and Martin, M.A. (1998). Transcriptional activation of the integrated chromatin-associated human immunodeficiency virus type 1 promoter. Mol Cell Biol *18*, 2535-2544.

el Kharroubi, A., and Verdin, E. (1994). Protein-DNA interactions within DNase I-hypersensitive sites located downstream of the HIV-1 promoter. J Biol Chem *269*, 19916-19924.

Endoh, M., Zhu, W., Hasegawa, J., Watanabe, H., Kim, D.K., Aida, M., Inukai, N., Narita, T., Yamada, T., Furuya, A.*, et al.* (2004). Human Spt6 stimulates transcription elongation by RNA polymerase II in vitro. Mol Cell Biol *24*, 3324-3336.

Epie, N., Ammosova, T., Sapir, T., Voloshin, Y., Lane, W.S., Turner, W., Reiner, O., and Nekhai, S. (2005). HIV-1 Tat interacts with LIS1 protein. Retrovirology *2*, 6.

Farnet, C.M., and Bushman, F.D. (1997). HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes in vitro. Cell *88*, 483-492.

Faulkner, N.E., Lane, B.R., Bock, P.J., and Markovitz, D.M. (2003). Protein phosphatase 2A enhances activation of human immunodeficiency virus type 1 by phorbol myristate acetate. J Virol *77*, 2276-2281.

Felice, B., Cattoglio, C., Cittaro, D., Testa, A., Miccio, A., Ferrari, G., Luzi, L., Recchia, A., and Mavilio, F. (2009). Transcription factor binding sites are genetic determinants of retroviral integration in the human genome. PLoS One *4*, e4571.

Ferrari, A., Pellegrini, V., Arcangeli, C., Fittipaldi, A., Giacca, M., and Beltram, F. (2003). Caveolae-mediated internalization of extracellular HIV-1 tat fusion proteins visualized in real time. Mol Ther *8*, 284-294.

Fittipaldi, A., Ferrari, A., Zoppe, M., Arcangeli, C., Pellegrini, V., Beltram, F., and Giacca, M. (2003). Cell membrane lipid rafts mediate caveolar endocytosis of HIV-1 Tat fusion proteins. J Biol Chem *278*, 34141-34149.

Fittipaldi, A., and Giacca, M. (2005). Transcellular protein transduction using the Tat protein of HIV-1. Adv Drug Deliv Rev *57*, 597-608.

Fogal, V., Gostissa, M., Sandy, P., Zacchi, P., Sternsdorf, T., Jensen, K., Pandolfi, P.P., Will, H., Schneider, C., and Del Sal, G. (2000). Regulation of p53 activity in nuclear bodies by a specific PML isoform. EMBO J *19*, 6185-6195.

Fong, Y.W., and Zhou, Q. (2000). Relief of two built-In autoinhibitory mechanisms in P-TEFb is required for assembly of a multicomponent

transcription elongation complex at the human immunodeficiency virus type 1 promoter. Mol Cell Biol *20*, 5897-5907.

Frankel, A.D., and Pabo, C.O. (1988). Cellular uptake of the tat protein from human immunodeficiency virus. Cell *55*, 1189-1193.

Fraser, P. (2006). Transcriptional control thrown for a loop. Curr Opin Genet Dev *16*, 490-495.

Fraser, P., and Bickmore, W. (2007). Nuclear organization of the genome and the potential for gene regulation. Nature *447*, 413-417.

Furger, A., Monks, J., and Proudfoot, N.J. (2001). The retroviruses human immunodeficiency virus type 1 and Moloney murine leukemia virus adopt radically different strategies to regulate promoter-proximal polyadenylation. J Virol *75*, 11735-11746.

Garber, M.E., Wei, P., KewalRamani, V.N., Mayall, T.P., Herrmann, C.H., Rice, A.P., Littman, D.R., and Jones, K.A. (1998). The interaction between HIV-1 Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein. Genes Dev *12*, 3512-3527.

Garrus, J.E., von Schwedler, U.K., Pornillos, O.W., Morham, S.G., Zavitz, K.H., Wang, H.E., Wettstein, D.A., Stray, K.M., Cote, M., Rich, R.L.*, et al.* (2001). Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. Cell *107*, 55-65.

Gaspar, H.B., Parsley, K.L., Howe, S., King, D., Gilmour, K.C., Sinclair, J., Brouns, G., Schmidt, M., Von Kalle, C., Barington, T.*, et al.* (2004). Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. Lancet *364*, 2181-2187.

Gautier, V.W., Gu, L., O'Donoghue, N., Pennington, S., Sheehy, N., and Hall, W.W. (2009). In vitro nuclear interactome of the HIV-1 Tat protein. Retrovirology *6*, 47.

Gaynor, R. (1992). Cellular transcription factors involved in the regulation of HIV-1 gene expression. AIDS *6*, 347-363.

Geijtenbeek, T.B., Kwon, D.S., Torensma, R., van Vliet, S.J., van Duijnhoven, G.C., Middel, J., Cornelissen, I.L., Nottet, H.S., KewalRamani, V.N., Littman, D.R.*, et al.* (2000). DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. Cell *100*, 587-597.

Geyer, M., Fackler, O.T., and Peterlin, B.M. (2001). Structure--function relationships in HIV-1 Nef. EMBO Rep *2*, 580-585.

Giacca, M. (2004). The HIV-1 Tat protein: a multifaceted target for novel therapeutic opportunities. Curr Drug Targets Immune Endocr Metabol Disord *4*, 277-285.

Giacca, M. (2005). HIV-1 Tat, apoptosis and the mitochondria: a tubulin link? Retrovirology *2*, 7.

Giacca, M., Zentilin, L., Norio, P., Diviacco, S., Dimitrova, D., Contreas, G., Biamonti, G., Perini, G., Weighardt, F., Riva, S.*, et al.* (1994). Fine mapping of a replication origin of human DNA. Proc Natl Acad Sci U S A *91*, 7119-7123.

Gibellini, D., Re, M.C., La Placa, M., and Zauli, G. (2002). Differentially expressed genes in HIV-1 tat-expressing CD4(+) T-cell line. Virus Res *90*, 337-345.

Gilboa, E., Mitra, S.W., Goff, S., and Baltimore, D. (1979). A detailed model of reverse transcription and tests of crucial aspects. Cell *18*, 93-100.

Goila-Gaur, R., and Strebel, K. (2008). HIV-1 Vif, APOBEC, and intrinsic immunity. Retrovirology *5*, 51.

Gondor, A., and Ohlsson, R. (2009). Chromosome crosstalk in three dimensions. Nature *461*, 212-217.

Goodrich, J.A., and Tjian, R. (1994). Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. Cell *77*, 145-156.

Green, C.M., and Almouzni, G. (2002). When repair meets chromatin. First in series on chromatin dynamics. EMBO Rep *3*, 28-33.

Green, M., and Loewenstein, P.M. (1988). Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. Cell *55*, 1179-1188.

Greger, I.H., Demarchi, F., Giacca, M., and Proudfoot, N.J. (1998). Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. Nucleic Acids Res *26*, 1294-1301.

Grewal, S.I., and Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. Science *301*, 798-802.

Groth, A., Rocha, W., Verreault, A., and Almouzni, G. (2007). Chromatin challenges during DNA replication and repair. Cell *128*, 721-733.

Hacein-Bey, S., Basile, G.D., Lemerle, J., Fischer, A., and Cavazzana-Calvo, M. (1998). gammac gene transfer in the presence of stem cell factor, FLT-3L, interleukin-7 (IL-7), IL-1, and IL-15 cytokines restores T-cell differentiation from gammac(-) X-linked severe combined immunodeficiency hematopoietic progenitor cells in murine fetal thymic organ cultures. Blood *92*, 4090-4097.

Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K.*, et al.* (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J Clin Invest *118*, 3132-3142.

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E.*, et al.* (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science *302*, 415-419.

Hajkova, P., Ancelin, K., Waldmann, T., Lacoste, N., Lange, U.C., Cesari, F., Lee, C., Almouzni, G., Schneider, R., and Surani, M.A. (2008). Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature *452*, 877-881.

Han, Y., Lassen, K., Monie, D., Sedaghat, A.R., Shimoji, S., Liu, X., Pierson, T.C., Margolick, J.B., Siliciano, R.F., and Siliciano, J.D. (2004). Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. J Virol *78*, 6122-6133.

Han, Y., Lin, Y.B., An, W., Xu, J., Yang, H.C., O'Connell, K., Dordai, D., Boeke, J.D., Siliciano, J.D., and Siliciano, R.F. (2008). Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. Cell Host Microbe *4*, 134-146.

Harrich, D., Garcia, J., Wu, F., Mitsuyasu, R., Gonazalez, J., and Gaynor, R. (1989). Role of SP1-binding domains in in vivo transcriptional regulation of the human immunodeficiency virus type 1 long terminal repeat. J Virol *63*, 2585-2591.

Harrich, D., Ulich, C., Garcia-Martinez, L.F., and Gaynor, R.B. (1997). Tat is required for efficient HIV-1 reverse transcription. EMBO J *16*, 1224-1235.

Hartley, P.D., and Madhani, H.D. (2009). Mechanisms that specify promoter nucleosome location and identity. Cell *137*, 445-458.

Hauber, J., Malim, M.H., and Cullen, B.R. (1989). Mutational analysis of the conserved basic domain of human immunodeficiency virus tat protein. J Virol *63*, 1181-1187.

He, G., and Margolis, D.M. (2002). Counterregulation of chromatin deacetylation and histone deacetylase occupancy at the integrated promoter of human immunodeficiency virus type 1 (HIV-1) by the HIV-1 repressor YY1 and HIV-1 activator Tat. Mol Cell Biol *22*, 2965-2973.

He, J., Choe, S., Walker, R., Di Marzio, P., Morgan, D.O., and Landau, N.R. (1995). Human immunodeficiency virus type 1 viral protein R (Vpr) arrests

cells in the G2 phase of the cell cycle by inhibiting p34cdc2 activity. J Virol *69*, 6705-6711.

Henderson, A., Holloway, A., Reeves, R., and Tremethick, D.J. (2004). Recruitment of SWI/SNF to the human immunodeficiency virus type 1 promoter. Mol Cell Biol *24*, 389-397.

Herrmann, C.H., and Rice, A.P. (1993). Specific interaction of the human immunodeficiency virus Tat proteins with a cellular protein kinase. Virology *197*, 601-608.

Holowaty, M.N., Zeghouf, M., Wu, H., Tellam, J., Athanasopoulos, V., Greenblatt, J., and Frappier, L. (2003). Protein profiling with Epstein-Barr nuclear antigen-1 reveals an interaction with the herpesvirus-associated ubiquitin-specific protease HAUSP/USP7. J Biol Chem *278*, 29987-29994.

Horike, S., Cai, S., Miyano, M., Cheng, J.F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. Nat Genet *37*, 31-40.

Hottiger, M.O., and Nabel, G.J. (1998). Interaction of human immunodeficiency virus type 1 Tat with the transcriptional coactivators p300 and CREB binding protein. J Virol *72*, 8252-8256.

Howe, S.J., Mansour, M.R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempski, H., Brugman, M.H., Pike-Overzet, K., Chatters, S.J., de Ridder, D., *et al.* (2008). Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. J Clin Invest *118*, 3143-3150.

Huang, J., Wang, F., Argyris, E., Chen, K., Liang, Z., Tian, H., Huang, W., Squires, K., Verlinghieri, G., and Zhang, H. (2007). Cellular microRNAs contribute to HIV-1 latency in resting primary CD4+ T lymphocytes. Nat Med *13*, 1241-1247.

Ikegami, K., Horigome, D., Mukai, M., Livnat, I., MacGregor, G.R., and Setou, M. (2008). TTLL10 is a protein polyglycylase that can modify nucleosome assembly protein 1. FEBS Lett *582*, 1129-1134.

Ishimi, Y., Hirosumi, J., Sato, W., Sugasawa, K., Yokota, S., Hanaoka, F., and Yamada, M. (1984). Purification and initial characterization of a protein which facilitates assembly of nucleosome-like structure from mammalian cells. Eur J Biochem *142*, 431-439.

Ishimi, Y., and Kikuchi, A. (1991). Identification and molecular cloning of yeast homolog of nucleosome assembly protein I which facilitates nucleosome assembly in vitro. J Biol Chem *266*, 7025-7029.

Ishimi, Y., Sato, W., Kojima, M., Sugasawa, K., Hanaoka, F., and Yamada, M. (1985). Rapid purification of nucleosome assembly protein (AP-I) and production of monoclonal antibodies against it. Cell Struct Funct *10*, 373-382.

Ishimi, Y., Yasuda, H., Hirosumi, J., Hanaoka, F., and Yamada, M. (1983). A protein which facilitates assembly of nucleosome-like structures in vitro in mammalian cells. J Biochem (Tokyo) *94*, 735-744.

Ito, T., Bulger, M., Kobayashi, R., and Kadonaga, J.T. (1996). Drosophila NAP-1 is a core histone chaperone that functions in ATP-facilitated assembly of regularly spaced nucleosomal arrays. Mol Cell Biol *16*, 3112-3124.

Ito, T., Ikehara, T., Nakagawa, T., Kraus, W.L., and Muramatsu, M. (2000). p300-mediated acetylation facilitates the transfer of histone H2A-H2B dimers from nucleosomes to a histone chaperone. Genes Dev *14*, 1899-1907.

Jacque, J.M., and Stevenson, M. (2006). The inner-nuclear-envelope protein emerin regulates HIV-1 infectivity. Nature *441*, 641-645.

Janardhan, A., Swigut, T., Hill, B., Myers, M.P., and Skowronski, J. (2004). HIV-1 Nef binds the DOCK2-ELMO1 complex to activate rac and inhibit lymphocyte chemotaxis. PLoS Biol *2*, E6.

Jeang, K.T., Xiao, H., and Rich, E.A. (1999). Multifaceted activities of the HIV-1 transactivator of transcription, Tat. J Biol Chem *274*, 28837-28840.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Jones, K.A., Kadonaga, J.T., Luciw, P.A., and Tjian, R. (1986). Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1. Science *232*, 755-759.

Jordan, A., Bisgrove, D., and Verdin, E. (2003). HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. EMBO J *22*, 1868-1877.

Jordan, A., Defechereux, P., and Verdin, E. (2001). The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. EMBO J *20*, 1726-1738.

Ju, S.M., Song, H.Y., Lee, J.A., Lee, S.J., Choi, S.Y., and Park, J. (2009). Extracellular HIV-1 Tat up-regulates expression of matrix metalloproteinase-9 via a MAPK-NF-kappaB dependent pathway in human astrocytes. Exp Mol Med *41*, 86-93.

Kaehlcke, K., Dorr, A., Hetzer-Egger, C., Kiermer, V., Henklein, P., Schnoelzer, M., Loret, E., Cole, P.A., Verdin, E., and Ott, M. (2003). Acetylation of Tat

defines a cyclinT1-independent step in HIV transactivation. Mol Cell *12*, 167-176.

Kalantari, P., Narayan, V., Natarajan, S.K., Muralidhar, K., Gandhi, U.H., Vunta, H., Henderson, A.J., and Prabhu, K.S. (2008). Thioredoxin reductase-1 negatively regulates HIV-1 transactivating protein Tat-dependent transcription in human macrophages. J Biol Chem *283*, 33183-33190.

Kalpana, G.V., Marmon, S., Wang, W., Crabtree, G.R., and Goff, S.P. (1994). Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. Science *266*, 2002-2006.

Kameoka, M., Morgan, M., Binette, M., Russell, R.S., Rong, L., Guo, X., Mouland, A., Kleiman, L., Liang, C., and Wainberg, M.A. (2002). The Tat protein of human immunodeficiency virus type 1 (HIV-1) can promote placement of tRNA primer onto viral RNA and suppress later DNA polymerization in HIV-1 reverse transcription. J Virol *76*, 3637-3645.

Kamine, J., Elangovan, B., Subramanian, T., Coleman, D., and Chinnadurai, G. (1996). Identification of a cellular protein that specifically interacts with the essential cysteine region of the HIV-1 Tat transactivator. Virology *216*, 357-366.

Kanazawa, S., Okamoto, T., and Peterlin, B.M. (2000). Tat competes with CIITA for the binding to P-TEFb and blocks the expression of MHC class II genes in HIV infection. Immunity *12*, 61-70.

Kang, E.M., Choi, U., Theobald, N., Linton, G., Long Priel, D.A., Kuhns, D., and Malech, H.L. (2009). Retrovirus gene therapy for X-linked chronic granulomatous disease can achieve stable long term correction of oxidase activity in peripheral blood neutrophils. Blood.

Kao, S.Y., Calman, A.F., Luciw, P.A., and Peterlin, B.M. (1987). Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature *330*, 489-493.

Kato, H., Horikoshi, M., and Roeder, R.G. (1991). Repression of HIV-1 transcription by a cellular protein. Science *251*, 1476-1479.

Kawase, H., Okuwaki, M., Miyaji, M., Ohba, R., Handa, H., Ishimi, Y., Fujii-Nakata, T., Kikuchi, A., and Nagata, K. (1996). NAP-I is a functional homologue of TAF-I that is required for replication and transcription of the adenovirus genome in a chromatin-like structure. Genes Cells *1*, 1045-1056.

Kellogg, D.R., and Murray, A.W. (1995). NAP1 acts with Clb1 to perform mitotic functions and to suppress polar bud growth in budding yeast. J Cell Biol *130*, 675-685.

Kepert, J.F., Mazurkiewicz, J., Heuvelman, G.L., Toth, K.F., and Rippe, K. (2005). NAP1 modulates binding of linker histone H1 to chromatin and induces an extended chromatin fiber conformation. J Biol Chem *280*, 34063-34072.

Kiernan, R., Bres, V., Ng, R.W., Coudart, M.P., El Messaoudi, S., Sardet, C., Jin, D.Y., Emiliani, S., and Benkirane, M. (2003). Post-activation turn-off of NF-kappa B-dependent transcription is regulated by acetylation of p65. J Biol Chem *278*, 2758-2766.

Kiernan, R.E., Emiliani, S., Nakayama, K., Castro, A., Labbe, J.C., Lorca, T., Nakayama Ki, K., and Benkirane, M. (2001). Interaction between cyclin T1 and SCF(SKP2) targets CDK9 for ubiquitination and degradation by the proteasome. Mol Cell Biol *21*, 7956-7970.

Kiernan, R.E., Vanhulle, C., Schiltz, L., Adam, E., Xiao, H., Maudoux, F., Calomme, C., Burny, A., Nakatani, Y., Jeang, K.T.*, et al.* (1999). HIV-1 tat transcriptional activity is regulated by acetylation. EMBO J *18*, 6106-6118.

Kim, H.J., Seol, J.H., Han, J.W., Youn, H.D., and Cho, E.J. (2007). Histone chaperones regulate histone exchange during transcription. Embo J *26*, 4467-4474.

Kinoshita, S., Chen, B.K., Kaneshima, H., and Nolan, G.P. (1998). Host control of HIV-1 parasitism in T cells by the nuclear factor of activated T cells. Cell *95*, 595-604.

Koues, O.I., Dudley, R.K., Mehta, N.T., and Greer, S.F. (2009). The 19S proteasome positively regulates histone methylation at cytokine inducible genes. Biochim Biophys Acta.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P.*, et al.* (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637-643.

Kumar, A., and Jeang, K.T. (2008). Insights into cellular microRNAs and human immunodeficiency virus type 1 (HIV-1). J Cell Physiol *216*, 327-331.

Kumar, P.P., Bischof, O., Purbey, P.K., Notani, D., Urlaub, H., Dejean, A., and Galande, S. (2007). Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. Nat Cell Biol *9*, 45-56.

Kumar, P.P., Purbey, P.K., Ravi, D.S., Mitra, D., and Galande, S. (2005). Displacement of SATB1-bound histone deacetylase 1 corepressor by the human immunodeficiency virus type 1 transactivator induces expression of interleukin-2 and its receptor in T cells. Mol Cell Biol *25*, 1620-1633.

Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J., and Hendrickson, W.A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature *393*, 648-659.

Lamond, A.I., and Spector, D.L. (2003). Nuclear speckles: a model for nuclear organelles. Nat Rev Mol Cell Biol *4*, 605-612.

Landry, S., Halin, M., Lefort, S., Audet, B., Vaquero, C., Mesnard, J.M., and Barbeau, B. (2007). Detection, characterization and regulation of antisense transcripts in HIV-1. Retrovirology *4*, 71.

Lankenau, S., Barnickel, T., Marhold, J., Lyko, F., Mechler, B.M., and Lankenau, D.H. (2003). Knockout targeting of the Drosophila nap1 gene and examination of DNA repair tracts in the recombination products. Genetics *163*, 611-623.

Lassen, K., Han, Y., Zhou, Y., Siliciano, J., and Siliciano, R.F. (2004). The multifactorial nature of HIV-1 latency. Trends Mol Med *10*, 525-531.

Lassot, I., Latreille, D., Rousset, E., Sourisseau, M., Linares, L.K., Chable-Bessia, C., Coux, O., Benkirane, M., and Kiernan, R.E. (2007). The proteasome regulates HIV-1 transcription by both proteolytic and nonproteolytic mechanisms. Mol Cell *25*, 369-383.

Lee, D., Ezhkova, E., Li, B., Pattenden, S.G., Tansey, W.P., and Workman, J.L. (2005). The proteasome regulatory particle alters the SAGA coactivator to enhance its interactions with transcriptional activators. Cell *123*, 423-436.

Lenasi, T., Contreras, X., and Peterlin, B.M. (2008). Transcriptional interference antagonizes proviral gene expression to promote HIV latency. Cell Host Microbe *4*, 123-133.

Levchenko, V., and Jackson, V. (2004). Histone release during transcription: NAP1 forms a complex with H2A and H2B and facilitates a topologically dependent release of H3 and H4 from the nucleosome. Biochemistry *43*, 2359-2372.

Lewinski, M.K., Bisgrove, D., Shinn, P., Chen, H., Hoffmann, C., Hannenhalli, S., Verdin, E., Berry, C.C., Ecker, J.R., and Bushman, F.D. (2005). Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. J Virol *79*, 6610-6619.

Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. Cell *128*, 707-719.

Li, C.J., Friedman, D.J., Wang, C., Metelev, V., and Pardee, A.B. (1995). Induction of apoptosis in uninfected lymphocytes by HIV-1 Tat protein. Science *268*, 429-431.

Li, Y.P. (1997). Protein B23 is an important human factor for the nucleolar localization of the human immunodeficiency virus protein Tat. J Virol *71*, 4098-4102.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O.*, et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

Lin, C.W., and Engelman, A. (2003). The barrier-to-autointegration factor is a component of functional human immunodeficiency virus type 1 preintegration complexes. J Virol *77*, 5030-5036.

Lomvardas, S., Barnea, G., Pisapia, D.J., Mendelsohn, M., Kirkland, J., and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. Cell *126*, 403-413.

Lorch, Y., Maier-Davis, B., and Kornberg, R.D. (2006). Chromatin remodeling by nucleosome disassembly in vitro. Proc Natl Acad Sci U S A *103*, 3090-3093.

Loyola, A., and Almouzni, G. (2004). Histone chaperones, a supporting role in the limelight. Biochim Biophys Acta *1677*, 3-11.

Luban, J. (2007). Cyclophilin A, TRIM5, and resistance to human immunodeficiency virus type 1 infection. J Virol *81*, 1054-1061.

Lusic, M., Marcello, A., Cereseto, A., and Giacca, M. (2003). Regulation of HIV-1 gene expression by histone acetylation and factor recruitment at the LTR promoter. Embo J *22*, 6550-6561.

Maertens, G., Cherepanov, P., Pluymers, W., Busschots, K., De Clercq, E., Debyser, Z., and Engelborghs, Y. (2003). LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. J Biol Chem *278*, 33528-33539.

Mahmoudi, T., Parra, M., Vries, R.G., Kauder, S.E., Verrijzer, C.P., Ott, M., and Verdin, E. (2006). The SWI/SNF chromatin-remodeling complex is a cofactor for Tat transactivation of the HIV promoter. J Biol Chem *281*, 19960-19968.

Maison, C., and Almouzni, G. (2004). HP1 and the dynamics of heterochromatin maintenance. Nat Rev Mol Cell Biol *5*, 296-304.

Maldonado, E., Shiekhattar, R., Sheldon, M., Cho, H., Drapkin, R., Rickert, P., Lees, E., Anderson, C.W., Linn, S., and Reinberg, D. (1996). A human RNA polymerase II complex associated with SRB and DNA-repair proteins. Nature *381*, 86-89.

Malim, M.H., Tiley, L.S., McCarn, D.F., Rusche, J.R., Hauber, J., and Cullen, B.R. (1990). HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. Cell *60*, 675-683.

Mancebo, H.S., Lee, G., Flygare, J., Tomassini, J., Luu, P., Zhu, Y., Peng, J., Blau, C., Hazuda, D., Price, D.*, et al.* (1997). P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. Genes Dev *11*, 2633-2644.

Marban, C., Redel, L., Suzanne, S., Van Lint, C., Lecestre, D., Chasserot-Golaz, S., Leid, M., Aunis, D., Schaeffer, E., and Rohr, O. (2005). COUP-TF interacting protein 2 represses the initial phase of HIV-1 gene transcription in human microglial cells. Nucleic Acids Res *33*, 2318-2331.

Marban, C., Suzanne, S., Dequiedt, F., de Walque, S., Redel, L., Van Lint, C., Aunis, D., and Rohr, O. (2007). Recruitment of chromatin-modifying enzymes by CTIP2 promotes HIV-1 transcriptional silencing. EMBO J *26*, 412-423.

Marcello, A. (2006). Latency: the hidden HIV-1 challenge. Retrovirology *3*, 7.

Marcello, A., Cinelli, R.A., Ferrari, A., Signorelli, A., Tyagi, M., Pellegrini, V., Beltram, F., and Giacca, M. (2001a). Visualization of in vivo direct interaction between HIV-1 TAT and human cyclin T1 in specific subcellular compartments by fluorescence resonance energy transfer. J Biol Chem *276*, 39220-39225.

Marcello, A., Ferrari, A., Pellegrini, V., Pegoraro, G., Lusic, M., Beltram, F., and Giacca, M. (2003). Recruitment of human cyclin T1 to nuclear bodies through direct interaction with the PML protein. EMBO J *22*, 2156-2166.

Marcello, A., Lusic, M., Pegoraro, G., Pellegrini, V., Beltram, F., and Giacca, M. (2004). Nuclear organization and the control of HIV-1 transcription. Gene *326*, 1-11.

Marcello, A., Zoppe, M., and Giacca, M. (2001b). Multiple modes of transcriptional regulation by the HIV-1 Tat transactivator. IUBMB Life *51*, 175-181.

Marheineke, K., and Krude, T. (1998). Nucleosome assembly activity and intracellular localization of human CAF-1 changes during the cell division cycle. J Biol Chem *273*, 15279-15286.

Martin, M., Cho, J., Cesare, A.J., Griffith, J.D., and Attardi, G. (2005). Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rRNA synthesis. Cell *123*, 1227-1240.

Martin-Serrano, J., Zang, T., and Bieniasz, P.D. (2001). HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. Nat Med *7*, 1313-1319.

Maruggi, G., Porcellini, S., Facchini, G., Perna, S.K., Cattoglio, C., Sartori, D., Ambrosi, A., Schambach, A., Baum, C., Bonini, C.*, et al.* (2009). Transcriptional enhancers induce insertional gene deregulation independently from the vector type and design. Mol Ther *17*, 851-856.

Marzio, G., Tyagi, M., Gutierrez, M.I., and Giacca, M. (1998). HIV-1 tat transactivator recruits p300 and CREB-binding protein histone acetyltransferases to the viral promoter. Proc Natl Acad Sci U S A *95*, 13519-13524.

Mateescu, B., Bourachot, B., Rachez, C., Ogryzko, V., and Muchardt, C. (2008). Regulation of an inducible promoter by an HP1beta-HP1gamma switch. EMBO Rep *9*, 267-272.

Michels, A.A., Fraldi, A., Li, Q., Adamson, T.E., Bonnet, F., Nguyen, V.T., Sedore, S.C., Price, J.P., Price, D.H., Lania, L.*, et al.* (2004). Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. EMBO J *23*, 2608-2619.

Miller, M.D., Farnet, C.M., and Bushman, F.D. (1997). Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. J Virol *71*, 5382-5390.

Miller, R.H. (1988). Human immunodeficiency virus may encode a novel protein on the genomic DNA plus strand. Science *239*, 1420-1422.

Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. Cell *128*, 787-800.

Mitchell, R.S., Beitzel, B.F., Schroder, A.R., Shinn, P., Chen, H., Berry, C.C., Ecker, J.R., and Bushman, F.D. (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol *2*, E234.

Mitola, S., Sozzani, S., Luini, W., Primo, L., Borsatti, A., Weich, H., and Bussolino, F. (1997). Tat-human immunodeficiency virus-1 induces human monocyte chemotaxis by activation of vascular endothelial growth factor receptor-1. Blood *90*, 1365-1372.

Miyaji-Yamaguchi, M., Kato, K., Nakano, R., Akashi, T., Kikuchi, A., and Nagata, K. (2003). Involvement of nucleocytoplasmic shuttling of yeast Nap1 in mitotic progression. Mol Cell Biol *23*, 6672-6684.

Miyauchi, K., Kim, Y., Latinovic, O., Morozov, V., and Melikyan, G.B. (2009). HIV enters cells via endocytosis and dynamin-dependent fusion with endosomes. Cell *137*, 433-444.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science *303*, 343-348.

Montano, M.A., Kripke, K., Norina, C.D., Achacoso, P., Herzenberg, L.A., Roy, A.L., and Nolan, G.P. (1996). NF-kappa B homodimer binding within the HIV-1 initiator region and interactions with TFII-I. Proc Natl Acad Sci U S A *93*, 12376-12381.

Morita, E., and Sundquist, W.I. (2004). Retrovirus budding. Annu Rev Cell Dev Biol *20*, 395-425.

Mosammaparast, N., Del Rosario, B.C., and Pemberton, L.F. (2005). Modulation of histone deposition by the karyopherin kap114. Mol Cell Biol *25*, 1764-1778.

Mosammaparast, N., Ewart, C.S., and Pemberton, L.F. (2002). A role for nucleosome assembly protein 1 in the nuclear transport of histones H2A and H2B. Embo J *21*, 6527-6538.

Moshkin, Y.M., Kan, T.W., Goodfellow, H., Bezstarosti, K., Maeda, R.K., Pilyugin, M., Karch, F., Bray, S.J., Demmers, J.A., and Verrijzer, C.P. (2009). Histone chaperones ASF1 and NAP1 differentially modulate removal of active histone marks by LID-RPD3 complexes during NOTCH silencing. Mol Cell *35*, 782-793.

Mousnier, A., Kubat, N., Massias-Simon, A., Segeral, E., Rain, J.C., Benarous, R., Emiliani, S., and Dargemont, C. (2007). von Hippel Lindau binding protein 1-mediated degradation of integrase affects HIV-1 gene expression at a postintegration step. Proc Natl Acad Sci U S A *104*, 13615-13620.

Muesing, M.A., Smith, D.H., and Capon, D.J. (1987). Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein. Cell *48*, 691-701.

Mujtaba, S., He, Y., Zeng, L., Farooq, A., Carlson, J.E., Ott, M., Verdin, E., and Zhou, M.M. (2002). Structural basis of lysine-acetylated HIV-1 Tat recognition by PCAF bromodomain. Mol Cell *9*, 575-586.

Mulky, A., Cohen, T.V., Kozlov, S.V., Korbei, B., Foisner, R., Stewart, C.L., and KewalRamani, V.N. (2008). The LEM domain proteins emerin and LAP2alpha are dispensable for human immunodeficiency virus type 1 and murine leukemia virus infections. J Virol *82*, 5860-5868.

Muto, S., Senda, M., Akai, Y., Sato, L., Suzuki, T., Nagai, R., Senda, T., and Horikoshi, M. (2007). Relationship between the structure of SET/TAF-Ibeta/INHAT and its histone chaperone activity. Proc Natl Acad Sci U S A *104*, 4285-4290.

Nabel, G., and Baltimore, D. (1987). An inducible transcription factor activates expression of human immunodeficiency virus in T cells. Nature *326*, 711-713.

Nakagawa, T., Bulger, M., Muramatsu, M., and Ito, T. (2001). Multistep chromatin assembly on supercoiled plasmid DNA by nucleosome assembly protein-1 and ATP-utilizing chromatin assembly and remodeling factor. J Biol Chem *276*, 27384-27391.

Nam, C.H., and Rabbitts, T.H. (2006). The role of LMO2 in development and in T cell leukemia after chromosomal translocation or retroviral insertion. Mol Ther *13*, 15-25.

Napolitano, G., Licciardo, P., Carbone, R., Majello, B., and Lania, L. (2002). CDK9 has the intrinsic property to shuttle between nucleus and cytoplasm, and enhanced expression of cyclin T1 promotes its nuclear localization. J Cell Physiol *192*, 209-215.

Negorev, D., and Maul, G.G. (2001). Cellular proteins localized at and interacting within ND10/PML nuclear bodies/PODs suggest functions of a nuclear depot. Oncogene *20*, 7234-7242.

Neil, S.J., Zang, T., and Bieniasz, P.D. (2008). Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature *451*, 425-430.

Nguyen, D.H., Gummuluru, S., and Hu, J. (2007). Deamination-independent inhibition of hepatitis B virus reverse transcription by APOBEC3G. J Virol *81*, 4465-4472.

Nguyen, V.T., Kiss, T., Michels, A.A., and Bensaude, O. (2001). 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. Nature *414*, 322-325.

O'Sullivan, J.M., Tan-Wong, S.M., Morillon, A., Lee, B., Coles, J., Mellor, J., and Proudfoot, N.J. (2004). Gene loops juxtapose promoters and terminators in yeast. Nat Genet *36*, 1014-1018.

Ohkuni, K., Shirahige, K., and Kikuchi, A. (2003). Genome-wide expression analysis of NAP1 in Saccharomyces cerevisiae. Biochem Biophys Res Commun *306*, 5-9.

Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. Genes Dev *10*, 2657-2683.

Orphanides, G., and Reinberg, D. (2002). A unified theory of gene expression. Cell *108*, 439-451.

Ott, M., Schnolzer, M., Garnica, J., Fischle, W., Emiliani, S., Rackwitz, H.R., and Verdin, E. (1999). Acetylation of the HIV-1 Tat protein by p300 is important for its transcriptional activity. Curr Biol *9*, 1489-1492.

Ott, M.G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., Glimm, H., Kuhlcke, K., Schilz, A., Kunkel, H.*, et al.* (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. Nat Med *12*, 401-409.

Pagans, S., Pedal, A., North, B.J., Kaehlcke, K., Marshall, B.L., Dorr, A., Hetzer-Egger, C., Henklein, P., Frye, R., McBurney, M.W.*, et al.* (2005). SIRT1 regulates HIV transcription via Tat deacetylation. PLoS Biol *3*, e41.

Parada, C.A., and Roeder, R.G. (1996). Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain. Nature *384*, 375-378.

Park, Y.J., Chodaparambil, J.V., Bao, Y., McBryant, S.J., and Luger, K. (2005). Nucleosome assembly protein 1 exchanges histone H2A-H2B dimers and assists nucleosome sliding. J Biol Chem *280*, 1817-1825.

Park, Y.J., and Luger, K. (2006a). Structure and function of nucleosome assembly proteins. Biochem Cell Biol *84*, 549-558.

Park, Y.J., and Luger, K. (2006b). The structure of nucleosome assembly protein 1. Proc Natl Acad Sci U S A *103*, 1248-1253.

Park, Y.J., and Luger, K. (2008). Histone chaperones in nucleosome eviction and histone exchange. Curr Opin Struct Biol *18*, 282-289.

Pearson, R., Kim, Y.K., Hokello, J., Lassen, K., Friedman, J., Tyagi, M., and Karn, J. (2008). Epigenetic silencing of human immunodeficiency virus (HIV) transcription by formation of restrictive chromatin structures at the viral long terminal repeat drives the progressive entry of HIV into latency. J Virol *82*, 12291-12303.

Pegoraro, G., Marcello, A., Myers, M.P., and Giacca, M. (2006). Regulation of adeno-associated virus DNA replication by the cellular TAF-I/set complex. J Virol *80*, 6855-6864.

Pereira, L.A., Bentley, K., Peeters, A., Churchill, M.J., and Deacon, N.J. (2000). A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. Nucleic Acids Res *28*, 663-668.

Perkins, K.J., Lusic, M., Mitar, I., Giacca, M., and Proudfoot, N.J. (2008). Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. Mol Cell *29*, 56-68.

Perkins, K.J., and Proudfoot, N.J. (2008). An ungracious host for an unwelcome guest. Cell Host Microbe *4*, 89-91.

Peterlin, B.M., and Price, D.H. (2006). Controlling the elongation phase of transcription with P-TEFb. Mol Cell *23*, 297-305.

Petersen-Mahrt, S.K., Estmer, C., Ohrmalm, C., Matthews, D.A., Russell, W.C., and Akusjarvi, G. (1999). The splicing factor-associated protein, p32, regulates RNA splicing by inhibiting ASF/SF2 RNA binding and phosphorylation. Embo J *18*, 1014-1024.

Peterson, C.L., and Cote, J. (2004). Cellular machineries for chromosomal DNA repair. Genes Dev *18*, 602-616.

Polo, S.E., and Almouzni, G. (2005). Histone metabolic pathways and chromatin assembly factors as proliferation markers. Cancer Lett *220*, 1-9.

Polo, S.E., and Almouzni, G. (2006). Chromatin assembly: a basic recipe with various flavours. Curr Opin Genet Dev *16*, 104-111.

Pomerantz, R.J., Trono, D., Feinberg, M.B., and Baltimore, D. (1990). Cells nonproductively infected with HIV-1 exhibit an aberrant pattern of viral RNA expression: a molecular model for latency. Cell *61*, 1271-1276.

Popovic, M., Sarngadharan, M.G., Read, E., and Gallo, R.C. (1984). Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. Science *224*, 497-500.

Raha, T., Cheng, S.W., and Green, M.R. (2005). HIV-1 Tat stimulates transcription complex assembly through recruitment of TBP in the absence of TAFs. PLoS Biol *3*, e44.

Raisner, R.M., Hartley, P.D., Meneghini, M.D., Bao, M.Z., Liu, C.L., Schreiber, S.L., Rando, O.J., and Madhani, H.D. (2005). Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. Cell *123*, 233-248.

Recchia, A., Bonini, C., Magnani, Z., Urbinati, F., Sartori, D., Muraro, S., Tagliafico, E., Bondanza, A., Stanghellini, M.T., Bernardi, M.*, et al.* (2006). Retroviral vector integration deregulates gene expression but has no consequence on the biology and function of transplanted T cells. Proc Natl Acad Sci U S A *103*, 1457-1462.

Regnard, C., Desbruyeres, E., Huet, J.C., Beauvallet, C., Pernollet, J.C., and Edde, B. (2000). Polyglutamylation of nucleosome assembly proteins. J Biol Chem *275*, 15969-15976.

Rehtanz, M., Schmidt, H.M., Warthorst, U., and Steger, G. (2004). Direct interaction between nucleosome assembly protein 1 and the papillomavirus E2 proteins involved in activation of transcription. Mol Cell Biol *24*, 2153-2168.

Richman, D.D., Margolis, D.M., Delaney, M., Greene, W.C., Hazuda, D., and Pomerantz, R.J. (2009). The challenge of finding a cure for HIV infection. Science *323*, 1304-1307.

Rodriguez, P., Pelletier, J., Price, G.B., and Zannis-Hadjopoulos, M. (2000). NAP-2: histone chaperone function and phosphorylation state through the cell cycle. J Mol Biol *298*, 225-238.

Rogner, U.C., Spyropoulos, D.D., Le Novere, N., Changeux, J.P., and Avner, P. (2000). Control of neurulation by the nucleosome assembly protein-1-like 2. Nat Genet *25*, 431-435.

Romani, B., Engelbrecht, S., and Glashoff, R.H. (2009). Functions of Tat: the versatile protein of human immunodeficiency virus type 1. J Gen Virol.

Rosen, C.A., Sodroski, J.G., and Haseltine, W.A. (1985). The location of cis-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. Cell *41*, 813-823.

Royer-Pokora, B., Loos, U., and Ludwig, W.D. (1991). TTG-2, a new gene encoding a cysteine-rich protein with the LIM motif, is overexpressed in acute T-cell leukaemia with the t(11;14)(p13;q11). Oncogene *6*, 1887-1893.

Royer-Pokora, B., Rogers, M., Zhu, T.H., Schneider, S., Loos, U., and Bolitz, U. (1995). The TTG-2/RBTN2 T cell oncogene encodes two alternative transcripts from two promoters: the distal promoter is removed by most 11p13 translocations in acute T cell leukaemia's (T-ALL). Oncogene *10*, 1353-1360.

Sabo, A., Lusic, M., Cereseto, A., and Giacca, M. (2008). Acetylation of conserved lysines in the catalytic core of cyclin-dependent kinase 9 inhibits kinase activity and regulates transcription. Mol Cell Biol *28*, 2201-2212.

Sasada, A., Takaori-Kondo, A., Shirakawa, K., Kobayashi, M., Abudu, A., Hishizawa, M., Imada, K., Tanaka, Y., and Uchiyama, T. (2005). APOBEC3G targets human T-cell leukemia virus type 1. Retrovirology *2*, 32.

Savarino, A., Mai, A., Norelli, S., El Daker, S., Valente, S., Rotili, D., Altucci, L., Palamara, A.T., and Garaci, E. (2009). "Shock and kill" effects of class I-selective histone deacetylase inhibitors in combination with the glutathione synthesis inhibitor buthionine sulfoximine in cell line models for HIV-1 quiescence. Retrovirology *6*, 52.

Sayah, D.M., Sokolskaja, E., Berthoux, L., and Luban, J. (2004). Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. Nature *430*, 569-573.

Schones, D.E., Cui, K., Cuddapah, S., Roh, T.Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. Cell *132*, 887-898.

Schreck, R., Rieber, P., and Baeuerle, P.A. (1991). Reactive oxygen intermediates as apparently widely used messengers in the activation of the NF-kappa B transcription factor and HIV-1. EMBO J *10*, 2247-2258.

Schroder, A.R., Shinn, P., Chen, H., Berry, C., Ecker, J.R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. Cell *110*, 521-529.

Schulte, A., Czudnochowski, N., Barboric, M., Schonichen, A., Blazek, D., Peterlin, B.M., and Geyer, M. (2005). Identification of a cyclin T-binding domain in Hexim1 and biochemical analysis of its binding competition with HIV-1 Tat. J Biol Chem *280*, 24968-24977.

Schwarzwaelder, K., Howe, S.J., Schmidt, M., Brugman, M.H., Deichmann, A., Glimm, H., Schmidt, S., Prinz, C., Wissler, M., King, D.J.*, et al.* (2007). Gammaretrovirus-mediated correction of SCID-X1 is associated with skewed vector integration site distribution in vivo. J Clin Invest *117*, 2241-2249.

Seger, R.A. (2008). Modern management of chronic granulomatous disease. Br J Haematol *140*, 255-266.

Selby, M.J., Bain, E.S., Luciw, P.A., and Peterlin, B.M. (1989). Structure, sequence, and position of the stem-loop in tar determine transcriptional elongation by tat through the HIV-1 long terminal repeat. Genes Dev *3*, 547-558.

Selliah, N., Zhang, M., DeSimone, D., Kim, H., Brunner, M., Ittenbach, R.F., Rui, H., Cron, R.Q., and Finkel, T.H. (2006). The gammac-cytokine regulated transcription factor, STAT5, increases HIV-1 production in primary CD4 T cells. Virology *344*, 283-291.

Sexton, T., Schober, H., Fraser, P., and Gasser, S.M. (2007). Gene regulation through nuclear organization. Nat Struct Mol Biol *14*, 1049-1055.

Sharma, N., and Nyborg, J.K. (2008). The coactivators CBP/p300 and the histone chaperone NAP1 promote transcription-independent nucleosome eviction at the HTLV-1 promoter. Proc Natl Acad Sci U S A *105*, 7959-7963.

Sheehy, A.M., Gaddis, N.C., Choi, J.D., and Malim, M.H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature *418*, 646-650.

Sheridan, P.L., Mayall, T.P., Verdin, E., and Jones, K.A. (1997). Histone acetyltransferases regulate HIV-1 enhancer activity in vitro. Genes Dev *11*, 3327-3340.

Shibahara, K., and Stillman, B. (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. Cell *96*, 575-585.

Shikama, N., Chan, H.M., Krstic-Demonacos, M., Smith, L., Lee, C.W., Cairns, W., and La Thangue, N.B. (2000). Functional interaction between nucleosome assembly proteins and p300/CREB-binding protein family coactivators. Mol Cell Biol *20*, 8933-8943.

Shintomi, K., Iwabuchi, M., Saeki, H., Ura, K., Kishimoto, T., and Ohsumi, K. (2005). Nucleosome assembly protein-1 is a linker histone chaperone in Xenopus eggs. Proc Natl Acad Sci U S A *102*, 8210-8215.

Shun, M.C., Daigle, J.E., Vandegraaff, N., and Engelman, A. (2007). Wild-type levels of human immunodeficiency virus type 1 infectivity in the absence of cellular emerin protein. J Virol *81*, 166-172.

Siekevitz, M., Josephs, S.F., Dukovich, M., Peffer, N., Wong-Staal, F., and Greene, W.C. (1987). Activation of the HIV-1 LTR by T cell mitogens and the trans-activator protein of HTLV-I. Science *238*, 1575-1578.

Simic, R., Lindstrom, D.L., Tran, H.G., Roinick, K.L., Costa, P.J., Johnson, A.D., Hartzog, G.A., and Arndt, K.M. (2003). Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. Embo J *22*, 1846-1856.

Simon, H.U., Mills, G.B., Kozlowski, M., Hogg, D., Branch, D., Ishimi, Y., and Siminovitch, K.A. (1994). Molecular characterization of hNRP, a cDNA encoding a human nucleosome-assembly-protein-I-related gene product involved in the induction of cell proliferation. Biochem J *297 (Pt 2)*, 389-397.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet *38*, 1348-1354.

Simonis, M., Kooren, J., and de Laat, W. (2007). An evaluation of 3C-based methods to capture DNA interactions. Nat Methods *4*, 895-901.

Singh, B.N., and Hampsey, M. (2007). A transcription-independent role for TFIIB in gene looping. Mol Cell *27*, 806-816.

Sivakumaran, H., van der Horst, A., Fulcher, A.J., Apolloni, A., Lin, M.H., Jans, D.A., and Harrich, D. (2009). Arginine Methylation Increases the Stability of HIV-1 Tat. J Virol.

Smith, S.M., Pentlicky, S., Klase, Z., Singh, M., Neuveut, C., Lu, C.Y., Reitz, M.S., Jr., Yarchoan, R., Marx, P.A., and Jeang, K.T. (2003). An in vivo

replication-important function in the second coding exon of Tat is constrained against mutation despite cytotoxic T lymphocyte selection. J Biol Chem *278*, 44816-44825.

Song, S.H., Hou, C., and Dean, A. (2007). A positive role for NLI/Ldb1 in long-range beta-globin locus control region function. Mol Cell *28*, 810-822.

Stauber, R.H., and Pavlakis, G.N. (1998). Intracellular trafficking and interactions of the HIV-1 Tat protein. Virology *252*, 126-136.

Steer, W.M., Abu-Daya, A., Brickwood, S.J., Mumford, K.L., Jordanaires, N., Mitchell, J., Robinson, C., Thorne, A.W., and Guille, M.J. (2003). Xenopus nucleosome assembly protein becomes tissue-restricted during development and can alter the expression of specific genes. Mech Dev *120*, 1045-1057.

Steger, D.J., and Workman, J.L. (1997). Stable co-occupancy of transcription factors and histones at the HIV-1 enhancer. EMBO J *16*, 2463-2472.

Strack, B., Calistri, A., Craig, S., Popova, E., and Gottlinger, H.G. (2003). AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. Cell *114*, 689-699.

Stremlau, M., Owens, C.M., Perron, M.J., Kiessling, M., Autissier, P., and Sodroski, J. (2004). The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. Nature *427*, 848-853.

Suhasini, M., and Reddy, T.R. (2009). Cellular proteins and HIV-1 Rev function. Curr HIV Res *7*, 91-100.

Suzuki, Y., and Craigie, R. (2007). The road to chromatin - nuclear entry of retroviruses. Nat Rev Microbiol *5*, 187-196.

Tan-Wong, S.M., French, J.D., Proudfoot, N.J., and Brown, M.A. (2008). Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. Proc Natl Acad Sci U S A *105*, 5160-5165.

Tasciotti, E., and Giacca, M. (2005). Fusion of the human immunodeficiency virus type 1 tat protein transduction domain to thymidine kinase increases bystander effect and induces enhanced tumor killing in vivo. Hum Gene Ther *16*, 1389-1403.

Taylor, N., Uribe, L., Smith, S., Jahn, T., Kohn, D.B., and Weinberg, K. (1996). Correction of interleukin-2 receptor function in X-SCID lymphoblastoid cells by retrovirally mediated transfer of the gamma-c gene. Blood *87*, 3103-3107.

Thambirajah, A.A., Li, A., Ishibashi, T., and Ausio, J. (2009). New developments in post-translational modifications and functions of histone H2A variants. Biochem Cell Biol *87*, 7-17.

Titz, B., Thomas, S., Rajagopala, S.V., Chiba, T., Ito, T., and Uetz, P. (2006). Transcriptional activators in yeast. Nucleic Acids Res *34*, 955-967.

Tiwari, V.K., Cope, L., McGarvey, K.M., Ohm, J.E., and Baylin, S.B. (2008). A novel 6C assay uncovers Polycomb-mediated higher order chromatin conformations. Genome Res *18*, 1171-1179.

Toledano, M.B., Ghosh, D., Trinh, F., and Leonard, W.J. (1993). N-terminal DNA-binding domains contribute to differential DNA-binding specificities of NF-kappa B p50 and p65. Mol Cell Biol *13*, 852-860.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell *10*, 1453-1465.

Torok, D., Ching, R.W., and Bazett-Jones, D.P. (2009). PML nuclear bodies as sites of epigenetic regulation. Front Biosci *14*, 1325-1336.

Toth, K.F., Mazurkiewicz, J., and Rippe, K. (2005). Association states of nucleosome assembly protein 1 and its complexes with histones. J Biol Chem *280*, 15690-15699.

Towers, G.J. (2007). The control of viral infection by tripartite motif proteins and cyclophilin A. Retrovirology *4*, 40.

Treand, C., du Chene, I., Bres, V., Kiernan, R., Benarous, R., Benkirane, M., and Emiliani, S. (2006). Requirement for SWI/SNF chromatin-remodeling complex in Tat-mediated activation of the HIV-1 promoter. EMBO J *25*, 1690-1699.

Triboulet, R., Mari, B., Lin, Y.L., Chable-Bessia, C., Bennasser, Y., Lebrigand, K., Cardinaud, B., Maurin, T., Barbry, P., Baillat, V.*, et al.* (2007). Suppression of microRNA-silencing pathway by HIV-1 during virus replication. Science *315*, 1579-1582.

Tsubota, T., Berndsen, C.E., Erkmann, J.A., Smith, C.L., Yang, L., Freitas, M.A., Denu, J.M., and Kaufman, P.D. (2007). Histone H3-K56 acetylation is catalyzed by histone chaperone-dependent complexes. Mol Cell *25*, 703-712.

Tyagi, M., and Karn, J. (2007). CBF-1 promotes transcriptional silencing during the establishment of HIV-1 latency. EMBO J *26*, 4985-4995.

Tyagi, M., Rusnati, M., Presta, M., and Giacca, M. (2001). Internalization of HIV-1 tat requires cell surface heparan sulfate proteoglycans. J Biol Chem *276*, 3254-3261.

Tyler, J.K., Adams, C.R., Chen, S.R., Kobayashi, R., Kamakaka, R.T., and Kadonaga, J.T. (1999). The RCAF complex mediates chromatin assembly during DNA replication and repair. Nature *402*, 555-560.

Vallabhapurapu, S., and Karin, M. (2009). Regulation and function of NF-kappaB transcription factors in the immune system. Annu Rev Immunol *27*, 693-733.

Van Duyne, R., Easley, R., Wu, W., Berro, R., Pedati, C., Klase, Z., Kehn-Hall, K., Flynn, E.K., Symer, D.E., and Kashanchi, F. (2008). Lysine methylation of HIV-1 Tat regulates transcriptional activity of the viral LTR. Retrovirology *5*, 40.

Van Lint, C., Amella, C.A., Emiliani, S., John, M., Jie, T., and Verdin, E. (1997). Transcription factor binding sites downstream of the human immunodeficiency virus type 1 transcription start site are important for virus infectivity. J Virol *71*, 6113-6127.

Van Lint, C., Emiliani, S., Ott, M., and Verdin, E. (1996). Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. EMBO J *15*, 1112-1120.

Varthakavi, V., Heimann-Nichols, E., Smith, R.M., Sun, Y., Bram, R.J., Ali, S., Rose, J., Ding, L., and Spearman, P. (2008). Identification of calcium-modulating cyclophilin ligand as a human host restriction to HIV-1 release overcome by Vpu. Nat Med *14*, 641-647.

Verdin, E., Paras, P., Jr., and Van Lint, C. (1993). Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. Embo J *12*, 3249-3259.

von Lindern, M., van Baal, S., Wiegant, J., Raap, A., Hagemeijer, A., and Grosveld, G. (1992). Can, a putative oncogene associated with myeloid leukemogenesis, may be activated by fusion of its 3' half to different genes: characterization of the set gene. Mol Cell Biol *12*, 3346-3355.

von Schwedler, U.K., Stuchell, M., Muller, B., Ward, D.M., Chung, H.Y., Morita, E., Wang, H.E., Davis, T., He, G.P., Cimbora, D.M.*, et al.* (2003). The protein network of HIV budding. Cell *114*, 701-713.

Wahlers, A., Kustikova, O., Zipfel, P.F., Itoh, K., Koester, M., Heberlein, C., Li, Z., Schiedlmeier, B., Skerka, C., Fehse, B.*, et al.* (2002a). Upstream conserved sequences of mouse leukemia viruses are important for high transgene expression in lymphoid and hematopoietic cells. Mol Ther *6*, 313-320.

Wahlers, A., Zipfel, P.F., Schwieger, M., Ostertag, W., and Baum, C. (2002b). In vivo analysis of retroviral enhancer mutations in hematopoietic cells: SP1/EGR1 and ETS/GATA motifs contribute to long terminal repeat specificity. J Virol *76*, 303-312.

Walfridsson, J., Khorosjutina, O., Matikainen, P., Gustafsson, C.M., and Ekwall, K. (2007). A genome-wide role for CHD remodelling factors and Nap1 in nucleosome disassembly. EMBO J *26*, 2868-2879.

Walter, P.P., Owen-Hughes, T.A., Cote, J., and Workman, J.L. (1995). Stimulation of transcription factor binding and histone displacement by nucleosome assembly protein 1 and nucleoplasmin requires disruption of the histone octamer. Mol Cell Biol *15*, 6178-6187.

Wang, Q., Carroll, J.S., and Brown, M. (2005). Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. Mol Cell *19*, 631-642.

Wang, S., and Frappier, L. (2009). Nucleosome Assembly Proteins Bind to Epstein-Barr Nuclear Antigen 1 (EBNA1) and Affect its Functions in DNA Replication and Transcriptional Activation. J Virol.

Weeks, K.M., and Crothers, D.M. (1991). RNA recognition by Tat-derived peptides: interaction in the major groove? Cell *66*, 577-588.

Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H., and Jones, K.A. (1998). A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell *92*, 451-462.

Weissman, J.D., Brown, J.A., Howcroft, T.K., Hwang, J., Chawla, A., Roche, P.A., Schiltz, L., Nakatani, Y., and Singer, D.S. (1998). HIV-1 tat binds TAFII250 and represses TAFII250-dependent transcription of major histocompatibility class I genes. Proc Natl Acad Sci U S A *95*, 11601-11606.

Westendorp, M.O., Frank, R., Ochsenbauer, C., Stricker, K., Dhein, J., Walczak, H., Debatin, K.M., and Krammer, P.H. (1995). Sensitization of T cells to CD95-mediated apoptosis by HIV-1 Tat and gp120. Nature *375*, 497-500.

Williams, S.A., Chen, L.F., Kwon, H., Ruiz-Jarabo, C.M., Verdin, E., and Greene, W.C. (2006). NF-kappaB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. EMBO J *25*, 139-149.

Williams, S.A., and Greene, W.C. (2007). Regulation of HIV-1 latency by T-cell activation. Cytokine *39*, 63-74.

Workman, J.L. (2006). Nucleosome displacement in transcription. Genes Dev *20*, 2009-2017.

Wu, X., Li, Y., Crise, B., and Burgess, S.M. (2003). Transcription start regions in the human genome are favored targets for MLV integration. Science *300*, 1749-1751.

Wu, X., Luke, B.T., and Burgess, S.M. (2006). Redefining the common insertion site. Virology *344*, 292-295.

Wu, Y., and Marsh, J.W. (2001). Selective transcription and modulation of resting T cell activity by preintegrated HIV DNA. Science *293*, 1503-1506.

Wurtele, H., and Chartrand, P. (2006). Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. Chromosome Res *14*, 477-495.

Wysocka, J., Myers, M.P., Laherty, C.D., Eisenman, R.N., and Herr, W. (2003). Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. Genes Dev *17*, 896-911.

Xiao, H., Neuveut, C., Benkirane, M., and Jeang, K.T. (1998). Interaction of the second coding exon of Tat with human EF-1 delta delineates a mechanism for HIV-1-mediated shut-off of host mRNA translation. Biochem Biophys Res Commun *244*, 384-389.

Xie, B., Invernizzi, C.F., Richard, S., and Wainberg, M.A. (2007). Arginine methylation of the human immunodeficiency virus type 1 Tat protein by PRMT6 negatively affects Tat Interactions with both cyclin T1 and the Tat transactivation region. J Virol *81*, 4226-4234.

Yang, Z., Yik, J.H., Chen, R., He, N., Jang, M.K., Ozato, K., and Zhou, Q. (2005). Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. Mol Cell *19*, 535-545.

Yang, Z., Zhu, Q., Luo, K., and Zhou, Q. (2001). The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. Nature *414*, 317-322.

Yedavalli, V.S., Benkirane, M., and Jeang, K.T. (2003). Tat and trans-activation-responsive (TAR) RNA-independent induction of HIV-1 long terminal repeat by human and murine cyclin T1 requires Sp1. J Biol Chem *278*, 6404-6410.

Yeung, F., Hoberg, J.E., Ramsey, C.S., Keller, M.D., Jones, D.R., Frye, R.A., and Mayo, M.W. (2004). Modulation of NF-kappaB-dependent transcription and cell survival by the SIRT1 deacetylase. EMBO J *23*, 2369-2380.

Yik, J.H., Chen, R., Nishimura, R., Jennings, J.L., Link, A.J., and Zhou, Q. (2003). Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. Mol Cell *12*, 971-982.

Yik, J.H., Chen, R., Pezda, A.C., Samford, C.S., and Zhou, Q. (2004). A human immunodeficiency virus type 1 Tat-like arginine-rich RNA-binding domain is

essential for HEXIM1 to inhibit RNA polymerase II transcription through 7SK snRNA-mediated inactivation of P-TEFb. Mol Cell Biol *24*, 5094-5105.

Yu, L., Loewenstein, P.M., Zhang, Z., and Green, M. (1995). In vitro interaction of the human immunodeficiency virus type 1 Tat transactivator and the general transcription factor TFIIB with the cellular protein TAP. J Virol *69*, 3017-3023.

Zack, J.A., Arrigo, S.J., Weitsman, S.R., Go, A.S., Haislip, A., and Chen, I.S. (1990). HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. Cell *61*, 213-222.

Zaiss, A.K., Son, S., and Chang, L.J. (2002). RNA 3' readthrough of oncoretrovirus and lentivirus: implications for vector safety and efficacy. J Virol *76*, 7209-7219.

Zhao, R., Bodnar, M.S., and Spector, D.L. (2009). Nuclear neighborhoods and gene expression. Curr Opin Genet Dev *19*, 172-179.

Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U.*, et al.* (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet *38*, 1341-1347.

Zhou, M., Deng, L., Lacoste, V., Park, H.U., Pumfery, A., Kashanchi, F., Brady, J.N., and Kumar, A. (2004). Coordination of transcription factor phosphorylation and histone methylation by the P-TEFb kinase during human immunodeficiency virus type 1 transcription. J Virol *78*, 13522-13533.

Zhou, M., Halanski, M.A., Radonovich, M.F., Kashanchi, F., Peng, J., Price, D.H., and Brady, J.N. (2000). Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. Mol Cell Biol *20*, 5077-5086.

Zhou, M., Huang, K., Jung, K.J., Cho, W.K., Klase, Z., Kashanchi, F., Pise-Masison, C.A., and Brady, J.N. (2009). Bromodomain protein Brd4 regulates human immunodeficiency virus transcription through phosphorylation of CDK9 at threonine 29. J Virol *83*, 1036-1044.

Zhou, Q., Chen, D., Pierstorff, E., and Luo, K. (1998). Transcription elongation factor P-TEFb mediates Tat activation of HIV-1 transcription at multiple stages. EMBO J *17*, 3681-3691.

Zhou, Q., and Yik, J.H. (2006). The Yin and Yang of P-TEFb regulation: implications for human immunodeficiency virus gene expression and global

control of cell growth and differentiation. Microbiol Mol Biol Rev *70*, 646-659.

Zlatanova, J., Seebart, C., and Tomschik, M. (2007). Nap1: taking a closer look at a juggler protein of extraordinary skills. FASEB J *21*, 1294-1310.

Zou, Z., Anisowicz, A., Hendrix, M.J., Thor, A., Neveu, M., Sheng, S., Rafidi, K., Seftor, E., and Sager, R. (1994). Maspin, a serpin with tumor-suppressing activity in human mammary epithelial cells. Science *263*, 526-529.

Images of the cover: Fabrice Hyber (www.hyber.tv).

APPENDIX

Vardabasso C., Manganaro L., Lusic M., Marcello A., and Giacca M. (2008). The histone chaperone proteinNucleosome Assembly Protein-1 (hNAP-1) binds HIV-1 Tat and promotes viral transcription. *Retrovirology 2008, 5:8.*
PAPER ATTACHED

De Marco A., Biancotto C., Knezevich A., Maiuri P., Vardabasso C., and Marcello A. (2008). Intragenic transcriptional *cis*-activation of the human immunodeficiency virus 1 does not result in allele-specific inhibition of the endogenous gene. *Retrovirology 2008, 5:98.*
PAPER ATTACHED

Agostini S., Vardabasso C., Fittipaldi A., Tasciotti E., Cereseto A., and Giacca M. Cellular Na^+/K^+ ATPase mediates extracellular release of HIV-1 Tat and Tat-fusion proteins.
PAPER IN PREPARATION

Retrovirology

BioMed Central

Research

Open Access

The histone chaperone protein Nucleosome Assembly Protein-1 (hNAP-1) binds HIV-1 Tat and promotes viral transcription

Chiara Vardabasso[1], Lara Manganaro[1], Marina Lusic[1], Alessandro Marcello[2] and Mauro Giacca*[1]

Address: [1]Molecular Medicine Laboratory, International Centre for Genetic Engineering and Biotechnology (ICGEB), Padriciano 99, 34012 Trieste, Italy and [2]Molecular Virology Laboratory, International Centre for Genetic Engineering and Biotechnology (ICGEB), Padriciano 99, 34012 Trieste, Italy

Email: Chiara Vardabasso - vardabas@icgeb.org; Lara Manganaro - manganar@icgeb.org; Marina Lusic - lusic@icgeb.org; Alessandro Marcello - marcello@icgeb.org; Mauro Giacca* - giacca@icgeb.org

* Corresponding author

Published: 28 January 2008

Retrovirology 2008, **5**:8 doi:10.1186/1742-4690-5-8

Received: 1 October 2007
Accepted: 28 January 2008

This article is available from: http://www.retrovirology.com/content/5/1/8

Abstract

Background: Despite the large amount of data available on the molecular mechanisms that regulate HIV-1 transcription, crucial information is still lacking about the interplay between chromatin conformation and the events that regulate initiation and elongation of viral transcription. During transcriptional activation, histone acetyltransferases and ATP-dependent chromatin remodeling complexes cooperate with histone chaperones in altering chromatin structure. In particular, human Nucleosome Assembly Protein-1 (hNAP-1) is known to act as a histone chaperone that shuttles histones H2A/H2B into the nucleus, assembles nucleosomes and promotes chromatin fluidity, thereby affecting transcription of several cellular genes.

Results: Using a proteomic screening, we identified hNAP-1 as a novel cellular protein interacting with HIV-1 Tat. We observed that Tat specifically binds hNAP1, but not other members of the same family of factors. Binding between the two proteins required the integrity of the basic domain of Tat and of two separable domains of hNAP-1 (aa 162–290 and 290–391). Overexpression of hNAP-1 significantly enhanced Tat-mediated activation of the LTR. Conversely, silencing of the protein decreased viral promoter activity. To explore the effects of hNAP-1 on viral infection, a reporter HIV-1 virus was used to infect cells in which hNAP-1 had been either overexpressed or knocked-down. Consistent with the gene expression results, these two treatments were found to increase and inhibit viral infection, respectively. Finally, we also observed that the overexpression of p300, a known co-activator of both Tat and hNAP-1, enhanced hNAP-1-mediated transcriptional activation as well as its interaction with Tat.

Conclusion: Our study reveals that HIV-1 Tat binds the histone chaperone hNAP-1 both in vitro and in vivo and shows that this interaction participates in the regulation of Tat-mediated activation of viral gene expression.

Background

Efficient packaging of DNA in a highly organized chromatin structure inside the cell is one of the most remarkable characteristics of all eukaryotic organisms. Chromatin assembly and disassembly are dynamic biological processes that increase chromatin fluidity and regulate the accessibility of the genome to all DNA transactions, including transcription, DNA replication and DNA repair. The basic structural unit of eukaryotic chromatin is the nucleosome, formed by the wrapping of DNA around an octamer of core histone proteins. By restricting the access to DNA-binding factors and impeding elongation by RNA polymerase II (RNAPII), the nucleosome is not only a structural unit of the chromosome, but perhaps the most important regulator of gene expression (for recent reviews, see refs. [1,2]). Chromatin structure is modulated by the covalent modifications of the N-termini of the core histones in nucleosomes and by the action of ATP-dependent chromatin remodeling complexes. In particular, histone acetylation at the promoter of genes, mediated by histone acetyltransferases (HATs), has been shown to be necessary, albeit not sufficient, for transcriptional activation [2,3].

Chromatin assembly is a stepwise process which requires histone chaperones to deposit histones on forming nucleosomes (reviewed in refs. [4-7]). The Nucleosome Assembly Protein-1 (NAP-1) is one of the major histone chaperones involved in this process. This factor belongs to the NAP family of proteins, which is characterized by the presence of a NAP domain [8]. NAP-1 is conserved in all eukaryotes from yeast to humans [9-12], and is responsible for the incorporation of two histone H2A-H2B dimers to complete the nucleosome (reviewed in ref. [7]). The protein acts as a nucleo-cytoplasmic shuttling factor that delivers H2A-H2B dimers from cytoplasm to the chromatin assembly machinery in the nucleus [13]. In addition, NAP-1 has been involved in the regulation of cell-cycle progression [14-16], incorporation and exchange of histone variants [17-19], and promotion of nucleosome sliding [20].

Most relevant to the regulation of gene expression, the chromatin-modifying activity of histone chaperones also facilitates transcription. In particular, recent information suggests that HAT complexes as well as ATP-dependent chromatin remodeling complexes cooperate with histone chaperones in altering chromatin structure during transcriptional activation [21-24]. In addition, NAP proteins have been reported to interact with the histone acetyltransferase (HAT) and transcriptional coactivator p300/CBP [25-27], suggesting that NAPs may augment activation by all the transcription factors that use p300/CBP as a co-activator. Accordingly, a yeast two-hybrid screen revealed that hNAP-1 forms a complex with the HPV E2 transcription factor, and a complex formed by hNAP-1, E2 and p300 proved able to activate transcription in vitro [28].

One of the promoters that show exquisite sensitivity to regulation by chromatin structure and its modifications is the long terminal repeat (LTR) of the Human Immunodeficiency Virus type 1 (HIV-1) (reviewed in ref. [29]). Following infection of susceptible cells, the HIV-1 provirus becomes integrated into the host genome and, for still poorly understood reasons, the LTR promoter enters a latent state and becomes silenced by chromatin conformation [29,30]. Independent of the site of integration, two distinct nucleosomes are precisely positioned in the 5' LTR, separated by a nuclease-hypersensitivity region containing the enhancer and basal promoter elements [31-34]. Genomic footprinting experiments performed in either activated or latently infected cells have revealed that most of the critical protein-DNA interactions in the promoter region are preserved, independent from the LTR activation state [35,36]. This observation first indicated that the transcriptional activation of the integrated LTR is not primarily restricted by DNA target site accessibility, but occurs through the modulation of chromatin conformation. Indeed, Nuc-1, which is positioned near the viral mRNA start site, appears to exert a repressive role on transcription; this nucleosome becomes remodelled when HIV-1 transcription is activated [37,38]. Which are the factors involved in chromatin remodelling during transcriptional activation, besides the recruitment of several HATs [39], is a still poorly addressed question.

One of the key factors involved in transcriptional activation of the provirus is the HIV-1 Tat protein, a highly unusual transactivator that binds an RNA element (TAR) positioned at the 5' end of the primary proviral transcript [40]. Tat activates HIV-1 transcription by promoting the assembly of transcriptionally active complexes at the LTR by multiple protein-protein interactions. Over the last few years, a number of cellular proteins have been reported to interact with Tat and to mediate or modulate its activity. Among these interacting partners, a major role can be ascribed to the P-TEFb complex [41-43] and to several cellular HATs, including p300/CBP, P/CAF and GCN5 [44-47]. P-TEFb promotes processive transcription by phosphorylating the RNAPII carboxy-terminal domain (CTD) [48,49], while HATs induce the activation of chromatinized HIV-1 LTR through the acetylation of histones [39]. Of interest, optimal Tat-mediated activation of viral gene expression also requires the function of ATP-dependent chromatin-remodelling complexes [50].

In this work we address the issue of identifying novel cellular interactors of Tat through a proteomic screening. We identify human NAP-1 as a major Tat partner and show

that the interaction between the two proteins is important for Tat-mediated transcriptional activation and for efficient viral infection.

Results

Identification of cellular factors binding to HIV-1 Tat by proteomic analysis

With the aim of identifying cellular partners of HIV-1 Tat through a proteomic approach, we used an expression vector encoding the open reading frame of full length Tat (101 aa) fused with a C-terminal Flag tag. This epitope-tagged version of Tat was active in HIV-1 LTR transactivation similar to the wild type protein (data not shown). Extracts from HEK 293T cells transfected with Flag-Tat101, as well as from mock-transfected cells, were immunoprecipitated with M2 Flag antibody conjugated to agarose beads. Affinity purified Tat-Flag protein and co-purifying cellular factors were subsequently eluted with an excess of Flag peptide, run on a 6–15% gradient SDS-PAGE gel and stained with silver stain (Figure 1). Individual bands that were apparent only in the sample from Tat-Flag transfected cells were excised and their identification attempted by ESI-MS/MS (Electrospray tandem Mass Spectrometry) analysis of peptides obtained after trypsin digestion. Five bands were unequivocally identified, as shown in Figure 1. One corresponding to Tat-Flag itself; B23/nucleophosmin, a nucleolar protein possibly associated with ribosome assembly and/or transport [51]; the p32 protein, an inhibitor of the ASF/SF2 splicing regulator [52], also known as Tat-associated protein (TAP) [53,54]; ribosomal protein S4 and the histone chaperone NAP-1 (Nucleosome Assembly Protein-1). The proteomic analysis was repeated and the results were also confirmed by sequencing proteins directly from the Flag beads, rather than from gel-excised bands.

Since overexpressed Tat is known to accumulate in the nucleoli, probably due to its unspecific RNA binding capacity, and given the observation that the same proteomic assay resulted in the identification of a number of other ribosomal proteins when performed in the absence of RNase (data not shown), no further work was performed on the B23/nucleophosmin and ribosomal S4 proteins. In this respect, other investigators have already shown that Tat binds B23/nucleophosmin when both proteins are overexpressed [55] and that B23/nucleophosmin protein is required for Tat nucleolar localization but not for promoter transactivation [56]. The rest of our research was therefore focused on the characterization of the hNAP-1/Tat interaction.

HIV-1 Tat interacts with hNAP-1 in vivo

A schematic representation of hNAP-1 is shown in Figure 2A. The protein has 391 amino acids, contains three acidic domains and has a long KIX-binding domain. This

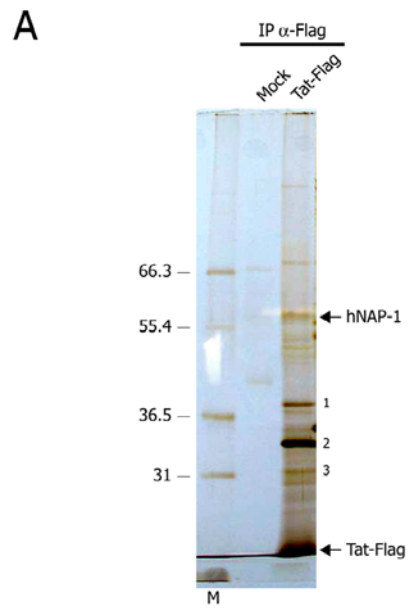


Figure 1
Identification of Tat-interacting proteins by mass spectrometry. A. Flag-immunoprecipitated material from Tat-Flag- and mock-transfected HEK 293T cells was resolved by 6–10% gradient SDS-PAGE gel, followed by silver staining. Protein bands present exclusively in the sample transfected with Tat-Flag were excised from the gel and their identification attempted by ESI-MS/MS. The identified proteins, in addition to hNAP-1 and Tat-Flag, are indicated (1: B23/nucleophosmin; 2: pre-mRNA splicing factor SF2p32 – Tat-associated protein TAP; 3: ribosomal protein S4). **B.** Amino acid sequence of the human NAP-1 protein (locus NP_631946) – 391 aa. The underlined amino acid sequences correspond to peptides obtained from MS/MS analysis of three independent preparations ($P = 7.8 \times 10^{-19}$).

domain and the C-terminal acidic domain are very conserved in other members of the NAP family of histone chaperones, including SET-TAF-I (47% and 68% amino acid homology in the two regions respectively [57,58];

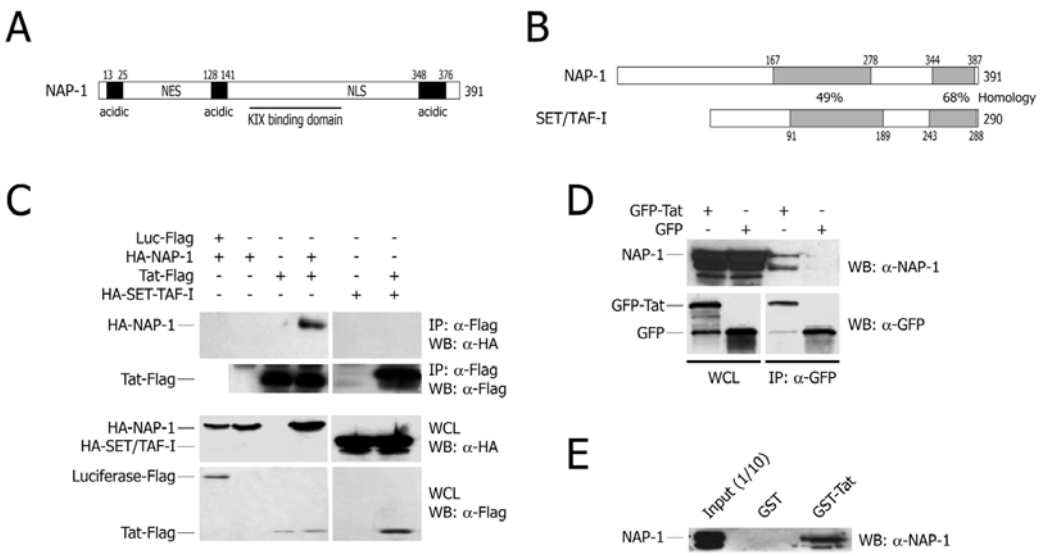


Figure 2

Co-immunoprecipitation of Tat with transfected and endogenous hNAP-1. A. Schematic representation of hNAP-1 structure. The acidic domains of the protein are shown by black boxes, with the indication of their boundary amino acids. The localization of nuclear export and nuclear localization signals (NES and NLS respectively) are indicated. **B.** Schematic representation of the regions of amino acid homology between hNAP-1 and hSET/TAF-I. **C.** Co-immunoprecipitation of transfected hNAP-1 with Tat. The plasmids indicated on top of the figure were transfected into HEK 293T cells. The upper two panels show western blots with the indicated antibodies after immunoprecipitation using an anti-Flag antibody; the lower two panels show western blotting controls from whole cell lysates (WCL) from transfected cells to show the levels of expression of the transfected proteins. **D.** Co-immunoprecipitation of endogenous hNAP-1 with Tat. The experiment was performed by transfecting HEK 293T cells with plasmids encoding GFP-Tat or GFP alone, followed by co-immunoprecipitation with anti-GFP antibody. GFP-Tat retains full transcriptional and trafficking capacities as wt Tat [69, 74, 75]. **E.** GST-pulldown experiment using GST-Tat and HEK 293T whole cell lysates. GST-Tat, but not control GST protein, pulled down endogenous hNAP-1.

Figure 2B).

The interaction between HIV-1 Tat and hNAP-1 was confirmed by co-immunoprecipitation analysis. When expression vectors for Tat-Flag and for an N-terminal HA-tagged version of hNAP-1 (HA-NAP-1) were transfected into HEK 293T, HA-NAP-1 was co-immunoprecipitated with Tat using anti-Flag antibody (Figure 2C). The specificity of interaction of the two proteins is underlined by the observation that no co-immunoprecipitation was observed when Tat was co-expressed with HA-hSET/TAF-I, despite its sequence homology with hNAP-1 (Figure 2C).

Tat was also found to bind endogenous hNAP-1. As shown in Figure 2D, an anti-GFP antibody was able to precipitate endogenous hNAP-1, as detected with an anti-hNAP-1 antibody, from extracts of cells transfected with GFP-Tat but not from extracts of cells transfected with control GFP.

Finally, a bacterially expressed and purified GST-Tat recombinant protein was also able to pull-down endogenous hNAP-1 from a HEK 293T cell extract (Figure 2E).

Binding domain analysis

The domains within hNAP-1 and HIV-1 Tat that were responsible for the interaction were defined by in vitro GST-pulldown assays. A series of N- and C-terminal deletion mutants of hNAP-1 (Figure 3A) was expressed after fusion to GST, and incubated with ^{35}S-labeled full-length HIV-1 Tat obtained by in vitro translation. All deletants lacking the N-terminus of the protein up to aa 161 bound Tat as efficiently as the full length protein; in contrast, binding was impaired when the hNAP-1 domain from residues 163 to 289 as well as the C-terminal region from

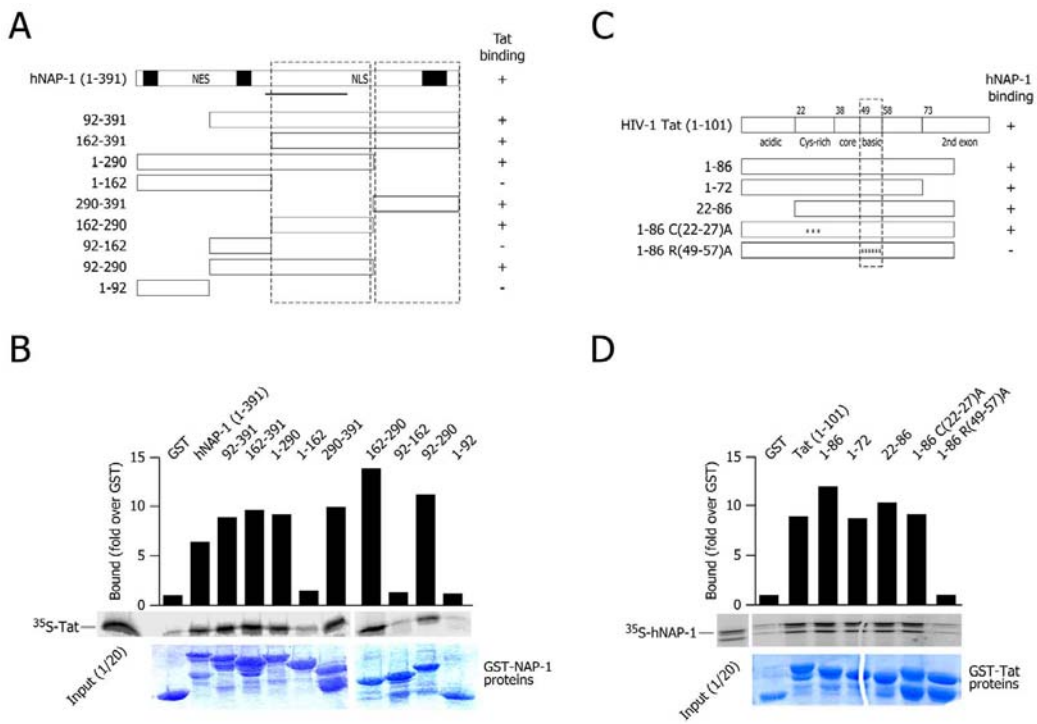


Figure 3

Mapping of hNAP-1 and Tat interacting domains. A. Schematic representation of hNAP-1 protein and of its deletion mutants obtained as GST fusion proteins. The capacity of binding to Tat – see experiment in panel B – is indicated on the right side of each mutant. The two dotted boxes indicate the hNAP-1 domains interacting with Tat. **B.** Representative GST pull-down experiment using the indicated hNAP-1 mutants and radiolabelled Tat101 protein. The autoradiography shows the amount of Tat binding to each mutant; the histogram on top shows densitometric quantification of data, expressed as fold binding with respect to background binding to GST alone (set as 1). The lower panel shows the Coomassie stained gel at the end of the binding experiment. The experiment was repeated at least three times with similar results. **C.** Schematic representation of HIV-1 Tat protein and of its mutants obtained as GST fusion proteins. The capacity of binding to hNAP-1 – see experiment in panel D – is indicated on the right side of each mutant. The dotted box corresponds to the basic domain of Tat, which binds hNAP-1. **D.** Representative GST pulldown experiment using the indicated Tat mutants (obtained as GST fusion proteins) and in vitro transcribed and translated hNAP-1 protein. The autoradiography shows the amount of hNAP-1 binding to each mutant; the histogram on top shows densitometric quantification of data, expressed as fold binding with respect to background binding to GST alone (set as 1). The lower panel shows the Coomassie stained gel at the end of the binding experiment. The experiment was repeated at least three times with similar results.

residues 290 to 391 were deleted (Figure 3B). These results indicate that Tat binds two separable domains within hNAP-1, one internal from amino acids 162 to 290 and one C-terminal from residues 290 to 391.

Next we analyzed the domains of Tat responsible for the interaction with hNAP-1. GST pull-down experiments were performed using wild type Tat (101 aa), Tat72 (lacking the second exon), Tat86 (HXB2 clone), and mutated derivatives of Tat86 carrying cysteine to alanine mutations at positions 22, 25 and 27 in the cysteine-rich domain or arginine to alanine mutations at positions 49, 52, 53, 55, 56 and 57 in the basic domain (Tat86 C(22–27)A and R(49–57)A respectively); Figure 3C. These proteins, obtained as C-terminal fusions to GST, were used to pull-down ^{35}S-methionine-labelled hNAP-1 obtained by in

vitro transcription/translation. The results obtained demonstrated that hNAP-1 bound the basic domain of HIV-1 Tat (Figure 3D).

hNAP-1 and Tat cooperate in the activation of HIV-1 gene expression

One of the essential molecular events that parallel Tat-driven transcriptional activation is the modification of chromatin structure at the HIV-1 promoter [34,39]. We therefore investigated whether NAP-1 might contribute to Tat transactivation. A reporter construct containing the U3 and R sequences of the HIV-1 LTR upstream of the luciferase gene was co-transfected into HeLa cells, together with vectors for HA-tagged hNAP-1 and HIV-1 Tat. As shown in Figure 4A, hNAP-1, when co-transfected with Tat, significantly enhanced Tat-mediated transactivation of the LTR; hNAP-1 alone had no effect on promoter activity.

To test the requirement for endogenous hNAP-1 protein in Tat-mediated HIV-1 LTR transactivation, luciferase assays were performed with HeLa cells in which expression of hNAP-1 was down-regulated by RNAi. A specific siRNA oligonucleotide was designed which was able to silence ~80% of the expression of its target from 48 hours after transfection onward, as assessed by western blot analysis (Figure 4B). In hNAP-1-knock down cells, Tat transactivation of the HIV-1 LTR was significantly impaired, compared to cells treated with a control siRNA.

Collectively, the results of these experiments indicate that hNAP-1 participates in Tat-mediated control of HIV-1 gene expression.

p300, hNAP-1 and Tat synergistically activate HIV-1 transcription

Previous work has indicated that NAP-1 interacts with the cellular transcriptional co-activator and histone acetyltransferase p300 [25-27]. Since p300 is also an essential co-factor for Tat transactivation, we investigated the effects of hNAP-1 and p300 on Tat-mediated transactivation. For this purpose, HeLa cells were transfected with an LTR-luciferase reporter plasmid and expression vectors for p300 and hNAP-1 together with Tat. As previously described [47], p300 enhanced Tat-driven transcriptional activation; when hNAP-1 was co-transfected, transcription was further increased (~3.5 fold Tat+hNAP-1+p300 over Tat alone; Figure 4C).

As shown in the co-immunoprecipitation experiment in Figure 4D, the overexpression of p300 in the same experimental conditions did not affect the levels of expression of NAP-1 or Tat proteins (as shown in the anti-Flag immunoblot). However, in cells overexpressing p300, the amount of hNAP-1 protein co-immunoprecipitating with Tat was markedly increased, a result that is consistent with the possibility that p300 might stabilize the formation of the Tat-hNAP-1 complex in vivo.

Effect of hNAP-1 on HIV-1 infection

To further examine the effect of hNAP-1 on viral replication, we used an HIV vector in which a portion of nef had been replaced by the firefly luciferase gene; two frameshifts inactivate vpr and env in this clone, thus blocking subsequent rounds of viral replication. Infectious virus, pseudotyped with VSV-G, was produced by transfections of HEK 293T cells, and used to infect HeLa cells in which hNAP-1 had been earlier either overexpressed or knocked down by RNAi. As shown in Figure 5A, the overexpression of hNAP-1 (as assessed by western blot analysis) resulted in a 5-fold increase of luciferase activity in HA-hNAP-1-transfected cells compared to mock-transfected cells. Conversely, in cells in which the levels of hNAP-1 had been reduced to <20% by RNAi, viral luciferase activity was reduced 3-fold compared to control-treated cells (Figure 5B).

Taken together, these results support the conclusion that hNAP-1 also plays an important activating role in the context of HIV-1 infection.

Discussion

Activation of the HIV-1 LTR is a complex event involving the coordinated function of several cellular proteins acting by both releasing the negative inhibition that chromatin imposes on the promoter and inducing the recruitment of elongation-competent RNPII-containing complexes. Tat appears to exert an essential activating function for both these processes. In the last decade, a number of laboratories have reported the identification of various cellular factors that mediate Tat function. These factors fall in several broad categories, including members of the basal transcriptional machinery, among which RNAPII itself, ubiquitous transcription factors, transcriptional co-activators, histone-acetyltransferases, and others [29,59,60]. Our proteomic screening led to the identification of yet another cellular partner, hNAP-1, that appears to be essentially involved in mediating Tat function. We could confirm the interaction between Tat and hNAP-1 both in vitro and inside the cells, and demonstrate its specificity by showing that Tat was not able to co-precipitate hSET/TAF I, another member of the NAP family of proteins. The relevance of the detected interaction between Tat and hNAP-1 was further reinforced by the observations that the overexpression of hNAP-1 stimulated Tat-mediated transactivation of the LTR as well as increased HIV-1 infection. Conversely, the down-regulation of the protein by RNAi impaired both transcription and viral infection. To our knowledge, this is the first demonstration of an interaction between Tat and a histone chaper-

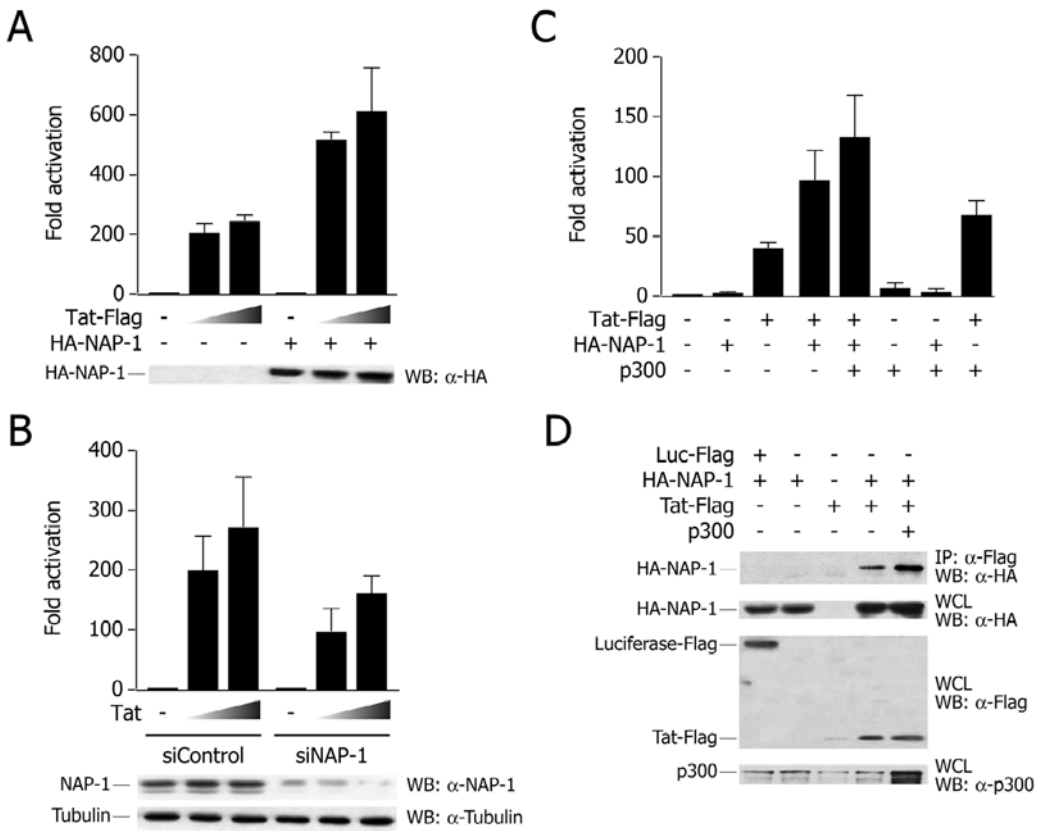


Figure 4

hNAP-1 cooperates with Tat in LTR transactivation. A. hNAP-1 synergizes with Tat in transcriptional activation. HeLa cells were cotransfected with a reporter construct containing the HIV-1 LTR upstream of the luciferase gene, and with vectors for HA-tagged hNAP-1 (100 ng) and HIV-1 Tat (5 and 25 ng), as indicated. The histogram shows mean ± s.d. of at least three independent experiments; the results are shown as fold transactivation over LTR-luciferase reporter alone. The co-expression of hNAP-1 significantly increased Tat transactivation of the LTR promoter. The western blot at the bottom shows the levels of transfected hNAP-1 protein in a representative experiment. **B.** hNAP-1 knock down decreases Tat transactivation. HeLa cells were transfected with a specific siRNA against hNAP-1 or a control siRNA, and then transfected with the LTR-luciferase reporter together with Tat (5 and 25 ng). The histogram shows mean ± s.d. of at least three independent experiments; the results are shown as fold transactivation over LTR-luciferase reporter alone. The western blot at the bottom shows the levels of endogenous hNAP-1 protein and of tubulin as a control in a representative experiment. **C.** hNAP-1, Tat and the acetyltransferase p300 synergistically activate viral transcription. HeLa cells were transfected with LTR-luciferase reporter plasmid and with vectors for HIV-1 Tat (5 ng), HA-hNAP-1 (100 ng) and p300 (100 ng), as indicated. After 24 h from transfection, luciferase assays were performed. The histogram shows mean ± s.d. of at least three independent experiments; the results are shown as fold transactivation over LTR-luciferase reporter alone. **D.** p300 enhances Tat-hNAP-1 interaction in vivo. The plasmids indicated on top of the figure were transfected into HEK 293T cells. The upper panel shows western blots with the indicated antibodies after immunoprecipitation using an anti-Flag antibody; the lower three panels show western blotting controls from whole cell lysates (WCL) from transfected cells to show the levels of expression of the transfected proteins.

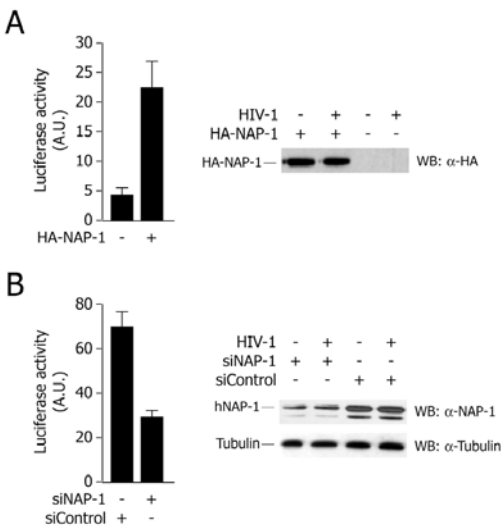


Figure 5

Effect of hNAP-1 on HIV-1 infection. A. Overexpression of hNAP-1 enhances LTR transcription upon HIV-1 infection. HeLa cells were transfected with an expression vector for HA-hNAP-1 or with a control vector, and then infected with VSG-luciferase HIV-1 vector. Luciferase activity was measured after 24 h post-infection. The mean ± s.d. of at least three different experiments is shown. The panel on the right side shows anti-HA western blottings to assess HA-hNAP-1 expression in a representative experiment. **B.** Silencing of hNAP-1 impairs LTR transcription upon HIV-1 infection. HeLa cells were treated with an siRNA directed against hNAP-1 or a control siRNA. Forty-eight hours after the beginning of siRNA treatment, cells were infected with the luciferase reported virus, and luciferase assays were performed on cell lysates 24 hours later. The mean ± s.d. of at least three different experiments is shown. The panel on the right side shows anti-hNAP-1 western blottings to assess the levels of endogenous hNAP-1 and tubulin expression in a representative experiment.

one and a first proof of the involvement of this class of proteins in the regulation of proviral transcription.

Of notice, and in contrast to our expectations, our proteomic screening did not detect several of the cellular proteins previously reported to associate with Tat and to mediate some of its functions. There are several possible explanations for this outcome. Our proteomic screening was conducted by immunoprecipitating a Flag epitope-tagged version of Tat (which was fully active transcriptionally) followed by RNase/DNase treatment, elution with a Flag peptide and resolution of Tat-associated proteins by gradient gel electrophoresis. In particular, we found that RNase treatment was essential to avoid the purification of a vast number of RNA-binding proteins unspecifically co-immunoprecipitating with Tat (data not shown). It might well be envisaged, however, that this clearing step might also affect the binding of Tat to some of its known partners, the interaction of which is strengthened by RNA bridging. In addition, RNA removal also frees the basic domain of Tat, thus rendering this region available for the interaction with hNAP-1. An additional explanation for the lack of other known Tat partners in our screening relates to the relative abundance of hNAP-1 in the cells, compared to other proteins such as p300 and P/CAF HATs, or Cyclin T1. Since our method relied on the identification of protein bands in silver-stained gels, a likely possibility is that we missed the detection of lower abundance proteins. Finally, it is worth however noting that other proteomic screenings aimed at the identification of cellular partners to other proteins also failed in identifying obvious candidates, while successfully discovering new factors essential for the function of the investigated proteins (see, among others, refs. [53,61]).

The basic region of Tat was found to bind two separable domains within hNAP-1, one internal from amino acids 162 to 290 and one C-terminal from residues 290 to 391. These domains correspond to a series of alternate α helix/β sheet regions known to be involved in the interaction with histones and other cellular proteins (see ref. [8,62] and citations therein). Of notice, the observation that Tat does not bind the highly acidic protein hSET/TAF I, another member of the NAP family with high structural and functional homology to hNAP-1 [57,58], argues in favor of a specific interaction between Tat and hNAP-1 which is not merely based on electrostatic interactions.

There is growing evidence that hNAP-1 plays important roles during transcriptional activation [21-24]. In particular, hNAP-1 and other histone chaperones both cooperate with ATP-dependent chromatin remodeling complexes [25,63] and participate in the formation of protein complexes also containing p300/CBP [25-28]. Taken together, these observations clearly suggest that hNAP-1 may serve as an interaction hub between transcriptional coactivators and chromatin. As far as p300/CBP is specifically concerned, p300 has been shown to directly bind the C-terminus of hNAP-1, namely the same region that is also involved in binding to Tat. Since the basic domain of Tat is also involved in binding to p300 [47], we cannot rule out the possibility that p300 might act as a scaffold for the simultaneous interaction with the two proteins. While further biochemical studies are clearly needed to ascertain this possibility, it is of interest to observe that the overexpression of all the three proteins together determined an increase in the levels of LTR transcription that is higher

than those obtained by overexpression of either p300 or hNAP-1 alone together with Tat. In addition, expression of p300 did not affect the levels hNAP-1 or Tat proteins, but markedly increased their binding in vivo. This observation is again in favor of the possibility that p300 might exert a stabilizing role on the Tat-hNAP-1 interaction. This possibility would be consistent with the proposed function for hNAP-1 in regulating transcription in all p300-dependent promoters [27,28].

What might be the actual mechanism by which hNAP-1 might facilitate Tat transactivation? First, overexpression of hNAP-1 significantly increases the overall levels of Tat inside the cells. This result is consistent with the possibility that the interaction with hNAP-1 might increase the stability of Tat. Second, and more relevant to a specific and direct role of hNAP-1 on the LTR promoter, previous results have indicated that the acetylation of histones by p300 helps transfer histones H2A and H2B from nucleosomes to hNAP-1 [26], and that, at least in vitro, the absence of these histones correlates with increased gene activity, probably by decreasing the level of chromatin folding [64,65]. On the basis of these observations, we can speculate that hNAP-1 and p300, brought to the LTR promoter through their interaction with Tat, might cooperate in the creation of an open-chromatin environment, favorable for gene expression. Of interest, a recent genome-wide analysis in fission yeast has revealed that chromatin remodeling factors and NAP-1 colocalize within promoter regions, where they disassemble nucleosomes near the transcriptional start site, an event that is linked to changes in the levels of histone acetylation [24].

Conclusion

In conclusion, this proteomic study reveals that the histone chaperone hNAP-1 is an important cellular factor specifically binding HIV-1 Tat. The interaction between the two proteins is involved in the regulation of Tat-mediated activation of viral gene expression, exerting a positive role on transcription. In particular, our findings indicate that HIV-1 Tat, hNAP-1 and p300 functionally cooperate to induce transcriptional activation of the HIV-1 LTR promoter.

Methods

Protein purification and identification

Twenty-four hours after transfection, ≈2 × 10^8 HEK 293T cells were washed once in phosphate-buffered saline (PBS) and lysed on ice in lysis buffer (150 mM NaCl/20 mM HEPES pH 7.9/0.5% NP-40/1 mM EDTA/1 mM DTT/protease inhibitor cocktail-Roche). The cell extract was sonicated once and then centrifuged for 15' at 14,000 rpm at 4°C. An aliquot of the cleared extract was kept as input, while the rest was incubated with 100 μl of packed and pre-equilibrated Flag M2 agarose beads overnight at 4°C. Beads were rinsed twice in lysis buffer, before treatment with DNAse I (Invitrogen, according to manufacturer's instructions) and RNAse A (150 mM NaCl/10 mM Tris HCl pH 7.5/5 mM EDTA/10 units RNAse A, for 30' at 37°C) and then washed in the same buffer three times. Immunocomplexes were eluted by adding 500 μg/ml Flag peptide (Sigma) in lysis buffer. The eluate was concentrated by standard trichloroacetic acid precipitation and resuspended in 1X sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) protein loading buffer. Proteins were then subjected to 6–15% gradient SDS-PAGE and then stained with silver stain. Stained proteins were excised and processed for in-gel trypsin digestion following standard protocols. The resulting peptides were extracted and purified on C18-Ziptips (Millipore) according to the manufacturer's protocol and resuspended in 10 μl of 30% methanol, 0.5% acetic acid. Protein identification was performed by the ICGEB Proteomics Facility by analyzing the purified peptides by MALDI-TOF mass spectrometry using an ABI 4800 TOF/TOF instrument (Applied Biosystems). The remaining sample was analyzed by LC-MS/MS using an LCQDeca mass spectrometer (Thermo-Finnigan).

Cell cultures, plasmids and siRNAs

HeLa and HEK 293T cells were cultured in Dulbecco's modified Eagle's medium with Glutamax (Life Technologies, Inc.) supplemented with 10% fetal bovine serum (Life Technologies, Inc.) and gentamicin (100 μg/ml) at 37°C in a humidified 95% air-5% CO_2 incubator.

All hNAP-1 encoding plasmids (wild type and mutants) were a kind gift by G. Steger [28]. All other plasmids used have already been described elsewhere [47,66-69].

RNA interference (RNAi) with hNAP-1 was performed against the target sequence 5' AAGGAACACGAUGAACC UAUU 3'. An siRNA targeted against the GFP RNA was used as a control (5' GGCTACGTCCAGGAGCGCACC 3'). Synthetic double-stranded RNA oligonucleotides were purchased by Dharmacon.

Co-immunoprecipitation

For co-immunoprecipitation analyses, HEK 293T cells where transfected with the indicated plasmids using the standard calcium phosphate coprecipitation method. Twenty-four hours after transfection cells were washed once in PBS and lysed on ice in 1 ml/dish lysis buffer (150 mM NaCl/20 mM HEPES pH 7.9/0.5% NP-40/1 mM EDTA/1 mM DTT/protease inhibitor cocktail-Roche). After sonication, cleared cell extracts were incubated with pre-equilibrated Flag M2 agarose beads on a rotating wheel for 4 hours at 4°C. Beads were washed twice with 1 ml of lysis buffer, treated with DNase I (Invitrogen, according to manufacturer's instructions) and RNAse A

(150 mM NaCl/10 mM Tris HCl pH 7.5/5 mM EDTA/10 units RNAse A, for 30' at 37°C) and then washed in the same buffer three times.

Antibodies

Anti-hNAP-1 mouse monoclonal antiserum was a kind gift from Y. Ishimi [70]. Mouse monoclonal anti-Flag M2 antibody, mouse monoclonal anti-tubulin, and mouse monoclonal anti-Flag M2 agarose-conjugated beads were purchased from Sigma. Rat monoclonal anti-HA high affinity (3F10) antibody was purchased from Roche diagnostics. Rabbit polyclonal anti-GFP antibody SC8334 was purchased from Santa Cruz Biotechnology.

Recombinant proteins

Glutathione S-transferase (GST), GST-Tat, GST-hNAP-1, GST-Tat mutants and GST-hNAP-1 mutants were prepared as already described [71]. Plasmids pcDNA3-Tat101 and pcDNA3-HA-NAP-1 were used as templates to produce the in vitro ^{35}S-labeled Tat and hNAP-1 proteins, respectively, by using the TNT Reticulocyte Lysate System (Promega) according to the manufacturer's protocol.

GST pull-down assay

GST and GST-Tat recombinant proteins immobilized on agarose beads were pre-treated with nucleases (see below). HEK293T cells were lysed in 150 mM NaCl/20 mM HEPES pH 7.9/0.5% NP-40/1 mM EDTA/1 mM DTT/ protease inhibitors (Roche). Recombinant proteins and cell extracts were incubated 1 hour and 30 minutes at 4°C, and washed four times in lysis buffer.

In vitro binding assay

To remove contaminant bacterial nucleic acids, recombinant proteins were pretreated with nucleases (0.25 U/μl DNase I and 0.2 μg/μl RNase) for 1 hour at 25°C in 50 mM Tris HCl, pH 8.0/5 mM $MgCl_2$/2.5 mM $CaCl_2$/100 mM NaCl/5% glycerol/1 mM DTT. Subsequently, GST fusion proteins immobilized on agarose beads were washed and resuspended in NETN buffer (20 mM Tris HCl, pH 7.5/100 mM NaCl/1 mM EDTA/0.5% NP-40/1 mM DTT/1 mM PMSF) supplemented with 0.2 mg/ml ethidium bromide to block the possible formation of non-specific interactions between residual DNA and proteins. ^{35}S-labeled hNAP-1 or Tat101 proteins (400 cpm) were added and incubated at 4°C on a rotating wheel. After 90 min, bound proteins were washed twice with 0.3 ml of NETN with ethidium bromide, three times with 0.3 ml of NETN without ethidium bromide and once with 0.3 ml of 10 mM Tris HCl pH 8.0/100 mM NaCl. Finally, bound proteins were separated by electrophoresis on a 12% SDS-polyacrylamide gel. Gels were stained and fixed for 1 hour with 10% acetic acid/40% methanol/0.1% Coomassie Brilliant blue G250, and destained with 10% acetic acid/40% methanol. Dried gels were quantitated by Instant Imager (Packard).

Luciferase assay

Reporter gene assays were performed using pLTR-luciferase plasmid as a reporter and pcDNA3-Tat101 as an effector in the presence or absence of plasmids pcDNA3-hNAP-1 and pCMV-p300. HeLa cells were transfected using Effectene Reagent (Quiagen, according to manufacturer's protocol), with 100 ng of pLTR-luciferase, 50 ng of pcDNA3-hNAP-1 and 5 or 25 ng of pcDNA3-Tat101. A Renilla luciferase expression plasmid, in which reporter gene expression was driven by the CMV promoter, was cotransfected to standardize each experiment for the efficiency of gene transfer. Cells were harvested 48 hours post transfection, and luciferase activity was measured with Luciferase assay kit (Promega). The measured activities were standardized by the activities of Renilla, and transactivation was expressed as fold activation compared with the basal activity of LTR-luciferase without effectors. All experiments were performed in duplicate and repeated at least three times.

For the transactivation experiments following RNAi, siRNAs were transfected using Oligofectamin Reagent (Invitrogen, according to manufacturer's protocol). After 36 hours from the beginning of siRNA treatment, cells were transfected with LTR-luciferase and CMV-Renilla plasmids and increasing amounts of pcDNA3-Tat101. Thirty-six hours later luciferase assays were performed on cell lysates.

In the case of infection with VSV-G-luciferase vectors, luciferase assays were performed 24 hours after the beginning of infection. For the gene-silencing experiments, cells were infected 48 hours after siRNA transfection. To normalize luciferase measures, protein concentrations in the lysates were determined with Bradford reagent (BioRad, according to manufacturer's protocol).

Virus production and infections

To produce VSV-G-luciferase vectors, HEK 293T cells were transfected with pNL4.3-luciferase plasmid [72,73] and VSV-G encoding plasmid at a ratio 3:1, according to a standard calcium phosphate coprecipitation method. Supernatants were collected 48 hours after the beginning of transfections, centrifuged and filtered with a 45 μm syringe.

Infections with viral supernatants was carried out for 6 hours in the presence of polybrene (Sigma) at a final concentration of 5 μg/ml.

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

LM, and AM carried out proteomic analysis; CV performed all other experiments. AM and ML participated in the experimental design and data analysis. MG contributed to the experimental design and coordination of the study, data analysis, as well as to writing the manuscript. All Authors have read and approved the final manuscript.

Acknowledgements

The authors wish to thank the ICGEB proteomic facility (Alessandro Vindigni, Federico Odreman) for MS analysis and, in particular, Mike Myers, for his outstanding technical and scientific advice. The authors are grateful to G. Steger for the kind gift of the GST-hNAP-1 deletion mutants and Y. Ishimi for the antibody against hNAP-1. This work was supported by a grant from the Italian National Research Programme on AIDS of the Istituto Superiore di Sanità, Italy to MG.

References

1. Polo SE, Almouzni G: **Chromatin assembly: a basic recipe with various flavours.** *Curr Opin Genet Dev* 2006, **16(2):**104-111.
2. Li B, Carey M, Workman JL: **The role of chromatin during transcription.** *Cell* 2007, **128(4):**707-719.
3. Kouzarides T: **Chromatin modifications and their function.** *Cell* 2007, **128(4):**693-705.
4. Akey CW, Luger K: **Histone chaperones and nucleosome assembly.** *Curr Opin Struct Biol* 2003, **13(1):**6-14.
5. Loyola A, Almouzni G: **Histone chaperones, a supporting role in the limelight.** *Biochim Biophys Acta* 2004, **1677(1-3):**3-11.
6. Mello JA, Almouzni G: **The ins and outs of nucleosome assembly.** *Curr Opin Genet Dev* 2001, **11(2):**136-141.
7. Tyler JK: **Chromatin assembly. Cooperation between histone chaperones and ATP-dependent nucleosome remodeling machines.** *Eur J Biochem* 2002, **269(9):**2268-2274.
8. Park YJ, Luger K: **The structure of nucleosome assembly protein 1.** *Proc Natl Acad Sci U S A* 2006, **103(5):**1248-1253.
9. Ishimi Y, Hirosumi J, Sato W, Sugasawa K, Yokota S, Hanaoka F, Yamada M: **Purification and initial characterization of a protein which facilitates assembly of nucleosome-like structure from mammalian cells.** *Eur J Biochem* 1984, **142(3):**431-439.
10. Ishimi Y, Kikuchi A: **Identification and molecular cloning of yeast homolog of nucleosome assembly protein I which facilitates nucleosome assembly in vitro.** *J Biol Chem* 1991, **266(11):**7025-7029.
11. Ito T, Bulger M, Kobayashi R, Kadonaga JT: **Drosophila NAP-1 is a core histone chaperone that functions in ATP-facilitated assembly of regularly spaced nucleosomal arrays.** *Mol Cell Biol* 1996, **16(6):**3112-3124.
12. Simon HU, Mills GB, Kozlowski M, Hogg D, Branch D, Ishimi Y, Siminovitch KA: **Molecular characterization of hNRP, a cDNA encoding a human nucleosome-assembly-protein-I-related gene product involved in the induction of cell proliferation.** *Biochem J* 1994, **297 (Pt 2):**389-397.
13. Mosammaparast N, Ewart CS, Pemberton LF: **A role for nucleosome assembly protein 1 in the nuclear transport of histones H2A and H2B.** *Embo J* 2002, **21(23):**6527-6538.
14. Altman R, Kellogg D: **Control of mitotic events by Nap1 and the Gin4 kinase.** *J Cell Biol* 1997, **138(1):**119-130.
15. Kellogg DR, Kikuchi A, Fujii-Nakata T, Turck CW, Murray AW: **Members of the NAP/SET family of proteins interact specifically with B-type cyclins.** *J Cell Biol* 1995, **130(3):**661-673.
16. Kellogg DR, Murray AW: **NAP1 acts with Clb1 to perform mitotic functions and to suppress polar bud growth in budding yeast.** *J Cell Biol* 1995, **130(3):**675-685.
17. Levchenko V, Jackson V: **Histone release during transcription: NAP1 forms a complex with H2A and H2B and facilitates a topologically dependent release of H3 and H4 from the nucleosome.** *Biochemistry* 2004, **43(9):**2359-2372.
18. Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C: **ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex.** *Science* 2004, **303(5656):**343-348.
19. Okuwaki M, Kato K, Shimahara H, Tate S, Nagata K: **Assembly and disassembly of nucleosome core particles containing histone variants by human nucleosome assembly protein I.** *Mol Cell Biol* 2005, **25(23):**10639-10651.
20. Park YJ, Chodaparambil JV, Bao Y, McBryant SJ, Luger K: **Nucleosome assembly protein 1 exchanges histone H2A-H2B dimers and assists nucleosome sliding.** *J Biol Chem* 2005, **280(3):**1817-1825.
21. Ohkuni K, Shirahige K, Kikuchi A: **Genome-wide expression analysis of NAP1 in Saccharomyces cerevisiae.** *Biochem Biophys Res Commun* 2003, **306(1):**5-9.
22. Rogner UC, Spyropoulos DD, Le Novere N, Changeux JP, Avner P: **Control of neurulation by the nucleosome assembly protein-1-like 2.** *Nat Genet* 2000, **25(4):**431-435.
23. Steer WM, Abu-Daya A, Brickwood SJ, Mumford KL, Jordanaires N, Mitchell J, Robinson C, Thorne AW, Guille MJ: **Xenopus nucleosome assembly protein becomes tissue-restricted during development and can alter the expression of specific genes.** *Mech Dev* 2003, **120(9):**1045-1057.
24. Walfridsson J, Khorosjutina O, Matikainen P, Gustafsson CM, Ekwall K: **A genome-wide role for CHD remodelling factors and Nap1 in nucleosome disassembly.** *Embo J* 2007, **26(12):**2868-2879.
25. Asahara H, Tartare-Deckert S, Nakagawa T, Ikehara T, Hirose F, Hunter T, Ito T, Montminy M: **Dual roles of p300 in chromatin assembly and transcriptional activation in cooperation with nucleosome assembly protein 1 in vitro.** *Mol Cell Biol* 2002, **22(9):**2974-2983.
26. Ito T, Ikehara T, Nakagawa T, Kraus WL, Muramatsu M: **p300-mediated acetylation facilitates the transfer of histone H2A-H2B dimers from nucleosomes to a histone chaperone.** *Genes Dev* 2000, **14(15):**1899-1907.
27. Shikama N, Chan HM, Krstic-Demonacos M, Smith L, Lee CW, Cairns W, La Thangue NB: **Functional interaction between nucleosome assembly proteins and p300/CREB-binding protein family coactivators.** *Mol Cell Biol* 2000, **20(23):**8933-8943.
28. Rehtanz M, Schmidt HM, Warthorst U, Steger G: **Direct interaction between nucleosome assembly protein 1 and the papillomavirus E2 proteins involved in activation of transcription.** *Mol Cell Biol* 2004, **24(5):**2153-2168.
29. Marcello A, Lusic M, Pegoraro G, Pellegrini V, Beltram F, Giacca M: **Nuclear organization and the control of HIV-1 transcription.** *Gene* 2004, **326(Feb 4):**1-11.
30. Greene WC, Peterlin BM: **Charting HIV's remarkable voyage through the cell: Basic science as a passport to future therapy.** *Nat Med* 2002, **8(7):**673-680.
31. El Kharroubi A, Piras G, Zensen R, Martin MA: **Transcriptional activation of the integrated chromatin-associated human immunodeficiency virus type 1 promoter.** *Mol Cell Biol* 1998, **18:**2535-2544.
32. Sheridan PL, Mayall TP, Verdin E, Jones KA: **Histone acetyltransferases regulate HIV-1 enhancer activity in vitro.** *Genes Dev* 1997, **11(24):**3327-3340.
33. Van Lint C, Emiliani S, Ott M, Verdin E: **Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation.** *EMBO J* 1996, **15(5):**1112-1120.
34. Verdin E: **DNase I-hypersensitive sites are associated with both long terminal repeats and with the intragenic enhancer of integrated human immunodeficiency virus type 1.** *J Virol* 1991, **65(12):**6790-6799.
35. Demarchi F, D'Agaro P, Falaschi A, Giacca M: **Probing protein-DNA interactions at the long terminal repeat of human immunodeficiency virus type 1 by in vivo footprinting.** *J Virol* 1992, **66:**2514-2518.
36. Demarchi F, D'Agaro P, Falaschi A, Giacca M: **In vivo footprinting analysis of constitutive and inducible protein-DNA interactions at the long terminal repeat of human immunodeficiency virus type 1.** *J Virol* 1993, **67:**7450-7460.
37. Jackson PK, Eldridge AG, Freed E, Furstenthal L, Hsu JY, Kaiser BK, Reimann JD: **The lore of the RINGs: substrate recognition and**

catalysis by ubiquitin ligases. *Trends Cell Biol* 2000, **10(10):**429-439.
38. Jordan A, Bisgrove D, Verdin E: **HIV reproducibly establishes a latent infection after acute infection of T cells in vitro.** *EMBO J* 2003, **22(8):**1-10.
39. Lusic M, Marcello A, Cereseto A, Giacca M: **Regulation of HIV-1 gene expression by histone acetylation and factor recruitment at the LTR promoter.** *EMBO J* 2003, **22(24):**6550-6561.
40. Berkhout B, Silverman RH, Jeang KT: **Tat trans-activates the human immunodeficiency virus through a nascent RNA target.** *Cell* 1989, **59(2):**273-282.
41. Mancebo HS, Lee G, Flygare J, Tomassini J, Luu P, Zhu Y, Peng J, Blau C, Hazuda D, Price D, Flores O: **P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro.** *Genes Dev* 1997, **11(20):**2633-2644.
42. Wei P, Garber ME, Fang SM, Fisher WH, Jones KA: **A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA.** *Cell* 1998, **92:**451-462.
43. Zhu Y, Pe'ery T, Peng J, Ramanathan Y, Marshall N, Marshall T, Amendt B, Mathews MB, Price DH: **Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro.** *Genes Dev* 1997, **11(20):**2622-2632.
44. Benkirane M, Chun RF, Xiao H, Ogryzko VV, Howard BH, Nakatani Y, Jeang KT: **Activation of integrated provirus requires histone acetyltransferase. p300 and P/CAF are coactivators for HIV-1 Tat.** *J Biol Chem* 1998, **273(38):**24898-24905.
45. Col E, Caron C, Seigneurin-Berny D, Gracia J, Favier A, Khochbin S: **The histone acetyltransferase, hGCN5, interacts with and acetylates the HIV transactivator, Tat.** *J Biol Chem* 2001, **276(30):**28179-28184.
46. Hottiger MO, Nabel GJ: **Interaction of human immunodeficiency virus type 1 Tat with the transcriptional coactivators p300 and CREB binding protein.** *J Virol* 1998, **72:**8252-8256.
47. Marzio G, Tyagi M, Gutierrez MI, Giacca M: **HIV-1 Tat transactivator recruits p300 and CREB-binding protein histone acetyltransferases to the viral promoter.** *Proc Natl Acad Sci USA* 1998, **95(23):**13519-13524.
48. Fujinaga K, Cujec TP, Peng J, Garriga J, Price DH, Grana X, Peterlin BM: **The ability of positive transcription elongation factor B to transactivate human immunodeficiency virus transcription depends on a functional kinase domain, cyclin T1, and Tat.** *J Virol* 1998, **72(9):**7154-7159.
49. Zhou Q, Chen D, Pierstorff E, Luo K: **Transcription elongation factor P-TEFb mediates Tat activation of HIV-1 transcription at multiple stages.** *EMBO J* 1998, **17(13):**3681-3691.
50. Treand C, du Chene I, Bres V, Kiernan R, Benarous R, Benkirane M, Emiliani S: **Requirement for SWI/SNF chromatin-remodeling complex in Tat-mediated activation of the HIV-1 promoter.** *Embo J* 2006, **25(8):**1690-1699.
51. Chan PK, Chan FY, Morris SW, Xie Z: **Isolation and characterization of the human nucleophosmin/B23 (NPM) gene: identification of the YY1 binding site at the 5' enhancer region.** *Nucleic Acids Res* 1997, **25(6):**1225-1232.
52. Petersen-Mahrt SK, Estmer C, Ohrmalm C, Matthews DA, Russell WC, Akusjarvi G: **The splicing factor-associated protein, p32, regulates RNA splicing by inhibiting ASF/SF2 RNA binding and phosphorylation.** *Embo J* 1999, **18(4):**1014-1024.
53. Berro R, Kehn K, de la Fuente C, Pumfery A, Adair R, Wade J, Colberg-Poley AM, Hiscott J, Kashanchi F: **Acetylated Tat regulates human immunodeficiency virus type 1 splicing through its interaction with the splicing regulator p32.** *J Virol* 2006, **80(7):**3189-3204.
54. Yu L, Loewenstein PM, Zhang Z, Green M: **In vitro interaction of the human immunodeficiency virus type 1 Tat transactivator and the general transcription factor TFIIB with the cellular protein TAP.** *J Virol* 1995, **69(5):**3017-3023.
55. Li YP: **Protein B23 is an important human factor for the nucleolar localization of the human immunodeficiency virus protein Tat.** *J Virol* 1997, **71(5):**4098-4102.
56. Stauber RH, Pavlakis GN: **Intracellular trafficking and interactions of the HIV-1 Tat protein.** *Virology* 1998, **252(1):**126-136.
57. Kawase H, Okuwaki M, Miyaji M, Ohba R, Handa H, Ishimi Y, Fujii-Nakata T, Kikuchi A, Nagata K: **NAP-I is a functional homologue of TAF-I that is required for replication and transcription of the adenovirus genome in a chromatin-like structure.** *Genes Cells* 1996, **1(12):**1045-1056.
58. von Lindern M, van Baal S, Wiegant J, Raap A, Hagemeijer A, Grosveld G: **Can, a putative oncogene associated with myeloid leukemogenesis, may be activated by fusion of its 3' half to different genes: characterization of the set gene.** *Mol Cell Biol* 1992, **12(8):**3346-3355.
59. Brady J, Kashanchi F: **Tat gets the "green" light on transcription initiation.** *Retrovirology* 2005, **2:**69.
60. Giacca M: **The HIV-1 Tat protein: a multifaceted target for novel therapeutic opportunities.** *Curr Drug Targets Immune Endocr Metabol Disord* 2004, **4(4):**277-285.
61. Janardhan A, Swigut T, Hill B, Myers MP, Skowronski J: **HIV-1 Nef binds the DOCK2-ELMO1 complex to activate rac and inhibit lymphocyte chemotaxis.** *PLoS Biol* 2004, **2(1):**E6.
62. Park YJ, Luger K: **Structure and function of nucleosome assembly proteins.** *Biochem Cell Biol* 2006, **84(4):**549-558.
63. Nakagawa T, Bulger M, Muramatsu M, Ito T: **Multistep chromatin assembly on supercoiled plasmid DNA by nucleosome assembly protein-1 and ATP-utilizing chromatin assembly and remodeling factor.** *J Biol Chem* 2001, **276(29):**27384-27391.
64. Baer BW, Rhodes D: **Eukaryotic RNA polymerase II binds to nucleosome cores from transcribed genes.** *Nature* 1983, **301(5900):**482-488.
65. Chan HM, La Thangue NB: **p300/CBP proteins: HATs for transcriptional bridges and scaffolds.** *J Cell Sci* 2001, **114(Pt 13):**2363-2373.
66. Dorr A, Kiermer V, Pedal A, Rackwitz HR, Henklein P, Schubert U, Zhou MM, Verdin E, Ott M: **Transcriptional synergy between Tat and PCAF is dependent on the binding of acetylated Tat to the PCAF bromodomain.** *EMBO J* 2002, **21(11):**2715-2723.
67. Emiliani S, Van Lint C, Fischle W, Paras P Jr., Ott M, Brady J, Verdin E: **A point mutation in the HIV-1 Tat responsive element is associated with postintegration latency.** *Proc Natl Acad Sci USA* 1996, **93(13):**6377-6381.
68. Pegoraro G, Marcello A, Myers MP, Giacca M: **Regulation of adeno-associated virus DNA replication by the cellular TAF-I/set complex.** *J Virol* 2006, **80(14):**6855-6864.
69. Marcello A, Cinelli RA, Ferrari A, Signorelli A, Tyagi M, Pellegrini V, Beltram F, Giacca M: **Visualization of in vivo direct interaction between HIV-1 Tat and human cyclin T1 in specific subcellular compartments by fluorescence resonance energy transfer.** *J Biol Chem* 2001, **276:**39220-39225.
70. Ishimi Y, Sato W, Kojima M, Sugasawa K, Hanaoka F, Yamada M: **Rapid purification of nucleosome assembly protein (AP-I) and production of monoclonal antibodies against it.** *Cell Struct Funct* 1985, **10(4):**373-382.
71. Demarchi F, d'Adda di Fagagna F, Falaschi A, Giacca M: **Activation of transcription factor NF-kB by the Tat protein of human immunodeficiency virus-1.** *J Virol* 1996, **70:**4427-4437.
72. Connor RI, Chen BK, Choe S, Landau NR: **Vpr is required for efficient replication of human immunodeficiency virus type-1 in mononuclear phagocytes.** *Virology* 1995, **206(2):**935-944.
73. He J, Choe S, Walker R, Di Marzio P, Morgan DO, Landau NR: **Human immunodeficiency virus type 1 viral protein R (Vpr) arrests cells in the G2 phase of the cell cycle by inhibiting p34cdc2 activity.** *J Virol* 1995, **69(11):**6705-6711.
74. Marcello A, Ferrari A, Pellegrini V, Pegoraro G, Lusic M, Beltram F, Giacca M: **Recruitment of human cyclin T1 to nuclear bodies through direct interaction with the PML protein.** *EMBO J* 2003, **22(9):**2156-2166.
75. Fittipaldi A, Ferrari A, Zoppe M, Arcangeli C, Pellegrini V, Beltram F, Giacca M: **Cell membrane lipid rafts mediate caveolar endocytosis of HIV-1 Tat fusion proteins.** *J Biol Chem* 2003, **278(36):**34141-34149.

Retrovirology

BioMed Central

Research

Open Access

Intragenic transcriptional *cis*-activation of the human immunodeficiency virus 1 does not result in allele-specific inhibition of the endogenous gene

Alex De Marco[3,4], Chiara Biancotto[2,4], Anna Knezevich[4], Paolo Maiuri[4], Chiara Vardabasso[1,4] and Alessandro Marcello*[4]

Address: [1]Laboratory of Molecular Medicine, International Centre for Genetic Engineering and Biotechnology (ICGEB), Padriciano, 99, 34012 Trieste, Italy, [2]IFOM-IEO, Milan, Italy, [3]EMBL, Heidelberg, Deutschland, Germany and [4]Laboratory of Molecular Virology, International Centre for Genetic Engineering and Biotechnology (ICGEB), Padriciano, 99, 34012 Trieste, Italy

Email: Alex De Marco - alex.demarco@embl.de; Chiara Biancotto - chiara.biancotto@ifom-ieo-campus.it; Anna Knezevich - knezevich@icgeb.org; Paolo Maiuri - maiuri@icgeb.org; Chiara Vardabasso - vardabasso@icgeb.org; Alessandro Marcello* - marcello@icgeb.org

* Corresponding author

Published: 4 November 2008

Retrovirology 2008, **5**:98 doi:10.1186/1742-4690-5-98

Received: 6 October 2008
Accepted: 4 November 2008

This article is available from: http://www.retrovirology.com/content/5/1/98

Abstract

Background: The human immunodeficiency virus type 1 (HIV-1) favors integration in active genes of host chromatin. It is believed that transcriptional interference of the viral promoter over the endogenous gene or vice versa might occur with implications in HIV-1 post-integrative transcriptional latency.

Results: In this work a cell line has been transduced with a HIV-based vector and selected for Tat-inducible expression. These cells were found to carry a single silent integration in sense orientation within the second intron of the *HMBOX1* gene. The HIV-1 Tat transactivator induced the viral LTR and repressed *HMBOX1* expression independently of vector integration. Instead, single-cell quantitative *in situ* hybridization revealed that allele-specific transcription of *HMBOX1* carrying the integrated provirus was not affected by the transactivation of the viral LTR in *cis*.

Conclusion: A major observation of the work is that the HIV-1 genome has inserted in genes that are also repressed by Tat and this could be an advantage for the virus during transcriptional reactivation. In addition, it has also been observed that transcription of the provirus and of the endogenous gene in which it is integrated may coexist at the same time in the same genomic location.

Background

Retroviruses, such as human immunodeficiency virus type 1 (HIV-1) require reverse transcription and integration into host chromatin to establish a provirus as an obligatory replication step. The choice of the integration site is a crucial intermediate of the virus life cycle. The chromatin context determines the efficiency of viral transcription and is involved in the establishment of post-integrative latency that is the major obstacle to HIV-1 eradication with current antiviral therapies [1-3]. In addition, insertion of a provirus in the human genome can cause several adverse effects [4]. For example, insertion of the retrovirus

close to a proto-oncogene may induce transformation of the cell. Gene therapy approaches suffer most from these effects and recently it has been demonstrated that the activation of an oncogene caused transformation in several children treated with a therapeutic retroviral vector [5]. In principle, insertion of an ectopic transcription unit within a gene may also result either in disruption of exonic sequences, introduction of alternative splicing or transcriptional interference. Clearly, these negative effects would increase in importance relative to the increasing unbalance of the endogenous gene expression between alleles.

Integration site selection by retroviruses is not sequence-specific but also not random. HIV-1 favors integration within active transcription units [6-8]. Additional features are the requirement of host factors such as the lens epithelium-derived growth factor LEDGF/p75 for efficient targeting of active transcription units [9] and a DNA substrate wrapped around nucleosomes. Indeed, integration of HIV-1 is linked to nucleosomal markers of active transcription (H3/H4 acetylation, H3K4 methylation) and negatively correlated with inhibitory modifications (H3K27 trimethylation and DNA CpG methylation) [10]. Subtle differences in the integration site choice exist among retroviruses. Murine leukemia virus (MLV) integrates within highly active promoters at ± 5 kb from the transcription start sites [7,11]. HIV-1 instead, although also favoring active genes, does not show a preference for promoter-proximal integration. Rather, the virus inserts throughout the transcriptional unit with a bias towards intronic sequences: this is the likely result of the greater size of introns compared to exons within a gene [6].

A crucial aspect of HIV-1 pathogenesis is the control of provirus transcription. In particular the ability of the virus to maintain a reservoir of transcriptionally silent proviruses in resting memory T cells for long periods of time. Multiple mechanisms have been postulated to concur in these processes. Host factors, for example, may be limiting the activity of the Tat transactivator. Tat interacts with a *cis*-acting RNA element (trans-activation-responsive region; TAR) present at the 5' end of each viral transcript [12]. Through this interaction, the protein activates HIV-1 transcription by promoting the assembly of transcriptionally active complexes at the LTR through multiple protein-RNA and protein-protein interactions [13]. Tat interacts with the core RNA polymerase II [14,15], the TATA-binding protein associated factor (TAFII) [16], TFIIH [17], cyclin-dependent protein kinase 7 [18], SP1 [19], nuclear factor of activated T cells (NFAT) [20], several histone acetyltransferases [21-23] and cyclin T1 [24]. On the other hand, the chromatin context at the site of integration should determine whether the provirus is transcriptionally active, poised for activation or inactive [25]. Early studies showed that latency involved integration into regions of heterochromatin [26,27]. More recent systematic genome-wide analysis of the chromosomal features negatively associated to HIV-1 transcription revealed that low levels of LTR-driven expression correlated with integration in gene deserts and in centromeric heterochromatin, but also in highly expressed cellular genes [28]. Furthermore, HIV-1 has been found in intronic regions of actively transcribed genes in resting memory CD4+ cells derived from patient on highly active antiretroviral treatment [29]. The paradox of HIV-1 integration in active genes while being transcriptionally silent requires molecular investigation of the phenomenon. Unfortunately most cellular models of HIV-1 post-integrative latency harbor the provirus outside of transcribed genes [3]. In this work a cell-line that carries a single repressed provirus integrated within the active transcription unit of the *HMBOX1* gene has been generated. Tat-mediated induction of provirus transcription resulted in the inhibition of *HMBOX1* expression. However, this effect could be ascribed to Tat expression and not to activation of the viral LTR. Indeed, a subset of activated cells showed bi-allelic expression of *HMBOX1* together with expression of the provirus within one of the alleles. These results are discussed in light both of HIV-1 pathogenesis and of the potential use of lentiviral vectors for gene therapy applications.

Results

Generation and characterization of a cell line carrying a stably integrated lentiviral vector

The HIV-Intro-MS2 × 24-ECFPskl-IRES-TK lentiviral vector (for simplicity: HIV-Intro) has been engineered to contain the elements required for RNA production: the 5' LTR, the major splice donor (SD1), the packaging signal Ψ, the RRE, the splice acceptor A7 and the 3' LTR that drives 3'-end formation (Figure 1). The construct carries also an array of 24 repeats of the MS2 phage coat protein within the intron, to increase specific detection of nascent mRNA, a reporter of gene expression (ECFP) fused to the peroxisome localization signal Ser-Lys-Leu (skl) and the selectable marker thymidine kinase (TK) of herpes simplex type 1.

In order to characterize this construct extensively before transduction, HeLa cells were transfected with plasmid HIV-Intro together with a plasmid expressing a monomeric DsRed-tagged Tat Figure 2A, top panels). As expected from previous studies showing transcribed nascent RNA by MS2-tagging [30,31], bright yellow spots appeared within the nucleus. Each spot corresponds to several plasmids clustered together that express viral RNA [32]. As expected, Tat was found at transcription sites because it binds the 5'-end of each transcript. The reporter of gene expression ECFPskl was found in the cytoplasm.

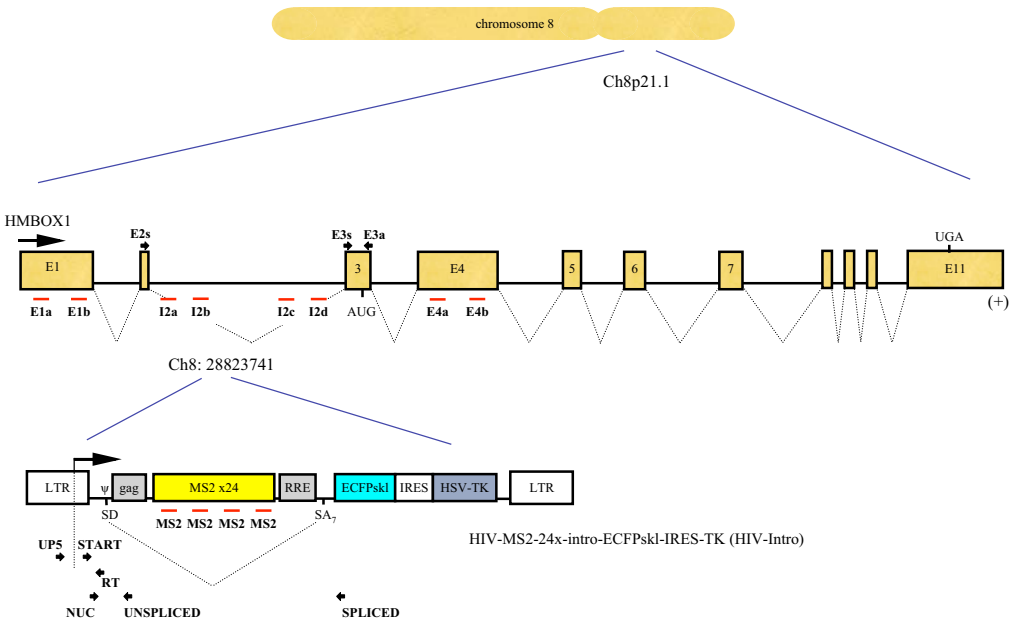


Figure 1
Genomic organization of the HMBOX1 gene and of the HIV-intro construct in HOS_A4 cells. Position of the RT-PCR primers are indicated by black arrows. Positions of the FISH probes are indicated by red bars.

When a plasmid expressing a DsRed-tagged Rev was co-transfected together with Tat (without tag), the unspliced RNA was found in the cytoplasm, consistent with its Rev-mediated export (Figure 2A, bottom panels). These results are mirrored by the behavior in RT-PCR using a set of primers that distinguish pre-mRNA from spliced RNA. As shown in figure 2B, basal transcription is up-regulated by Tat with a higher proportion of spliced over unspliced RNA. Co-transfection of a plasmid encoding pEYFP-MS2nls does not affect the splicing reaction, ruling out perturbation of the system by such a strong RNA binding protein. Expression of Rev instead increased the proportion of unspliced RNA, consistent with its role in RRE-containing RNA stabilization and export.

A key question that arose while doing these experiments was the real nature of these yellow spots in the nucleus (Figure 2A, top panels). To confirm that these where sites of HIV-Intro transcription we incubated the cells with inhibitors such as Actinomicin D, α-Amanitin or Flavopiridol. As shown in Figure 2C, a rapid decrease of the number of transcription spots was observed with all three inhibitors. Hence, RNA-dependent accumulation of RNA at these sites was dependent on RNAPII activity.

Next a strategy was designed to express the HIV-Intro construct from a single chromatinized location in a Tat-inducible way. Osteosarcoma HOS_143b cells, that are negative for thymidine kinase activity (TK-), were transduced with the HIV-Intro vector pseudotyped with the VSV-G envelope. To select for clones that carry an inducible integrated provirus, cells that constitutively expressed high levels of HSV-TK were selected against by treatment with 50 μg/ml ganciclovir. Surviving cells, that were either non-transduced, or transduced but with a low level of TK expression, were treated with GST-Tat and briefly selected for inducible HSV-TK expression in hypoxanthine, aminopterin and thymidine (HAT) medium. Clonal populations were obtained by limiting dilutions and colonies were visually scored for low basal level of ECFP expression in the cytoplasm and to be highly inducible by GST-Tat by fluorescence microscopy. The HOS_A4 cell clone showed a robust and homogenous induction of ECFPskl in the cytoplasm upon treatment with GST-Tat (Figure 3A). These

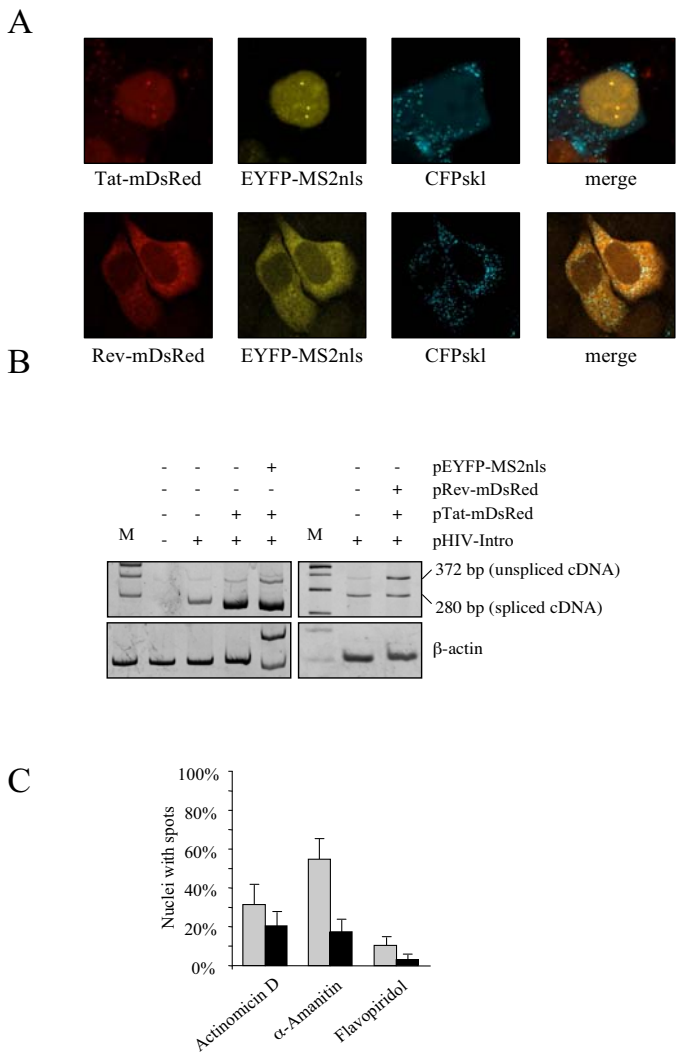


Figure 2

A) HeLa cells were cotransfected with pHIV-Intro, pEYFP-MS2nls and either mDsRed tagged Tat (top) or Tat and mDsRed tagged Rev (bottom). Yellow spots in the nucleus correspond to nascent RNA from transfected plasmids. Cyan spots in the cytoplasm correspond to ECFPskl localized to peroxisomes. **B)** RT-PCR on HeLa cells transfected as indicated. Three primers were used, their position is shown in Figure 1. Resulting bands correspond to the unspliced and spliced HIV-Intro RNAs. Bottom panels: β-actin loading control (M = molecular weight marker). **C)** Effect of RNAPII inhibitors on HIV-Intro transcription in transfected HeLa cells transfected as indicated in Figure 2A, top panels. Nuclei showing transcription spots were scored 1 hour (gray bars) an 6 hours (black bars) after treatment with Actinomicin D (10 μg/ml), α-Amanitin (10 μg/ml) or Flavopiridol (500 μM).

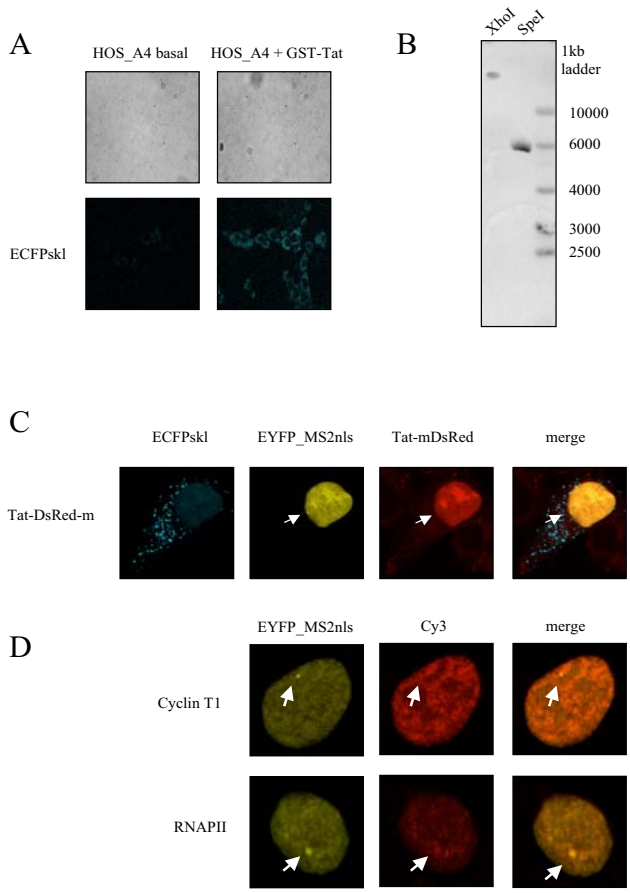


Figure 3

A) Generation of HOS_A4 cells by transduction with HIV-Intro and selection as described in the text. Tat induction induced the expression of ECFPskl in the cytoplasm. Top panels: phase contrast. Bottom panels: ECFP channel. **B)** Southern blot analysis of HOS_A4 cells shows the presence of a single integration event. Genomic DNA was digested with XhoI or SpeI and hybridized with a probe encompassing ECFP. **C)** Effect of Tat-mDsRed on HOS_A4 cells. Co-localization of Tat and HIV-Intro RNA is shown on the single transcription spot present in HOS_A4 cells. Correct gene expression is demonstrated by the ECPFskl signal in the cytoplasm. **D)** Co-localization of RNAPII and Cyclin T1 on HOS_A4 transcription spots. Cells were transfected with pEYFP-MS2nls and Tat, fixed and Cyclin T1 (top panels) or RNAPII (bottom panels) detected by immunofluorescence as described in [30].

cells were transfected with plasmids encoding EYFP-MS2nls and Tat-mDsRed. As shown in figure 3C, HOS_A4 showed ECFPskl in the cytoplasm and presented one single bright yellow spot in the nucleus compatible with a single site of HIV-Intro transcription that co-localized with Tat-mDsRed. Immuonfluorescence with antisera against Cyclin T1, the P-TEFb component recruited directly by Tat on the viral RNA, or RNAPII demonstrated enrichment of such factors at these sites (Figure 3D).

These results are compatible with one integration event of the HIV vector in HOS_A4 cells. Indeed, analysis by southern blotting (Figure 3B) and cloning of the integration sites by inverse PCR revealed that the provirus lay within the *HMBOX1* (homeobox containing 1) cellular gene.

Allele-specific expression of HMBOX1 *following HIV-Intro-MS2 × 24-ECFPskl-IRES-TK transactivation*

Human *HMBOX1* is composed of 11 exons, spanning about 160 kb within chromosome 8 p21.1 (Figure 1). *HMBOX1* is believed to encode for a transcription factor involved in the transcriptional regulation of key eukaryotic developmental processes. *HMBOX1* is widely expressed in pancreas and the expression of this gene can also be detected in pallium, hippocampus and hypothalamus [33]. In HOS_A4 cells the HIV-Intro lentiviral vector integrated within the second giant intron of *HMBOX1* (Figure 1). In order to assess expression of *HMBOX1* in HOS cells RT-PCR was performed with primers specific for Exon3 or encompassing Exon2/Exon3 junction (Figure 4A). Both parental HOS 143b and HOS_A4 expressed *HMBOX1* at similar levels. However, this assay was neither quantitative nor specific for the *HMBOX1* allele carrying the integrated vector. A similar approach was also employed for the HIV-Intro transcript. As shown in Figure 4B, a basal level of HIV-Intro expression was detected in HOS_A4 that could be up regulated by Tat transfection.

In order to detect allele-specific transcription in the *HMBOX1* locus carrying the integrated provirus a quantitative RT-PCR was developed according to the protocol of Han and collaborators [29]. As shown in the diagram of Figure 1, RT-PCR primers were designed to detect also *HMBOX1* transcripts containing HIV sequences upstream of the viral transcription start site. RNA from HOS_A4 cells was reverse-transcribed and the resulting cDNA was amplified with two primers that share the HIV_RT primer. HIV_UP5 amplifies only HIV-Intro sequences produced as a result of transcription of *HMBOX1* reading through the HIV-Intro genome that is inserted into the gene. Because the forward primer is located upstream of the transcription start site and the reverse primer is located downstream of the LTR, only RNA species initiating upstream of the HIV-1 transcription start site could be amplified. HIV_START instead is able to amplify any HIV-Intro transcript that has initiated at the viral start site. To prevent amplification from HIV-1 DNA, isolated RNA was treated with DNase before RT-PCR. In addition, control reactions from which RT was omitted were included in each experiment and were invariably negative. A positive control of HIV transactivation involved a set of primers for the HIV-1 spliced RNA product (primers HIV_nuc and HIV_spliced). PCR amplification was conducted in the presence of the dye CyberGreen for relative quantification of PCR products. Transfection of Tat induced HIV-1 transcripts several folds with both primer sets detecting HIV-1 transcripts (Figure 4C and Figure 4D). This result is perfectly in line with the well-known Tat-transactivation of the viral LTR and with RT PCR data shown in Figure 4B[34]. Allele-specific detection of *HMBOX1* RNA instead showed a marked decrease in response to Tat (Figure 4E). This result would be explained by negative interference with the expression of *HMBOX1* due to activation of a strong promoter embedded within the gene. However, when the analysis was conducted on two primer sets specific for the *HMBOX1* gene both in HOS_A4 and the parental HOS_143b we realized that expression of *HMBOX1* was affected by the presence of Tat *per se* and not by the activation of the viral LTR (Figure 4F and Figure 4G). This effect was not a general effect on transcription since the *GAPDH* gene was not affected (Figure 4H).

Single-cell analysis of HMBOX1 *and HIV-1 expression*

Ensemble-averaged analysis such as RT-PCR that relies on the evaluation of a number of cells does not allow distinction between expression of each *HMBOX1* allele. In fact, although we analyzed allele-specific expression of the allele carrying the provirus, still we don't know whether *HMBOX1* expression was balanced between alleles or not.

In order to evaluate the simultaneous expression of both *HMBOX1* alleles and of the integrated provirus, transcripts were detected by quantitative fluorescent *in situ* RNA hybridization (FISH). The amount of RNA on the transcription spot is determined by the rate of transcription and the rate of RNA processing. At steady state it could be derived from the intensity of the fluorescence signal compared with the intensity of the signal from a known reference as described previously [30]. For this purpose for each probe and each acquisition we prepared a calibration curve spotting different amounts of probe on a coverslip in a constant volume. The probes were acquired and deconvoluted using the same conditions used for the samples (see Methods). The z-projection sum of all planes was averaged and this value represent the signal emitted by each amount of probe. Therefore, the number of probes for each voxel (a volume pixel in a three-dimensional image) could be calculated for each point of the calibration curve. In the case of HIV-Intro

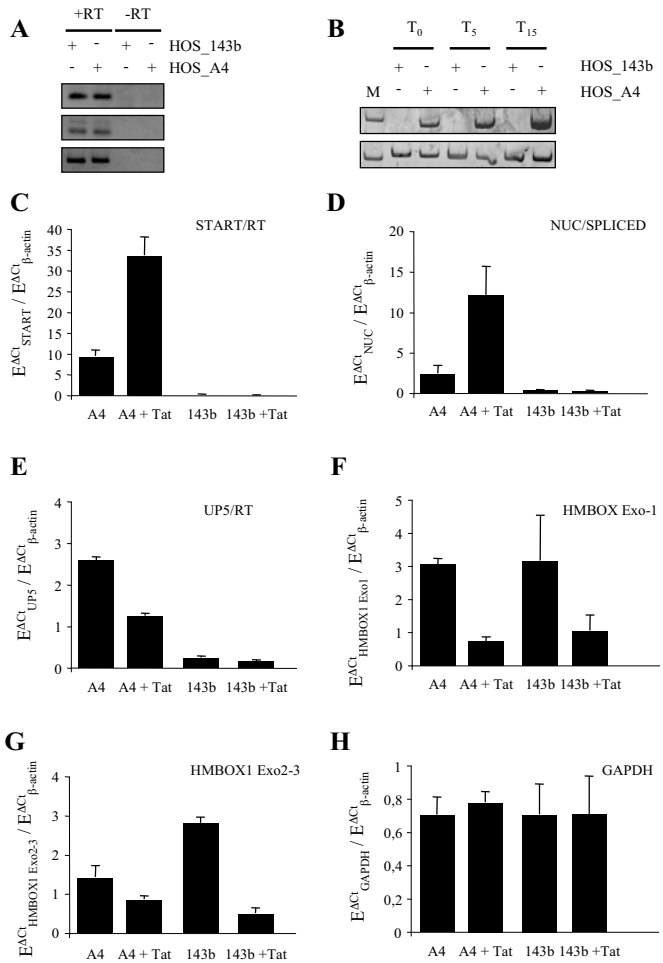


Figure 4

A) RT-PCR analysis of *HMBOX1* expression in parental HOS_143b and clone HOS_A4 using primers for *HMBOX1* Exon 3 (79 bp, top panel) and Exon 2/Exon 3 splicing (86 bp, middle panel). Bottom panel: β-actin control (230 bp). **B)** RT-PCR analysis of HIV-Intro expression (280 bp top panel) in parental HOS_143b and clone HOS_A4 using primers HIV_SPLICED and HIV_NUC (Table 1). Bottom panel: β-actin control (M = molecular weight marker; T = hours after Tat induction). C-H) Quantitative RT PCR for HIV-Intro, *HMBOX1* and *GAPDH* expression using the indicated primers shown in Figure 1 and Table 1. Each histogram is the mean of three experiments normalized for β-actin expression and corrected for primer efficiency ($E = 10^{-1/slope}$).

transcripts, the number of RNAs on the transcription spot in the presence of Tat was calculated to be 17 ± 4.

HMBOX1 expression was low and could not be detected by a single probe. Therefore a mixture of eight oligonucleotides, distributed in the first and fourth exon and in the second intron before and after the integration (Figure 1), were designed to detect the nascent unprocessed RNA transcripts of *HMBOX1* (Table 1). As shown in figure 5A, in parental HOS-143b two spots of equal intensity were clearly visible in 37.2% of nuclei indicating that the *HMBOX1* gene is expressed from both alleles, but only in a fraction of the asynchronous population of cells. This is not surprising since there is ample variation of the number of alleles/nucleus detected by this method, depending on how robust is gene expression and how efficient is the processing of the RNA; both contribute to the level of RNA at the site of transcription at steady state [35]. It is however unlikely that detection was lowered by scarce accessibility of the probe to the RNA since positive cells showed invariably two alleles of equal intensity, where in the case of technical problems there would be a higher proportion of single-allele expressing cells. The same was observed in derivative HOS_A4 cells where 36.8% of nuclei showed biallelic expression of the *HMBOX1* gene (Figure 5C). The number of nascent RNAs present on the transcription site at steady-state was calculated to be significantly similar in both cell lines: 4.13 ± 1.02 for HOS_143b and 4.11 ± 0.91 for HOS_A4 (p value = 0.18) (Figure 6B). Most importantly, the ratio between the intensity of the signal of the two loci was invariably close to 1 in both cell lines demonstrating bi-allelic expression of the *HMBOX1* gene with comparable levels (Figure 6C).

Next we investigated the effect of Tat transfection. Consistently with what has been observed with RT PCR analysis, the number of Tat-transfected nuclei showing expression of the two *HMBOX1* loci decreased (Figure 6A). This difference was significant in both cell lines (p value = 1.34×10^{-23} for HOS_143b and p value = 2.63×10^{-4} for HOS_A4). Interestingly, in those cells where both spots were detected, the number of RNAs and the ratio of the *HMBOX1* alleles were not affected (Figure 6B and Figure 6C). HIV transcripts instead were present in most (93%) of Tat-EGFP transfected cells, consistent with transactivation of the viral LTR in the clonal population (Figure 7). Interestingly, most cells that express the vector do not show expression of *HMBOX1* in both alleles, indicating that the effect on the expression of *HMBOX1* was dependent on the expression of Tat and not on the transactivation of the provirus. Even more strikingly, in 11% of Tat-EGFP transfected nuclei both *HMBOX1* alleles and the proviral transcript were active (Figure 7). In this subset of cells the ratio between *HMBOX1* alleles was also close to 1 and the intensity of the proviral signal comparable to that of cells where there was no *HMBOX1* expression. In fact, as shown in figure 6D, transcription of HIV-1 was not significantly affected by allele-specific *HMBOX1* transcription (p value = 0.91). Hence, there are conditions where intragenic transcription of HIV-1 can occur in the presence of transcription of the host gene.

Discussion

Integration of HIV-1 in host chromatin is a crucial event for viral pathogenesis. Chromatin control of provirus gene expression has been postulated to be a major determinant of post-integration latency that is the cause of failure to eradicate HIV-1 infection by current antiretroviral regimens [1-3]. In addition, development of lentiviral vectors for gene therapy requires that endogenous genes shouldn't be affected by the integration event. However, recent evidence suggests that down-modulation of HIV-1 expression occurs also within active genes, in the absence of a repressive chromatin context. A cell line harboring a Tat-inducible HIV-1 vector integrated within the *HMBOX1* endogenous gene has been engineered in this work. This allowed the detailed investigation of the reciprocal influence of HIV-1 and *HMBOX1* expression both with or without Tat induction.

As a result of the double selection procedure, HOS_A4 showed a basal level of HIV-1 RNA by RT PCR that could be assigned to *HMBOX1* read-through transcription across the silent HIV-1 provirus since neither RNA was evident in FISH (Figure 5C) nor ECFPskl could be detected in the cytoplasm (Figure 3A). Tat overexpression, while increasing HIV-1 expression as expected, also reduced the level of expression of *HMBOX1* in both alleles. Besides its essential role in *trans*-activating HIV-1 transcription, Tat is known to regulate key host cell functions, primarily at the level of transcription. For example, Tat down-regulates MHC class II by preventing the interaction of cyclin T1 with the class II transactivator CIITA [36]. It is conceivable that Tat, being able to interact with a variety of host factors required for HIV transactivation [13,37], at the same time pulls these factors away from specific host genes, altering transcription from these promoters. Genome-wide expression profiling indeed revealed that Tat overexpression resulted in down-modulation of many cellular genes, possibly through targeting of general factors such as the SWI/SNF chromatin remodeling complex and the p300 acetyltransferase [38,39]. Hence, *HMBOX1* adds to the list of genes being down modulated by Tat overexpression. It is possible that during the establishment of the HOS_A4 cell line there has been a positive selection of integration loci where Tat induced repression of transcription. If expression of the endogenous gene interferes in *cis* with the expression of the provirus, the net result of Tat induction would have been of increased LTR-driven expression due also in part to the decrease of *HMBOX1* expression.

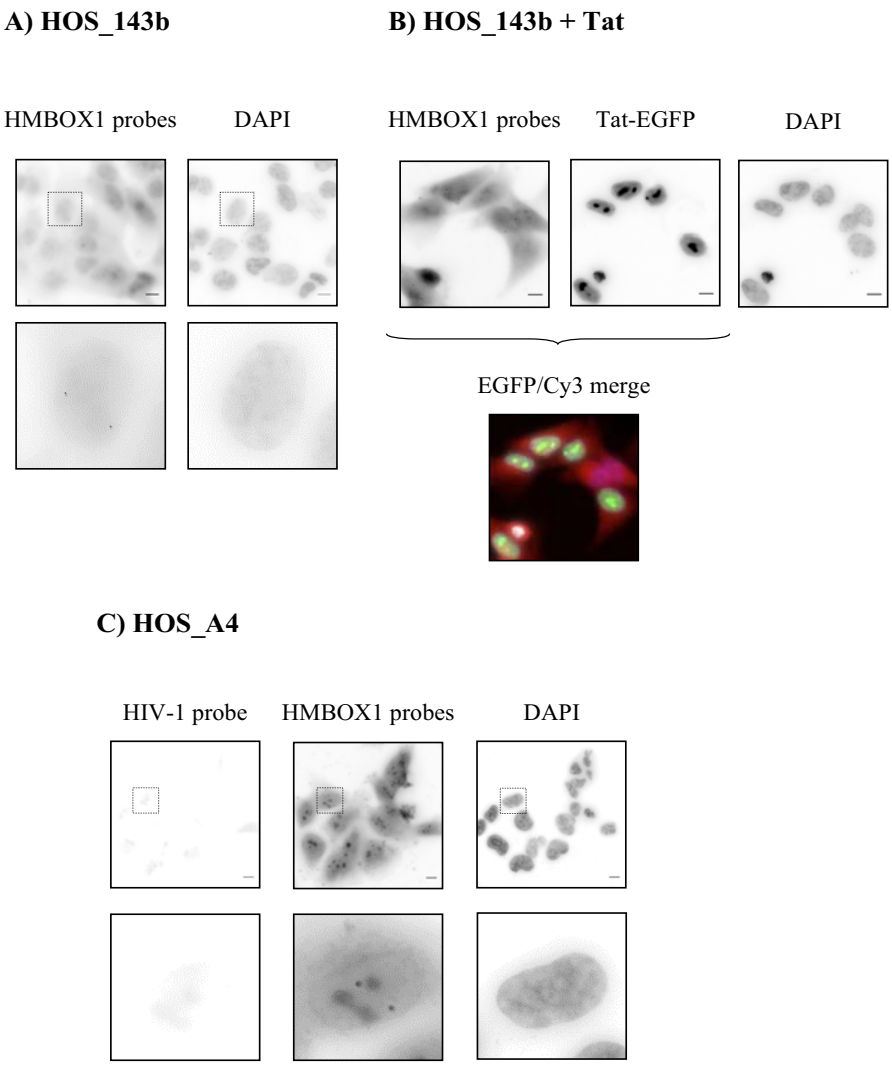


Figure 5
A) FISH for *HMBOX1* RNA on parental HOS_143b cells. Top panel: large field image (bar = 10 μm). Bottom panels: single cell from the figure above (inset). Two distinct hybridization signals per nucleus demonstrate bi-allelic expression of the *HMBOX1* gene. **B)** Same as A after transfection of Tat-EGFP. Also the signal for Tat-EGFP is shown (middle panels). Bottom panel: merge of *HMBOX1* hybridization and Tat-EGFP expression. **C)** FISH for HIV-Intro RNA on HOS_A4 cells. Top panel: large field image (bar = 10 μm). Bottom panels: single cell from the figure above (inset). Absence of the hybridization signal with the HIV probe is due to silencing of the gene without Tat (left panels). Two distinct hybridization signals per nucleus demonstrate bi-allelic expression of the HMBOX1 gene (middle panels).

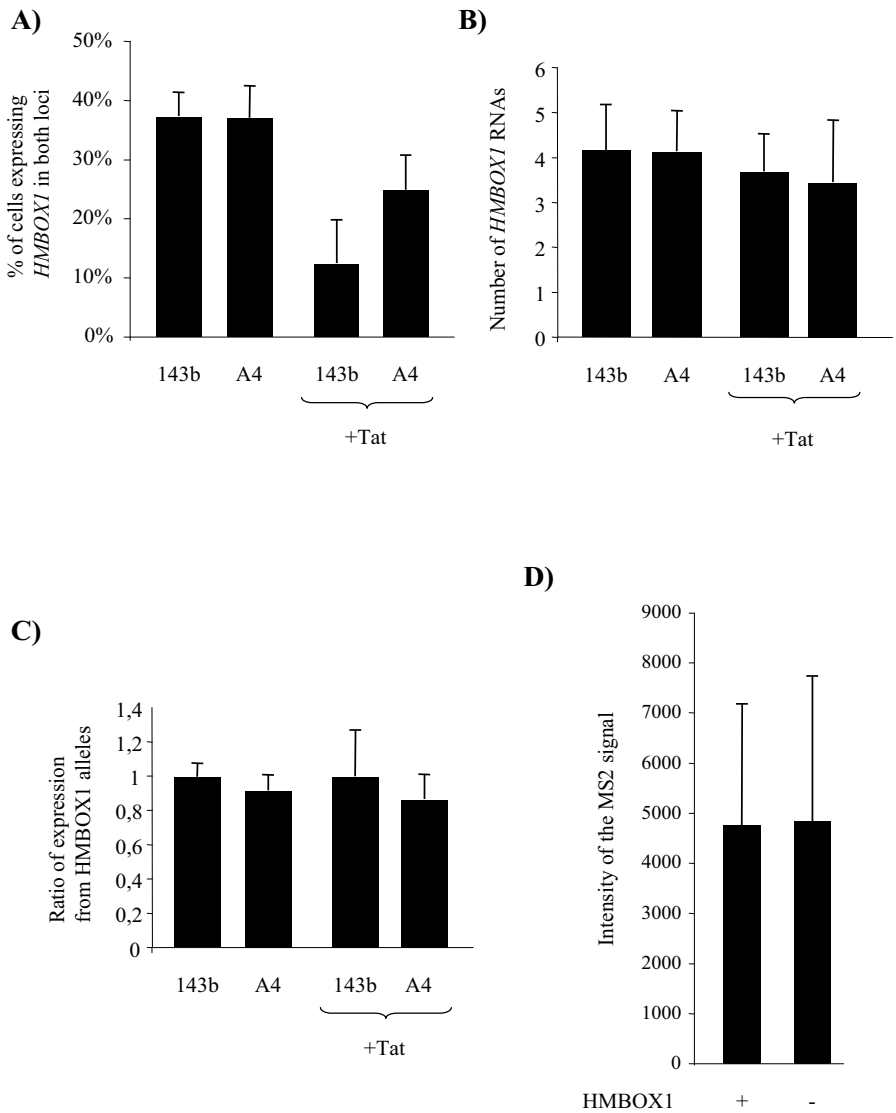


Figure 6

A) The number of cells expressing *HMBOX1* in both loci before and after Tat-EGFP transfection is shown. B) The number of *HMBOX1* RNAs on the transcription spots before and after Tat-EGFP transfection is shown. **C)** The ratio between the intensity of the hybridization signals on the transcription spots of the *HMBOX1* alleles is shown. **D)** The intensity of the MS2 hybridization signal on HIV-Intro transcription spots is shown both on a *HMBOX1* active and inactive background.

Table 1: Primers for RT-PCR and probes for RNA FISH used in this work.

Name	Sequence 5' > 3'
PRIMERS FOR RT-PCR	
HIV_UP5	GGCGAGCCCTCAGATCCTGC
HIV_START	GGGTCTCTCTGGTTAGACCAGATCTGAGCC
HIV_RT	AGTCGCCGCCCCTCGCCTCCTGC
HIV_SPLICED	GGATTAACTGCGAATCGTTCTAGC
HIV_NUC	CGTCTGTTGTGTGACTCTGGTAACT
HMBOX_E1s	GTCTCTTTCCTCACTTCTTTCT
HMBOX_E1a	TCACAGTTTCCAGAACTCCAC
HMBOX_E2s	GGAATGGAACAGTGAAGAAGCA
HMBOX_E3s	AATGGTAGATAACGCAGATCATC
HMBOX_E3a	ACCACTGGAAAGGAACTAAGCA
PROBES FOR RNA FISH	
HIV_MS2	AxGTCGACCTGCAGACAxGGGTGATCCTCAxGTTTTCTAGGCAATxA
HMBOX_E1a	AxAGTTTCCAGAAxTCCACACCGGAGACCCCACxTCCAGGATTCAAACCxT
HMBOX_E1b	CxCAGCGTCCGCxCACTTCCTCCCCAAAACCCCxCCAAAAAAATTGTTxT
HMBOX_I2a	AxGGTTGGTATAAACACAxAAAGCATGGTGGTxGTCTGGAGCTGGGGTTxA
HMBOX_I2b	AxTGCAGTGAGCCATGAxCACACCACAGTACxACAGCCTGGGTGATGAAxA
HMBOX_I2c	AxATTGCTGTCCTAAxCAGACTGCACCTGTGGxGTGGCTCTGACTGGTxA
HMBOX_I2d	AxGGTATGGTGGCAAAxCGACTCCCCCAGxACAACCACCAGAATATCAGxA
HMBOX_E4a	AxACGCCGAAGTCGCxGAAGCAGATCTATCTGCxCTATGGTAAATCTGGxA
HMBOX_E4b	AxATAACTGTTGCTAGGxGACGGGGACATTCCCGAAxGCTGCGTCTGTxA

Modified aminoallyl thymidines in the probes are indicated with "x". The sequence of one binding site for MS2 is underlined in the HIV_MS2 probe.

However, the number of HIV-1 nascent RNAs in Tat-transfected HOS_A4 cells that were negative for *HMBOX1* expression was similar to the number of those where instead *HMBOX1* was active (Figure 6D).

Another important finding of this work was that in some cells expression of the endogenous *HMBOX1* gene and of the provirus coexisted at the same transcription site. This finding could not be anticipated since it was believed that transcriptional interference should have occurred. Transcriptional interference is the suppressive influence of one transcriptional process directly and in *cis*, on a second transcriptional process [40,41]. Several combinations of the disposition of the two transcription units produce different effects [42]. For example, two promoters firing in opposite orientation would end up in collision of the two converging transcription elongation complexes. In HOS_A4 cells instead, the two promoters elongate in the same direction allowing a possible transcriptional interference through occlusion of the downstream promoter. In such model transcription from the upstream *HMBOX1* promoter should transiently preclude the occupation by RNAPII and/or associated transcription factors of the downstream LTR promoter. Alternatively, the LTR could pose a roadblock to the progress of the transcription-elongation complex firing from the *HMBOX1* promoter resulting in its inhibition of expression from one allele.

The observation that in the absence of Tat transcription both *HMBOX1* alleles were equally expressed while expression from the viral LTR remained undetectable might indicate that transcription elongation of RNAPII across the *HMBOX1* gene occluded the viral LTR. The situation changed dramatically in the presence of Tat. Transcription from the viral LTR is switched on as shown in RT PCR as well as demonstrated by the appearance of the viral transcript in RNA FISH. As discussed previously, Tat repressed *HMBOX1* expression while allowing provirus transcription. However, in a subset of cells, expression of both genes coexisted. Recent data also analyzed HIV proviral gene expression from within a cellular gene [43]. Work by Peterlin's group showed transcriptional interference occurring from elongating polymerase firing from the host gene [44], whereas work from Silicano's laboratory showed that elongating polymerase from the host gene could enhance HIV transcription when orientated in the same direction [45]. It is difficult to compare directly these data with those presented here since different cell lines and host cell genes were studied. However the findings presented in this work show an alternative situation that will require further analysis to understand the molecular basis of the phenomenon.

Conclusion

In this work HIV-1 gene expression was studied from within the endogenous gene *HMBOX1*. Transcriptional Tat-transactivation of the viral LTR resulted in up-regulation of HIV-1 transcription while it repressed HMBOX1 gene expression in both alleles, independent of vector integration. Hence, it could be proposed that HIV-1

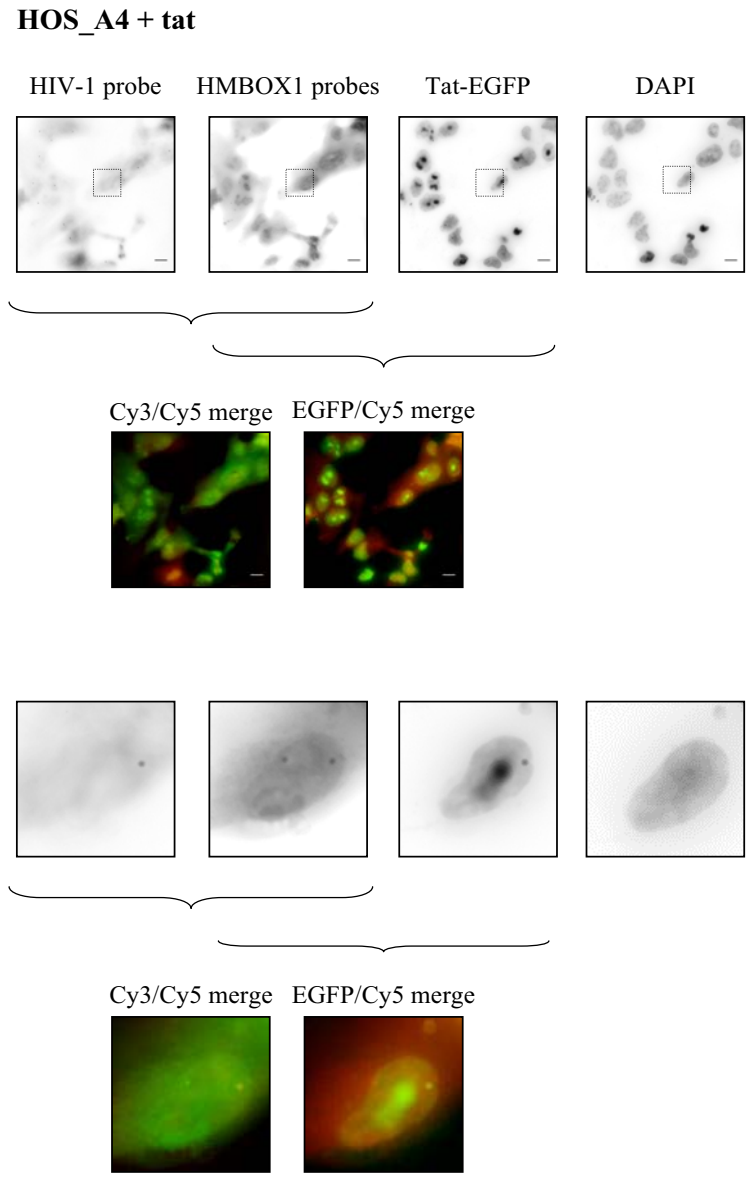


Figure 7
FISH for *HMBOX1* and HIV-Intro RNA on HOS_A4 cells after transfection of Tat-EGFP. Merged images are shown in color both for the large fields and for the zoomed figures (bar = 10 μm).

genome insertion in genes repressed by Tat could be an advantage for the virus allowing its transactivation from a low-transcribing endogenous gene. It has also been observed that both HIV-1 and *HMBOX1* gene expression may occur at the same genomic location in the same cell, allowing speculation on the lack mutual interference between transcription units.

Methods

Plasmids and cells

Plasmid pHIV-Intro was derived from the plasmid pEV731 [27] by cloning 24 MS2 repeats into the NotI restriction site. A cassette encoding for ECFP with the peroxisome localization signal Ser-Lys-Leu (skl) was inserted between the ClaI and XhoI sites, the IRES from EMCV and the thymidine kinase from HSV-1 were cloned in the XhoI site [46]. To obtain cells transduced by the HIV-Intro lentiviral vector expressing the HSV-TK gene we exploited a protocol for negative and positive selection of TK- cells. HEK-293T cells were transfected with the vector plasmid pHIV-Intro together with the packaging plasmids as previously described [26]. The supernatant was filtered and used to transduce human bone osteosarcoma TK- cells (HOS 143b, ECACC n. 91112502). Next day cells were treated with ganciclovir at 50 μg/ml. Surviving cells were expanded and then treated with 2.5 μg/ml of GST-Tat to induce LTR expression as described previously [34,47]. The following day cells were incubated in hypoxanthine, aminopterin and thymidine (HAT) medium and then cloned by colony picking and expansion in complete DMEM medium. Colonies were visually scored for low basal level of ECFP expression and to be highly inducible by GST-Tat by fluorescence microscopy. The number of integrations was assessed by southern blotting with a probe for ECFP and the cloning of the integration sites was obtained by inverse PCR essentially as described [26]. Briefly, genomic DNA was digested with BamHI (cleavage site within HIV-Exo-MS2 × 24), and the resulting products were circularized with DNA ligase. The product of two nested PCRs performed with primers pointing outwards from the vector was cloned and sequenced.

Plasmids expressing Tat and MS2 have been described previously [30,48]. Tat and Rev were fused to pDsRed-Monomer-N1 (Clontech) by PCR.

Allele-specific RT-PCR

Total RNA was extracted as described by the kit (Qiagen, RNA easy). Reverse transcription was performed with M-MLV RT (Invitrogen) using random primers. Amplification of the cDNA was conducted in the presence of CyberGreen™ (Applied Biosystems) and monitored on AbiPrism 7000 (Applied Biosystems). Specific primers are shown in Table 1.

Quantitative RNA fluorescent in situ hybridization (RNA FISH)

In situ hybridization was performed essentially as previously described [30]. Cells were fixed with 4% PFA pH7.4 and permeabilized overnight in 70% ethanol. Formamide concentration was 40% for *HMBOX1* probes and 20% for MS2. The amino-allyl thymidine modified oligonucleotide probes (Table 1) have been synthesized by J-M Escudier (Platforme de synthèse d'Oligonucléotides modifiés de l'Interface Chimie Biologie de l'ITAV, Toulouse, France). Probes were labeled with Cy3 or Cy5 (Cyn monoreactive dye, Amersham). For quantitative measurements stacks of 21 planes were acquired at bin = 2 with steps of 0.5 μm in the z-axis using a wide-field Leica DMRI inverted microscope (63× objective, NA 1.3) controlled by Metamorph (Universal Imaging). Digital images were collected using a CoolSnap K CCD camera (Roper scientific). The three-dimensional deconvolution and reconstruction was performed with the ImageJ plug-in "Iterative Deconvolve 3D". The total light intensity at the transcription site was calculated and divided for the number of planes and the number of molecule were computed from a calibration curve of the probes in solution [30].

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

ADM and CB carried out the RT-PCR, ADM and PM carried out the *in situ* hybridization and quantitative analysis, AK, CV and AM prepared and characterized the cell line, AM contributed to the experimental design and coordination of the study, data analysis, as well as to writing the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

We thank Edouard Bertrand (IGMM, Montpellier, France) for the gift of reagents and useful suggestions and Gianluca Pegoraro (NIH, Bethesda, US) for critically reading the manuscript. This work was supported in part by a HFSP Young Investigators Grant, by the Italian FIRB program of the "Ministero dell'Istruzione, Universita' e Ricerca" of Italy and by the AIDS Program of the "Istituto Superiore di Sanità" of Italy.

References

1. Han Y, Wind-Rotolo M, Yang HC, Siliciano JD, Siliciano RF: **Experimental approaches to the study of HIV-1 latency.** *Nat Rev Microbiol* 2007, **5:**95-106.
2. Bisgrove D, Lewinski M, Bushman F, Verdin E: **Molecular mechanisms of HIV-1 proviral latency.** *Expert Rev Anti Infect Ther* 2005, **3:**805-814.
3. Marcello A: **Latency: the hidden HIV-1 challenge.** *Retrovirology* 2006, **3:**7.
4. Daniel R, Smith JA: **Integration site selection by retroviral vectors: molecular mechanism and clinical consequences.** *Hum Gene Ther* 2008, **19:**557-568.
5. Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, Lim A, Osborne CS, Pawliuk R, Morillon E, Sorensen R, Forster A, Fraser P, Cohen JI, de Saint Basile G, Alexander I, Wintergerst U, Frebourg T, Aurias A, Stoppa-Lyonnet D,

Romana S, Radford-Weiss I, Gross F, Valensi F, Delabesse E, Macintyre E, Sigaux F, Soulier J, Leiva LE, Wissler M, Prinz C, Rabbitts TH, Le Deist F, Fischer A, Cavazzana-Calvo M: **LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1.** *Science* 2003, **302:**415-419.
6. Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F: **HIV-1 integration in the human genome favors active genes and local hotspots.** *Cell* 2002, **110:**521-529.
7. Wu X, Li Y, Crise B, Burgess SM: **Transcription start regions in the human genome are favored targets for MLV integration.** *Science* 2003, **300:**1749-1751.
8. Mitchell RS, Beitzel BF, Schroder AR, Shinn P, Chen H, Berry CC, Ecker JR, Bushman FD: **Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences.** *PLoS Biol* 2004, **2:**E234.
9. Ciuffi A, Llano M, Poeschla E, Hoffmann C, Leipzig J, Shinn P, Ecker JR, Bushman F: **A role for LEDGF/p75 in targeting HIV DNA integration.** *Nat Med* 2005, **11:**1287-1289.
10. Wang GP, Ciuffi A, Leipzig J, Berry CC, Bushman FD: **HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications.** *Genome Res* 2007, **17:**1186-1194.
11. De Palma M, Montini E, Santoni de Sio FR, Benedicenti F, Gentile A, Medico E, Naldini L: **Promoter trapping reveals significant differences in integration site selection between MLV and HIV vectors in primary hematopoietic cells.** *Blood* 2005, **105:**2307-2315.
12. Berkhout B, Silverman RH, Jeang KT: **Tat trans-activates the human immunodeficiency virus through a nascent RNA target.** *Cell* 1989, **59:**273-282.
13. Marcello A, Zoppe M, Giacca M: **Multiple modes of transcriptional regulation by the HIV-1 Tat transactivator.** *IUBMB Life* 2001, **51:**175-181.
14. Cujec TP, Cho H, Maldonado E, Meyer J, Reinberg D, Peterlin BM: **The human immunodeficiency virus transactivator Tat interacts with the RNA polymerase II holoenzyme.** *Mol Cell Biol* 1997, **17:**1817-1823.
15. Mavankal G, Ignatius Ou SH, Oliver H, Sigman D, Gaynor RB: **Human immunodeficiency virus type 1 and 2 Tat proteins specifically interact with RNA polymerase II.** *Proc Natl Acad Sci USA* 1996, **93:**2089-2094.
16. Chiang CM, Roeder RG: **Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators.** *Science* 1995, **267:**531-536.
17. Blau J, Xiao H, McCracken S, O'Hare P, Greenblatt J, Bentley D: **Three functional classes of transcriptional activation domain.** *Mol Cell Biol* 1996, **16:**2044-2055.
18. Cujec TP, Okamoto H, Fujinaga K, Meyer J, Chamberlin H, Morgan DO, Peterlin BM: **The HIV transactivator TAT binds to the CDK-activating kinase and activates the phosphorylation of the carboxy-terminal domain of RNA polymerase II.** *Genes Dev* 1997, **11:**2645-2657.
19. Jeang KT, Chun R, Lin NH, Gatignol A, Glabe CG, Fan H: **In vitro and in vivo binding of human immunodeficiency virus type 1 Tat protein and Sp1 transcription factor.** *J Virol* 1993, **67:**6224-6233.
20. Macian F, Rao A: **Reciprocal modulatory interaction between human immunodeficiency virus type 1 Tat and transcription factor NFAT1.** *Mol Cell Biol* 1999, **19:**3645-3653.
21. Marzio G, Tyagi M, Gutierrez MI, Giacca M: **HIV-1 tat transactivator recruits p300 and CREB-binding protein histone acetyltransferases to the viral promoter.** *Proc Natl Acad Sci USA* 1998, **95:**13519-13524.
22. Benkirane M, Chun RF, Xiao H, Ogryzko VV, Howard BH, Nakatani Y, Jeang KT: **Activation of integrated provirus requires histone acetyltransferase. p300 and P/CAF are coactivators for HIV-1 Tat.** *J Biol Chem* 1998, **273:**24898-24905.
23. Hottiger MO, Nabel GJ: **Interaction of human immunodeficiency virus type 1 Tat with the transcriptional coactivators p300 and CREB binding protein.** *J Virol* 1998, **72:**8252-8256.
24. Wei P, Garber ME, Fang SM, Fischer WH, Jones KA: **A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA.** *Cell* 1998, **92:**451-462.
25. Marcello A, Lusic M, Pegoraro G, Pellegrini V, Beltram F, Giacca M: **Nuclear organization and the control of HIV-1 transcription.** *Gene* 2004, **326:**1-11.
26. Jordan A, Bisgrove D, Verdin E: **HIV reproducibly establishes a latent infection after acute infection of T cells in vitro.** *Embo J* 2003, **22:**1868-1877.
27. Jordan A, Defechereux P, Verdin E: **The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation.** *Embo J* 2001, **20:**1726-1738.
28. Lewinski MK, Bisgrove D, Shinn P, Chen H, Hoffmann C, Hannenhalli S, Verdin E, Berry CC, Ecker JR, Bushman FD: **Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription.** *J Virol* 2005, **79:**6610-6619.
29. Han Y, Lassen K, Monie D, Sedaghat AR, Shimoji S, Liu X, Pierson TC, Margolick JB, Siliciano RF, Siliciano JD: **Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes.** *J Virol* 2004, **78:**6122-6133.
30. Boireau S, Maiuri P, Basyuk E, de la Mata M, Knezevich A, Pradet-Balade B, Backer V, Kornblihtt A, Marcello A, Bertrand E: **The transcriptional cycle of HIV-1 in real-time and live cells.** *J Cell Biol* 2007, **179:**291-304.
31. Darzacq X, Shav-Tal Y, de Turris V, Brody Y, Shenoy SM, Phair RD, Singer RH: **In vivo dynamics of RNA polymerase II transcription.** *Nat Struct Mol Biol* 2007, **14:**796-806.
32. Binnie A, Castelo-Branco P, Monks J, Proudfoot NJ: **Homologous gene sequences mediate transcription-domain formation.** *J Cell Sci* 2006, **119:**3876-3887.
33. Chen S, Saiyin H, Zeng X, Xi J, Liu X, Li X, Yu L: **Isolation and functional analysis of human HMBOX1, a homeobox containing protein with transcriptional repressor activity.** *Cytogenet Genome Res* 2006, **114:**131-136.
34. Lusic M, Marcello A, Cereseto A, Giacca M: **Regulation of HIV-1 gene expression by histone acetylation and factor recruitment at the LTR promoter.** *Embo J* 2003, **22:**6550-6561.
35. Levsky JM, Shenoy SM, Pezo RC, Singer RH: **Single-cell gene expression profiling.** *Science* 2002, **297:**836-840.
36. Kanazawa S, Okamoto T, Peterlin BM: **Tat competes with CIITA for the binding to P-TEFb and blocks the expression of MHC class II genes in HIV infection.** *Immunity* 2000, **12:**61-70.
37. Jeang KT, Xiao H, Rich EA: **Multifaceted activities of the HIV-1 transactivator of transcription, Tat.** *J Biol Chem* 1999, **274:**28837-28840.
38. de la Fuente C, Santiago F, Deng L, Eadie C, Zilberman I, Kehn K, Maddukuri A, Baylor S, Wu K, Lee CG, Pumfery A, Kashanchi F: **Gene expression profile of HIV-1 Tat expressing cells: a close interplay between proliferative and differentiation signals.** *BMC Biochem* 2002, **3:**14.
39. Gibellini D, Re MC, La Placa M, Zauli G: **Differentially expressed genes in HIV-1 tat-expressing CD4(+) T-cell line.** *Virus Res* 2002, **90:**337-345.
40. Mazo A, Hodgson JW, Petruk S, Sedkov Y, Brock HW: **Transcriptional interference: an unexpected layer of complexity in gene regulation.** *J Cell Sci* 2007, **120:**2755-2761.
41. Shearwin KE, Callen BP, Egan JB: **Transcriptional interference – a crash course.** *Trends Genet* 2005, **21:**339-345.
42. Eszterhas SK, Bouhassira EE, Martin DI, Fiering S: **Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position.** *Mol Cell Biol* 2002, **22:**469-479.
43. Perkins KJ, Proudfoot NJ: **An ungracious host for an unwelcome guest.** *Cell Host Microbe* 2008, **4:**89-91.
44. Lenasi T, Contreras X, Peterlin BM: **Transcriptional interference antagonizes proviral gene expression to promote HIV latency.** *Cell Host Microbe* 2008, **4:**123-133.
45. Han Y, Lin YB, An W, Xu J, Yang HC, O'Connell K, Dordai D, Boeke JD, Siliciano JD, Siliciano RF: **Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough.** *Cell Host Microbe* 2008, **4:**134-146.
46. Marcello A, Giaretta I: **Inducible expression of herpes simplex virus thymidine kinase from a bicistronic HIV1 vector.** *Res Virol* 1998, **149:**419-431.
47. du Chéné I, Basyuk E, Lin YL, Triboulet R, Knezevich A, Chable-Bessia C, Mettling C, Baillat V, Reynes J, Corbeau P, Bertrand E, Marcello A, Emiliani S, Kiernan R, Benkirane M: **Suv39H1 and HP1gamma are**

responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. *Embo J* 2007, **26:**424-435.

48. Marcello A, Cinelli RA, Ferrari A, Signorelli A, Tyagi M, Pellegrini V, Beltram F, Giacca M: **Visualization of in vivo direct interaction between HIV-1 TAT and human cyclin T1 in specific subcellular compartments by fluorescence resonance energy transfer.** *J Biol Chem* 2001, **276:**39220-39225.