GCN5-Dependent Acetylation of HIV-1 Integrase Enhances Viral Integration

Mariaelena Terreni

Ph.D Thesis in Molecular Biology

Supervisor: Dott.ssa Anna Cereseto

Scuola Normale Superiore

2012

Alla mia mamma

CONTENTS

CHAPTER 1

AIM OF THE THESIS

Integration of the reverse transcribed viral cDNA into the host cellular genome is an essential event during the replication cycle of HIV-1. This step is catalyzed by the viral integrase (IN) protein, which has been identified as a substrate for acetylation by the cellular histone acetyltransferase (HAT) p300 (Cereseto et al., 2005; Topper et al., 2007). The modified residues have been mapped to the IN C-terminal domain (CTD), corresponding to lysines 264, 266 and 273 (Cereseto et al., 2005; Topper et al., 2007). While IN acetylation by p300 was shown to increase both IN binding affinity to DNA and its strand transfer activity *in vitro* (Cereseto et al., 2005), conflicting results were obtained regarding the relevance of this post-translational modification (PTM) of IN during HIV-1 replication, with one study reporting a severe replication impairment for a mutant viral clone in which lysine-to-arginine substitutions were introduced at the IN sites targeted by p300-mediated acetylation (Cereseto et al., 2005), and a subsequent report claiming that the virus expressing the acetylation-defective mutant IN was able to replicate with kinetics similar to the wild type virus (Topper et al., 2007).

Based on these premises, the aim of the present thesis was that of investigating the role of IN acetylation by cellular HATs during HIV-1 integration and how this PTM influences the viral replication cycle. Since proteins modified by acetylation are often substrates for multiple HATs, we set out to explore whether IN might be acetylated by enzymes other than p300 and indeed found the existence of a physical and functional interaction between IN and the member of the GNAT family of HATs GCN5, resulting in the acetylation of four lysines in the IN CTD, at positions 258, 264, 266 and 273. Given that IN lysines 264, 266 and 273 proved to be common substrates of GCN5 and p300, we decided to perform a comparative analysis of the roles played by these enzymes during HIV-1 integration and replication. With the present study we obtained results supporting the notion that HIV-1 integration is modulated by an acetylation balance that controls the activity of the IN enzyme and involves at least two different cellular HATs, GCN5 and p300. Finally, the identification of a novel pathway of IN regulation unveils a new mechanism that can be exploited for future therapeutic development.

CHAPTER 2

INTRODUCTION

HIV-1 and AIDS

The Human Immunodeficiency Virus (HIV) is a member of the genus *Lentivirus* in the *Retroviridae* family. Retroviruses are so called because, upon entry into the host cell, their RNA genome is reverse transcribed into a DNA copy that is stably inserted into the cellular chromosomal DNA. The integrated form of the viral DNA, the provirus, then serves as template for the formation of viral RNAs and proteins that constitute the progeny virions. The reverse flow of genetic information from RNA to DNA and the establishment of the viral DNA in an integrated form in the host cell genome are the defining hallmarks of retroviruses.

Since its initial description in 1983, HIV type 1 (HIV-1) was identified as the etiological agent of the acquired immunodeficiency syndrome (AIDS) (Barre-Sinoussi et al., 1983; Broder and Gallo, 1984; Gallo et al., 1984), a clinical condition characterized by profound immunosuppression with associated opportunistic infections and malignant tumors, wasting, and central nervous system (CNS) degeneration.

Even if the AIDS epidemic was first identified only in the 1980s, the degree of morbidity and mortality caused by HIV and the global impact of HIV infection on health care resources and economics are already enormous. Although HIV incidence has been steadily declining since the late 1990s, levels of new infections are still high, and, with significant reductions in mortality due to the scale up of antiretroviral therapy over the past few years, the number of people living with HIV worldwide has increased: UNAIDS estimated that there were about 33 million people living with HIV infection and AIDS worldwide at the end of 2009 (UNAIDS, 2010).

The structure of HIV-1

The morphologic structure of the viral particle

HIV-1 particles are around 120 nm in diameter and roughly spherical. The viral genome consists of two copies of a single stranded RNA about 9500 nucleotides (nt) long and is tightly associated with the nucleocapsid protein p7. This ribonucleoprotein complex is enclosed by a conical capsid made up of the viral protein p24. The capsid environment also contains other viral proteins, such as

reverse transcriptase and integrase. The capsid is surrounded by a layer of matrix protein, p17, which is in turn anchored to the inside of the viral lipoprotein membrane, the envelope. This is formed when the virus buds from the infected cell, taking some of the host plasma membrane with it. The major HIV-1 proteins associated with the envelope are gp120 and gp41. The gp41 transmembrane (TM) subunit traverses the envelope, while the gp120 surface (SU) subunit is present on the outer surface and is non-covalently linked to gp41, thus forming a complex that mediates the attachment to target cells. Also enclosed within the virion are the viral protease and the accessory proteins Vif, Vpr and Nef. The structure of a HIV-1 mature virion is schematically represented in Figure 1.

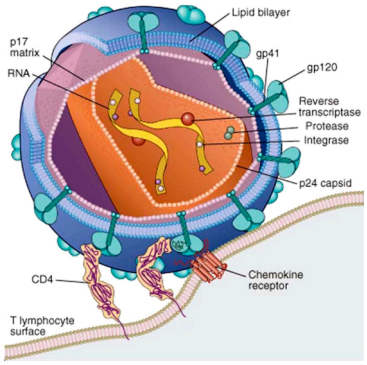

Figure 1. Schematic representation of a mature HIV-1 virion. (Abbas, 2010).

The organization of the viral genome

The HIV-1 genome is comprised of nine open reading frames (ORFs). The *gag* (group-specific antigen), *pol* (polymerase) and *env* (envelope) ORFs encode the classical structural and enzymatic precursor polyproteins, which are subsequently cleaved into the individual components common to all retroviruses (Figure 2). The four Gag proteins, p17 matrix (MA), p24 capsid (CA), p7 nucleocapsid (NC) and p6, together with the two Env proteins, gp120 and gp41, are the structural components that make up the core of the virion and the outer membrane envelope. The three *pol*-encoded proteins, protease (PR), reverse transcriptase (RT) and integrase (IN), provide essential enzymatic functions and are also included within the viral particle. In addition to the products of the *gag*, *pol* and *env* genes, HIV-1

genome encodes six additional proteins (Figure 2), often called accessory proteins: two expression enhancing factors, Tat (transactivator of transcription) and Rev (regulator of expression of the virion), as well as four proteins that modulate cell systems and functions for the advantage of the virus, Vif (viral infectivity factor), Vpu (viral protein U), Nef (negative factor), and Vpr (viral protein R).

Long terminal repeats (LTRs) at each end of the viral genome are generated during the process of reverse transcription and, therefore, exist only as repeats in the viral DNA. In the context of the provirus, the major function of the LTRs consists in the regulation of viral RNA synthesis. The 5' LTR normally acts as the HIV-1 transcriptional promoter, whose function is to recruit the RNA polymerase II (RNA pol II) holoenzyme to the start site of viral RNA synthesis, while the 3' LTR acts in transcription termination and polyadenylation.

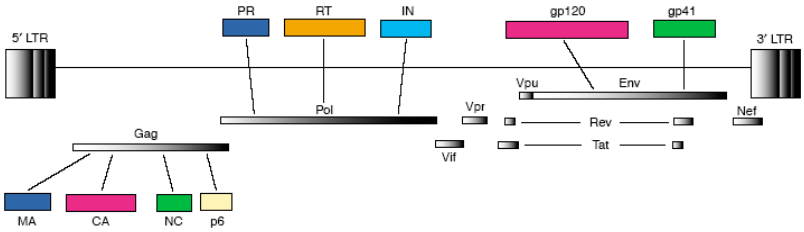

Figure 2. Schematic representation of the proteins encoded by HIV-1 proviral DNA. (Freed, 2004).

The HIV-1 replication cycle: an overview

The replication cycle of HIV-1 is a complex multistep process that depends on both viral and host cell factors (Figure 3) (Fields et al., 2007). The first step in HIV-1 infection involves the interaction between the virion SU gp120 protein and the primary receptor CD4 expressed on T-lymphocytes, T-cell precursors within the bone-marrow and thymus, monocytes and macrophages, eosinophils, dendritic cells and microglial cells of the CNS (Wyatt and Sodroski, 1998). CD4 binding induces conformational changes in gp120, resulting in the exposure of the coreceptor-binding determinants (Sullivan et al., 1998). The two major coreceptors required for entry of HIV-1 are the chemokine receptor molecules CCR5 and CXCR4 (Alkhatib et al., 1996; Choe et al., 1996; Deng et al., 1996; Feng et al.,

1996), which are used by monocyte/macrophage (M)-tropic strains (R5 isolates) and T-cell (T)-tropic viruses (X4 isolates), respectively (Berger et al., 1998). The final step for viral entry requires the gp41-mediated fusion of the viral envelope components with the host cell membrane. The current model to explain membrane fusion assumes that, after the formation of a ternary gp120-CD4-coreceptor complex, conformational changes occurr in gp41 that lead to the exposure of its N-terminal part. This part, also known as the fusion-peptide, is then inserted into the target lipid bilayer (Eckert and Kim, 2001), triggering the fusion of the HIV-1 envelope with the host cell membrane and virus internalization.

Following entry of the HIV-1 core into the cytoplasm of the target cell, reverse transcription of the viral genomic RNA into a linear double-stranded cDNA copy takes place within a large macromolecular complex, known as the reverse transcription complex (RTC), which comprises proteins of both cellular and viral origin (Fassati and Goff, 2001; Karageorgos et al., 1993). The enzyme that catalyzes the reaction is RT. RT possesses three catalytic activities essential for replication of the virus: RNA-dependent DNA polymerase (i.e. reverse transcriptase), RNase H (i.e. for degradation of the genomic RNA in RNA/DNA hybrids during DNA synthesis) and DNA-dependent DNA polymerase activity (i.e. for synthesis of the second strand of the viral cDNA) (Fields et al., 2007).

The next step in the replication cycle is the translocation of the viral double-stranded cDNA into the nucleus as a component of a large nucleoprotein complex called the pre-integration complex (PIC), which is composed of both viral and cellular proteins (Bukrinsky et al., 1993b; Farnet and Haseltine, 1991; Miller et al., 1997). To cross the intact nuclear envelope and enter the nucleus, the PIC must be actively transported through the nuclear pore complex (NPC). Many attempts have been made to determine the viral and cellular factors mediating nuclear import of the HIV-1 PIC. Various viral karyophilic proteins, such as MA, Vpr and IN have been suggested to actively translocate the PIC into the host cell nucleus (Bukrinsky et al., 1993b; Depienne et al., 2000; Farnet and Haseltine, 1991; Haffar et al., 2000; Jenkins et al., 1998; Miller et al., 1997; Sherman and Greene, 2002; Zennou et al., 2000). The cellular protein lens epithelium-derived growth factor p75 (LEDGF/p75), along with a short section of triple-stranded DNA present within the

viral cDNA and known as the DNA flap, have also been implicated in promoting translocation of the PIC into the nuclei of infected cells (Cherepanov et al., 2003; Llano et al., 2004b; Zennou et al., 2000). In addition, CA recently emerged to play a key role in mediating HIV-1 nuclear import (Yamashita and Emerman, 2004) (Yamashita et al., 2007). With regard to cellular cofactors, several nuclear-import receptors, such as the importin α/β heterodimer (Ao et al., 2005; Armon-Omer et al., 2004; Gallay et al., 1997; Hearps and Jans, 2006; Levin et al., 2009), importin 7 (Ao et al., 2007; Zaitseva et al., 2009) and transportin-SR2 (TRN-SR2, TNPO3, transportin 3) (Brass et al., 2008; Christ et al., 2008; Konig et al., 2008), as well as several components of the nuclear pore complex, among which Nup 98, Nup153 and RanBP2 (Nup358) (Brass et al., 2008; Ebina et al., 2004; Konig et al., 2008; Woodward et al., 2009; Zhang et al.), were all proposed to be required for efficient translocation of the HIV-1 PIC into the host cell nucleus. However, at present, the exact mechanism governing nuclear import of the HIV-1 PIC still awaits further clarification.

Following the PIC arrival in the nucleus, the viral IN protein catalyzes the integration of HIV-1 cDNA into the host cellular genome. The integration reaction and the host factors participating in the process will be discussed in full detail in the next section.

The integrated viral DNA is then transcribed, thus generating full-length progeny viral RNA and a number of spliced mRNA transcripts to be translated in the cytoplasm (Coffin et al., 1997). Transcription of HIV-1 proviral DNA is broadly similar to host gene transcription and involves a large array of cellular transcriptional activator and repressor proteins, as well as the entire RNA Pol II machinery. An exception is represented by the enhancement of expression from the HIV-1 LTR promoter, which is mediated by the viral regulatory protein Tat. Tat is an atypical transcriptional activator, since it does not bind to DNA, but to a RNA stem-loop structure, the transactivation response region (TAR), which forms at the 5′-end of nascent HIV-1 transcripts (Berkhout et al., 1989; Dingwall et al., 1989). Through its interaction with TAR, Tat activates transcription from the HIV-1 LTR by at least two different molecular mechanisms. The first mechanism consists in the recruitment of the cellular cofactor P-TEFb (positive transcription elongation factor

b), composed of the subunits cyclin T1 and CDK9 kinase. Tat interaction with cyclin T1 (Wei et al., 1998) increases its binding affinity to the TAR element, thus resulting in the formation of a ternary Tat/P-TEFb/TAR complex, which brings the kinase activity of P-TEFb in the proximity of the HIV-1 promoter. The Cdk9 subunit of P-TEFb then phosphorylates the carboxyl-terminal domain (CTD) of the host cell RNA pol II, which stimulates the elongation process and thereby the overall transcriptional efficiency (Bieniasz et al., 1999; Chen et al., 1999; Fujinaga et al., 1998; Garber et al., 1998; Ivanov et al., 1999). The second mechanism of Tat-induced transcriptional activation involves the recruitment of chromatin-modifying complexes to the HIV-1 promoter (Benkirane et al., 1998; Col et al., 2001; Hottiger and Nabel, 1998; Marzio et al., 1998)((Agbottah et al., 2006; Mahmoudi et al., 2006), which results in an increased DNA accessibilty to transcription factors and activators, thus promoting transcription initiation and efficient elongation.

The nuclear export of unspliced or partially spliced HIV-1 messages to the cytoplasm requires the action of the viral regulatory protein Rev. In the absence of Rev function, incompletely spliced HIV-1 mRNAs, primarily encoding viral structural proteins, are retained in the nucleus by cellular mRNA-processing factors, known as splicing commitment factors (Chang and Sharp, 1989; Emerman et al., 1989; Malim et al., 1989). Although these commitment factors can prevent incompletely spliced mRNAs from accessing the canonical cellular mRNA-export pathway, they cannot prevent the nuclear export of HIV-1 mRNAs bound by Rev. Rev contains an RNA binding motif that directly interacts with a stem-loop secondary structure, termed the Rev responsive element (RRE), which is present in all incompletely spliced viral mRNAs. The RRE can accommodate the binding of at least 8 Rev molecules and, at a certain threshold concentration of Rev protein in the nucleus, functional Rev/RRE complexes are formed, which greatly stimulate the export of unspliced and singly spliced mRNAs to the cytoplasm. The nuclear export of Rev with its RNA cargo is mediated by the cellular nucleocytoplasmic transport factor Crm1 (chromosome region maintenance 1), which interacts with the Rev nuclear export signal (NES) sequence (Fornerod et al., 1997; Fukuda et al., 1997; Stade et al., 1997). Once in the cytoplasm, the complex disassembles (Gorlich and Mattaj, 1996; Nigg, 1997) and Rev is released along with its unspliced or partially spliced

HIV-1 RNA cargo, allowing the latter to be incorporated into progeny virions or translated into viral proteins. The nuclear import of Rev is then achieved through the direct interaction of its nuclear localization signal (NLS) motif with the cellular cofactor importin β (Henderson and Percipalle, 1997; Szebeni et al., 1997), which results in the translocation of the complex through the nuclear pore.

HIV-1 assembly takes place in lipid rafts, plasma membrane microdomains enriched in cholesterol and glycosphingolipid (Ono and Freed, 2005), and is directed by Gag. The MA domain of Gag appears to be responsible for both directing the association of the polyprotein with the lipid bilayer and determining the specificity of membrane targeting (Facke et al., 1993) (Freed et al., 1994; Ono et al., 2000), while the CA domain seems to play a key role in driving the viral particle assembly (Ehrlich et al., 1992; Gross et al., 1997; von Schwedler et al., 1998). The third major domain synthesized as part of the Gag precursor is NC, whose main function in assembly involves the encapsidation of full-length, unspliced HIV-1 genomic RNA into virions through specific interaction with a cis-acting target known as packaging signal or ψ-site (Berkowitz et al., 1996).

Virus particle production is completed upon budding of the nascent virion from the plasma membrane (Adamson and Freed, 2007). To allow virus release from the infected cell, the p6 domain of Gag uses its highly conserved Pro-Thr-Ala-Pro (PTAP) motif (Huang et al., 1995), also known as “late domain” (since point mutations in this motif block the virus release at a late stage), to recruit components of the cellular endosomal sorting machinery, which normally function to promote the budding of vesicles into late endosome to form multivesicular bodies (MVBs) (Bieniasz, 2009; Demirov and Freed, 2004; Fujii et al., 2007; Morita and Sundquist, 2004). Concomitant with virus release, PR cleaves the Gag and Gag-Pol precursors into their respective protein domains (Coffin et al., 1997), leading to virion maturation, a reassembly event that results in the formation of the characteristic cone-shaped core and is essential for the acquisition of infectivity (Adamson and Freed, 2007; Ganser-Pornillos et al., 2008). Only after maturation is the virus ready for another round of infection.

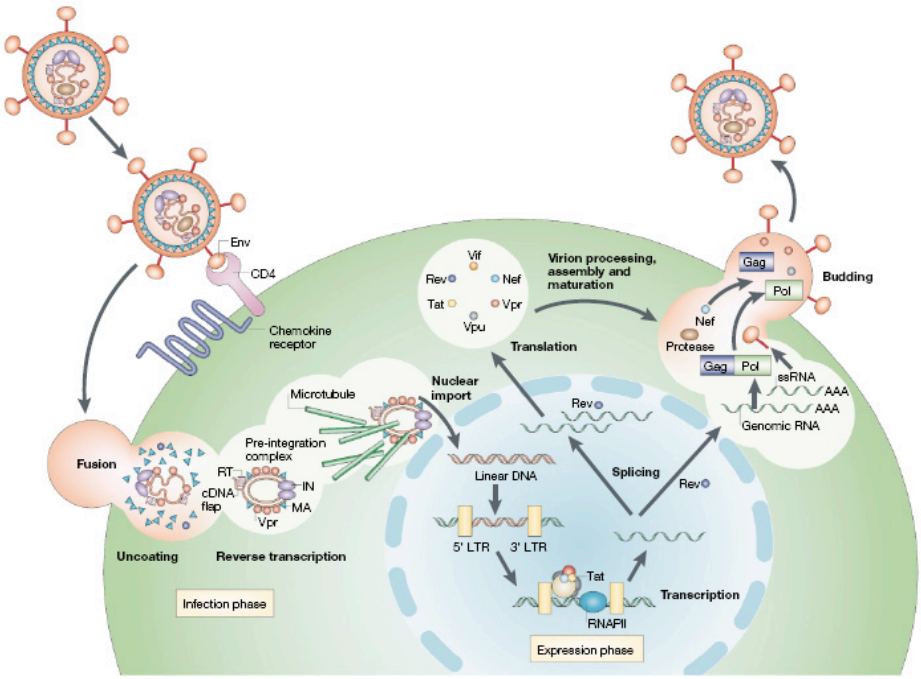

Figure 3. Schematic representation of the HIV-1 replication cycle. (Peterlin and Trono, 2003).

HIV-1 integration

A distinguishing feature of retroviral replication is the insertion of a DNA copy of the genomic RNA into a chromosome of the host cell following reverse transcription. The provirus then serves as the template for the synthesis of viral RNAs and is maintained as part of the host cell genome for the lifetime of the infected cell. Retroviral mutants deficient in the ability to integrate generally fail to establish a productive infection.

The integration reaction is catalyzed by the viral IN enzyme, a 32 kDa protein of 288 aminoacids generated by PR-mediated cleavage of the C-terminal portion of the HIV-1 Gag-Pol polyprotein.

IN structural overview

HIV-1 IN is made up of three structurally and functionally distinct domains: the N-terminal domain (NTD) spanning residues 1 to 50, the protease-resistant central catalytic core domain (CCD) from amino acids 50 to 212, and the C-terminal

domain (CTD) from residues 212 to 288 (Chiu and Davies, 2004; Jaskolski et al., 2009) (**Figure 4**). Among retroviral INs, the NTD and CCD share regions with a high degree of sequence similarity, while the CTD is much less conserved.

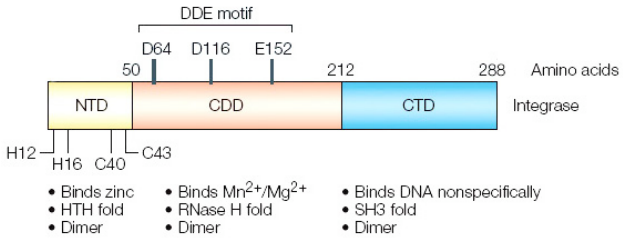

Figure 4. HIV-1 IN domain organization. Adapted from (Pommier et al., 2005).

It is not yet possible to crystallize the entire 288-aminoacid HIV-1 IN protein because of its low solubility and propensity to aggregate. The structure of all three domains, however, has been solved independently by x-ray crystallography or nuclear magnetic resonance (NMR) methods. In addition, structures are available for HIV- 1 IN fragments containing the NTD plus CCD (Wang et al., 2001), or the CCD plus CTD (Chen et al., 2000).

The NTD contains a conserved zinc-binding motif HHCC (His-12, His-16, Cys-40 and Cys-43) that coordinates one zinc atom (Zheng et al., 1996), though the structure of this region does not resemble that of zinc finger domains involved in DNA-binding (Lodi et al., 1995). Consistently, one known function of the HIV-1 IN NTD is to contribute to protein multimerization.

The CCD contains the catalytic DDE motif, which is conserved among all retroviral INs and consists of the active site residues Asp-64, Asp-116 and Glu-152 (Engelman and Craigie, 1992; Kulkosky et al., 1992; Rowland and Dyke, 1990). Mutation of any one of these three residues is sufficient to inactivate IN (Engelman and Craigie, 1992; Leavitt et al., 1993; van Gent et al., 1992). Crystal structures of the HIV-1 IN CCD were only obtained after extensive mutagenesis studies, which identified mutants with enhanced solubility (Bujacz et al., 1996; Dyda et al., 1994; Goldgur et al., 1999; Jenkins et al., 1995). These structures show that the HIV-1 IN CCD binds one Mg^{2+} ion between Asp-64 and Asp-116. Since Avian Sarcoma Virus (ASV) IN binds an additional Zn^{2+} or Cd^{2+} ion between Asp-64 and Glu-157

(the ortholog of Glu-152) (Bujacz et al., 1997), it is likely that the HIV-1 IN active site also binds two metal ions (Mg^{2+} or Mn^{2+}) when complexed with the ends of the viral DNA during the cleavage and joining reactions. Another structural feature of the CCD is the 10-aminoacid flexible loop encompassed between residues Gly-140 and Gly-149. These two residues potentially act as hinges for the overall movement of the loop that may serve as a clamp for the binding of the viral DNA ends to the catalytic site of IN. Consistent with this possibility, Gln-148, one of the flexible loop residues, has been shown to selectively bind to the penultimate cytosine at the 5'-end of the viral DNA (Johnson et al., 2006). Gln-148 is also a key residue for IN catalytic activity and resistance to IN inhibitors raltegravir and elvitegravir (Marinello et al., 2008).

The CTD is considerably less conserved among retroviral INs than are the other domains of the protein. The structure of the isolated CTD of HIV-1 IN was solved using NMR (Eijkelenboom et al., 1995; Lodi et al., 1995) and showed a folding topology that is reminiscent of Src homology 3 (SH3) domains found in many proteins that interact either with other proteins or with nucleic acids (Musacchio et al., 1992; Yu et al., 1992), although no sequence similarity to SH3 proteins could be detected. The HIV-1 IN CTD possesses non-specific DNA binding activity and seems to be important for the binding of the viral DNA ends, thus conferring to IN the ability to remain associated with the viral genome during reverse trancription (Esposito and Craigie, 1998; Heuer and Brown, 1997; Heuer and Brown, 1998; Jenkins et al., 1997).

Direct physical measurements of purified IN, as well as cross-linking experiments and *in vitro* complementation between defective variants of IN, provided compelling evidence that retroviral IN is a multimeric enzyme (Ellison et al., 1995; Engelman et al., 1993; Grandgenett et al., 1978; Jones et al., 1992; van Gent et al., 1993). While each of the HIV-1 IN domains forms dimers, IN appears to function as a tetramer (Cherepanov et al., 2003; Li et al., 2006; Wang et al., 2001). Indeed, a tetramer would be the minimal IN multimer to provide a pair of active sites with the expected spacing for the concerted integration of two viral DNA ends, and the two-domain NTD-CCD construct has been shown to crystallize in a tetrameric form, best described as a dimer-of-dimers (Hare et al., 2009; Wang et al., 2001). The

crystal structure of full-length IN from the prototype foamy virus (PFV) in complex with its cognate viral DNA also showed that the organization of the minimal functional complex involving viral DNA and IN, the so-called intasome, comprises an IN tetramer tightly associated with a pair of viral DNA ends (Hare et al.).

The mechanism of integration

The integration reaction biochemically occurs in two steps (**Figure 5**) following reverse transcription and binding of IN to the viral DNA ends in the context of the PIC (Vandegraaff and Engelman, 2007). Assembly of the PIC onto the 5'-GCAGT-3' sequence of the viral DNA ends forms a complex competent for integration, although, *in vitro*, only IN and DNA are required. The first step, named 3'-end processing, is an endonucleolytic cleavage, which results in the removal of two nucleotides from the 3'-ends of each viral DNA strand immediately 3' from the conserved CA dinucleotide (5'-GCAGT-3'); a 5'-GT-3' dinucleotide is generally released, resulting in recessed 3'-OH ends at each terminus of the viral DNA. The phosphodiester bond in the viral DNA strand is hydrolyzed by water as the nucleophile in the presence of Mg^{2+} (the presumably physiological metal) or Mn^{2+} (Esposito and Craigie, 1998). Following nuclear translocation of the PIC, IN catalyzes the second transesterification reaction, known under the name of strand transfer, where the recessed 3'-OH ends of the viral DNA act as nucleophiles to attack the DNA phosphodiester backbone of a host chromosome. HIV-1 IN positions the two viral ends to attack both strands of the host acceptor DNA, with a five base pair (bp) gap on each strand. The result is a staggered insertion that is subsequently sealed by host DNA repair enzymes. The HIV-1 provirus is flanked by a 5 bp direct repeat and terminates with the dinucleotides 5′-TG and CA-3′. The direct repeat results from the duplication of the cellular target sequence.

In addition to the linear, double-stranded viral cDNA that serves as the substrate for integration, two types of extrachromosomal viral DNA molecules are observed in the nucleus of HIV-1 acutely infected cells. These are circular DNA forms: 1-LTR circles, generated through homologous DNA recombination, and 2-LTR circles, produced by the nonhomologous end-joining pathway (Brown et al., 1989; Bukrinsky et al., 1993a). Although 1-LTR and 2-LTR circles are not substrates for

integration, they provide useful markers for nuclear import of the viral DNA, since their formation requires the enzymatic machinery present in the nucleus.

Our understanding of the chemistry of the retroviral integration reaction, as well as the screening of HIV-1 IN inhibitors, have been greatly assisted by the development of *in vitro* integration assays, where purified recombinant IN carries out 3'-end processing and strand transfer reactions in the presence of short synthetic oligonucleotides that mimic the viral DNA ends and a divalent metal ion (Mg^{2+} or Mn^{2+}) (Craigie et al., 1990; Debyser et al., 2001; Katz et al., 1990; Marchand et al., 2001).

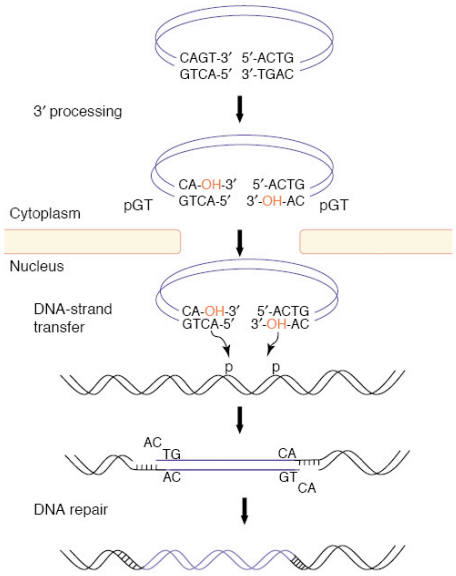

Figure 5. Outline of the integration reaction. (Van Maele et al., 2006).

Integration site selection by HIV-1

To replicate, a retrovirus must integrate a DNA copy of its genomic RNA into a chromosome of the host cell. The selection of cellular integration acceptor sites is crucial for both the retrovirus and the host (Bushman, 2001; Coffin et al., 1997). For the virus, selection of favorable integration target sites is required for efficient viral gene expression (Bisgrove et al., 2005; Jordan et al., 2003; Lewinski et al., 2005). For the host, integration can cause adverse events such as activation of proto-

oncogenes or inactivation of required cellular genes. Indeed, insertional activation of oncogenes has been observed in a human gene therapy trial, in which integration of a therapeutic retroviral vector near the *LMO2* proto-oncogene contributed to malignant transformation in several patients (Hacein-Bey-Abina et al., 2003a; Hacein-Bey-Abina et al., 2003b). These adverse events have focused intense interest on the mechanisms mediating retroviral integration site selection. The cellular DNA sequences hosting integration events show modest similarity to one another, indicating that retroviral DNA integration is not tightly sequence-specific (Berry et al., 2006; Carteau et al., 1998; Holman and Coffin, 2005; Stevens and Griffith, 1996; Wu et al., 2005).

Studies using *in vitro* integration assays allowed the identification of factors influencing integration site selection in simplified models. DNA binding proteins bound to the target DNA block integration by obstructing the access of PICs to the target DNA (Bushman, 1994; Pryciak and Varmus, 1992). In contrast, DNA bending proteins, such as nucleosomes, actually promote integration (Pruss et al., 1994a; Pruss et al., 1994b; Pryciak et al., 1992; Pryciak and Varmus, 1992). On the nucleosome, the positions of maximal DNA distortion are particularly favored for integration (Pruss et al., 1994a; Pruss et al., 1994b), probably because IN distorts its DNA substrates during the reaction cycle (Bushman and Craigie, 1992; Katz et al., 2001; Katz et al., 1998; Scottoline et al., 1997). Consistent with this notion, prior distortion of the integration target DNA promotes the integration reaction. Thus, wrapping of DNA around nucleosomes alone does not inhibit integration (Bushman and Craigie, 1992; Scottoline et al., 1997). However, it should be noted that the nucleosomal templates used in the above-mentioned experiments were all presumed to be in the 10 nm “beads on a string” conformation. The consequences on integration deriving from the incorporation of such structures into the 30 nm chromatin fibers, or the still higher-order structures that comprise the chromosomes, are unknown.

Several groups have investigated HIV-1 integration targeting *in vivo* by sequencing the junctions between viral and human DNA and analysing their positions in the human genome. Following large-scale sequence analysis, the position of integration sites in the human genome was compared with that of other annotated

features. In the first such study, the distribution of HIV-1 integration sites in the chromosomes of a human lymphoid cell line, SupT1, was investigated (Schroder et al., 2002). This study revealed that genes are favoured targets for HIV-1 integration, and subsequent experiments performed in other cell types reached the same conclusion (Mitchell et al., 2004; Wu et al., 2003). Further studies investigated whether any preferences exist in the location of HIV-1 integration sites along transcription units (TUs) (Mitchell et al., 2004; Schroder et al., 2002; Wu et al., 2003), but no biases have been found. Evidently, the positive influence of TUs on HIV-1 integration extends across their entire length. Gene-rich regions in the human genome are depleted in certain classes of repeat DNA sequences (such as human endogenous retroviruses, HERVs, and long interspersed nuclear elements, LINEs) and enriched in others (such as Alu elements). Therefore, a bias in the rates of integration in these different classes of DNA repeats was expected. Indeed, HIV-1 integration appears to be disfavoured in HERV elements, consistently with favoured HIV-1 integration in genes (Schroder et al., 2002). In addition, HIV-1 integration is strongly disfavoured in alphoid repeats (which, in humans, are composed of α-satellite DNA). This indicates that centromeric heterochromatin, the location of most of the ⟨-satellite-containing DNA, is an unfavourable target for HIV-1 integration (Carteau et al., 1998; Schroder et al., 2002). Since centromeric heterochromatin is known to be tightly wrapped by specific DNA-binding proteins, it appears that the packing of DNA into centromeric heterochromatin renders it less accessible, consequently disfavouring integration.

Transcriptional profiling analysis has been carried out in some of the cell types studied as integration targets, allowing the influence of transcriptional activity on integration site selection to be assessed. Some of these transcriptional profiling studies were carried out on retrovirus-infected cells (Lewinski et al., 2005; Mitchell et al., 2004; Schroder et al., 2002), so that the results reflected the influence of infection on cellular gene activity (Corbeil et al., 2001; Mitchell et al., 2003; Schroder et al., 2002; van 't Wout et al., 2003). Analysis of the microarray data revealed that the median expression level of genes hosting HIV-1 integration events is consistently higher than that of all the genes assayed on the microarray. Transcriptional profiling studies have also been performed for integration of a HIV-

based vector in SupT1 cells, demonstrating that genes activated upon infection are favoured integration targets (Schroder et al., 2002).

Since gene-rich regions in the human genome correlate with high densities of CpG islands, an analysis of HIV-1 integration frequency near CpG islands was performed. CpG islands are chromosomal regions particularly enriched in the rare CpG dinucleotide, and usually correspond to gene-regulatory elements that contain clustered binding sites for transcription factors. The regions surrounding CpG islands appear to be disfavoured for integration by HIV-1 (Mitchell et al., 2004). Therefore, in the case of HIV-1, gene-dense regions seem to contain interleaved clusters of active genes favouring integration and unfavourable regions that include CpG islands. The mechanism by which CpG islands obstruct HIV-1 integration is unclear: there might be specific proteins bound at these sites that block integration, CpG islands might be located in a nuclear compartment resulting unfavourable for integration, or some other as-yet-unknown mechanism might be responsible.

What mechanisms direct retroviral integration target-site selection in the human genome? It has been proposed that retroviral integration is favoured in decondensed chromatin, which would be more accessible to the integration apparatus (Panet and Cedar, 1977). This notion is supported by genome-wide studies, as integration in TUs is favoured in all the data sets for retroviruses and other integrating elements (Lewinski et al., 2005; Mitchell et al., 2004; Nakai et al., 2003; Schroder et al., 2002; Yant et al., 2005). By contrast, integration in highly condensed centromeric heterochromatin is disfavoured (Carteau et al., 1998; Schroder et al., 2002). However, since the integration target preferences of HIV-1, MLV and ASV are so different, it seems unlikely that the accessibility of DNA is the only mechanism determining target-site selection.

Studies on the integration site selection by Ty retrotransponsons in yeast (Boeke and Devine, 1998; Bushman, 2003; Sandmeyer, 2003; Zhu et al., 2003), showing that Ty integration complexes are tethered to their favoured sites in the host genome by interactions with specific cellular proteins, led to the hypothesis that such a tethering mechanism might also operate for retroviruses (Bushman, 2003). Several studies tested *in vitro* the feasibility of directing integration into specific DNA sites by the use of fusion proteins consisting of HIV-1 IN and a sequence-

specific DNA-binding protein, such as the phage λ repressor (Bushman, 1994), the *Escherichia coli* LexA repressor (Goulaouic and Chow, 1996; Katz et al., 1996), the murine transcription factor Zif268 (Bushman and Miller, 1997), or the synthetic zinc finger protein E2C (Tan et al., 2004). The resulting chimeras proved able to direct integration by recognizing and binding to their cognate target sites on the DNA, causing integration to be mediated into the immediately adjacent regions. Subsequent work also demonstrated that the IN-E2C fusion protein, packaged in *trans* into virions, was capable of directing integration of retroviral DNA into a predetermined chromosomal region in cultured cells (Tan et al., 2006). All these results provided proof of principle that tethering of retroviral IN proteins at target sites can promote nearby integration. For HIV-1 PICs, any viral or cellular protein present in the complex could theoretically act as binding partner in a tethering interaction. Indeed, a number of studies performed in the past few years led to the notion that the target site preferences of HIV-1 integration are in part due to tethering by the host chromatin-binding protein Ledgf/p75, which interacts with HIV-1 IN (Cherepanov et al., 2003; Emiliani et al., 2005) and mediates its binding to chromatin (Llano et al., 2004b; Maertens et al., 2003). In the absence of Ledgf/p75, HIV-1 integration is severely compromised and preferential integration in TUs is diminished (Ciuffi et al., 2005; Marshall et al., 2007; Shun et al., 2007b). Recently, the tethering model for Ledgf/p75 function was strengthened by the finding that fusion proteins containing the IN-binding domain of Ledgf/p75 fused to alternative chromatin binding domains are able to efficiently retarget integration (Ferris et al., 2010; Gijsbers et al., 2010; Silvers et al., 2010). The role of LEDGF/p75 in HIV-1 integration site selection will be furtherly discussed in the following section.

Notably, the existence of a link between nuclear entry of the HIV-1 PIC and integration targeting was recently hypothesized, based on the finding that the cellular proteins TRN-SR2 (TNPO3, transportin 3) and RanBP2 (Nup358), which are components of the nuclear import machinery and of the nuclear pore, respectively, affect the distribution of HIV-1 integration sites in the host chromosomes by regulating the nuclear import of the PIC (Ocwieja et al., 2011). According to the proposed model, nuclear translocation through the pore would first place the PIC in regions of high gene density, and Ledgf/p75 would then tether

the PIC for integration to provide the final distribution in active TUs (Ocwieja et al., 2011).

HIV-1 integration and the host cell

Although purified recombinant IN is necessary and sufficient to carry out 3'-end processing and strand transfer reactions *in vitro*, numerous studies have demonstrated that the establishment of the integrated provirus in the infected cell also involves a variety of host proteins that can play a role during the different steps of the integration process, including nuclear import, IN catalysis, integration site selection and DNA gap repair.

Cellular cofactors of HIV-1 integration

INI1

Integrase Interactor 1 (INI1) is part of the mammalian SWI/SNF complex, which is implicated in ATP-dependent chromatin remodeling during transcriptional activation (Wang et al., 1996); it was first discovered as a binding partner for HIV-1 IN using the yeast two hybrid system (Kalpana et al., 1994). The interaction between the two proteins is HIV-1 specific (Yung et al., 2004), and INI1 was shown to be incorporated into virions (Yung et al., 2001) and to be a component of the reverse transcription and pre-integration complexes in the early steps of the viral replication cycle (Iordanskiy et al., 2006; Turelli et al., 2001). Furthermore, INI1 was found to stimulate the DNA strand transfer catalytic activity of IN *in vitro* (Kalpana et al., 1994).

There have been conflicting results regarding the role played by INI1 during HIV-1 infection and its necessity. One study found no effects on viral integration in cells depleted for INI1 (Boese et al., 2004), whereas other reports suggested that INI1 might participate in an anti-viral cellular response by interfering with the early steps of HIV-1 replication (Maroun et al., 2006; Turelli et al., 2001). Contrastingly, another work revealed that INI1 indirectly enhances proviral transcription (Ariumi et al., 2006), thus providing evidence of an infectivity-promoting effect of the protein. An ectopically expressed dominant negative mutant of INI1, termed S6, containing the minimal IN-interaction domain (amino acids 183-294), was found to potently inhibit HIV-1 assembly and particle production (Yung et al., 2001). S6 also proved

to be an efficient protective agent against HIV-1 infection when stably expressed before viral challenge. The inhibitory effects of S6 required the interaction with IN in the context of the Gag-Pol polyprotein (Yung et al., 2001). INI1 was also shown to be essential for the production of infectious HIV-1 virions (Sorin et al., 2006; Yung et al., 2001). In fact, viral particles produced in the absence of INI1 exhibited a drastic reduction in viral DNA synthesis, suggesting that the protein may be critical for the early events in reverse transcription (Sorin et al., 2006).

In conclusion, at present, INI1 appears to play dual roles in HIV-1 replication in that, while INI1 present in the producer cells seems to be required for HIV-1 replication, INI1 present in the target cells may be inhibitory. One possibility is that cellular INI1 may play a different role than the virally incorporated INI1, explaining the observed effects on HIV-1 replication.

BAF, Lap2α and Emerin

Barrier-to-Autointegration Factor (BAF) is a cellular component of HIV-1 PICs, which blocks suicidal autointegration and stimulates intermolecular integration activity *in vitro* (Chen and Engelman, 1998) (Lin and Engelman, 2003) (Lee and Craigie, 1998). This 89-aminoacid protein is highly conserved among multicellular eukaryotes and interacts with LEM domain proteins (Lin et al., 2000), which are components of the nuclear lamina (Foisner, 2001). It has been suggested that BAF mediates the association of the LEM proteins Lamina-associated polypeptide 2 α (LAP2α) and emerin to HIV-1 PICS, and that these two proteins, in their turn, target the PICs to chromatin, thus promoting integration (Jacque and Stevenson, 2006). Depletion of emerin from human cells impaired infection by HIV-1 expressing wild-type Env or the G glycoprotein from vesicular stomatitis virus (VSV-G), whereas depletion of LAP2α reduced infection by HIV-1 harboring wild-type Env, but not by HIV-1 pseudotyped with VSV-G (Jacque and Stevenson, 2006). Emerin, in particular, appeared to be critical for HIV-1 infection of primary macrophages. However, the requirement of LAP2α and emerin for HIV-1 replication has been questioned. In fact, subsequent reports demonstrated that cells from emerin knockout, LAP2α knockout, or emerin and LAP2α double knockout mice supported the same level of HIV-1 infectivity as cells from wild-type littermate mice, indicating

that these proteins do not function as universally important regulators of the PIC trafficking to chromatin for integration (Shun et al., 2007a) (Mulky et al., 2008). Thus, the roles played by emerin and LAP2α in HIV-1 pre-integration trafficking and integration still await further clarification.

HMGA1

High mobility group chromosomal protein A1 (HMGA1), a non-histone DNA binding protein that can modulate transcriptional regulation and chromatin structure (Farnet and Bushman, 1997), was proposed as a cofactor of HIV-1 integration due to its ability to restore the *in vitro* integration activity of salt-stripped PICs (Farnet and Bushman, 1997). Mechanistic studies pointed to a possible role of the DNA binding activity of HMGA1 in approximating the two LTR ends of the viral DNA molecule and facilitating IN binding by unwinding the LTR termini (Hindmarsh et al., 1999; Li et al., 2000). However, in coimmunoprecipitation experiments, no interaction between HMGA1 and HIV-1 IN could be detected (Hindmarsh et al., 1999). Consistently, chicken cells lacking HMGA1 are not deficient for retroviral integration, suggesting that HMGA1 is not absolutely required for integration (Beitzel and Bushman, 2003). In contrast, evidence was provided pointing to an important role of HMGA1 during transcription (Henderson et al., 2000; Henderson et al., 2004) and splice site regulation of HIV-1 (Tsuruno et al., 2011).

LEDGF/p75

Lens epithelium-derived growth factor (LEDGF/p75) was first identified as an interaction partner of HIV-1 IN by coimmunoprecipitation experiments in human cells expressing IN from a synthetic gene (Cherepanov et al., 2003). Independent screens exploiting the same method (Turlure et al., 2004) and subsequent use of the yeast two-hybrid assay (Emiliani et al., 2005) confirmed this result.

LEDGF/p75 is a ubiquitously expressed chromatin-associated protein that is thought to normally function in transcriptional regulation, stress response and apoptosis pathways (Ganapathy et al., 2003).

LEDGF/p75 is a multidomain protein (**Figure 6**), which not only comprises an IN binding domain (IBD) situated C-terminally (Cherepanov et al., 2004; Vanegas et al., 2005), but also several motifs involved in DNA and chromatin association

(Llano et al., 2006b). These include a NLS, two AT-hook elements (so named for their preferential binding to AT-rich DNA), and a PWWP motif that mediates chromatin binding (Llano et al., 2006b; Turlure et al., 2006).

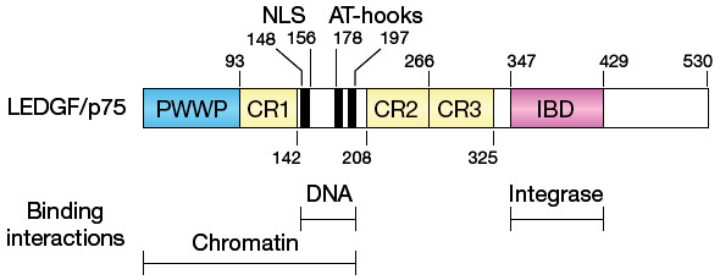

Figure 6. LEDGF/p75 domain organization. (Vandegraaff and Engelman, 2007).

Upon ectopic expression, HIV-1 IN localizes to the nucleus and tightly binds to mitotic chromosomes (Cherepanov et al., 2000; Devroe et al., 2003), properties that can be attributed to its interaction with endogenous LEDGF/p75. In fact, RNAi-mediated knockdown of LEDGF/p75 results in the disruption of IN nuclear localization, redistributing the vast majority of the protein to the cytoplasm (Maertens et al., 2003) (Llano et al., 2004b). Moreover, a mutant IN form, with reduced binding affinity to LEDGF/p75 *in vitro*, fails to interact with chromosomes (Emiliani et al., 2005). Based on these results, a possible role for LEDGF/p75 in chromosomal tethering of HIV-1 PICs has been proposed.

Initial studies on the effect of LEDGF/p75 depletion on HIV-1 infectivity produced conflicting and inconsistent results (Llano et al., 2004b) (Vandegraaff et al., 2006) (Vandekerckhove et al., 2006). However, more complete knockdowns, which were able to eradicate detectable LEDGF/p75 from chromatin, or the use of mouse embryo fibroblasts (MEFs) derived from LEDGF/p75 knockout mice, revealed a significant loss in HIV-1 infectivity in the absence of this cellular factor (Llano et al., 2006a; Shun et al., 2007b). Consistently, overexpression of LEDGF/p75 IBD, which competes with the endogenous protein for the interaction with IN but is unable to interact with chromatin, imposes a severe defect in HIV-1 infectivity (De Rijck et al., 2006). The defect imposed by both LEDGF/p75 depletion and LEDGF/p75 IBD overexpression specifically occurs at the step of integration (Llano et al., 2006a)

(De Rijck et al., 2006). Taken together, these results highlighted an essential role for the interaction between IN and LEDGF/p75 during the HIV-1 replication cycle, suggesting that this interaction might function by docking the PICs to chromatin.

Given its proposed role in the tethering of HIV-1 PICs to chromatin, the hypothesis was made that LEDGF/p75 could potentially function in integration target site selection. Indeed, analysis of the integration site distribution in human cells depleted of LEDGF/p75 by RNAi or in MEFs derived from LEDGF/p75 knockout mice revealed that the residual integration, which can still be found in these cells, is reduced in TUs (Ciuffi et al., 2005; Marshall et al., 2007; Shun et al., 2007b). Furthermore, analysis of the GC content of HIV-1 integration sites in LEDGF/p75-expressing cells indicated that AT-rich regions are favored for integration than GC-rich regions, consistently with the presence of an AT-hook motif in LEDGF/p75. Thus, LEDGF/p75 binding to AT-rich sites appears to direct integration into these sequences. Another approach to investigate the role of LEDGF/p75 in HIV-1 integration site selection examined the frequency of integration events in LEDGF/p75-modulated genes. Since LEDGF/p75 modulates gene expression by binding to DNA, LEDGF/p75-regulated genes should have more bound LEDGF/p75 than randomly selected genes, and thus be preferential integration targets. As expected, LEDGF/p75-modulated genes proved to be favored targets in cells expressing LEDGF/p75, while this preference was abolished in LEDGF/p75 depleted cells (Ciuffi et al., 2005). These data collectively suggested a role for LEDGF/p75 in promoting HIV-1 DNA integration in TUs and AT-rich regions, thus providing the first example of an apparent tethering protein for retroviral integration. Recently, the tethering model for Ledgf/p75 function was strengthened by the finding that expression of fusion proteins containing the IBD of Ledgf/p75 fused to alternative chromatin binding domains in LEDGF/p75-depleted cells redirects viral DNA integration to sites in the genome that reflect the chromatin binding patterns of the chimeric protein (Ferris et al., 2010; Gijsbers et al., 2010; Silvers et al., 2010).

The demonstrated importance of the association between IN and LEDGF/p75 in HIV-1 integration raised the possibility that this interaction could be exploited as an antiviral target. The proof-of-concept of the feasibility of this approach was provided

by the overexpression of the LEDGF/p75 IBD in human cells (De Rijck et al., 2006). In fact, this LEDGF/p75 fragment lacking the chromatin-binding domain proved able to efficiently compete with the endogenous cofactor, thus inhibiting HIV-1 replication and integration to nearly undetectable levels (De Rijck et al., 2006). Moreover, resistance to the LEDGF/p75 IBD fragment arose during virus passaging; this resistance was conferred by mutations in IN mapping to the IN/IBD binding interface. Notably, replication of the IBD-resistant virus was even more sensitive to LEDGF/p75 depletion than was that of the wild-type virus, indicating that the resistant mutant did not replicate in a LEDGF/p75-independent manner (Hombrouck et al., 2007). The development of small molecules that specifically disrupt the interaction between IN and LEDGF/p75 to block HIV-1 replication will be discussed in the following section.

TRN-SR2

The cellular protein transportin-SR2 (TRN-SR2, TNPO3, transportin 3) was recently identified as a cofactor of HIV-1 replication, which is required for nuclear import of the PIC both in cycling cell lines and in macrophages (Brass et al., 2008; Christ et al., 2008; Konig et al., 2008). TRN-SR2 is an importin-β-like karyopherin that plays the role of transporter of serine/arginine-rich splicing factors (SR proteins) into the nucleus (Kataoka et al., 1999). TRN-SR2 was also identified as a binding partner of HIV-1 IN by yeast two-hybrid screenings using IN as a bait, and rebound screenings using a HIV-1 library confirmed that TRN-SR2 interacts solely with IN and none of the others viral proteins (Christ et al., 2008; Rain et al., 2009). It was therefore proposed that TRN-SR2 mediates nuclear import of the HIV-1 PIC through the interaction with IN. However, whereas a direct interaction between HIV-1 IN and TRN-SR2 was confirmed by different studies (Christ et al., 2008; Krishnan et al.; Rain et al., 2009), its role during the PIC nuclear import is currently under debate. Given that MLV replication is independent of TRN-SR2 expression (Christ et al., 2008), the finding that a chimeric HIV/MLV clone carrying MLV CA is insensitive to TRN-SR2 knockdown (Krishnan et al., 2010), together with the identification of a HIV-1 clone bearing the N74D mutation in CA whose replication is not affected by TRN-SR2 depletion (Zhou et al., 2011), led to the hypothesis that the TRN-SR2 dependency of HIV-1 nuclear import is mediated by CA rather than

IN, although no evidence was provided that TRN-SR2 and CA physically interact. A subsequent report provided a potential explanation for the phenotype of the N74D CA mutant HIV-1 clone, showing that the mutant virus was insensitive to TRN-SR2 depletion only when pseudotyped with VSV-G, while the same mutant clone proved to be still dependent on TRN-SR2, although to a somewhat lesser extent, when retaining the HIV-1 envelope (Thys et al., 2011). It was therefore concluded that, while HIV-1 CA appears to modulate the nuclear entry of VSV-G pseudotyped viruses, its role in the nuclear import of viruses which retain their natural envelopes is less clear: CA may exert an indirect effect on the process of nuclear import, probably through its involvement in the steps of uncoating, trafficking and docking of the PICs to the nuclear pore, which all precede the interaction between IN and TRN-SR2 (Thys et al., 2011). However, a recent study aimed at assessing the role of the interaction between IN and TRN-SR2 during HIV-1 infection found that IN mutations disrupting the association with TRN-SR2, while impairing HIV-1 replication, do not significantly affect the amounts of 2-LTR circles formed in the nucleus of the target cells (Cribier et al., 2011). According to the authors, the lack of correlation between a defect in TRN-SR2 binding by IN and the formation of 2-LTR circles in the nucleus strongly supports the notion that the interaction between the two proteins is not directly involved in the nuclear import step of the viral replication cycle (Cribier et al., 2011). This latter study also raises the hypothesis that a direct interaction between IN and TRN-SR2 may take place after the entry of the PIC in the nucleus and proposes a model in which, beside its role in nuclear import, TRN-SR2 may additionally be involved in the nuclear events of HIV-1 replication. Indeed, another report recently proposed a role for TRN-SR2, together with NUP358/RanBP2, in regulating the passage of the PIC through the nuclear pore, so as to place the complex in gene-dense regions of the genome (Ocwieja et al., 2011), a process that may require a direct interaction between IN and TRN-SR2. Therefore, at present, both the mechanism of TRN-SR2-mediated nuclear import of HIV-1 PICs and the role played by the interaction between IN and TRN-SR2 during the viral replication cycle still await further clarification.

p300

p300, a HAT enzyme acting as a transcriptional coactivator (Fukuda et al., 2006),

was found to interact with HIV-1 IN both *in vitro* and *in vivo*, resulting in the acetylation of three lysine residues, located at positions 264, 266 and 273 in the CTD of the viral protein (Cereseto et al., 2005; Topper et al., 2007). Interestingly, a mutant IN carrying arginine substitutions at these residues also proved to be acetylated *in vivo*, thus suggesting that cellular HATs other than p300 might acetylate IN at different sites. Acetylation by p300 was shown to increase IN binding affinity to DNA and enhance its strand transfer activity *in vitro* (Cereseto et al., 2005). The role of IN acetylation at lysines 264, 266 and 273 during HIV-1 replication has been debated. One study first reported a severe replication impairment for a HIV-1 clone expressing the triple mutant Flag-tagged IN (Flag-IN K264,266,273R), due to a specific block at the step of integration (Cereseto et al., 2005). However, the results presented in a subsequent report, while confirming the acetylation of IN CTD by p300 *in vitro*, found no effect on viral replication when arginine substitutions were intoduced at the p300-targeted lysine positions of IN (Topper et al., 2007). The discrepancy was explained by noting that the earlier *in vivo* experiments were performed with a HIV-1 clone encoding a Flag-tag epitope fused downstream and in frame with the C-terminus of IN. Subsequently, in the context of a study aimed at developing HIV-based non-integrating lentiviral vectors (NILVs) as gene therapy tools, the integration efficiency of a GFP-expressing vector harboring mutations at IN lysines acetylated by p300 was determined (Apolonia et al., 2007). The IN K264,266,273R mutant vector exhibited a five-fold lower GFP expression level than the vector encoding wild type IN, which was paralleled by a fourteen-fold lower residual integration rate. The overall importance of IN CTD acetylation by p300 to HIV-1 integration and replication therefore awaits further clarification. For sure, given the observed *in vivo* acetylation of the IN triple mutant, an interesting issue for the future would be to determine whether additional IN lysines are targets for acetylation by HATs different from p300 and whether IN is subject to other post-translational modifications by as-yet-unidentified cellular cofactors.

Post-translational modifications of IN

HIV-1 IN is subject to several post-translational modifications that contribute to

modulate its activities.

The intracellular stability of IN is regulated through polyubiquitination and subsequent proteasome-dependent degradation (Devroe et al., 2003; Emiliani et al., 2005; Llano et al., 2004a; Mulder and Muesing, 2000; Tasaki et al., 2005). Notably, the IN cellular binding partners LEDGF/p75 (Llano et al., 2004a), hRad18 (Mulder et al., 2002) and Ku70 (Zheng et al., 2011) have all been reported to protect the viral protein from the host ubiquitin-proteasome system during HIV-1 replication.

Modulation of the efficiency of integration can be achieved through the reversible acetylation of IN, which has been found to enhance its strand transfer activity *in vitro* (Cereseto et al., 2005; Terreni et al.). Consistently, arginine substitutions introduced at the IN lysines targeted by acetylation have been shown to negatively affect the infectivity of the mutant viruses by specifically reducing their integration level (Cereseto et al., 2005; Terreni et al.).

SUMOylation of IN may also contribute to regulate the efficiency of HIV-1 integration. In fact, it has been recently reported that IN is SUMOylated and the functional relevance of this post-translational modification has been highlighted by the finding that mutant viruses harboring a SUMOylation-defective IN are less infectious than their wild-type counterparts (Zamborlini et al., 2011). However, differently from acetylation, SUMOylation of IN does not seem to affect the catalytic activity of the protein, since SUMOylation-defective IN mutants still retain wild-type levels of enzymatic activity. Therefore, the possibility is currently being explored that SUMOylation may regulate the binding affinity of IN to cofactors required for the efficient completion of the viral replication cycle.

Finally, IN phosphorylation proved to be involved in the regulation of the permissiveness of resting CD4+ T cells to HIV-1 infection (Manganaro et al., 2010). In activated T lymphocytes, IN is phosphorylated by c-Jun N-terminal kinase (JNK), and then becomes a substrate for the cellular peptidyl prolyl-isomerase enzyme Pin1, leading to a conformational change in the viral protein (Manganaro et al., 2010). These concerted enzymatic activities lead to increased IN stability and are required for efficient HIV-1 integration and infection (Manganaro et al., 2010). Since JNK is not expressed in resting CD4+ T cells, lack of the above-mentioned

modifications likely contributes to the restriction of HIV-1 infection observed in these lymphocytes.

Restriction of infection at the integration step: KAP1

A recent study extended the family of host factors able to restrict HIV-1 infection through the identification of a cellular protein belonging to the TRIM family, KAP1, which interferes with the integration step of the viral replication cycle (Allouch and Cereseto, 2011). KAP1 was identified in a yeast two-hybrid screening of host proteins that specifically interact with the acetylated form of HIV-1 IN (Allouch and Cereseto, 2011). Indeed, further characterization of the association between the two proteins showed that KAP1 binds with higher affinity to IN upon its acetylation, both *in vitro* and in cultured cells (Allouch et al., 2011). Investigation on the role of KAP1 during the HIV-1 replication cycle led to the conclusion that KAP1 inhibits viral infectivity by specifically interfering with the integration process, as evidenced by the finding that KAP1 downregulation causes an increase in level of integration, while KAP1 overexpression results in a reduction of provirus formation (Allouch et al., 2011). The antiviral mechanism proposed for KAP1 consists in the recruitment of the histone deacetylase HDAC1 to acetylated IN, thus leading to deacetylation of the viral protein and consequent reduction of the integration efficiency. Consistently, the activity of KAP1 against HIV-1 proved to be fully dependent on deacetylation, as it was found to be ineffective in the absence of HDAC1, even if KAP1 was overexpressed (Allouch et al., 2011). However, since KAP1 is ubiquitously expressed showing no major restriction of HIV-1 infectivity, further studies are required to elucidate the strategy adopted by the virus to counteract this newly discovered host defense mechanism.

IN inhibitors to treat HIV-1 infection

The development of RT and PR inhibitors and the subsequent introduction of combination drug regimens enhancing the overall efficacy and durability of antiretroviral therapy revolutionized the treatment of HIV-1 infection in the mid 1990's. As the last of the three essential HIV-1 enzymes, IN was considered an equally attractive target for antiretroviral drug development as RT and PR, but it is only a decade later that the first IN inhibitor, raltegravir (RAL, MK-0518) achieved

approval by the US Food and Drug Administration (FDA) (reviewed in (Cahn and Sued, 2007)), while elvitegravir (EVG, GS-9137, JTK303) has reached phase III clinical trials. Both RAL and EVG are derived from the diketo acid (DKA) family of IN inhibitors. DKAs were the first IN inhibitors reported with selectivity for the strand transfer step of the integration reaction (ST-inhibitors), high specificity for IN, and antiviral activity that could be related to IN inhibition (Hazuda et al., 2000). Their selective effect on the strand transfer reaction is a direct result of a mechanism of action in which the inhibitor: 1) only binds to the IN catalytic site in the presence of a viral DNA substrate and not to IN alone (Espeseth et al., 2000) and 2) chelates the two essential magnesium metal ion cofactors from the IN active site, thus rendering the metal-dependent phosphotransferase responsible for strand transfer inactive (Grobler et al., 2002; Marchand et al., 2003; Pais et al., 2002).

Despite its recent success in the clinic, emergence of resistance leading to treatment failure has already been reported for RAL. All the major mutations responsible for decreased susceptibility to RAL appear to localize within the IN active site proximal to the residues involved in coordinating the metal cofactors (Hazuda, 2010), consistent with the mechanism of inhibition based on metal chelation. Distinct subsets of IN mutations have been characterized in the viral sequences of patients who have failed the therapy with RAL. The “primary mutations” conferring high-level resistance to RAL have been identified as Q148R/H/K, N155H, and Y143R/C (Delelis et al., 2009; Kobayashi et al., 2008; Malet et al., 2009; Malet et al., 2008; Marinello et al., 2008; Shimura et al., 2008; Sichtig et al., 2009). Each tends to be associated with “secondary mutations” that appear to restore viral fitness in those primary mutants: for instance, the secondary mutation E92Q invariably clusters with the N155H mutation, while the G140S is mainly found in the Q148H mutants (Cooper et al., 2008; Delelis et al., 2009; Malet et al., 2009). The resistance profile of EVG closely resembles that of RAL, suggesting that EVG is unlikely to overcome resistance to RAL (Marinello et al., 2008). Current research efforts are therefore devoted to the discovery and development of novel compounds effective against viral strains resistant to IN ST-inhibitors. A promising approach that is currently being explored is that of targeting for antiviral therapy the interactions between IN and cellular cofactors, if

demonstrated to be essential for HIV-1 infection. In particular, the LEDGF/p75-IN interaction proved to be an attractive therapeutic target, since the crucial role of this cellular cofactor in HIV-1 replication was evidenced via mutagenesis, RNAi, transdominant overexpression of LEDGF/p75 IBD and knockout studies (Busschots et al., 2007; Cherepanov et al., 2003; De Rijck et al., 2006; Emiliani et al., 2005; Hombrouck et al., 2007; Llano et al., 2006a; Shun et al., 2007b; Vandekerckhove et al., 2006). Determination of the co-crystal structure of the LEDGF/p75 IBD bound to the IN CCD (Cherepanov et al., 2005) allowed structure-based drug design of a series of 2-(quinolin-3-yl)acetic acid derivatives that act as potent inhibitors of the LEDGF/p75-IN interaction and are known under the name of LEDGINs (which stands for LEDGF/p75-integrase inhibitors). LEDGINs were shown to inhibit HIV-1 replication by specifically blocking the integration step both in cultured cell lines and in peripheral blood lymphocytes from human donors. Most importantly, LEDGINs proved to remain fully active against viral strains resistant to IN ST-inhibitors, clearly demonstrating their different mode of action (Christ et al., 2010; Debyser and Christ, 2010). Based on the co-crystal structures determined for multiple LEDGINs in complex with the IN CCD, the initial compounds were optimized, leading to the development of molecules that exhibit potencies in the low nanomolar range and reduced toxicity in cell culture. Of note, these optimized LEDGIN compounds show a selectivity index, which is a measure for lack of cellular toxicity but potent antiviral activity, comparable to that of antiviral drugs currently used in the clinic (Debyser and Christ, 2010). Therefore, since LEDGINs are effective IN inhibitors that do not show any cross-resistance with ST-inhibitors, they have significant clinical potential as second-generation IN inhibitors.

Finally, another approach to inhibit HIV-1 IN function that is currently the object of intense research is the development of small molecules that can interfere with IN catalytic activity through allosteric modulation of the dynamic interaction existing between its monomeric, dimeric and tetrameric forms. The development of an assay for the screening of modulators of HIV-1 IN dimerization has recently allowed the identification of compounds that act as allosteric IN inhibitors (Demeulemeester et al., 2012). As a consequence of the strict correlation between IN catalytic activity and its specific oligomeric forms, these newly identified

modulators of the IN oligomerization equilibrium represent the first step towards an as-yet-unexplored promising antiviral strategy.

Acetylation

Post-translational modification (PTM) of proteins extends the range of possible molecular structures beyond the limits imposed by the twenty encoded amino acids and, if reversible, gives a means of control and signaling. Lysine acetylation consists in the transfer of an acetyl group from acetyl-coenzyme A (acetyl-CoA) to the ε-amino group of a specific lysine residue in the target protein. Lysine acetylation was first discovered with histones in 1968 (Gershey et al., 1968), but the responsible enzymes, histone acetyltransferases (HATs) and deacetylases (HDACs), were not identified until the mid-1990s (Bannister and Kouzarides, 1996; Brownell and Allis, 1995; Brownell et al., 1996; Kleff et al., 1995; Mizzen et al., 1996; Ogryzko et al., 1996; Parthun et al., 1996; Rundlett et al., 1996; Taunton et al., 1996; Yang et al., 1996a; Yang et al., 1996b). In the past decades, knowledge about this PTM has exploded, with targets rapidly expanding from histones to a wide variety of proteins of both cellular and viral origin (Caron et al., 2003; Cohen and Yao, 2004; Kouzarides, 2000; Sterner and Berger, 2000; Yang and Gregoire, 2007).

The diversity of HATs

According to sequence similarities, known HATs can be organized into different groups. Of the reported HAT families, the GNAT (GCN5-related N-acetyltransferase), MYST (MOZ, Ybf2, Sas2 and Tip60) and p300/CBP (Creb-binding protein) families are the best characterized, both structurally and functionally (Lee and Workman, 2007; Roth et al., 2001; Sterner and Berger, 2000).

GNAT family members include the highly homologous HATs GCN5 (general control nonderepressible 5) and PCAF (p300/CBP associated factor), as well as other more distantly related enzymes. GCN5 and PCAF appear to primarily function as histone-acetylating transcriptional coactivators (Carrozza et al., 2003; Nakatani, 2001; Roth et al., 2001), but also catalyze the acetylation of non-histone substrates, leading to changes in their activities (Glozak et al., 2005; Kouzarides,

2000; Polevoda and Sherman, 2002; Sterner and Berger, 2000).

p300 and CBP form a pair of paralogous transcriptional coactivators possessing intrinsic HAT activity (Bannister and Kouzarides, 1996; Ogryzko et al., 1996). Since these two proteins seem to be largely interchangeable in function, at least in cultured cells (Arany et al., 1995; Eckner et al., 1994; Lundblad et al., 1995; Shiama, 1997), they are commonly referred to as p300/CBP. Like PCAF and GCN5, p300/CBP is also known to acetylate and regulate various protein substrates other than histones (Glozak et al., 2005; Kouzarides, 2000; Polevoda and Sherman, 2002; Sterner and Berger, 2000).

The MYST family of proteins constitutes a third major group of HATs. The acronym MYST derives from its four founding members: human MOZ (monocytic leukemia zinc finger protein) (Borrow et al., 1996), yeast Ybf2/Sas3 (Reifsnyder et al., 1996; Takechi and Nakayama, 1999), yeast Sas2 (Reifsnyder et al., 1996) and mammalian Tip60 (HIV-1 Tat interacting 60 kDa protein) (Kamine et al., 1996; Ran and Pereira-Smith, 2000; Yamamoto and Horikoshi, 1997). As compared with the GCN5/PCAF and p300/CBP groups, the MYST family is larger and more heterogeneous in domain organization and biological function (Yang, 2004b).

Most HATs have been found within large multiprotein complexes *in vivo*. Importantly, these enzymes, upon incorporation into such complexes, often change their substrate specificities (Boudreault et al., 2003; Grant et al., 1997; Ogryzko et al., 1998; Sendra et al., 2000), suggesting that associated components regulate the activities of the respective catalytic subunits. Amazingly, one HAT can be the catalytic subunit of multiple complexes. For instance, metazoan GCN5 participates in the constitution of at least two distinct types of multiprotein complexes (Nagy and Tora, 2007), as will be furtherly discussed in the “Metazoan GCN5-containing complexes” section.

HDACs

Fortuitously, the same year that the first HATs were identified, the first HDAC, catalyzing the removal of acetyl groups from the N-terminal tails of nucleosomal histones, was also discovered (Taunton et al., 1996). Remarkably, this mammalian HDAC was homologous to the yeast transcriptional corepressor Rpd3p. In the

subsequent years, more corepressors were recognized to be HDACs, just as more coactivators were identified as HATs, providing strong support for the notion that hyperacetylated chromatin is transcriptionally active, while hypoacetylated chromatin is repressed.

Similarly to HATs, HDACs also occur in multiprotein complexes, which are recruited by DNA-binding factors to direct local transcriptional repression through histone deacetylation. In addition to histones, HDACs also deacetylate a variety of different substrates, including transcription factors and other cellular proteins, involved in the control of cell growth, differentiation and apoptosis (Glozak and Seto, 2007).

Aberrant activity of HDACs is well known to be associated with oncogenesis in humans, since it can lead to the suppression of the expression of various tumour suppressor genes (Cress and Seto, 2000; Timmermann et al., 2001), and numerous clinical trials involving HDAC inhibitors (HDACi) have been successfully conducted for the treatment of both haematological and solid tumours (Stimson et al., 2009).

Histone acetylation in the regulation of gene expression

Histones are extensively modified by a myriad of reactions; examples include acetylation, methylation, ubiquitination, SUMOylation, phosphorylation, and ADP-ribosylation. These modifications alter chromatin dynamics by influencing histone-DNA interactions, as well as the recruitment and binding of protein factors to chromatin (reviewed in (Taverna et al., 2007)). The tight correlation between histone modification patterns and their subsequent regulation of chromatin activity led to the hypothesis that PTM of histones encodes a language, the so-called “histone code”, that is read by chromatin-interacting proteins to bring about distinct downstream events crucial for both short- and long-term regulation of transcriptional activity (Strahl and Allis, 2000). These chromatin-binding proteins interact with chromatin through several distinct domains, each of which has the ability to recognize different histone modifications and can be responsible for the targeting of multisubunit protein complexes to specific chromatin regions.

Histone acetylation is a PTM that performs two known functions on chromatin. First, acetylation neutralizes the positive charge of lysine residues within the histone tails, thus reducing their ability to form electrostatic interactions with the negatively charged phosphodiester backbone of associated DNA. The acetylation of histones is therefore associated with chromatin decondensation, which results in an increased accessibility of transcription factors to their target sequences on DNA (Lee et al., 1993; Nightingale et al., 1998; Vettese-Dadey et al., 1996). The second known function of histone acetylation is to provide a platform for the recruitment of transcriptional regulatory proteins through acetyl-lysine interacting motifs, known as bromodomains (Dhalluin et al., 1999) (Jacobson et al., 2000).

Bromodomain structure and function

The bromodomain, found in HATs and chromatin-associated proteins, functions as the sole protein module known to bind acetyl-lysine motifs. Consistent with the reversible nature of lysine acetylation, the bromodomain/acetyl-lysine interaction is dynamic and likely transient in the cells. This notion is supported by the three-dimensional structure of the bromodomain, as first revealed with that of the human PCAF bromodomain (Dhalluin et al., 1999). Since then, a number of three-dimensional structures of bromodomains, either in the free form or bound to an acetyl-lysine containing peptide have been reported in literature, including those from human and yeast GCN5 (Hudson et al., 2000; Owen et al., 2000), human CBP (Mujtaba et al., 2004) and human SWI/SNF chromatin remodeling complex protein Brg1 (Brahma-related gene 1) (Shen et al., 2007). These structures, determined either in solution or crystalline state, confirmed the early prediction from the first bromodomain structure (Dhalluin et al., 1999), which is that most, if not all, bromodomains adopt a distinct structural fold, termed the “BRD fold”, consisting of a left-handed bundle of four helices (α_Z, α_A, α_B, α_C), with the inter-helical ZA and BC loops of variable length and sequence, and that the latter loops constitute a hydrophobic pocket serving for the interaction with the acetyl-lysine (Figure 7A). The structurally flexible ZA loop likely determines the dynamic nature of acetyl-lysine binding. The modular nature of the BRD fold enables the bromodomain to act as a functional unit within a protein, either individually or in combination with

other modules.

Despite the structurally conserved BRD fold, the overall sequence similarity of the large bromodomain family is not high, with major sequence variations in the ZA and BC loops (Jeanmougin et al., 1997). Nevertheless, the amino acids involved in acetyl-lysine recognition are among the most highly conserved residues in the bromodomain family and correspond to Tyr-1125, Tyr-1167 and Asn-1168 in CBP (**Figure 7**B) (Mujtaba et al., 2004; Zeng et al., 2008). The crystal structure of the yeast GCN5 bromodomain bound to a histone H4 peptide containing acetylated Lys-16 showed that, in addition to binding to the conserved Tyr-364 and Tyr-406 residues (corresponding to Tyr-1125 and Tyr-1167 in CBP, respectively), the acetyl-lysine residue forms a specific hydrogen-bond between the oxygen of the acetyl carbonyl group and the side chain amide nitrogen of the conserved asparagine residue, Asn-407 (corresponding to Asn-1168 in CBP) (Owen et al., 2000). A network of water-mediated hydrogen bonds involving carbonyl groups from the protein backbone at the base of the binding pocket also contributes to acetyl-lysine binding. The critical role of these three conserved amino acid residues in acetyl-lysine recognition was confirmed by mutagenesis studies (Dhalluin et al., 1999; Mujtaba et al., 2002; Mujtaba et al., 2004). The crystal structure of the yeast GCN5 bromodomain complexed with the histone H4-derived acetylated peptide also showed that the specificity for the sequence surrounding the acetyl-lysine is conferred by the binding of the peptide side chains at (K_{ac} +2) and (K_{ac} +3) positions, thus suggesting that the yeast GCN5 bromodomain may discriminate between different acetylated lysines, depending on the sequence context in which they are inserted.

Roles for bromodomains in at least four functions have been recognized. First, they are important for stable promoter occupancy by chromatin-modifying complexes in the absence of the transcriptional activators, which initially recruited them. This is exemplified by the observation that anchoring of the yeast SAGA coactivator complex on acetylated chromatin for persistent acetylation requires the bromodomain of its catalytic subunit, GCN5 (Hassan et al., 2002).

Second, bromodomains contribute to acetylation-dependent chromatin assembly and remodeling. This is evidenced by the finding that the bromodomain of

Swi2/Snf2, the ATPase subunit of the yeast chromatin-remodeling complex SWI/SNF, is required for both stable occupancy of acetylated promoter nucleosomes by SWI/SNF (Hassan et al., 2002) and functional activity of the complex on SAGA-acetylated nucleosomes (Hassan et al., 2006).

Third, bromodomains are involved in organizing chromosome or chromatin domains. Compared to nucleosome assembly and remodeling, this is a large-scale and long-range effect. For instance, Bdf1 (bromodomain factor 1) is a yeast protein containing two bromodomains, which, by binding to acetylated histone H4, is able to impose a physical barrier between euchromatin and heterochromatin, thus preventing heterochromatic spreading (Ladurner et al., 2003; Matangkasombut and Buratowski, 2003). Similarly to Bdf1, several mammalian bromodomain containing (Brd) proteins have been reported to exhibit acetylation-dependent binding to chromatin (Dey et al., 2003; Kanno et al., 2004; Peng et al., 2006; Pivot-Pajot et al., 2003), suggesting that proteins belonging to the Brd family may contribute to the translation of the so-called histone code (Crowley et al., 2002; Dey et al., 2003; Dey et al., 2000; Mattsson et al., 2002).

Fourth, bromodomains also recognize acetylated non-histone proteins, of both cellular and viral origin. Thus, bromodomain/acetyl-lysine interaction plays a direct role in mediating the consequential functions arising from acetylation of these proteins. For instance, transcriptional activation of HIV-1 proviral DNA is dependent on a molecular interaction between acetylated Tat and the bromodomain of the cellular coactivator PCAF (Dorr et al., 2002; Mujtaba et al., 2002). Again, human p53 transcriptional activity requires the binding of its C-terminal acetylated lysine 382 to the bromodomain of the coactivator CBP. This bromodomain/acetyl-lysine binding is responsible for p53 acetylation-dependent coactivator recruitment after DNA damage, a step essential for p53-induced transcriptional activation of its target genes, such as p21 (Mujtaba et al., 2004; Sachchidanand et al., 2006).

In conclusion, at present, available data are all consistent with the notion of the bromodomain being a protein module that can specifically process molecular information conveyed by lysine acetylation.

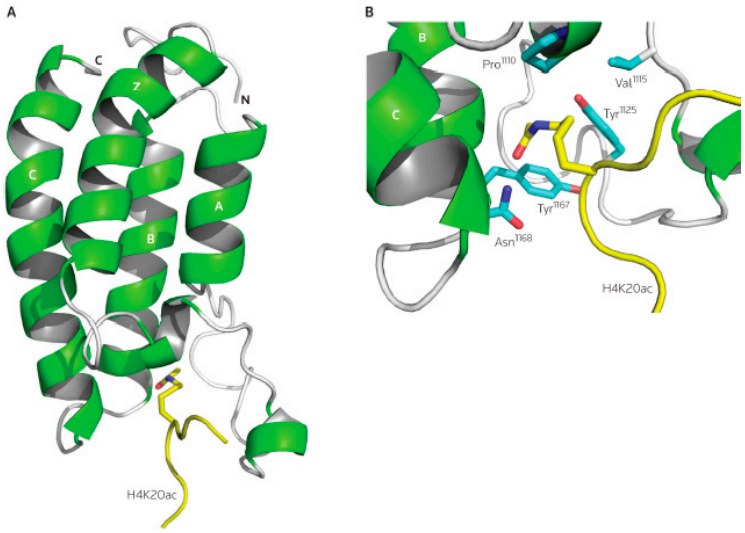

Figure 7. Structural basis of acetyl-lysine recognition by the bromodomain. (A) The 3D structure of the CBP bromodomain bound to a H4-K20ac peptide; and (B) the acetyl-lysine binding site, showing the key interactions between the CBP bromodomain and the H4-K20ac peptide. The peptide is yellow and the side chains of the protein residues are color-coded by atom type. (Sanchez and Zhou, 2009).

Acetylation of non-histone proteins

In addition to histones, HATs can also target non-histone substrates, such as transcription factors and other nuclear proteins, as well as structural components, metabolic enzymes and signalling regulators in the cytoplasm (Yang and Seto, 2007). Moreover, several viruses have been recently found to encode specific proteins that are substrates for cellular HATs. Acetylation has been reported to regulate the activity of target proteins by modulating DNA binding (Gu and Roeder, 1997; Marzio et al., 2000; Yao et al., 2001), interaction with partners (Bannister et al., 2000; Cohen et al., 2004; Kovacs et al., 2005; Yuan et al., 2005), subcellular localization (Blander et al., 2002; Kawaguchi et al., 2006; Kitamura et al., 2005; Thevenet et al., 2004) and stability (Bernassola et al., 2004; Jeong et al., 2002; Li et al., 2002). Protein acetylation is therefore emerging as an important signaling pathway involved in the control of both cellular and viral functions.

HAT autoacetylation

Reminiscent of kinase autophosphorylation, HATs themselves are acetylated. Prominent examples are represented by p300 and CBP, which are heavily

autoacetylated in an intermolecular fashion (Karanam et al., 2006), suggesting that dimerization may trigger autoacetylation. An attractive model proposes that DNA-binding proteins may recruit p300 and bring two or more p300 molecules into physical proximity, leading to intermolecular acetylation and consequent enzymatic activation. Consistent with this hypothesis, many transcription factors are able to stimulate the HAT activity of p300 and CBP, and the recruitment to DNA also seems to be important for p300 activation (Dornan et al., 2003).

Acetylation of trancription factors

A decade ago, HATs were first reported to acetylate the tumour suppressor and transcription factor p53, leading to the notion that HATs and HDACs are not specific to histones (Gu and Roeder, 1997). Since then, several transcription factors have been identified as substrates for HATs, including E2F (involved in the progression through the G1/S checkpoint) (Martinez-Balbas et al., 2000; Marzio et al., 2000), c-Jun (participating in the response to mitogens) (Vries et al., 2001), the erythroid Kruppel-like transcription factor (EKLF) (Zhang and Bieker, 1998; Zhang et al., 2001), the transcriptional coactivator GATA1, which is required for megakaryocyte and erythrocyte differentiation (Boyes et al., 1998), the muscle-specific differentiation regulator MyoD (Polesskaya et al., 2000; Sartorelli et al., 1999) and the oncoprotein c-Myc (Patel et al., 2004). Acetylation of trancription factors can alter their ability to bind DNA. For p53, E2F1, and GATA1, acetylation stimulates their sequence-specific DNA binding activity, consequently resulting in an increased activation of their target genes (Boyes et al., 1998; Gu and Roeder, 1997; Martinez-Balbas et al., 2000). Acetylation can also affect protein-protein interactions, forming docking sites for the recruitment of transcriptional coactivators. For instance, acetylation at lysine 382 increases p53 affinity for the bromodomain of CBP (Mujtaba et al., 2004), while acetylation of lysines 373 and 382 promotes the interaction with the tandem bromodomains of TAF1, a subunit of the basal transcription factor TFIID (Li et al., 2007). Similarly, MyoD acetylation triggers the recognition of the acetylated residues by the bromodomain of CBP, thus strengthening the interaction between the HAT and MyoD itself (Polesskaya et al., 2001). This event finally results in a more efficient recruitment of HAT

complexes to muscle-specific promoters, where they acetylate other substrates such as histones (Polesskaya et al., 2001). In addition, acetylation increases the half-life of p53 by preventing its Mdm2-mediated ubiquitination and subsequent degradation (Li et al., 2002). Since the lysines ubiquitinated by MDM2 are the same residues targeted for acetylation after DNA damage, acetylation of these lysines is believed to prevent ubiquitination at the same sites, leading to p53 stabilization (Li et al., 2002). Finally, acetylation is involved in the control of the nucleocytoplasmic shuttling of various transcription factors. For instance, the nuclear action of NF-κB (a heterodimer of RelA and p50 proteins) is controlled by regulating its interaction with the IκBα inhibitor, which normally binds and sequesters the transcription factor in the cytoplasm. Acetylation of the RelA subunit of NF-κB results in the disruption of the inhibitory interaction with IκBα (Chen et al., 2001). Acetylated RelA is subsequently deacetylated, promoting effective binding to IκBα and leading to IκBα-dependent nuclear export of the complex (Chen et al., 2001). Deacetylation of RelA thus acts as an intranuclear molecular switch that both controls the duration of the NF-κB transcriptional response and contributes to the replenishment of the depleted cytoplasmic pool of inactive NF-κB/IκBα complexes.

Acetylation of HMG proteins

The high mobility group (HMG) proteins are abundant and ubiquitous nuclear proteins that bind to DNA, nucleosomes and other multiprotein complexes in a dynamic and reversible fashion to regulate DNA processing in the context of chromatin. Three distinct families of HMG proteins have been defined, including HMG-AT-hook (HMGA), HMG-box (HMGB) and HMG-nucleosome binding (HMGN) families (Bustin, 2001b).

The ordered acetylation of lysine residues in the HMGA1a protein governs the accurate execution of the *interferon-β* gene transcriptional switch (Munshi et al., 2001). CBP and PCAF are both able to acetylate HMGA1a (Munshi et al., 2001; Munshi et al., 1998). However, the acetylation mediated by these two enzymes has opposite biological outcomes: CBP-acetylated HMGA1a destabilizes the enhanceosome, a higher-order nucleoprotein complex formed in response to viral

infection, while PCAF-induced acetylation potentiates the gene transcription of *interferon-β* by stabilizing the enhanceosome and preventing HMGA1a from CBP-mediated acetylation.

Although the PTMs of HMGB proteins have not been investigated as extensively as those of HMGA1 proteins, growing evidence indicates the remarkable biological significance of HMGB acetylation. Acetylated HMGB1, isolated from cells grown in the presence of the HDACi butyrate, exhibits a significantly enhanced ability to recognize ultraviolet (UV) light- or cisplatin-damaged DNA and four-way junctions (Ugrinova et al., 2001), thus suggesting a critical role for acetylation in structure-specific DNA binding by HMGB1 and its potential function in DNA repair. In addition, although HMGB1 localization is predominantly nuclear, it was shown that this protein can be either passively leaked out of cells during necrosis (Scaffidi et al., 2002), or secreted by monocytes and macrophages (Muller et al., 2001). A related study reported that, in the latter cell types, HMGB1 relocation from the nucleus to the cytoplasm and its subsequent secretion are dependent on the acetylation state of the protein (Bonaldi et al., 2003).

HMGN proteins are the only nuclear proteins that specifically recognize the nucleosomal core particle (Bustin et al., 1990; Sandeen et al., 1980) and are present in the nuclei of all mammalian and most vertebrate cells (Bustin, 1999). The enzymes responsible for acetylation of HMGN proteins are p300 and PCAF (Bergel et al., 2000; Bustin, 2001a; Herrera et al., 1999). The main consequence of both p300- and PCAF-mediated acetylation of HMGN proteins is the reduction of their binding affinity affinity to nucleosomes (Bergel et al., 2000; Herrera et al., 1999).

Acetylation of nuclear import factors

HATs also target nuclear proteins other than transcription factors. A screen of a large set of proteins involved in different cellular processes resulted in the identification of importin α, an adaptor protein involved in nuclear import, as a substrate for p300/CBP-mediated acetylation (Bannister et al., 2000). Importin α is acetylated on a single lysine residue, located at position 22. Acetylation has been found to promote the interaction of importin α with importin β *in vitro* (Bannister et

al., 2000), which is required for the transport of the bound cargo through the nuclear pore complex. Acetylation of importin α by p300 is also critical for the nuclear import of HuR, a shuttling RNA binding protein that plays an important role in regulating mRNA stability. Consistently, mutation of lysine 22 to arginine results in the impairment of HuR nuclear import (Wang et al., 2004). Taken together, these findings highlight protein acetylation as a mechanism involved in the regulation of nuclear import.

Acetylation of non-nuclear proteins: α-tubulin and cortactin acetylation

Microtubules are cylindrical cytoskeletal structures found in almost all eukaryotic cell types and involved in a great variety of cellular processes, including mitosis, ciliary and flagellar motility, intracellular transport of vesicles and organelles, and possibly in determining the morphology of certain cells (Luduena, 1998). The structural subunit of microtubules is the 100 kDa tubulin protein, which exists in the α and β isoforms that form heterodimeric complexes. Several types of PTMs have been found to affect tubulin function, including acetylation, phosphorylation, polyglutamation, polyglycylation, and detyrosination (MacRae, 1997). Acetylation is unique among the known tubulin modifications in that it occurs at lysine 40 of α-tubulin, which is predicted to reside on the luminal face of microtubules. Therefore, it is unclear how the acetylating/deacetylating enzymes would have access to this site. It is also unclear how this luminal modification might influence microtubule-based functions occurring on the cytoplasmic face of the microtubule. The acetylating enzymes have not been identified, but two enzymes have been shown to deacetylate α-tubulin both *in vitro* and *in vivo*, namely HDAC6 and Sirt2 (Hubbert et al., 2002; Matsuyama et al., 2002). Acetylation of α-tubulin at lysine 40 is fairly common and can be detected on stable microtubules in most cell types. Recent work has suggested that this modification plays a positive role in motor-based trafficking in mammals (Bulinski, 2007; Dompierre et al., 2007; Reed et al., 2006). Consistently, the motor protein Kinesin-1 binds with higher affinity to acetylated microtubules *in vitro* (Reed et al., 2006), and hyperacetylation of microtubules in neuronal cells treated with HDAC6 inhibitors causes Kinesin-1-mediated transport

of JIP1 to be redirected from a subset to the majority of neuritis, as well as enhancing the anterograde and retrograde transport of Brain Derived Nerve Factor (BDNF) vesicles (Dompierre et al., 2007; Reed et al., 2006).

Acetylation of cortactin, a cytoplasmic protein involved in the regulation of cell motility, has been recently reported (Zhang et al., 2007). Cortactin undergoes acetylation at up to ten lysine residues, eight of which are located in a region of the protein, known as the repeat region, consisting of six and a half tandem repeats of a unique 37-amino acid sequence (Zhang et al., 2007). Acetylation of cortactin impedes its interaction with F-Actin and inhibits its translocation to the cell periphery, leading to aberrant cell motility. The lysines targeted for acetylation are separated in the primary sequence of the protein, but form two “charged patches” in the repeat region at the three-dimensional level. Available data indicate that acetylation at these sites reduces the affinity of cortactin for F-actin in a graded manner by altering the charge of the cortactin repeat region: once some threshold number of lysines becomes acetylated, the F-actin binding affinity of cortactin is decreased to a level that inhibits the interaction between the two proteins (Zhang et al., 2007). This would suggest that the acetylation of a particular lysine in the patch is of lesser importance than the total number of lysines that are acetylated.

Acetylation of viral proteins

Recent investigations revealed that several virally encoded proteins are substrates for acetylation by cellular HATs (Alfonso et al., 2007; Cereseto et al., 2005; Col et al., 2001; Kiernan et al., 1999; Madison et al., 2002; Marzio et al., 2000; Mu et al., 2004; Ott et al., 1999; Shimazu et al., 2006; Topper et al., 2007; Xie et al., 2002; Zhang et al., 2000). For instance, Adenovirus E1A and HIV-1 Tat are both acetylated by cellular HATs, resulting in a wide variety of functional effects. The oncoprotein E1A is known to play a key role in the dramatic alteration of various essential cellular activities, leading to cell transformation and tumorigenicity (Gallimore and Turnell, 2001). E1A was found to be acetylated by p300/CBP and PCAF at a single lysine residue located at position 239 in the C-terminal domain (Madison et al., 2002; Zhang et al., 2000), which is involved in the interaction with the transcriptional corepressor C-terminal binding protein (CtBP) (Boyd et al.,

1993). The functional outcome of E1A acetylation has been debated. One study showed that acetylation at lysine 239 inhibits the interaction between E1A and CtBP, leading to the loss of CtBP-mediated transcriptional repression and to the increase in the transforming potential of E1A (Zhang et al., 2000). Conversely, a subsequent report claimed that, rather than interfering with CtBP recruitment, acetylation at lysine 239 prevents the nuclear import of E1A by abrogating its interaction with importin 〈 (Madison et al., 2002). In such a case, acetylation may act to either attenuate the nuclear functions of E1A or redirect a portion of the protein to cytoplasmic targets.

HIV-1 Tat can be acetylated by three different HATs at three specific lysine residues: lysine 28 is targeted by PCAF, while lysines 50 and 51 are substrates for p300/CBP and GCN5 (Col et al., 2001; Kiernan et al., 1999; Ott et al., 1999). Most interestingly, the acetylation of each individual lysine differently affects Tat functionality at the molecular level. Lysine 28 acetylation enhances the ability of Tat to recruit the P-TEFb complex (Kiernan et al., 1999), while modification of lysine 50 leads to Tat dissociation from TAR RNA (Deng et al., 2000). In addition, the acetylation at lysine 50, but not that at lysine 28, creates a high-affinity binding site for the bromodomain of PCAF (Dorr et al., 2002; Mujtaba et al., 2002). The accepted model explaining the role of Tat acetylation in the regulation of its activity is as follows (Nakatani, 2002) (Figure 8). Tat acetylated at lysine 28 would efficiently recruit P-TEFb and interact with TAR to enhance the processivity of RNA Pol II. Acetylation of Tat at lysine 50 would then help the recycling of Tat by inducing its release from TAR. This release seems to be promoted by the bromodomain of PCAF, which competes with TAR for the binding to acetylated lysine 50. In agreement with the proposed model, the three HATs capable of acetylating Tat all efficiently cooperate to stimulate transcription from the 5′ LTR of HIV-1 proviral DNA. In addition to Tat, the IN protein of HIV-1 has been recently identified as a target for acetylation by at least two different HATs, p300 and GCN5 (Cereseto et al., 2005; Terreni et al.; Topper et al., 2007). The functional consequences arising from p300-mediated acetylation of IN have been previously discussed in the section concerning the cellular cofactors of HIV-1 integration, while the role of HIV-1 IN acetylation by GCN5 represents the argument of the

present treatise.

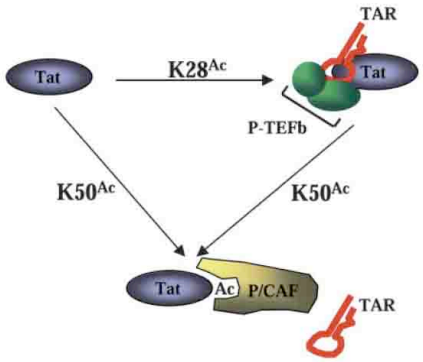

Figure 8. Acetylation signaling and Tat function. The acetylation of Tat K28 enhances the recruitment of the P-TEFb complex. The binding of Tat/P-TEFb to TAR RNA increases RNA Pol II processivity and stimulates the production of viral RNA transcripts. Acetylation of K50, which is in contact with the RNA, creates a high affinity binding site for the PCAF bromodomain, which then mediates the release of Tat from TAR. (Caron et al., 2003).

Another viral factor functionally modulated by acetylation is the Simian virus 40 large T antigen (SV40 T-Ag), a multifunctional protein possessing cellular transforming activity. In this case, acetylation seems to be involved in the regulation of the protein stability, likely enhancing T-Ag degradation through a pathway independent of the ubiquitin-proteasome system (Shimazu et al., 2006).

Acetylation has also been reported to play a role in the replication cycle of Hepatitis delta virus (HDV). In fact, the small hepatitis delta antigen (S-HDAg), a viral nucleocapsid protein essential for HDV RNA replication, can be acetylated *in vitro* at lysine 72 by the HAT domain of p300 (Mu et al., 2004). Most importantly, substitution of lysine 72 with arginine causes mutant S-HDAg to redistribute from the nucleus, where HDV replication takes place, to the cytoplasm and finally results in the reduction of the accumulation of HDV genomic RNA.

The human T-cell leukemia virus type 1 (HTLV-1) transcriptional activator Tax has been recently identified as the target of a complex regulatory program involving multiple PTMs (Lodewick et al., 2009). Leukemogenesis induced by HTLV-1 has been linked to the ability of Tax to constitutively activate the NF-κB pathway (Chu et al., 1999; Harhaj and Sun, 1999; Huang et al., 2000; Kehn et al., 2004; Malek et al., 2001; Ratner et al., 2000; Yamamoto and Gaynor, 2004), which is known to depend on Tax phosphorylation, ubiquitination and sumoylation (Bex et al., 1999; Durkin et al., 2006; Lamsoul et al., 2005; Nasr et al., 2006; Peloponese et al., 2004; Shembade et al., 2007; Yu et al., 2008). Another level of complexity was

added to the post-translational regulation of Tax functionality by the finding that the nuclei of Tax-expressing cells, including HTLV-1 transformed T-lymphocytes, contain a pool of Tax molecules acetylated at a single lysine residue by p300 (Lodewick et al., 2009). A functional crosstalk seems to occur between the different PTMs to which Tax is subjected. In fact, phosphorylation of Tax is a prerequisite for its localization in the nucleus and correlates with its subsequent acetylation by p300, whereas sumoylation, resulting in the formation of Tax nuclear bodies in which p300 is recruited, favors Tax acetylation (Lodewick et al., 2009). Overexpression of p300 markedly increases Tax acetylation and the ability of a wild type HTLV-1 provirus, but not of a mutant provirus encoding an acetylation deficient Tax protein, to activate gene expression from an integrated NF-κB-controlled promoter (Lodewick et al., 2009). Thus, the ability of Tax to constitutively activate the NF-κB pathway critically depends on a hierarchical sequence of PTMs, which differentially controls Tax intracellular localization and its transcriptional activity.

Acetylation of non-histone substrates and the protein modification code

It is now clear that lysine acetylation occurs in proteins from across the nucleus to the plasma membrane. Acetylation patterns and the mechanistic impact vary from one protein to another (Glozak et al., 2005). As exemplified by the above-cited cases of p53 and HIV-1 Tat, the effects of acetylation can be multifaceted even for one protein. In general, different acetylation events can be roughly categorized into three groups. First, acetylation occurs at one or a few residues, and thus functions as a simple on/off switch. For example, acetylation of acetyl-CoA synthase (Starai et al., 2002) and nitric oxide synthase (Mattagajasingh et al., 2007) inactivates the enzymes, whereas ⟨-tubulin acetylation generates a simple mark for cargo transport (Dompierre et al., 2007; Reed et al., 2006). Second, acetylation occurs at clusters of lysine residues that form “charged patches”. The number of acetylated residues rather than acetylation of an individual lysine seems to be important for the functional impact. In this case, acetylation events at different sites can have additive effects, thereby producing a linear output and regulating protein function in a quantitative manner, or provide a fail-safe mechanism. For example, cortactin is

acetylated at eight residues in its repeat region (Zhang et al., 2007), with each repeat being sufficient for F-actin binding and containing at least one lysine targeted by acetylation. In a third category, acetylation interacts with other PTMs and imparts effects in a site-specific manner. By analogy with the “histone code” hypothesis, a “post-translational code” model was proposed for p53, whose activity is regulated by numerous stress-induced PTMs (Appella and Anderson, 2000; Chuikov et al., 2004). Such a model has subsequently been extended to p300/CBP (Legube and Trouche, 2003), ⟨-tubulin (Westermann and Weber, 2003), the RNA pol II CTD (Buratowski, 2003; Corden, 2007), nuclear receptor coregulators (Lonard and O'Malley B, 2007), as well as HTLV-1 Tax protein (Lodewick et al., 2009). Conceptually, a general term like “post-translational code” can be used to refer to proteins with multisite PTMs. The meaning of the word “code” here is clearly more complicated than that in the “genetic code”, but recent studies strongly suggest that multisite PTMs are somewhat codified and act in sequential and combinatory manners (Corden, 2007; Kurash et al., 2008; Sampath et al., 2007; Walter et al., 2008; Winter et al., 2008). As in the case of histones (Berger, 2007), multisite modifications of p53 and other tightly regulated proteins form sophisticated programs for “intramolecular and intermolecular signaling” (Yang, 2005). Within these programs, different PTMs interact with one another and various combinations yield distinct functional outcomes. Dependent on the context, lysine acetylation can interplay with other PTMs in an agonistic or antagonistic manner. The complexity of such programs in a given protein is proportional to its biological importance and the complexity of the corresponding organism. Therefore, lysine acetylation, along with other PTMs, exerts diverse cellular effects in a context-dependent manner.

GCN5

The first histone-specific HAT, p55, was isolated from *Tetrahymena* and shown to be the homolog of the yeast GCN5 putative transcriptional coactivator and, thus, was the first enzyme to link histone acetylation and transcriptional activation (Brownell et al., 1996). Since then, homologs of GCN5 have been cloned and sequenced from numerous divergent organisms, such as human (Candau et al.,

1996), mouse (Xu et al., 1998), and *Drosophila melanogaster* (Smith et al., 1998), suggesting that its function is highly conserved throughout the eukaryotes.

To date, yeast GCN5 (also referred to as yGCN5) is the best characterized of the HATs, both structurally and functionally and both *in vivo* and *in vitro*. Various studies have mapped and characterized the functional domains of yGCN5 (Candau and Berger, 1996; Candau et al., 1997) (Figure 9). These include a C-terminal bromodomain, an Ada2 interaction motif and the HAT domain, which, by use of truncation mutants, was found to be required for the function of transcriptional coactivator *in vivo* (Candau et al., 1997). The substrate specificity of yGCN5 has also been investigated. *In vitro*, recombinant yGCN5 was shown to acetylate histone H3 strongly and H4 weakly in a free histone mixture (although histone H4 was well acetylated individually). Protein sequence analysis of these reaction products revealed that the primary sites of acetylation were lysine 14 on histone H3 and lysines 8 and 16 on histone H4 (Kuo et al., 1996). Although recombinant yGCN5 proved able to acetylate free histones efficiently, it was unable to acetylate nucleosomal histones (Grant et al., 1997; Kuo et al., 1996; Ruiz-Garcia et al., 1997), the more physiological substrates, except under special conditions and at high enzyme concentrations (Tse et al., 1998). The lack of nucleosomal HAT activity suggested that other factors are required to regulate either the recognition of histone substrates within the nucleosome or the catalytic activity of the enzyme in this altered context.

Early studies involving yGCN5 demonstrated its ability to functionally interact both *in vitro* and *in vivo* with two adaptor proteins, Ada2 and Ada3, to form a trimeric complex with the ability to promote transcriptional activation (Marcus et al., 1994) (Candau et al., 1996; Horiuchi et al., 1995). While it was subsequently shown that the heterotrimeric complex formed by GCN5, Ada2, and Ada3 has indeed the ability to acetylate nucleosomal histones (Balasubramanian et al., 2002), its native association with other proteins to form larger chromatin-modifying complexes had not yet been explored. The search for these complexes possessing HAT activity resulted in the identification of four distinct complexes with the ability to acetylate nucleosomal histones in yeast (Grant et al., 1997). Two of these four complexes were shown to contain GCN5 and the adaptor protein Ada2, suggesting that they

possessed the necessary catalytic activity and targeting ability to function as multisubunit HAT complexes. Subsequent characterization of the highest molecular weight HAT complex, estimated at 1.8 MDa, identified several additional subunits, including members of the suppressor of Ty (Spt) protein family, Spt3, Spt7 and Spt20. The discovery of these subunits in association with the high molecular weight HAT complex led its naming as the Spt-Ada-GCN5 acetyltransferase (SAGA) complex. During initial characterization of the yeast SAGA complex it was shown that deletion of GCN5, Ada2, or Ada3 resulted in loss of HAT activity, indicating that Gcn5 was the sole HAT responsible for SAGA-mediated histone acetylation. Additionally, these experiments demonstrated that the adaptor proteins Ada2 and Ada3 were required for mediating the activity of Gcn5 in a physiologically relevant context. While GCN5, Ada2, and Ada3 were not required for the overall integrity of the complex, they were shown to be critical for its ability to recognize and acetylate nucleosomal substrates and defined the GCN5-Ada2-Ada3 heterotrimer as the catalytic core of the SAGA complex (Grant et al., 1997). It was later determined that there was a second GCN5-containing protein complex in yeast that functioned in stress-response pathways and shared a similar subunit composition and histone substrate specificity with SAGA (Pray-Grant et al., 2002; Sterner et al., 2002). Due to its similarity to SAGA, this second GCN5-containing complex was named the SAGA-like (SLIK) complex. Both the SAGA and SLIK complexes represent examples of how histone acetylation by GCN5 is controlled through its interactions with accessory factors in the context of multiprotein chromatin-modifying complexes. The composition and functions of these yeast HAT complexes has served as the as the basis for the identification and characterization of similar complexes in higher eukaryotes.

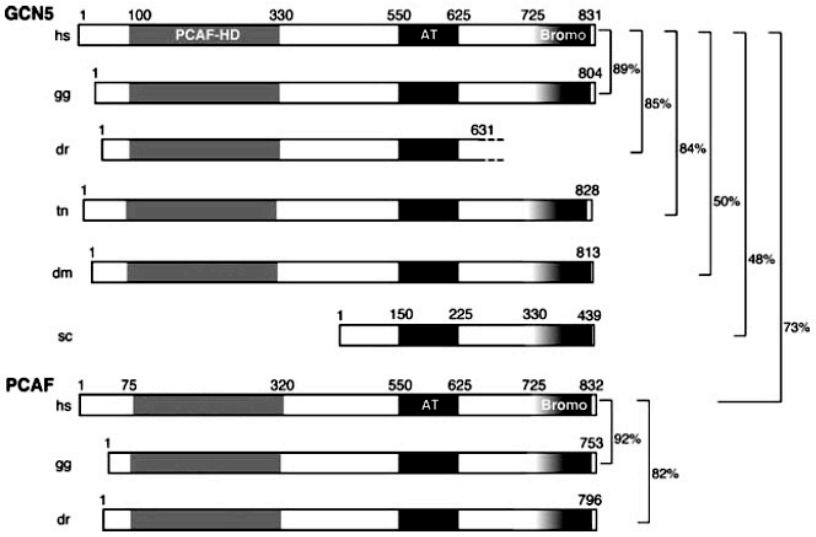

Figure 9. The overall structure of the GCN5 and PCAF enzymes in vertebrates, *Drosophila* and yeast. Schematic representation and domain organization of the GCN5 and PCAF proteins from human (hs; *Homo sapiens*), chicken (gg; *Gallus gallus*), zebrafish (dr; *Danio rerio*), pufferfish (tn; *Tetraodon nigroviridis*), *Drosophila melanogaster* (dm) and yeast (sc; *Saccharomyces cerevisiae*) are shown. The PCAF homology domain (PCAF-HD) is shown in grey, the acetyltransferase (AT) domain is shown in black and the bromodomain (Bromo) is shaded. The numbers over the boxes indicate amino acid positions. The identity between the different factors is indicated in % on the right of the horizontal lines, representing the pair wise comparisons. (Nagy and Tora, 2007).

Unlike yeast, metazoans possess two GCN5 isoforms that are believed to arise from the alternative splicing of a single gene product (Smith et al., 1998). The lower molecular weight isoform (GCN5-S) contains both a HAT domain and a single bromodomain and is similar in size and function to the single form of Gcn5 found in yeast. The second, higher molecular weight isoform (GCN5-L) has been shown to contain an additional N-terminal domain that presents high similarity to the N-terminal domain of the GCN5 homolog PCAF (Xu et al., 1998) (**Figure 9**).

The function of human GCN5 (also known as hGCN5) has been investigated *in vitro* and *in vivo*, and it was found to carry out transcriptional coactivator roles analogous to those of yGCN5 (Candau et al., 1996). Further studies revealed that hGCN5 possesses HAT activity *in vitro* (Yang et al., 1996b), and its HAT domain can successfully substitute for that of yGCN5 *in vivo*, indicating the evolutionary conservation of this HAT function (Wang et al., 1997). However, in contrast to yGCN5, the larger GCN5 isoform found in higher eukaryotes is able to acetylate

histones H3 and H4 in both the free and nucleosome-incorporated states *in vitro*, without the presence of accessory proteins (Xu et al., 1998). The ability of free GCN5-L protein to acetylate nucleosomal histones in more complex organisms is presumably dependent on the presence of the additional N-terminal domain, which is absent from yeast GCN5.

GCN5 isoforms serve as the catalytic HAT subunits in several multiprotein complexes in human cells, including the STAGA, TFTC and ATAC complexes, which will be the argument of the following section. Interestingly, the vertebrate-specific HAT PCAF, identified as a paralog of human GCN5 (Yang et al., 1996b), can also function as the catalytic subunit of these HAT complexes.

Metazoan GCN5-containing complexes

Analyses of protein samples purified from either *Drosophila* or human cell nuclear extracts separated by size suggested that, in metazoans, GCN5 is the catalytic subunit of at least two types of multiprotein complexes, one having a molecular weight of 2 MDa, and a second type with a size of about 700 kDA. The 2 MDa GCN5-containing complexes in metazoans are the STAGA (SPT3-TAF9-GCN5 acetyltransferase) and the TFTC (TBP-free-TAF) complexes (Brand et al., 1999; Martinez et al., 1998; Wieczorek et al., 1998). In addition to GCN5, all of them contain a set of Tata binding protein (TBP)-associated factors (TAFs) and several human homologs of proteins that were earlier identified in yeast as necessary for either correct transcription initiation site selection by RNA Pol II (the Spt group of proteins (Winston and Sudarsanam, 1998)) or transcriptional activation (the Ada group of adaptor proteins (Berger et al., 1992; Pina et al., 1993; Roberts and Winston, 1996)). The human 2 MDa complexes also contain a 400 kDa protein, TRRAP (transactivation/transformation domain associated protein), which was originally isolated as a Myc-associated transcriptional coactivator (McMahon et al., 1998). Since their original identification, a lot of information has been published on the subunit composition and function of the STAGA and TFTC complexes (Brand et al., 2001; Cavusoglu et al., 2003; Demeny et al., 2007; Helmlinger et al., 2004; Martinez et al., 2001; Palhan et al., 2005). As a result, today these 2MDa GCN5-containing complexes appear to be very similar and highly homologous to the

yeast SAGA complex. Genetic and biochemical analysis in *Drosophila* led to the conclusion that the 2MDa GCN5-containing complexes play multiple roles in the cells. The first would be the deposition of histone H3 acetylation marks along the entire chromosomes, while the second would be the targeted acetylation of histones at specific loci. Evidence exists that the equivalent vertebrate complexes also play a role in global chromosome acetylation, since total histone preparations from GCN5-lacking cells show very drastic changes in several differently deposited acetyl marks (Kikuchi et al., 2005). If such changes occurred only at certain specific gene loci, they would not be detectable in total histone preparations. Conversely, increased levels of GCN5 alone can strongly augment global nuclear acetylation, indicating that GCN5 has widespread global effects on the chromatin acetylation state, independent of its recruitment by activators (Knoepfler et al., 2006). Another study also described the role of STAGA/TFTC in global chromatin decondensation (Helmlinger et al., 2006). SCA7 (spinocerebellar ataxia 7) is one of the several neurodegenerative disorders caused by a polyglutamine (polyQ) expansion, but it is the only one in which the retina is affected. Vertebrate Ataxin-7, homologue of yeast Sca7 SAGA component, is a subunit of the STAGA/TFTC-type complexes (Helmlinger et al., 2004; Palhan et al., 2005). Studying the effect of polyQ-expanded Ataxin-7 on STAGA/TFTC activity revealed an interesting connection between the function of these complexes and global histone acetylation. Careful examination of mouse photoreceptors, in which polyQ-expanded Ataxin-7 had been exogenously expressed, indicated a dramatic increase in the photoreceptor nuclear volume, probably due to chromatin decondensation (Helmlinger et al., 2006). Histone hyperacetylation caused by the aberrant HAT function of the polyQ-containing STAGA/TFTC complexes has been hypothesized to account for the global chromatin decondensation (Helmlinger et al., 2006). Thus, it seems that hyperacetylation and the consequent dramatic chromatin decondensation disrupts the normally highly compacted chromatin architecture of the differentiated photoreceptor nuclei.

Many metazoan transcriptional activators and cell cycle regulatory proteins directly contact and recruit GCN5-containing complexes to specific promoters (Anafi et al., 2000; Flinn et al., 2002; Kahata et al., 2004; Lebedeva et al., 2005; Liu et al., 2003;

McMahon et al., 1998; Paulson et al., 2002; Yanagisawa et al., 2002). According to the most accepted model (Carrozza et al., 2003), these activators would directly contact the TRRAP subunit of the STAGA/TFTC complexes; however, preceding histone modifications created by other chromatin-modifying complexes may also influence the recruitment. Once targeted, STAGA/TFTC complexes mainly acetylate histone H3 in the proximity of promoters and, in such a way, they are thought to stabilize their own binding to these promoter regions. In addition, it appears that the targeted binding of ATP-dependent chromatin-remodeling complexes, containing subunits with bromodomains, also requires GCN5-dependent acetylation (Carrozza et al., 2003; Hassan et al., 2001).

Similarly to the yeast SAGA, STAGA/TFTC complexes also contain a subunit functioning as H2B ubiquitin protease, USP22 in humans (corresponding to the Ubp8 SAGA component), which can specifically remove the monoubiquitin moiety from histone H2B (Daniel et al., 2004; Henry et al., 2003; Weake et al., 2008; Zhao et al., 2008). Given the regulatory effect that monoubiquitinated histone H2B exerts on transcriptional initiation and subsequent elongation of the nascent transcript by RNA pol II in yeast (Wyce et al., 2007), it appears that the SAGA complex and its counterparts found in higher eucaryotes can facilitate transcriptional activation and elongation not only through acetylation of chromatin, but also through the modulation of histone ubiquitination.

The STAGA/TFTC-type complexes are also believed to play a role in linking chromatin modifications and DNA repair mechanisms. Since STAGA/TFTC was shown to possess higher binding affinity to UV-irradiated DNA and enhanced acetylation activity on nucleosomes formed on irradiated DNA (Brand et al., 2001), the possibility was raised that these complexes may be involved in making the DNA damage more accessible to the repair machinery in the context of chromatin. According to the proposed model, STAGA/TFTC, by binding to UV lesions, would deliver HAT activity to the damaged sites, consequently loosening up the chromatin structure around these regions and thus facilitating the rapid access of the DNA repair machinery to the UV lesions. This is consistent with the observation that DNA repair synthesis is enhanced on hyperacetylated nucleosomes, and that nucleotide excision repair (NER) is facilitated by ATP-dependent chromatin-

remodeling factors, which may be recruited after or during the acetylation signal has been deposited by GCN5 (Ura et al., 2001).

Finally, metazoan 2 MDa GCN5-containing complexes are known to function as cofactors for several oncoproteins. For instance, in human cells, following mitogen-mediated activation, the Myc oncoprotein recruits STAGA/TFTC-type complexes as transcriptional coactivators to Myc-dependent promoters through the interaction with their TRRAP subunit (Frank et al., 2001). Consistently, a naturally occurring truncated form of Myc, which does not interact with purified STAGA complex, shows reduced transcription activation potential and cannot transform primary cells (Spotts et al., 1997). In addition, Myc-dependent transcription and cell proliferation requires the integrity of the STAGA complex (Liu et al., 2008). These results suggest that Myc activates transcription of target genes via mechanisms that involve modification of the chromatin structure partly through TRRAP-mediated recruitment of GCN5-containing HAT complexes. Therefore, the oncogenic transformation carried out by Myc is not a result of the protein *per se*, but rather depends on its physical and functional contact with multisubunit GCN5-containing complexes.

In the past few years, an additional GCN5-containing complex has been described from metazoans, called ATAC (ADA-two-a-containing) (Guelman et al., 2009; Guelman et al., 2006; Suganuma et al., 2008; Wang et al., 2008). The first indication of its existence came from the discovery that the yeast Ada2 protein has two orthologues in metazoans: ADA2a and ADA2b (Kusch et al., 2003; Muratoglu et al., 2003). In fly embryo extracts, ADA2a co-fractionates with GCN5 in the 600-700 kDa size range, while ADA2b copurifies with the 2 MDa STAGA-type complexes (Kusch et al., 2003; Muratoglu et al., 2003). After this finding, the affinity purification of the *Drosophila* 700 kDa ATAC complex (dATAC) was reported (Guelman et al., 2006). dATAC has been shown to possess different substrate specificity than dSTAGA, as it mainly acetylates histone H4 (Ciurciu et al., 2006; Guelman et al., 2006; Suganuma et al., 2008). Interestingly, further characterization of dATAC established that a second HAT enzyme, dATAC2, is also part of the complex and seems to be responsible for the H4-specific acetylation activity (Suganuma et al., 2008). The differential histone acetylation

patterns induced by the activity of the two complexes seem to determine distinct functions in changing the chromatin structure and also in consequent downstream regulatory events. Given the high conservation of the known subunits of both dSTAGA and dATAC through evolution (Lee and Workman, 2007; Nagy and Tora, 2007), the identification of an ATAC-like complex in vertebrates was expected. Indeed, purification and characterization of the endogenous human ATAC complex (hATAC) were recently reported (Nagy et al., 2010), thus validating previous results obtained by overexpressing one of the ATAC subunits and using it for subsequent immunopurification (Guelman et al., 2009; Wang et al., 2008). The hATAC complex shares a common core with STAGA, composed of GCN5-ADA3-SGF29 and either ADA2a (in ATAC) or ADA2b (in SAGA). Besides these subunits, hATAC contains a second putative HAT, called hATAC2, as well as several other subunits involved in transcriptional regulation (hNC2β), nucleosome remodelling (hWDR5), cell growth (hYEATS2) and potential DNA binding (hZZZ3) (Guelman et al., 2009; Nagy et al., 2010; Wang et al., 2008). However, since the analysis of endogenous or Flag-purified hATAC highlighted a surprisingly large and variable size of this complex (from 2 MDa to about 600 kDa) (Guelman et al., 2009), while the sum of the masses of the already identified components is only about 800 kDa, the future identification of new subunits is expected. With regard to the substrate specificity of hATAC, differently from what observed in *Drosophila*, in humans both the STAGA and ATAC complexes exhibit the same specificity towards histones H3 and H4 (Guelman et al., 2009; Nagy et al., 2010; Wang et al., 2008). What is then the biological relevance of having two GCN5-containing HAT complexes, STAGA and ATAC, in humans? A potential explanation came from a recent study, where the genome-wide binding map of the two GCN5-containing HAT complexes, STAGA and ATAC, was determined in two human cell lines (Krebs et al., 2011). The results of this study showed that the recruitment of the two complexes is highly specific, as their binding sites show little overlap and they bind to distinct functional elements. Infact, ATAC is recruited to enhancers and promoters, while STAGA is mainly recruited to promoters. STAGA and ATAC bound genes are mostly active, yet they represent only a subset of active genes, suggesting that they function as coactivators for specific expression programs. Interestingly, the same study also

defined an as-yet-undescribed role for ATAC in the regulation of a limited set of enhancers, which differ from all previously known ones, since they are not bound by p300 (Krebs et al., 2011). Finally, another novel and specific function of ATAC has been recently identified in controlling the activity of the Cyclin A/Cdk2 complex, and consequently the cell cycle progression, through direct acetylation of Cyclin A, which in turn targets the protein for degradation (Orpinell et al., 2010).

GCN5 and PCAF exhibit distinct, yet partially overlapping functions

While most of the metazoan genomes code for one GCN5-type factor, vertebrates have a second gene encoding PCAF, which is highly homologous to GCN5 (Nagy and Tora, 2007; Sterner and Berger, 2000; Yang, 2004b) (Figure 9). Available data (Gamper and Roeder, 2008; Nagy et al.) suggest that both GCN5 and PCAF form STAGA- and ATAC-type complexes. The presence of GCN5 or PCAF in these complexes appears to be mutually exclusive, and even all the four possibilities may coexist in one cell (meaning that GCN5-containing STAGA and ATAC complexes can be present together with PCAF-containing STAGA and ATAC complexes).

To better understand the physiological roles of GCN5 and PCAF, knockout mice were generated harboring deletions of the GCN5 locus, the PCAF locus, or both, and the effects of such deletions on the development and adult phenotype of the transgenic models were examined. Despite the high similarity between the two proteins, their control over embryogenesis was shown to be quite disparate. Surprisingly, homozygous deletion of the PCAF gene resulted in a lack of physical impairment in adult mice (Yamauchi et al., 2000). Loss of PCAF expression, however, led to neurological dysfunction in young adult mice, including memory deficit and an exaggerated response to externally applied stress (Maurice et al., 2008). In contrast, despite a lack of phenotype in heterozygous GCN5 knockout mice, loss of both GCN5 alleles led to embryonic lethality that was associated with an increase in apoptosis and a failure to complete neural tube development (Xu et al., 2000; Yamauchi et al., 2000). Homozygous deletion of both PCAF and GCN5 alleles led to embryonic lethality occurring at earlier stages in development than that observed in GCN5-null mice, thus suggesting that the two proteins are required for functionally distinct processes. Not surprisingly, considering the distinct

phenotypes of GCN5- and PCAF-null mice, GCN5 mRNA was found to be highly expressed until later stages during embryonic development, while PCAF mRNA was barely detectable until mice reached adulthood (Xu et al., 2000). Additional analysis of tissues from these mouse models indicated that expression of GCN5 increased in response to loss of PCAF (Yamauchi et al., 2000). Coupled with the more severe developmental phenotype observed in mice lacking both PCAF and GCN5, these findings indicated that these two HATs exhibit partially overlapping, yet distinct functions.

The idea that GCN5 and PCAF have partially overlapping functions has been further supported by experiments performed in cultured chicken cells, which demonstrated a robust upregulation of PCAF expression in response to loss of GCN5 (Kikuchi et al., 2005). The observed compensation for loss of GCN5 was PCAF-specific, since the expression of other HATs was not significantly altered. Similarly, in nuclear extracts prepared from MEFs derived from GCN5 knockout mice, PCAF expression was significantly up-regulated (2- to 3-fold), suggesting that, when GCN5 is inactivated, cells try to compensate the loss of its activity with that of its paralogue (Nagy et al., 2010). However, altered patterns of histone H3 and H4 acetylation were observed in GCN5-lacking cells and gene expression analyses revealed that the GCN5-deficiency specifically affected the transcription of G1/S checkpoint-related genes and also that of some apoptosis-related ones (Kikuchi et al., 2005), thus indicating that the functions of GCN5 and PCAF are partially overlapping, yet not redundant in a given cell, and that GCN5 has unique functions in regulating either global histone acetylation and/or specific gene expression.

CHAPTER 3

MATERIALS AND METHODS

Plasmids

Construction of pGEX-IN has previously been described (Cereseto et al., 2005). pcDNA3-HA-IN was obtained by subcloning IN sequence from pGEX-IN plasmid into pcDNA3-HA vector. pGEX-IN and pcDNA3-HA-IN fragments were produced by PCR amplification of IN with primers specific to the deleted forms. pASK-IBA37-IN was constructed by subcloning IN sequence from pGEX-IN plasmid into pASK-IBA37 vector (IBA GmbH, Göttingen, DE). pFlag-IN codon optimized (c.o.) was kindly provided by A. Engelman. pASK-IBA37-IN point mutants and pFlag-IN c.o. K264,266,273R and K258,264,266,273R were obtained by PCR-based site-directed mutagenesis starting from the corresponding plasmids encoding wild type IN.

pGEX-GCN5 was a kind gift of M. Benkirane. pcDNA3-HA-GCN5 was constructed by subcloning GCN5 sequence from pGEX-GCN5 plasmid into pcDNA3-HA vector. pcDNA3-HA-GCN5 (Y260A/F261A) (Paulson et al., 2002) was obtained by PCR-based site-directed mutagenesis starting from the plasmid encoding wild type GCN5.

For production of IN-GCN5 tethered catalysis constructs, the sequence coding for the 6-300 amino acid region of GCN5 was amplified by PCR from pcDNA3-HA-GCN5 or pcDNA3-HA-GCN5 (Y260A/F261A) and cloned into pASK-IBA37 vector downstream and in frame with c.o. IN. The sequence encoding TEV protease recognition site was inserted by PCR between IN and GCN5 cDNAs.

pGIPZ and pGIPZGCN5 lentiviral vectors were purchased from Open Biosystems (Huntsville, AL). The sequence of GCN5 shRNAmir inserted into the pGIPZGCN5 vector is as follows: 5′-CCCATTCATTCCCTGGCATTAATAGTGAAGCCACAG ATGTATTAATGCCAGGGAATGAATGGT-3′. For production of pGIPZGCN5 mut vector, four point mutations were introduced in the shRNAmir cassette of pGIPZGCN5, obtaining the following sequence: 5′-CCCATTCAAAGGCTGGCA TTAATAGTGAAGCCACAGATGTATTAATGCCAGCCTTTGAATGGT-3′, where mutated nucleotides are underlined.

The NL4.3-Luc *env*-deleted virus expressing the luciferase reporter gene was produced from the pNL4.3.Luc.R-E- molecular clone obtained from the AIDS

Research and Reference Reagent Program, Division of AIDS, NIAID, NIH. For construction of pNL4.3.Luc.R-E-/IN WT and pNL4.3/IN WT plasmids, IN sequence was subcloned from the molecular clone pHXB2. The IN mutations in pNL4.3.Luc.R-E-/IN K264,266,273R, pNL4.3.Luc.R-E-/IN K258,264,266,273R, pNL4.3/IN K264,266,273R and pNL4.3/IN K258,264,266,273R were introduced by PCR-based site-directed mutagenesis using the corresponding viral clones encoding wild type IN as templates.

The envelope plasmid pMDG and the packaging plasmid pCMVΔR8.91 were kindly provided by Z. Debyser.

Recombinant proteins production and proteolytic processing

GST fusion proteins were purified from *Escherichia Coli* BL21 cells transformed with the respective plasmids. Bacterial cultures were grown at 37 °C in terrific broth supplemented with 100 mg/ml ampicillin until reaching an absorbance of 0.6 OD_{600}. Protein expression was induced with 1 mM IPTG and the incubation was continued for 3 h at 30 °C. Bacterial cells were pelleted, resuspended in cold lysis buffer (1× PBS pH 7.4, 50mM EDTA pH 8.0, 1% Triton X-100, 2 mM DTT, 1 mM PMSF), and sonicated by 4 pulses of 15 s each. Cleared lysates were mixed with a 50% slurry of glutathione Sepharose beads, and GST fusion proteins were allowed to bind to the resin at 4 °C on a rotating wheel. After 2 h of incubation, the beads were spun down by centrifugation and washed for three times with 10 bed volumes of lysis buffer. For the elution of GST fusion proteins, 1 bed volume of elution buffer (50 mM Tris-HCl pH 8.0, 25 mM reduced glutathione) was added to the resin. After incubation of the mixture at 4 °C on a rotating wheel for 15 min, the beads were spun down by centrifugation and the supernant (containing the eluted GST fusion proteins) was collected. The elution procedure was repeated for three times. Eluted proteins were dyalised overnight at 4 °C against one liter of dyalisis buffer (20 mM Tris-HCl, pH 8.0, 100 mM KCl, 20% glycerol).

N-terminal 6×His-tagged IN proteins were expressed in *Escherichia Coli* BL21 strain and purified by metal ion affinity chromatography (BD TALON Metal Affinity Resin, BD Biosciences, Palo Alto, CA) according to a previously reported protocol

(Bushman et al., 1993), with minor modifications. Briefly, bacterial pellets were resuspended in cold lysis buffer (20 mM Tris-HCl, pH 8.0, 1 M NaCl, 0.5% Triton X-100, 1 mM PMSF, 5mM imidazole) and sonicated by 4 pulses of 15 s each. Cleared lysates were incubated with a 50% BD TALON Resin slurry at 4 °C on a rotating wheel for 2 h. The lysate/BD Talon Resin suspension was loaded into an empty plastic column (Bio-Rad, Richmond, CA), letting the unbound proteins to pass through, and the resin was washed for three times with 10 bed volumes of wash buffer (20 mM Tris-HCl, pH 8.0, 1 M NaCl, 0.5% Triton X-100, 10 mM imidazole). 6×His-tagged proteins were eluted with 1 bed volume of elution buffer (20 mM Tris-HCl, pH 8.0, 1 M NaCl, 0.1% Triton X-100, 200 mM imidazole), repeating the procedure for four times. Eluted proteins were dyalised against one liter of dialysis buffer (20 mM Tris-HCl, pH 8.0, 1 M NaCl, 10% glycerol, 1 mM DTT) overnight at 4 °C.

Proteolytic processing of IN-GCN5 chimeras was performed by incubating 20 μg of fusion protein with 30 U of TEV protease (AcTEV Protease, Invitrogen, Inc., Carlsbad, CA) in a buffer containing 50 mM Tris-HCl, pH 8.0, 0.5 mM EDTA, 0.1 M NaCl, 1 mM DTT and 10% glycerol, overnight at 4°C. 6×His-tagged IN was then recovered from the reaction mixture by adsorption on BD TALON Resin. The purity and integrity of purified proteins were checked by SDS-PAGE followed by Coomassie Blue staining.

In vitro acetylation assay

In vitro acetylation assays were performed as previously reported with minor modifications (Cereseto et al., 2005). Briefly, GST or 6×His-tag fusion proteins were incubated with GST-GCN5 and [^{14}C]-acetyl-CoA in HAT buffer (50 mM Tris-HCl, pH 8.0, 5% glycerol, 0.1 M EDTA, 50 mM KCl, 2 mM sodium butyrate) in a final volume of 30 µl for 45 min at 30°C. Following SDS-PAGE resolution of the reaction mixtures, acetylated proteins were visualized by phosphoimaging (Cyclone).

In vitro binding assay

[^{35}S]-labeled IN proteins used for *in vitro* binding assays were produced from the corresponding pcDNA3-HA plasmids by using the TNT Reticulocyte Lysate System (Promega Corp., Madison, WI). Analysis of *in vitro* binding between GST fusion proteins and [^{35}S]-IN or [^{35}S]-IN fragments was performed as previously described (Cereseto et al., 2005). Briefly, GST fusion proteins (1 μg) immobilized on glutathione Sepharose beads, after pre-treatment in a solution containing DNase I 0.25 U/μl and RNase H 0.25 U/μl, were incubated for 1 h at 4 °C with 600 c.p.m. of *in vitro* translated [^{35}S]-proteins in NETN buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1mM EDTA, 0.5% NP-40, 1mM DTT, 1mM PMSF) supplemented with 0.2 mg/ml ethidium bromide. Following extensive washes, the reaction mixtures were resolved by SDS-PAGE and radiolabeled proteins visualized by phosphoimaging (Cyclone).

In vitro integration assays

Oligonucleotide substrates for *in vitro* integration assays were as follows: 70 (5′-GTGTGGAAAATCTCTAGCAGT-3′), 71 (5′-GTGTGGAAAATCTCTAGCA-3′), and 72 (5′-ACTGCTAGAGATTTTCCACAC-3′) (Parissi et al., 2001). 70 or 71 oligonucleotides were labeled with ^{32}P using polynucleotide kinase and annealed to the complementary 72 oligonucleotide. 3′ end processing reaction was performed in a buffer containing 25 mM MOPS, 5% PEG-8000, 5% DMSO, 0.05% NP-40, 30 mM NaCl and 10 mM DTT, in the presence of the 70/72 blunt substrate. Strand transfer reaction was carried out in 20 mM Hepes, pH 7.2, 7.5 mM $MnCl_2$, 0.05% NP-40 and 10 mM DTT, in the presence of the 71/72 substrate. For both assays, [^{32}P]-labeled duplex DNAs (1 pmol) were incubated in a final volume of 20 μl with 100 or 200 ng of recombinant IN derived from GCN5 wild type or mutant fusion protein at 37 °C for 1 h. The reaction products were resolved by electrophoresis on a 15% polyacrylamide gel with 7M urea in Tris-Borate-EDTA buffer, pH 7.6, and then visualized by phosphoimaging (Cyclone).

Cell culture and transfection

HeLa and HEK 293T cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal calf serum (FCS), 100 U/ml penicillin and 100 μg/ml streptomycin. HEK 293T cells stably transduced with pGIPZ lentiviral vectors were grown with the addition of puromycin 2 μg/ml. CEM T cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 10% FCS, 2 mM glutamine, 100 U/ml penicillin and 100 μg/ml streptomycin.

Transfections were performed by the standard calcium phosphate coprecipitation procedure, or by using the polyethylenimine (PEI) reagent (MW 25000, Sigma, Inc., St Louis, MO) according to a previously reported protocol (Durocher et al., 2002).

Immunoprecipitation and Western blotting

For immunoprecipitation, cell pellets were lysed 36 h after transfection in RIPA buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% Triton X-100, 0.1% SDS, 0.5% deoxycolic acid) containing 10 mM sodium butyrate (Sigma, Inc.) and protease inhibitors (Complete Protease Inhibitor Cocktail Tablets, Roche Diagnostics, Indianapolis, IN). The protein concentration of cell extracts was determined by Bradford assay (Bio-Rad). Anti-Flag M2 affinity resin or rat monoclonal anti-HA antibody were incubated overnight at 4 °C with the cell lysates (2 mg of total proteins for coimmunoprecipitation or 4 mg for *in vivo* acetylation experiments). The HA-immune complexes were precipitated by incubation with UltraLink Immobilized Protein G (Pierce Biotechnology, Inc., Rockford, IL). The precipitated complexes were then extensively washed and analyzed by Western blotting using the appropriate antibodies.

Antibodies

The following primary antibodies were used: rabbit anti-acetylated-lysine (Cell Signaling Technology, Inc., Danvers, MA); mouse anti-Flag M2 (Sigma, Inc.), either free or bound to agarose beads; rat anti-HA Clone 3F10 (Roche Diagnostics); mouse anti-IN 8G4, obtained from the AIDS Research and Reference Reagent Program; rabbit anti-GCN5 H-75 (Santa Cruz Biotechnology, Inc., Santa Cruz, CA); and mouse anti-α-tubulin Clone B-5-1-2 (Sigma, Inc.).

For the production of a polyclonal, anti-acetylated IN antibody, three rabbits were immunized with a peptide corresponding to amino acids 261-280 of the IN sequence, chemically acetylated at lysines 264, 266 and 273, after conjugation with Maleimide Activated mcKLH (Pierce Biotechnology, Inc.). The IgG fraction was obtained from collected sera with the use of ImmunoPure (A) IgG Purification Kit (Pierce Biotechnology, Inc.). The purified samples were then passed over a column conjugated with the unmodified IN peptide to remove the antibody cross-reacting with non-acetylated IN.

Secondary horseradish peroxidase (HRP)-conjugated antibodies against mouse or rabbit Igs were purchased by Santa Cruz Biotechnology, Inc. For detection of IN acetylation with anti-acetylated-lysine antibody, Biotin-SP-conjugated AffiniPure $F(ab')_2$ Fragment Donkey Anti-Rabbit IgG (H+L) (Jackson ImmunoResearch Laboratories, Inc., West Grove, PA) and ECL Streptavidin-HRP conjugate (Amersham Biosciences Corp., Piscataway, NJ) were employed.

Virus and viral vector production

For production of *env* deleted, VSV-G pseudotyped NL4.3-Luc viral stocks, 6×10^6 HEK 293T cells were seeded in 100 mm dishes in Optimem (Invitrogen, Inc.) containing 2% FCS. After 24 h of incubation, the medium was replaced by Optimem without FCS and cells were transfected with 20 μg of pNL4.3.Luc.R-E- (either wild type or mutated) and 5 μg of the envelope plasmid pMDG using the PEI reagent (Sigma, Inc.). The cell culture supernatant was collected 48 h after transfection and filtered through a 0.45 μM pore size filter.

NL4.3 replication competent viruses were prepared as described for NL4.3-Luc viral clones, using 25 μg of pNL4.3 plasmid (wild type or mutated) for transfections.

For the generation of viral vector stocks, HEK 293T cells were transfected with 10 μg of the packaging plasmid pCMVΔR8.91, 5 μg of pMDG, and 20 μg of the gene transfer plasmid (pGIPZ, pGIPZGCN5, or pGIPZGCN5 mut) following the protocol used for virus production. The cell culture supernatant was collected twice, at 48 and 72 h after transfection, filtered through a 0.45 μM pore size filter and concentrated by ultracentrifugation at 110000 × g for 2 h at 4 °C.

Both viruses and viral vectors were titered by quantification of p24 antigen in cell culture supernatants with an enzyme-linked immunoabsorbent assay (Innogenetics, Gent, Belgium).

Transient and stable knockdown of GCN5 expression

GCN5-targeting siRNA (Dharmacon Research, Boulder, CO) had the following plus-strand sequence: 5′-AACCAUGGAGCUGGUCAAUGAAA-3′. As a non-silencing control, Dharmacon ON-TARGET*plus* si*CONTROL* Non-Targeting Pool was employed.

HeLa cells, seeded in 6-well plates the day before treatment (1.5×10^6 cells/well), were transfected twice at a 24 h interval with 150 nM siRNA using Gene Silencer reagent as recommended by the manufacturer (Gene Therapy Systems, Inc., San Diego, CA). Cells trypsinized after 20 h were either plated for infections, or lysed for Western blot analysis.

For production of stably silenced cell lines, HEK 293T cells, seeded in 24-well plates the day before transduction (5×10^4 cells/well), were incubated for 4 h with 1 μg p24 antigen of the appropriate pGIPZ lentiviral vector (encoding GCN5 shRNAmir, GCN5 shRNAmir mut or lacking a silencing insert). Two days after transduction, selection with 2 μg/ml puromycin was initiated. For selection of GCN5 knockdown cell clones, the clones with the highest GFP expression levels, as determined by fluorescence-activated cell sorting analysis (FACS), were chosen.

HIV-1 infectivity assays

For single-round viral replication assays, siRNA-treated HeLa cells (2.5×10^6/well) or HEK 293T cells (5×10^6/well) were seeded in 6-well plates and incubated for 3 h, in a total volume of 500 μl, with 50 or 100 ng p24 antigen of NL4.3-Luc virus (wild type or mutated), respectively. Cells were collected 48 h after infection and lysed for measurement of luciferase activity (Luciferase Assay System, Promega Corp.). Luminometer readouts, expressed as relative light units (RLU), were normalized with respect to protein concentration in each sample.

Viral stocks used in infections for measurement of HIV-1 DNA species by real time quantitative PCR (RT-Q-PCR) were pre-treated for 1 h at 37 °C with 160 U/ml Turbo DNase (Ambion, Inc., Austin, TX).

For multiple-round infections, 1×10^6 CEM T cells were incubated with 1 or 10 ng p24 antigen of NL4.3 virus (wild-type or mutated), in a total volume of 500 μl. After 3 h, cells were washed twice with PBS containing 1% FCS and put in culture in a 6-well plate at a density of 0.5×10^6 cells/ml. After 3 days, cells were counted, spun down by centrifugation and supernatant was collected. Cells were then resuspended in fresh medium and put in culture again at a density of 0.5×10^6 cells/ml. This procedure was repeated every 3 days over a period of 21 days. Viral titers in the collected supernatants were determined by a ^{32}P-based RT assay performed by standard procedures.

RT-Q-PCR analysis

Total DNA was extracted from HEK 293T cells with the DNeasy Tissue Kit (QIAGEN GmbH, Hilden, DE) at different time points after infection. Amplification reactions were performed with the Light Cycler 480 instrument (Roche Diagnostics). Quantification of total HIV-1 DNA was performed at 24 h post infection with a pair of primers (LucFw, LucRev) and a fluorogenic hybridization probe (LucProbe) annealing to the luciferase reporter gene of NL4.3-Luc viral clone (Table 1). Reaction mixtures contained 500 ng of total cellular DNA, 1× Light Cycler 480 Probe Master (Roche Diagnostics), 300 nM each forward and reverse primers and 200 nM probe in a total volume of 20 μl. After an initial denaturation step (95 °C for 10 min), the cycling profile was 40 cycles consisting of 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s. This protocol was also used for detection of integrated HIV-1 DNA at 14 days after infection in HEK 293T cells stably transduced with pGIPZ lentiviral vectors.

Quantification of proviral DNA at 48 h post infection in HEK 293T cells was performed by *Alu*-LTR nested PCR. In the first amplication step, two outward-facing primers annealing within the highly repeated chromosomal *Alu* element (*Alu* 1 and *Alu* 2) were used together with a HIV-1 LTR-specific primer containing a lambda phage-specific heel sequence at its 5′ end (L-M667) (Table 1). Alu-LTR

sequences were amplified from 500 ng of total cellular DNA in a 20 μl reaction mixture comprising 1× Light Cycler 480 Probe Master (Roche Diagnostics), 100 nM L-M667 primer, and 300 nM (each) primers *Alu* 1 and *Alu* 2. The first-round PCR cycle conditions were as follows: a denaturation step of 8 min at 95°C followed 12 cycles of amplification (95°C for 10 s, 60°C for 10 s, and 72°C for 170 s). In the second round of PCR, a lambda-specific primer (Lambda T) and a HIV-1 LTR-specific primer (AA55M) were used (Table 1), so that only LTR-containing products from the first-round PCR could be amplified. Nested PCR was performed on 1/10 of the first-round PCR product in a mixture comprising 1× Light Cycler 480 Probe Master (Roche Diagnostics), 300 nM Lambda T primer, 300 nM AA55M primer, and 200nM (each) hybridization probes LTR FL and LTR LC (Table 1). The nested-PCR cycling profile began with a denaturation step (95°C for 8 min), followed by 50 cycles of amplification (95°C for 10 s, 60°C for 10 s, and 72°C for 9 s).

Two-LTR circles were detected with primers spanning the LTR-LTR junction (HIV F and HIV R1, Table 1). Reaction mixtures contained 1× Light Cycler 480 Probe Master (Roche Diagnostics), 300 nM (each) forward and reverse primers, and 200 nM (each) fluorogenic hybridization probes HIV FL and HIV LC (Table 1) in a final volume of 20 μl. After an initial denaturation step (95°C for 8 min), the cycling profile was 15 cycles consisting of 95°C for 10 s, 66°C for 10 s, and 72°C for 10 s, followed by 35 cycles at the beginning of which the annealing temperature was decreased by 0.5°C per cycle to the secondary target temperature (59°C).

As an internal standard for normalizing the amount of cellular genomic DNA, the level of human β-globin DNA was quantified. The reaction was carried out using 1× Light Cycler 480 Probe Master (Roche Diagnostics), 400 nM of forward primer BGF, 400 nM of reverse primer BGR, and 200 nM of BGX-P fluorescent probe (Table 1). The amplification conditions included a hot start at 50 °C for 2 min and 95 °C for 10 min, followed by 40 cycles of denaturation at 95 °C for 15 s and extension at 60 °C for 1 min.

Statystical Analysis

Paired comparisons were carried out using the two-tailed Student's *t* test, assuming equal variance between samples to determine differences at the 5%

level. All data points (including outliers) were included in the analysis for significance.

Table 1. DNA sequences of primers and probes used in RT-Q-PCR assays.

Primer or probe[e]	Sequence (5′–3′)	Target
Luc Fw	GAAGAGATACGCCCTGGTTCC	Total HIV-1 DNA
Luc Rev	TGTGATTTGTATTCAGCCCATATCG	
Luc Probe*	TTCATAGCTTCTGCCAACCGAACGGACA[a]	
L-M667	ATGCCACGTAAGCGAAACTCTGGCTAACTAGGGAACCCACTG	Integrated HIV-1 DNA (first round PCR)
Alu 1	TCCCAGCTACTGGGGAGGCTGAGG	
Alu 2	GCCTCCCAAAGTGCTGGGATTACAG	
Lambda T	ATGCCACGTAAGCGAAACT	Integrated HIV-1 DNA (second round PCR)
AA55M	GCTAGAGATTTTCCACACTGACTAA	
LTR FL*	CACAACAGACGGGCACACACTACTTGA[b]	
LTR LC*	CACTCAAGGCAAGCTTTATTGAGGC[c]	
HIV F	GTGCCCGTCTGTTGTGTGACT	Two-LTR circle
HIV R1	ACTGGTACTAGCTTGTAGCACCATCCA	
HIV FL*	CCACACACAAGGCTACTTCCCTGA[b]	
HIV LC*	TGGCAGAACTACACACCAGGGC[c]	
BGF	CAACCTCAAACAGACACCATG	Human β-globin DNA
BGR	TCCACGTTCACCTTGCCC	
BGX-P*	CTCCTGAGGAGAAGTCTGCCGTTACTGCC[d]	

[a] Modified with 6-carboxyfluorescein (FAM) at the 5′ end and with BlackBerry Quencher (BBQ) at the 3′ end.
[b] Modified with fluorescein (FL) at the 3′ end.
[c] Phosphorylated at the 3′ end and modified with LC red 640 dye at the 5′ end.
[d] Modified with 6-carboxyfluorescein (FAM) at the 5′ end and with 6-carboxytetramethylrhodamine (TAMRA) at the 3′ end.
[e] Primers and probes were purchased from TIB MOLBIOL S.r.l (Genova, Italy). *, probe sequence.

CHAPTER 4

RESULTS

HIV-1 IN is acetylated by GCN5 acetyltransferase

IN is acetylated by GCN5 *in vitro*: identification of the acetylation sites

During the past few years, many reports have desribed the interaction of specific proteins encoded by a wide variety of viruses with cellular HATs. This interaction is often associated with the acetylation of the viral protein itself, which results in a modulation of its activity (Col et al., 2001; Deng et al., 2000; Kiernan et al., 1999; Lodewick et al., 2009; Madison et al., 2002; Mu et al., 2004; Ott et al., 1999; Poulin et al., 2004; Shimazu et al., 2006; Xie et al., 2002; Zhang et al., 2000). Indeed, a specific interaction between HIV-1 IN and p300 has been recently reported, leading to the acetylation of three lysine residues in the CTD of the viral protein (Cereseto et al., 2005; Topper et al., 2007). Since acetylation substrates are often targeted by multiple HATs, we sought to investigate whether IN might be acetylated by enzymes other than p300. It has already been reported that MOZ and PCAF (belonging to the MYST and GNAT families of HATs, respectively) are incapable to efficiently acetylate the IN CTD *in vitro* (Topper et al., 2007). Therefore, in the present study, another member of the GNAT family, GCN5, was examined.

To determine whether IN is a substrate for GCN5 catalytic activity, *in vitro* acetylation assays were performed with recombinant IN and GCN5, both purified as GST fusion proteins. Incubation of the single GST domain with GCN5 in the presence of [^{14}C]-labeled acetyl-CoA, and subsequent protein resolution by SDS-PAGE followed by autoradiography led to the appearance of a unique radiolabeled band at the same size as GST-GCN5, corresponding to the auto-acetylation product of the enzyme (Figure 10A, lane 1). Incubation of GST-IN with GST-GCN5 resulted in two major radiolabeled bands, the higher corresponding to auto-acetylated GST-GCN5 and the lower to GST-IN (Figure 10A, lane 2), clearly deriving from GCN5-mediated incorporation of [^{14}C]-acetyl groups into the viral protein. These results demonstrated that GCN5 specifically acetylates IN *in vitro*.

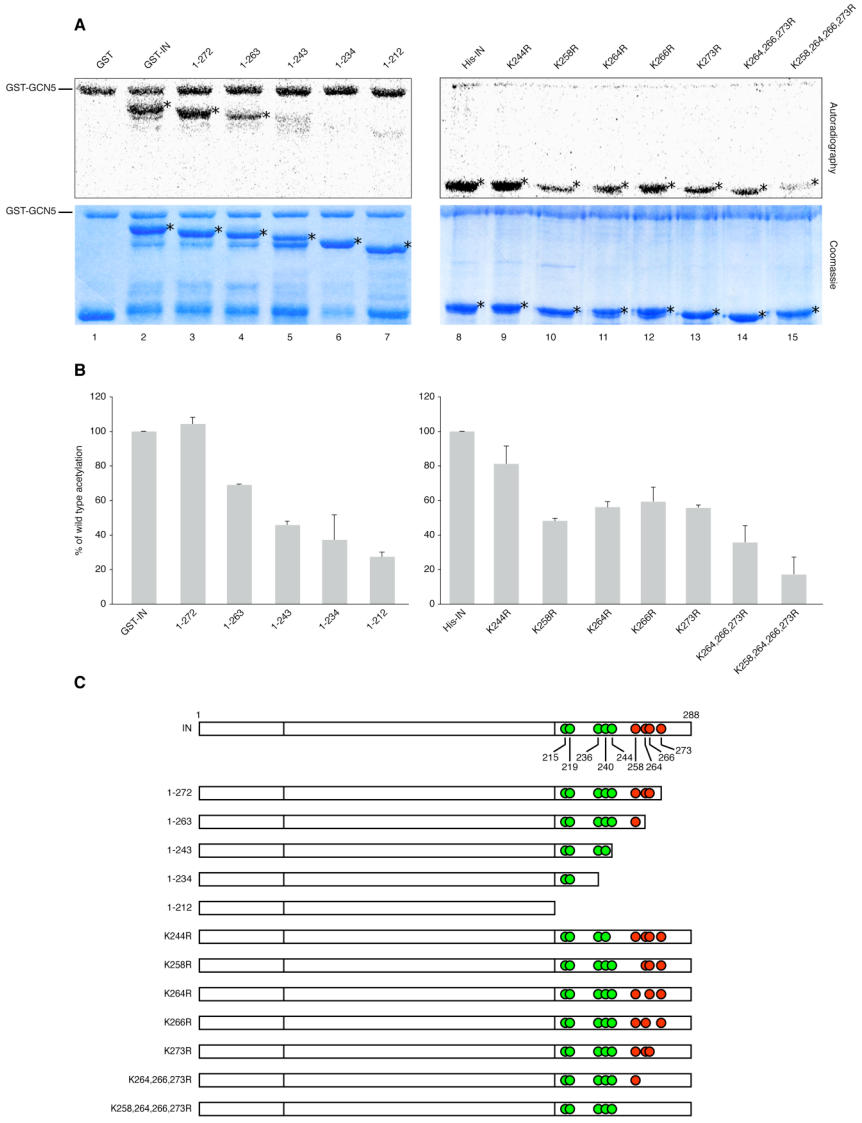

Figure 10. HIV-1 IN is acetylated by GCN5 *in vitro*. (A) Autoradiography (upper panels) and Coomassie Blue staining (lower panels) of *in vitro* acetylation assay with recombinant GST-GCN5 and IN wild type or mutant proteins. Lanes 1-7: GST fusion IN proteins; lanes 8-15: 6×His-tagged IN proteins. In the Coomassie panels, IN proteins used as acetylation substrates are indicated by asterisks; in the autoradiograms, IN proteins found positive for GCN5-mediated acetylation are indicated in the same way. Presented results are representative data from triplicate *in vitro* acetylation assay experiments. (B) Results of densitometric analysis of autoradiograms derived from three independent experiments (means ± standard errors of the means [SEM]) expressed as percent wild type IN acetylation. (C) Schematic representation of IN proteins used for the acetylation assays. The positions of lysines in the CTD of IN are indicated. Lysines positive for acetylation are shown in red.

To define which region of IN is acetylated by GCN5, GST-IN fragments with progressive deletions starting from the C-terminus (as schematized in Figure 10C) were used as substrates in *in vitro* acetylation assays and the corresponding acetylation signals in the autoradiograms evaluated by densitometric analysis (Figure 10B, left histogram). GST-IN fragment 1-272 was acetylated to a similar extent as full-length IN (Figure 10A, compare lanes 2 and 3, and Figure 10B, left histogram). Acetylation of fragment 1-263 (Figure 10A, lane 4) was reduced by 30% (Figure 10B, left histogram), while a more significant decrease in the radioactive signal (ranging from 60 to 70%) was detected by using shorter fragments (1-243, 1-234, and 1-212) (Figure 10A, lanes 5-7, and Figure 10B, left histogram). These results indicated that IN is acetylated by GCN5 within the region located between amino acid 244 and the C-terminus of the protein. As schematically represented in Figure 10C, this region contains five lysines at positions 244, 258, 264, 266, and 273 as possible targets for acetylation. Therefore, in order to identify the acetylated residues and exclude that the reduced acetylation level of the deleted IN forms derived from improper protein folding, each of these lysines was replaced with an arginine, an amino acid that conserves a positively charged side chain however cannot be acetylated. The resulting mutants were tested *in vitro* as substrates for GCN5 activity. In this experiment, IN was tagged at the N-terminus with a 6×His epitope in place of GST, in order to obtain better SDS-PAGE resolution between the signals deriving from acetylated GCN5 and IN (Figure 10A, lane 8). As reported in the right histogram of Figure 10B, densitometric analysis of radioactivity incorporation highlighted that mutation of individual lysines 258, 264, 266 and 273 (Figure 10A, lanes 10-13) caused a reduction in the acetylation level of IN ranging from 40 to 50%, while no significant decrease in the autoradiographic signal was detected upon mutation of lysine 244 (Figure 10A, lane 9). From these data we inferred that GCN5 acetylates IN at residues 258, 264, 266, and 273. Notably, previous reports demonstrated that another cellular HAT, p300, acetylates lysines 264, 266 and 273 in the CTD of IN (Cereseto et al., 2005; Topper et al., 2007). To confirm that GCN5 acetylates lysine 258 in addition to the above-mentioned residues, two mutant forms of IN were produced and assayed for *in vitro* acetylation: one containing mutations at the sites

acetylated by both GCN5 and p300 (IN K264,266,273R), and the other carrying the additional lysine-to-arginine substitution at position 258, specifically modified by GCN5 (IN K258,264,266,273R). Densitometric analysis of the resulting autoradiogram revealed that the decrease in the radioactive signal detected with IN K264,266,273R was similar to the one obtained with the single-mutated IN forms (compare lane 14 with lanes 10-13 in Figure 10A, and right histogram of Figure 10B), while the residual acetylation level of IN K258,264,266,273R dropped to 20 % with respect to wild type IN (Figure 10A, lane 15, and Figure 10B, right histogram), likely due to the mutation introduced at lysine 258. Taken together, the results of *in vitro* acetylation assays indicated that GCN5 acetylates lysines 264, 266 and 273 of IN, previously identified as p300 target sites, and lysine 258 as a specific site of modification.

IN is acetylated by GCN5 *in vivo*: detection of IN acetylation by a novel anti-acetylated IN antibody

Next, we investigated whether IN is also acetylated by GCN5 *in vivo*. To this aim, c.o. Flag-IN (Limon et al., 2002) was expressed in HEK 293T cells, alone or together with HA-GCN5 wild type or carrying two inactivating mutations in the catalytic HAT domain (Y260A/F261A) (Paulson et al., 2002). Immunoprecipitation of IN and subsequent Western blot analysis with an antibody specific to acetylated lysines revealed the highest acetylation signal in the sample corresponding to IN coexpressed with wild type GCN5 (Figure 11A, upper panel, lane 3). Conversely, expression of IN alone, or together with catalytically inactive GCN5, resulted in lower acetylation signals, likely deriving from the activity of endogenous HATs on the viral protein (Figure 11A, upper panel, lanes 2 and 4). In this experiment, the total amounts of immunoprecipitated IN and the expression levels of wild type and mutant GCN5 were verified by Western blot analysis with anti-Flag and anti-HA antibodies, respectively (Figure 11A, middle and lower panels).

To confirm the *in vitro* observation that IN is a substrate for both GCN5 and p300, an antibody specific to acetylated IN was produced by using for immunization a peptide corresponding to amino acids 261-280 of the IN sequence, chemically acetylated at lysines 264, 266 and 273, which are targeted in common by the two HATs (see the Material and Methods chapter for details). As shown in Figure 11B,

the purified antibody specifically recognized the acetylated IN peptide used as immunizing epitope in dot-blot experiments, while no cross-reactivity was detected with the unmodified IN peptide or acetylated BSA. By using this antibody for the analysis of IN immunoprecipitated from cell lysates, basal levels of IN acetylation due to the activity of endogenous HATs could be detected, as well as higher acetylation levels induced by exogenously expressed p300 (Figure 11C, top-left panel, lanes 1 and 2, respectively). This result is consistent with our previous study, showing that p300 catalyzes IN acetylation *in vivo* at positions 264, 266 and 273 (Cereseto et al., 2005). Conversely, no acetylation signal was detected with immunoprecipitated IN K264,266,273R, expressed in the cells either alone or together with p300 (Figure 11C, top-left panel, lanes 3 and 4), thus revealing the high specificity of the newly developed antibody. In this experiment, the amount of IN (wild type or mutated) immunoprecipitated in each sample was verified by Western blotting with an anti-Flag antibody (Figure 11C, bottom-left panel). The anti-acetylated IN antibody was also used for straight Western blot analysis of cell lysates. In these experimental conditions, basal levels of IN acetylation induced by endogenous HATs could not be revealed, likely due to lesser amounts of IN present in the cell lysates as compared to samples containing the immunoprecipitated protein (Figure 11C, top right panel, lane 5). However, the antibody was able to detect IN acetylation in the cell lysate corresponding to IN coexpressed with p300 (Figure 11C, top right panel, lane 6), thus exhibiting higher sensitivity than the standard anti-acetyl-lysine antibodies, which require an immunoprecipitation step to reveal IN acetylation, even upon coexpression of the acetylating enzyme. Given the high specificity and sensitivity of the anti-acetylated IN antibody, it was used to confirm the *in vivo* acetylation of IN by GCN5, as well as the mapping of the targeted lysines performed *in vitro*. As shown in the upper panel of Figure 11D, extracts from HEK 293T cells coexpressing wild type IN and GCN5 revealed a remarkable signal corresponding to IN acetylation (lane 4), while, consistently with the data reported in Figure 11C (top-right panel, lane 5), acetylation of the viral enzyme by endogenous HATs was almost undetectable (lane1). Conversely, no signal at all was detected upon immunoblotting of cell lysates containing triple- or quadruple-mutant IN, expressed either alone (lanes 2

and 3) or together with GCN5 (lanes 5 and 6). In this experiment, Western blot analysis of the cell extracts was also performed with anti-Flag and anti-HA antibodies to control the levels of exogenously expressed proteins. From these results we concluded that GCN5, in addition to p300, is able to acetylate IN *in vivo* at lysines 264, 266, and 273.

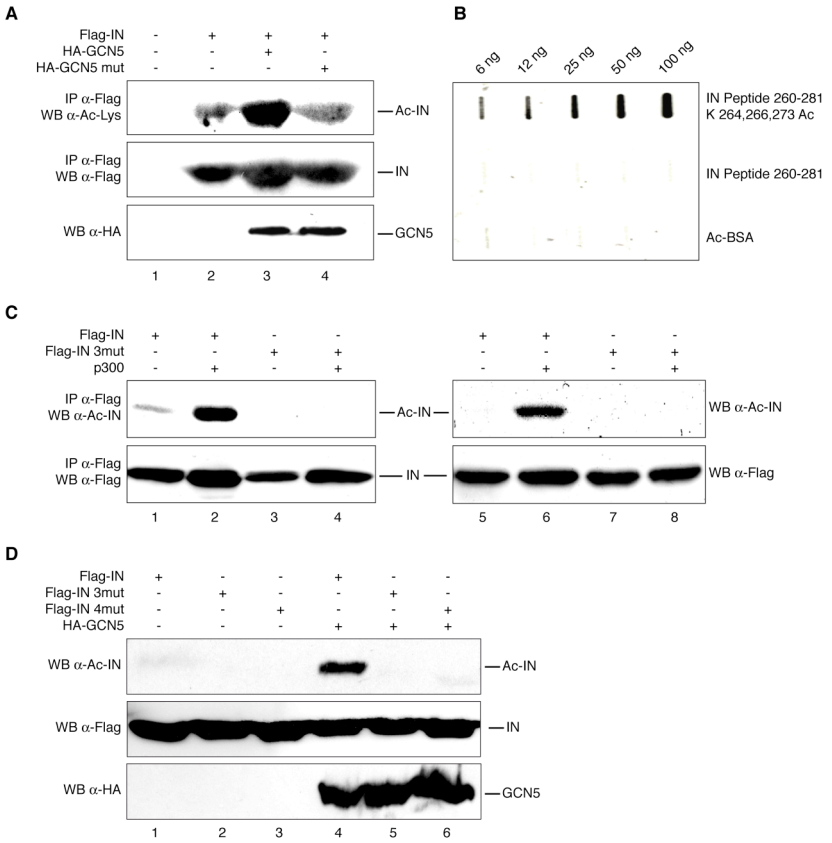

Figure 11. IN is acetylated by GCN5 *in vivo*. (A) Extracts from HEK 293T cells transfected with the indicated plasmids were immunoprecipitated using anti-Flag antibody and analyzed by Western blotting with anti-acetyl-lysine antibody (upper panel) or anti-Flag antibody (middle panel). Lower panel: cell extracts immunoblotted with anti-HA antibody. (B) Acetylated BSA and peptides corresponding to IN amino acids 260-281, either chemically acetylated at lysines 264, 266, and 273, or not acetylated, were blotted onto a nitrocellulose filter and incubated with anti-acetylated IN antibody. (C) Left panels (lanes 1-4): extracts from HEK 293T cells transfected with the indicated plasmids were immunoprecipitated using anti-Flag antibody and analyzed by Western blotting with anti-acetylated IN antibody (top panel) or anti-Flag antibody (bottom panel). Right panels (lanes 5-8): extracts from HEK 293T cells transfected with the indicated plasmids analyzed by Western blotting with anti-acetylated-IN antibody (top panel) or anti-Flag antibody (bottom panel). (D) Extracts from HEK 293T cells transfected with the indicated plasmids analyzed by Western blotting with anti-acetylated-IN antibody (upper panel), anti-Flag antibody (middle panel), or anti-HA antibody (lower panel).

In conclusion, data obtained from the *in vitro* and *in vivo* analysis of IN acetylation collectively indicated that IN is specifically modified by GCN5 at four lysine residues located in the C-terminal domain at positions 258, 264, 266, and 273.

IN interacts with GCN5

IN interacts with GCN5 *in vivo*

Since IN acetylated by GCN5, the interaction between the two proteins in living cells was investigated. To this aim, Flag-IN was expressed in HEK 293T cells together with HA-GCN5 wild type or mutated in the catalytic domain (Y260A/F261A) and immunoprecipitated using an anti-Flag antibody. As shown in the upper panel of Figure 12A, Western blot analysis of the immune complexes with an anti-HA antibody revealed that both wild type and mutant GCN5 specifically coprecipitated with Flag-IN (lanes 3 and 4), while they did not with the Flag epitope expressed from pFlag empty vector (lane 2). Accordingly, in the reciprocal experiment, where immunoprecipitation with an anti-HA antibody was followed by immunoblotting with an anti-Flag antibody, IN specifically coprecipitated with HA-GCN5 (both wild-type and mutant forms) (Figure 12B, upper panel, lanes 3 and 4), while it did not with the HA epitope expressed from pcDNA3-HA empty vector (Figure 12B, upper panel, lane 2). In both experiments, the total amounts of immunoprecipitated proteins and the expression levels of IN and GCN5 were verified by Western blotting (Figure 12A and B, middle and lower panels). Notably, both wild type GCN5 and its catalytically inactive mutant were able to coprecipitate with IN, thus indicating that the interaction between the two proteins does not depend on the HAT activity of GCN5.

IN interacts with GCN5 *in vitro*: mapping of the regions involved in the interaction

After the *in vivo* detection of the IN/GCN5 interaction, we set out to dissect the regions that mediate the binding between the two proteins.

To map the region of IN mediating the interaction with GCN5, GST pull-down assays were carried out between GST-GCN5 immobilized on glutathione-Sepharose beads and IN deleted forms labeled with [^{35}S]-Met by *in vitro* translation. The IN fragments tested for interaction with GCN5 contained

progressive deletions starting from the C-terminus of the protein, corresponding to those used in *in vitro* acetylation assays (and schematized in Figure 10C). As shown in Figure 12C, the affinity of IN fragments 1-272 and 1-263 to GST-GCN5 (13% binding efficiency) was similar to that of full-length IN (16% binding efficiency). Conversely, the interaction with GST-GCN5 was significantly decreased by using IN fragments (1-243 and 1-234) that contained further deletions towards the N-terminus ($p < 0.05$ and $p < 0.01$, respectively). These results indicated that the C-terminal region of IN located between amino acids 244 and 288, which contains the lysines targeted for acetylation, is involved in binding to GCN5.

To map the IN interacting regions on GCN5, the latter was divided into four fragments, which were subsequently expressed and purified as GST fusion proteins. Therefore, four GST-GCN5 deleted forms were produced (as schematized in Figure 12D), corresponding to regions encompassing amino acids 1-189 (N-terminal domain), 190-270 (HAT domain), 271-383 (intermediate domain), and 384-476 (bromodomain). These fragments, together with the full-length protein, were tested in GST pull-down assays for interaction with *in vitro* translated IN, labeled with [^{35}S]-Met. As shown in Figure 12E, IN bound to GCN5 fragments 1-189 and 271-283 with significantly lower efficiency ($p < 0,05$) than to the full-length protein, while no significant differences were detected in the affinity of IN to GCN5 HAT domain or bromodomain as compared to full-length GCN5, thus suggesting that these regions of the protein are mainly responsible for the interaction with IN. This result is consistent with the HAT domain being the catalytic portion of the enzyme, that has to contact the substrate in order to transfer the acetyl group, and the bromodomain being an amino acidic module with specific binding affinity to acetyl-lysine residues (Dhalluin et al., 1999; Hudson et al., 2000; Owen et al., 2000).

Taken together, data obtained in GST pull-down experiments indicated that the HAT domain and bromodomain are involved in mediating GCN5 binding to the CTD of IN within the region containing the lysines targeted for acetylation.

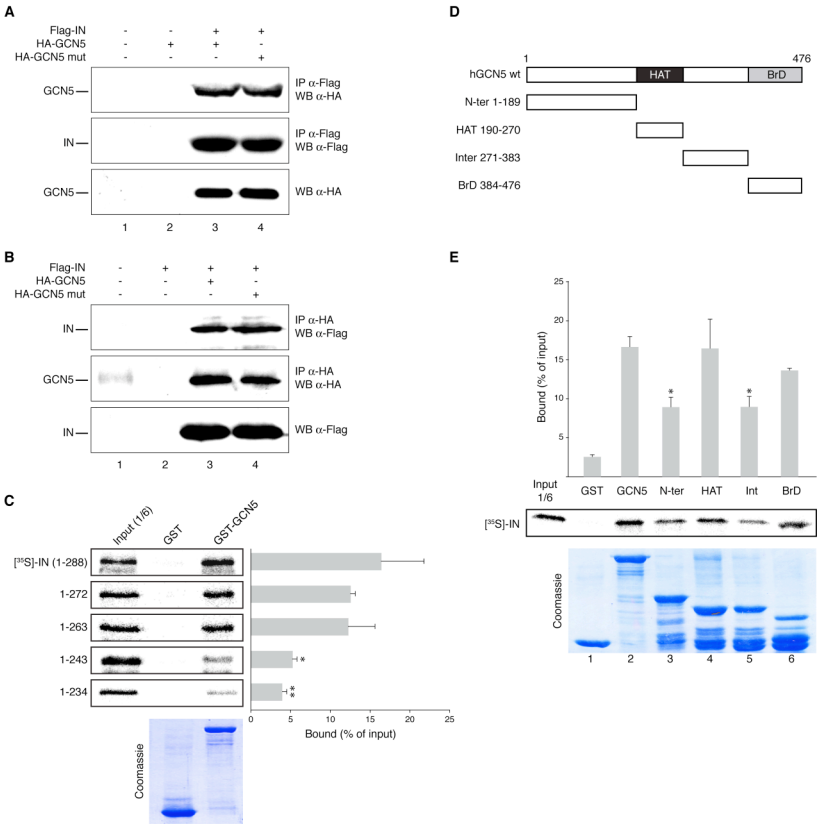

Figure 12. IN interacts with GCN5 both *in vitro* and *in vivo*. (A) Extracts from HEK 293T cells transfected with the indicated plasmids were immunoprecipitated using anti-Flag antibody and analyzed by Western blotting with anti-HA antibody (upper panel) or anti-Flag antibody (middle panel). Lower panel: extracts immunoblotted with anti-HA antibody. (B) Extracts from HEK 293T cells transfected with the indicated plasmids were immunoprecipitated using anti-HA antibody and analyzed by Western blotting with anti-Flag antibody (upper panel) or anti-HA antibody (middle panel). Lower panel: extracts immunoblotted with anti-Flag antibody. (C) Autoradiography and Coomassie Blue staining of *in vitro* binding assays with GST-GCN5 and ^{35}S-IN or the indicated ^{35}S-IN fragments. The histogram represents the results of three independent experiments (means ± SEM), where the amounts of bound proteins are expressed as percentages of the corresponding radiolabeled inputs. Statistical significance of the binding percentages was calculated by using the Student's two-sided *t* test. Asterisks directly above bars indicate differences in binding efficiency to GST-GCN5 between IN deleted forms and full-length IN. **, $P < 0,01$; *, $P < 0,05$. Conversely, where asterisks are not present, values obtained did not significantly differ ($P > 0,05$) from those obtained with full-length IN. (D) Schematic representation of the functional domains of human GCN5 (hGCN5 wt) used for the binding assay reported in (E). N-ter: N-terminal domain; HAT: HAT domain; Inter: Intermediate domain; BrD: Bromodomain. (E) Autoradiography (upper panel) and Coomassie Blue staining (lower panel) of *in vitro* binding assay with ^{35}S-IN and GST-GCN5 or the indicated GST-GCN5 fragments. The histogram represents the results of three independent experiments (means ± SEM), where the amounts of bound IN are expressed as percentages of the input of radiolabeled protein. Statistical significance of the binding percentages was calculated by using the Student's two-sided *t* test. Asterisks directly above bars indicate differences in binding efficiency to IN between GST-GCN5 deleted forms and GST-GCN5 full-length. *, $P < 0,05$. Conversely, where asterisks are not present, values obtained did not significantly differ ($P > 0,05$) from those obtained with GST-GCN5 full-length.

Acetylation by GCN5 enhances IN catalytic activity *in vitro*

Production of constitutively acetylated recombinant IN by tethered catalysis

To explore the effect of GCN5-mediated acetylation on the catalytic activity of IN, constitutively acetylated recombinant IN was produced by exploiting the "tethered catalysis" approach (Acharya et al., 2005; Guo et al., 2004). This method allows the production of a constitutively acetylated protein by tethering the factor of interest to the catalytic domain of a specific HAT enzyme. Based on this approach, as schematized in Figure 13A, a chimeric construct was generated where c.o. IN, tagged at the N-terminus with a 6×His epitope, was fused at its C-terminal end with the region of GCN5 encompassing amino acids 6-300, which includes the catalytic HAT domain. To obtain a control that cannot be acetylated, the same chimera was constructed by using the inactive mutant of GCN5 Y260A/F261A. In addition, a sequence encoding TEV protease recognition site was inserted between IN and GCN5 coding sequences to allow the separation of the two domains. Wild type and mutant GCN5 fusions were expressed in bacteria, purified from total cell lysates by metal ion affinity chromatography, and finally digested with TEV protease. The reaction mixtures were resolved by SDS-PAGE and Coomassie Blue staining of the resulting gel revealed, in the lanes corresponding to proteolytically processed samples, a band migrating at the same size as 6×His-tagged IN, clearly deriving from the removal of GCN5 moiety from the fusion protein (Figure 13B, compare lanes 3 and 5 with lanes 2 and 4). Most importantly, recombinant IN derived from the wild type GCN5 fusion resulted in a remarkable positive signal upon immunoblotting with an anti-acetyl-lysine antibody, while the mutant allele of GCN5 failed to generate a significant level of IN acetylation (Figure 13C, top panel, compare lanes 1 and 3 with lanes 2 and 4). In this experiment, the amounts of loaded proteins were verified by incubating the same blotted membrane with an antibody directed against IN (Figure 13C, bottom panel).

In vitro integration assays with constitutively acetylated recombinant IN

Constitutively acetylated recombinant IN and the non-acetylated control were tested *in vitro* for 3′-end processing and strand transfer activities. In the 3′-end processing reaction, recombinant IN was incubated with a 21-mer, [^{32}P]-labeled double stranded oligonucleotide mimicking one end of the viral DNA (S), in the presence of a divalent metal ion. The reaction mixture was analyzed by gel electrophoresis followed by autoradiography, and the excision of 2 nucleotides from the 3′-end of the labeled strand was evaluated by measuring the radioactive signal deriving from the 19-mer product (P). Acetylated IN (100 or 200 ng, lanes 1 and 3 of Figure 13D, respectively) proved to be two- to three-fold more active than the non-acetylated control (100 or 200 ng, lanes 2 and 4 of Figure 13D, respectively), as revealed by comparative densitometric analysis of the radiolabeled bands corresponding to the 3′-end processed template (Figure 13D, compare lanes 1 and 3 with lanes 2 and 4). IN strand transfer activity was evaluated by using the same [^{32}P]-labeled DNA substrate as in the 3′-end processing assay. Following separation of the reaction products from the substrate by gel electrophoresis and subsequent autoradiography, signals derived from the ladder of higher molecular weight species (P) generated in the reaction, migrating slower than the original 21-mer substrate, were used to quantify IN strand transfer activity. Consistent with the 3′-end processing results, densitometric analysis of the resulting autoradiogram highlighted that constitutively acetylated IN, used at two different doses (100 or 200 ng), was five- to ten-fold more active in catalyzing the strand transfer reaction than the corresponding non-acetylated controls (Figure 13E, compare lanes 1 and 3 with lanes 2 and 4). Taken together, these results indicated that GCN5-mediated acetylation enhances the catalytic activity of IN *in vitro*, both in 3′-end processing and strand transfer reactions.

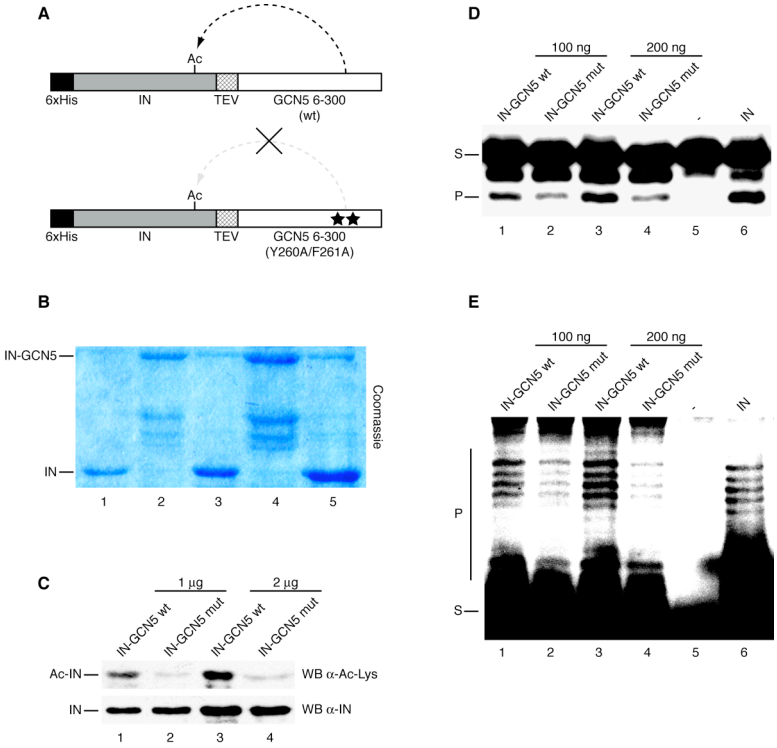

Figure 13. GCN5-mediated acetylation increases the catalytic activity of IN *in vitro*. (A) Schematic representation of IN-GCN5 tethered catalysis constructs. Full-length IN, tagged with a N-terminal 6×His epitope, is fused in frame with TEV proteolytic site and cloned upstream of the 6-300 amino acid region of wild type GCN5 (IN-GCN5 wt) or its catalytically inactive allele (IN-GCN5 mut). (B) Coomassie Blue staining of IN-GCN5 fusion proteins before and after TEV protease-mediated digestion. Lane 1: 6×His-tagged IN; lane 2: IN-GCN5 wt before proteolytic processing; lane 3: products derived from digestion of IN-GCN5 wt with TEV protease; lane 4: IN-GCN5 mut before proteolytic processing; lane 5: products derived from digestion of IN-GCN5 mut with TEV protease. (C) 1 μg and 2 μg of IN derived from IN-GCN5 wt (lanes 1 and 3, respectively), or 1 μg and 2 μg of IN derived from IN-GCN5 mut (lanes 2 and 4, respectively) were analyzed by Western blotting with anti-acetyl-lysine antibody (top panel) or anti-IN antibody (bottom panel). (D) 3′-end processing activity of IN derived from IN-GCN5 wt (lane 1: 100 ng; lane 3: 200 ng) or IN-GCN5 mut (lane 2: 100 ng; lane 4: 200 ng). Lane 5: DNA substrate; lane 6: DNA substrate with 40 ng of 6×His-tagged IN. (E) Strand transfer activity of IN derived from IN-GCN5 wt (lane 1: 100 ng; lane 3: 200 ng) or IN-GCN5 mut (lane 2: 100 ng; lane 4: 200 ng). Lane 5: DNA substrate; lane 6: DNA substrate with 40 ng of 6×His-tagged IN. In (C) and (D), the DNA substrate (S) and the catalytic products (P) are indicated.

HIV-1 infectivity is reduced in GCN5 knockdown cells

In order to evaluate the role of the interaction between IN and GCN5 and the consequent acetylation of IN during the HIV-1 replication cycle, viral infectivity upon GCN5 depletion in target cells was analyzed. Transient knockdown of GCN5 expression was obtained in HeLa cells by transfection with a chemically

synthesized siRNA targeting nucleotides 1364 to 1382 of GCN5 mRNA (siGCN5), while a non-targeting siRNA, unrelated to any human genomic sequence, was used as a negative control (Ctrl siRNA). For stable silencing of GCN5 expression, specific HEK 293T cell lines were generated by the use of lentiviral vector-encoded shRNAmirs. shRNAmirs represent a new generation of shRNAs, modeled after endogenous microRNAs (miRNAs) (Zeng et al., 2002). For our experiments, pGIPZ lentiviral vectors, purchased from Open Biosystems (Huntsville, AL), were employed. Since the addition of sequences from human microRNA-30 (miR30) has been reported to greatly increase the silencing potential of expressed hairpins (Silva et al., 2005), shRNAs are expressed from pGIPZ vectors as human miR30 primary transcripts. The shRNAmir insert thus comprises a hairpin stem made of 22 nt of double stranded DNA, specifically targeting the mRNA of interest, a 19 nt loop derived from human miR30, and 125 nt of the miR30 flanking sequence located on either side of the silencing hairpin (Figure 14A). In addition, the shRNAmir insert is cloned in frame with a puromycin drug resistance marker for the selection of stably transduced cells, and expressed as part of a bicistronic transcript together with TurboGFP, thus allowing the visual marking of shRNAmir-positive cells (Figure 14B).

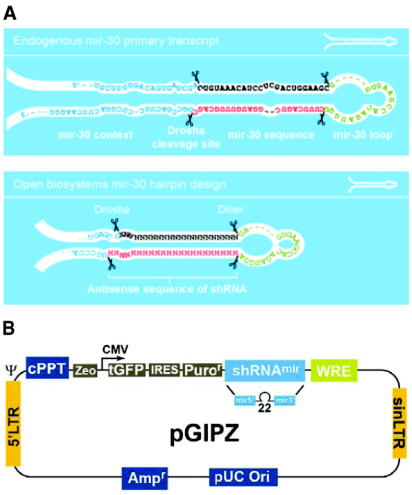

Figure 14. pGIPZ shRNAmir-containing lentiviral vectors. (A) Schematic illustration showing that silencing shRNAs are expressed from pGIPZ lentiviral vectors as human miR30 primary transcripts. (B) Schematic representation of pGIPZ lentiviral vectors, showing their main features.

HEK 293T cells were transduced with a pGIPZ vector encoding a shRNAmir specifically annealing to the 3′ untranslated region (3′ UTR) of GCN5 mRNA (GCN5 shRNAmir) and, after puromycin-mediated selection, the three cell clones with the highest TurboGFP expression levels, as determined by FACS analysis, were chosen. As non-silencing controls, two HEK 293T polyclonal cell lines were developed, one expressing a mismatched, non-targeting GCN5 shRNAmir (GCN5 shRNAmir mut), and the other carrying an empty pGIPZ vector, which lacks a silencing insert. As shown in Figure 15A, Western blot analysis of the cell lysates with an antibody directed against GCN5 revealed that the expression level of the protein was significantly reduced in both transiently and stably silenced cells with respect to controls (top panels), while tubulin expression was similar in all the samples (bottom panels). The efficiency of knockdown was similar in siRNA-treated and shRNAmir-expressing cells. Silenced cells were infected with an *env*-deleted, VSV-G pseudotyped NL4.3 virus expressing the luciferase reporter gene (indicated hereafter as NL4.3-Luc) and luciferase activity in the cell lysates was measured 48 hours after infection. As shown in Figure 15B, a two- to three-fold reduction in luciferase activity was detected in both transiently and stably silenced cells, thus indicating that knockdown of GCN5 expression in target cells reduces HIV-1 infectivity. To determine which step of viral replication was affected by GCN5 depletion, cells were harvested at various time points after infection and measurements of the different HIV-1 DNA species were performed by RT-Q-PCR. Total HIV-1 DNA was quantified with the use of primers annealing to the luciferase reporter gene, in order to avoid cross-reaction with the integrated pGIPZ lentiviral vectors present in stably transduced HEK 293T cell lines. In fact, primers most commonly used for the detection of total HIV-1 DNA anneal in the U5 region of the LTR and the 5′-end of the *gag* gene. Since analogous sequences are present in the pGIPZ lentiviral vectors used for the generation of stable GCN5 knockdown cell clones and the corresponding non-silenced controls, the use of these primers would result in a very high background “noise”, from which the signal deriving from infection with NL4.3-Luc could not be distinguished. As shown in Figure 15C, no significant alterations in total HIV-1 DNA levels were detected in cells either transiently or stably silenced, thus indicating that reverse transcription was not

affected by the reduction of GCN5 expression. To detect integrated HIV-1 DNA, siRNA-treated HeLa cells were analyzed 48 hours post infection by *Alu*-LTR nested PCR (see the Material and Methods chapter for details), while stably transduced HEK 293T cell lines were processed two weeks after infection using primers specific to the luciferase reporter gene. This was necessary in order to dilute non-integrated HIV-1 DNA and avoid cross-reaction with the integrated pGIPZ lentiviral vectors. Proviral DNA was about two-fold less in all GCN5 knockdown cells, either treated with siRNA or transduced with shRNAmir-encoding lentiviral vectors (Figure 15D). Finally, a two-fold increase in the amount of 2-LTR circles was detected in both stably and transiently silenced cells (Figure 15E). Since the increase in 2-LTR circles often correlates with a defect at the step of integration (Engelman, 1999), these data are collectively consistent with reduced integration efficiency in GCN5 knockdown cells. Therefore, we concluded that the presence of GCN5 in target cells enhances HIV-1 infectivity, likely by increasing the efficiency of provirus formation.

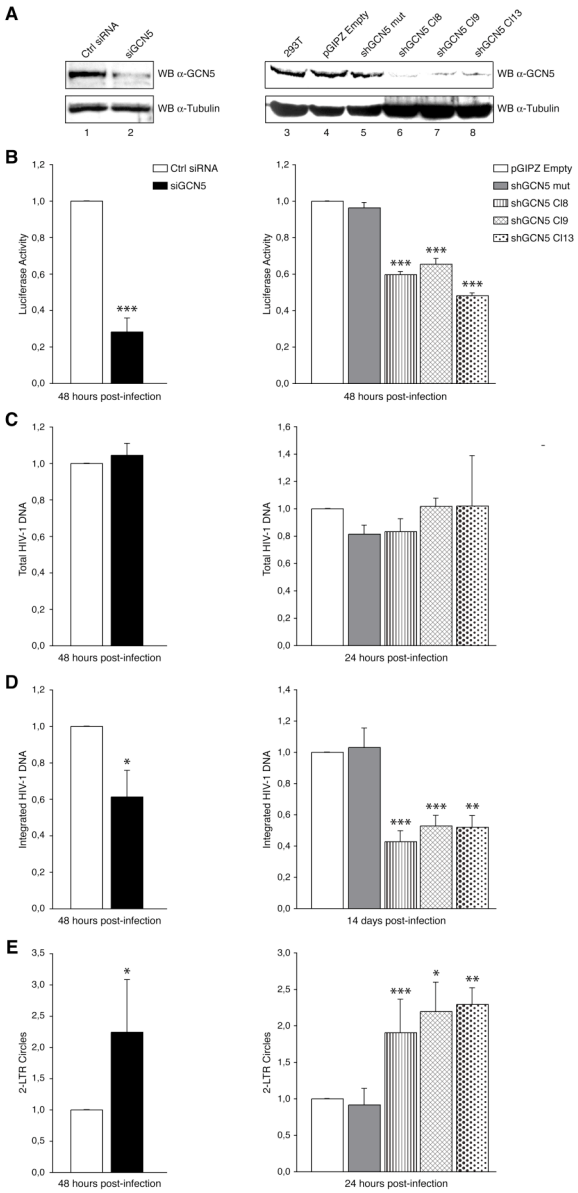

Figure 15. GCN5 depletion in infected cells reduces HIV-1 integration. (A) Left panels: extracts from siRNA-treated Hela cells analyzed by Western blotting with anti-GCN5 antibody (top) or anti-α-tubulin antibody (bottom). Lane 1: cells transfected with non-targeting siRNA (Ctrl siRNA); lane 2: cells transfected with GCN5-targeting siRNA (siGCN5). Right panels: extracts from stable GCN5 knockdown HEK 293T cell clones or control cell lines immunoblotted with anti-GCN5 antibody (top panel) or anti-α-tubulin antibody (bottom panel). Lane 3: untransduced HEK 293T cells; lane 4: HEK 293T cells carrying empty pGIPZ vector; lane 5: HEK 293T cells expressing mutant, non-targeting GCN5 shRNAmir; lanes 6-8: HEK 293T clones (Cl8, Cl9 and Cl13) expressing GCN5 shRNAmir. (B) siRNA-treated Hela cells (left histogram) or HEK 293T cells stably transduced with pGIPZ lentiviral vectors (right histogram) were infected with NL4.3-Luc and analyzed for luciferase activity 48 hours after infection. The histograms represent percentages of luciferase activity relative to control, non-silenced cells. Means ± SEM from three independent experiments are reported. (C-E) Total DNA extracted from siRNA-treated Hela cells (left histograms) or HEK 293T cells stably transduced with pGIPZ lentiviral vectors (right histograms) was analyzed by RT-Q-PCR for total HIV-1 DNA (C), integrated HIV-1 DNA (D) and 2-LTR circles (E). In (C-E), results are presented as percentages relative to control, non-silenced cells. Reported values are means ± SEM from three independent experiments. Statistical significance values shown in (B-E) were calculated by using the Student's two-sided *t* test. Asterisks directly above bars indicate differences between knockdown and control, non-silenced cells. ***, $P < 0,001$; **, $P < 0,01$; *, $P < 0,05$. Conversely, where asterisks are not present, values obtained did not significantly differ ($P > 0,05$) from those obtained with control, non-silenced cells.

Mutations at IN acetylation sites reduce HIV-1 integration efficiency and replication capacity

Reduced integration efficiency of HIV-1 viruses mutated at IN acetylation sites in single round infections

Since GCN5 depletion in target cells results in decreased HIV-1 infectivity, which specifically correlates with less efficient provirus formation, we decided to explore the effects of arginine substitutions introduced in the virus at the GCN5-targeted lysine positions of IN. To this aim, a NL4.3-Luc HIV-1 clone expressing quadruple-mutant IN (IN K258,264,266,273R) was generated (NL4.3-Luc 4mut). In addition, since IN lysines acetylated by GCN5 partially overlap with those targeted by p300, we set out to compare the replication properties of the IN quadruple-mutant virus with those of a corresponding HIV-1 clone carrying lysine-to-arginine substitutions at IN positions 264, 266, and 273, which are acetylated in common by the two HATs (NL4.3-Luc 3mut). Therefore, *env*-deleted, VSV-G pseudotyped NL4.3-Luc viral clones, bearing three or four mutations at the IN acetylation sites, were used to infect HEK 293T cells in single-round replication assays, together with the corresponding virus expressing wild type IN. Forty-eight hours after infection, measurements of luciferase activity in the cell lysates revealed that both mutant viruses were five-fold less infective than wild type (Figure 16A). To determine which step in the viral replication cycle was affected by the mutations introduced in the IN sequence, total DNA was extracted from cells harvested at various time points after infection, and the different HIV-1 DNA species were measured by RT-Q-PCR. By the use of primers annealing to the luciferase reporter gene present in NL4.3-Luc viral clones (wild type or mutated), similar levels of total HIV-1 DNA were detected at 24 hours post infection in cells infected with NL4.3-Luc 3mut or 4mut, as well as with wild type virus (Figure 16B). This finding indicated that the activity of HIV-1 RT was not perturbed by the mutations introduced in the IN sequence, so that the resulting viruses proved able to successfully accomplish reverse transcription. Integrated HIV-1 DNA was quantified at 48 hours post infection by *Alu*-LTR nested PCR (see the Materials and Methods chapter for detail). As a result, a five-fold reduction in the number of proviruses was detected upon infection with both mutant viruses with respect to wild type (Figure 16C).

These data indicated reduced HIV-1 integration efficiency upon mutation of IN lysines targeted by acetylation. Consistently, cells infected with NL4.3-Luc 3mut or 4mut exhibited a three-fold increase in the amount of 2-LTR circles as compared to cells infected with wild type virus (Figure 16D), thus confirming a specific defect in the viral replication cycle at the step of integration and no alterations during PICs nuclear import. Notably, the IN triple- and quadruple-mutant viruses exhibited almost identical reductions in infectivity and integration efficiency with respect to wild type. Therefore, IN lysines 264, 266 and 273, which are acetylated in common by GCN5 and p300, appear to be required for HIV-1 to integrate with maximal efficiency and be fully infectious, while IN lysine 258, although specifically modified by GCN5 *in vitro*, does not seem to play any significant role in the viral replication cycle.

HIV-1 clones mutated at IN acetylation sites exhibit delayed peaks of replication and reduced virus production in multiple round infections

Next, the role of IN acetylated lysines during HIV-1 replication in a T cell line was investigated. To this aim, two NL4.3-derived viral clones were generated, expressing either the triple- or quadruple-mutant IN (NL4.3 3mut and NL4.3 4mut, respectively). One million CEM T cells were infected with the resulting viruses using two different amounts of p24 antigen (10 or 1 ng). HIV-1 replication was monitored by measuring RT activity in the culture supernatants every three days over a period of 21 days. As shown in Figure 16E, cells infected with the higher viral load (10 ng p24 antigen) of wild type virus exhibited a peak of HIV-1 replication around day 9 post infection. Conversely, cells infected with same amounts of NL4.3 3mut or 4mut reached the RT production peak at day 12 post infection, thus with three days of delay as compared to cells infected with wild type virus. Notably, in these experimental conditions, the RT amounts produced by both mutant HIV-1 clones at the infectivity peaks were approximately half of the one obtained upon infection with the wild type virus. By using the lower viral load (1 ng p24 antigen), the replication curve of wild type NL4.3 started to raise quite steeply around day 12 post infection, while, for both mutant viruses, the replication curves started to appear at day 15. Detectable RT production was observed for both

NL4.3 3mut and 4mut at day 18, thus with 6 days of delay as compared to wild type virus (Figure 16F). In conclusion, as determined by tracking viral growth through culture, viruses carrying mutations at the IN acetylation sites exhibited delayed peaks of replication with respect to wild type in both the experimental conditions used (10 or 1 ng p24 antigen per 1×10^6 CEM T cells), as well as half reduced RT production at the infectivity peak upon infection with the higher viral load. Similarly to the results obtained in single-round infections, mutations introduced in the virus at IN sites targeted in common by GCN5 and p300 (K264, K266 and K273), or additional mutation at lysine 258, specifically acetylated by GCN5 *in vitro*, reduced the replication capacity of the virus to the same extent.

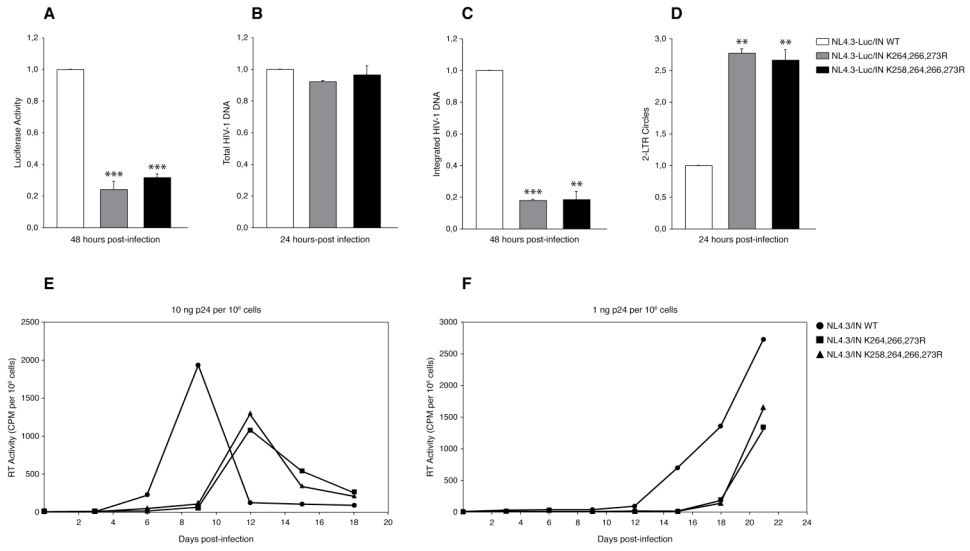

Figure 16. Mutations at IN acetylation sites in the virus cause a replication defect at the step of integration. (A) HEK 293T cells infected with NL4.3-Luc/IN WT, NL4.3-Luc/IN K264,266,273R, or NL4.3-Luc/IN K258,264,266,273R were analyzed for luciferase activity 48 hours after infection. (B-D) Total DNA extracted from HEK 293T cells infected with the same viral clones as in (A) was analyzed by RT-Q-PCR for total HIV-1 DNA at 24 hours after infection (B), integrated HIV-1 DNA at 48 hours after infection (C) and 2-LTR circles at 24 hours after infection (D). In (A-D), results are presented as percentages relative to cells infected with NL4.3-Luc/IN WT. Reported values are means ± SEM from three independent experiments. Statistical significance values shown in (A-D) were calculated by using the Student's two-sided *t* test. Asterisks directly above bars indicate differences between cells infected with mutant viruses and cells infected with wild type virus. ***, $P < 0{,}001$; **, $P < 0{,}01$. Conversely, where asterisks are not present, values obtained did not significantly differ ($P > 0{,}05$) from those obtained with cells infected with wild type virus. (E) RT activity detected in the culture supernatants of CEM cells at different time points after infection with NL4.3/IN WT, NL4.3/IN K264,266,273R, or NL4.3/IN K258,264,266,273R. (F) Infections performed as in (E), using 10-fold lower viral loads.

CHAPTER 5

DISCUSSION

Acetylation of the IN CTD by cellular HATs

After the finding that IN is a substrate for p300-mediated acetylation and the identification of three lysines in the IN CTD that are targeted by this PTM (Cereseto et al., 2005; Topper et al., 2007), the results of the present study reveal that GCN5 is an as-yet-unidentified cellular HAT catalyzing the acetylation of IN both *in vitro* and *in vivo*. Notably, the IN region acetylated by GCN5 partially overlaps with the one targeted by p300 in the CTD of the protein. Indeed, IN lysines 264, 266 and 273 appear to be common substrates for the two HATs. However, *in vitro* mutagenesis experiments performed here highlighted that recombinant GCN5 is able to acetylate an additional lysine residue in the CTD of IN, located at position 258. In fact, lysine-to-arginine substitutions at IN positions 264, 266 and 273 almost abolished p300-mediated acetylation *in vitro* (Cereseto et al., 2005; Topper et al., 2007). Conversely, *in vitro* acetylation assays performed with recombinant GCN5 revealed that IN K264,266,273R was acetylated to a similar extent as the single-mutated IN forms, while the additional lysine-to-arginine substitution introduced at position 258 significantly reduced the acetylation level of the protein (see Figure 10A-B in the Results chapter). From these data we concluded that four lysines in the CTD of IN, located at positions 258, 264, 266 and 273, are acetylated by recombinant GCN5 *in vitro*. Therefore, specific targeting of lysine 258 by GCN5 appears to distinguish the *in vitro* catalytic activities of the two HATs on IN.

The results of *in vitro* acetylation assays, showing that IN is an efficient substrate for both GCN5- and p300-catalyzed acetylation, were validated *in vivo* by exogenously expressing IN and the putative acetylating enzyme, either GCN5 or p300, in HEK 293T cells. Notably, the experiments performed in this study go beyond the analysis of the IN acetylation level in cell extracts with an anti-acetylated lysine antibody. In fact, by the use of an antibody developed *ad hoc* for the recognition of IN acetylated at lysines 264, 266 and 273, we were able to extend our *in vitro* observations about the specificity of the IN acetylation reactions catalyzed by GCN5 and p300 to *in vivo* circumstances, demonstrating that the three residues at positions 264, 266 and 273 are specifically targeted by both HATs also, and most importantly, in a physiological context (see Figure 11C-D in the Results chapter). This finding is of particular relevance, since it represents the

first demonstration that the IN CTD is acetylated in the cells at lysines 264, 266 and 273 by two different HATs, GCN5 and p300. A similar situation, where different cellular HATs modify the same sites on the target protein, has been reported for another HIV-1 factor, Tat, which is acetylated at lysines 50 and 51 by p300/CBP and GCN5, leading in both cases to an enhancement of Tat-dependent transcriptional activation of the viral LTR promoter (Col et al., 2001; Deng et al., 2000; Kiernan et al., 1999; Ott et al., 1999). The finding that, although divergent in sequence, belonging to different HAT families and involved in distinct signaling pathways, GCN5 and p300 converge to acetylate IN on the same sites may be attributed to the importance of the acetylation of lysines 264, 266 and 273 in the function of IN, since it would provide a "fail-safe" mechanism capable of maintaining adequate levels of IN acetylation at these key sites under varying cellular conditions. Accordingly, the fact that the three lysines, which are common substrates for acetylation by GCN5 and p300, are highly conserved among different HIV-1 strains (Lu et al., 2005) also suggests that their preservation is important for some aspect of viral replication.

IN acetylated lysines are located within the HAT-interacting region

Co-immunoprecipitation experiments performed in the present study revealed that IN and GCN5 specifically interact in the cells (see Figure 12A-B in the Results chapter). *In vitro* binding assays with GST-GCN5 and IN fragments containing progressive deletions starting from the C-terminus allowed the mapping of the region involved in the interaction with the HAT enzyme between residue 244 and the C-terminus of the viral protein (see Figure 12C in the Results chapter). Therefore, the portion of IN involved in the interaction with GCN5 also contains the acetylated residues (located at positions 258, 264, 266 and 273). This is true also in the case of p300, since it directly binds to IN within a region located in the CTD of the protein between amino acids 264 and 288 (Cereseto et al., 2005) and acetylates three residues, at positions 264, 266 and 273, that are included in this same region (Cereseto et al., 2005; Topper et al., 2007). Therefore, previous and present studies highlighted that, for both GCN5 and p300, the portion of IN

involved in the binding to the HAT enzyme also comprises the residues modified by acetylation.

Interestingly, our mapping of the IN regions interacting with the HAT enzymes, as a result of *in vitro* binding assays, is consistent with a recent report presenting two three-dimensional models of full-length IN complexed with GCN5 and p300 (Di Fenza et al., 2009). Both models predict that the IN C-terminal tail located between amino acids 271 and 288, due to its high flexibility, could easily adapt to the binding pocket of GCN5, as well as to that of p300 (Figure 17). Interestingly, IN lysine 273 is included in this unstructured region, and is therefore expected to be the residue most prone to acetylation. In fact, since IN lysines 264 and 266 are located in close proximity to a sandwich of two three-stranded antiparallel β sheets, their binding and acetylation would require a more complex unfolding of this stable secondary structure. Based on this model, we may hypothesize IN lysine 273 to be the first residue contacted and acetylated by the HAT enzyme, either GCN5 or p300. This event might in turn induce a conformational change in the C-terminal portion of IN, which could facilitate the modification of the other two lysines (as will be discussed later in this paragraph). This hypothesis is also compatible with the data reported by Topper and coworkers, demonstrating a hierarchy of reactivity between the three residues modified by p300, with lysine 273 as the primary target for acetylation (Topper et al., 2007).

The finding that the IN region interacting with GCN5 also contains the acetylated lysines is particularly interesting, especially when taking into consideration the results of our *in vitro* binding experiments aimed at the determination of the IN-interacting domains of GCN5. In fact, the HAT domain and bromodomain of GCN5 proved to be mainly responsible for the interaction with the C-terminal portion of IN (see Figure 12E in the Results chapter). Given the fact that the bromodomain functions as a protein module that specifically binds to acetyl-lysine-containing motifs (Yang, 2004a), the involvement of GCN5 bromodomain in the interaction with the IN region in which the acetylation sites are located raises the possibility that the IN acetylated residues could provide a recognition motif for the bromodomain of GCN5. A possible scenario is thus that, following acetylation of one or more residues on IN, the interaction between GCN5 bromodomain and the

acetylated lysines of IN provides additional links that strengthen the interaction between the two proteins. Several examples of such a mechanism have been reported in literature. For instance, the myogenic transcription factor MyoD is acetylated by p300/CBP and PCAF at lysines 99 and 102. MyoD acetylated residues are then recognized by the bromodomain of CBP, resulting in a strengthened interaction between the HAT and MyoD itself. This event finally results in a more efficient recruitment of HAT complexes to muscle-specific promoters, on which they acetylate other substrates, such as histones (Polesskaya et al., 2001). Given the hierarchy of reactivity that was established between IN lysines 264, 266 and 273, based both on experimental data (Topper et al., 2007) and three-dimensional models (Di Fenza et al., 2009), with residue at position 273 being the primary target for acetylation, the modification of IN at this site may trigger the recognition of the acetylated residue by the bromodomain of GCN5 and affect the interaction between the two proteins in such a way that promotes the modification of the other target lysines in the IN CTD. This kind of mechanism may also fit in with p300, since our previous experiments (data not shown) revealed that the p300 bromodomain is also involved in the interaction with the C-terminal portion of IN containing the acetylated lysines at positions 264, 266 and 273.

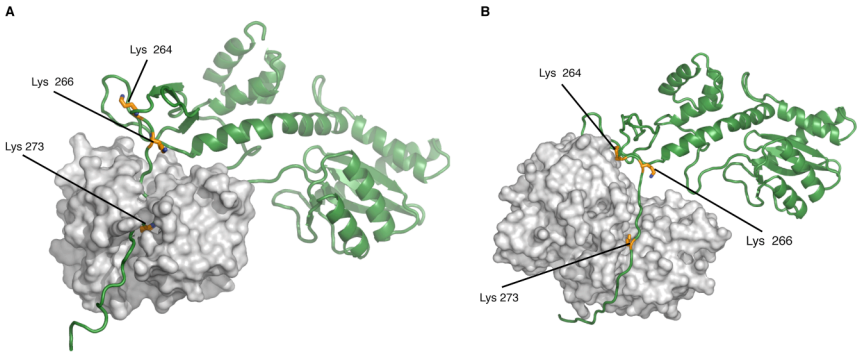

Figure 17. Three-dimensional models of IN complexes with GCN5 and p300. (A) Three-dimensional model of the IN/GCN5 complex. IN is represented in green and GCN5 in light grey. (B) Three-dimensional model of the IN/p300 complex. IN is represented in green and p300 in light grey. In (A) and (B), the three lysine residues in the CTD of IN that are acetylated by both GCN5 and p300 (Lys 264, Lys 266 and Lys 273) are shown in yellow. GCN5 and p300 are rendered as surfaces, while IN as a cartoon to highlight the C-terminal unfolded portion, which inserts in the binding pockets of the two HATs.

Comparative analysis of GCN5 and p300 function in HIV-1 integration

Since IN lysines 264, 266 and 273 are common substrates for GCN5 and p300 in *in vitro* acetylation assays, while lysine 258 appears to be a GCN5-specific site of modification, we performed a comparative analysis aimed at establishing the roles of the two HATs during the HIV-1 replication cycle. Viral clones mutated at the IN sites targeted in common by GCN5 and p300 (K264, K266 and K273), or carrying the additional lysine-to-arginine substitution at position 258 were generated. The resulting mutant viruses were used, together with their wild type counterpart, either to infect HEK 293T cells in single-round viral replication assays, or for multiple-round infections of CEM T cells. In all the experimental conditions tested, both mutant viruses proved to be less infective than the wild type. Quantification of the different HIV-1 DNA species in HEK 293T infected cells by RT-Q-PCR revealed that the infectivity decrease specifically correlated with less efficient provirus formation by the mutant viral clones as compared to wild type (see Figure 16A-D in the Results chapter). Notably, infections with IN triple- or quadruple-mutant viruses resulted in similar reductions in the levels of HIV-1 integrated DNA with respect to wild type, leading to identical decreases in infectivity. These results indicated that acetylation of IN C-terminal lysines 264, 266 and 273, which are substrates for both GCN5 and p300, is required for the virus to integrate with maximal efficiency and be fully infectious, while acetylation of IN lysine 258, although specifically catalyzed by GCN5 *in vitro*, does not seem to play any significant role during the HIV-1 replication cycle, since the additional amino acidic substitution introduced at this site in the context of the triple-mutant virus does not furtherly affect its replication capacity. However, it cannot be excluded that the apparent discrepancy between our *in vitro* and *in vivo* results (i.e. *in vitro* detection of GCN5-mediated acetylation of IN at lysine 258, with no functional outcome *in vivo*) may be due to an intrinsic limitation of *in vitro* acetylation assays performed with recombinant HAT enzymes. In fact, in such experiments, the activity of HATs is investigated as they would exist in the cell as individual proteins, while it is well established that HAT enzymes are subunits of large multiprotein complexes, and do not seem to exist individually in the cell. Moreover, it has been demonstrated that the different subunits of distinct

HAT-containing complexes can change the substrate specificity of the respective HAT enzymes. Therefore, it seems conceivable that the affinity towards the IN C-terminal lysines of the cellular GCN5-containing complexes may be slightly different from that of the isolated recombinant protein *in vitro*.

Notably, even if the comparative analysis of viral clones mutated at the IN sites acetylated by GCN5 or p300 points to redundant functions of the two HATs in regulating the activity of IN, the analysis of viral replication upon GCN5 depletion in target cells suggests that GCN5 and p300 play partially overlapping, yet distinct roles during the step of integration. In fact, the reduced infectivity, paralleled by lower provirus formation, which was detected upon infection of GCN5 knockdown cells (see Figure 15B and D in the Results chapter), indicates that endogenous p300 is not able to compensate for the loss of GCN5 so as to allow the virus to integrate with maximal efficiency and fully restore the sensitivity of GCN5-depleted cells to HIV-1 infection. The view of GCN5 and p300 having partially overlapping, yet not completely redundant roles in regulating IN function is also supported by the results of our *in vitro* integration assays, showing that GCN5-mediated acetylation enhances the catalytic activity of IN both in 3′-end processing and strand transfer reactions, while the acetylation of IN catalyzed by p300 stimulates the strand transfer, but doesn't affect the 3′-end processing activity of the viral enzyme (Cereseto et al., 2005).

The reduced integration efficiency of HIV-1 clones mutated at IN acetylated lysines is due to a lack of acetylation

The decreased HIV-1 infectivity detected in GCN5 knockdown cells, which is paralleled by a similar reduction in the level of integrated HIV-1 DNA (see Figure 15B and D in the Results chapter), indicates that the integration defect observed upon mutation of the IN acetylated lysines in the virus results from a loss of acetylation, rather than from functional changes in the protein induced by the amino acidic substitutions. In fact, the replication properties of wild type NL4.3-Luc virus in cells where GCN5 expression has been transiently or stably silenced closely resemble those of NL4.3-Luc 3mut and 4mut viral clones in non-silenced

cells, in the sense that, in both the experimental conditions, the replication cycle of HIV-1 exhibits a specific defect at the step of integration, while reverse transcription and PICs nuclear import remain unaffected, as revealed by RT-Q-PCR detection of the different HIV-1 DNA species at various time points after infection. In this respect, it is conceivable that the integration defect observed upon mutation of the IN acetylation sites in the virus is more severe than the one observed upon GCN5 depletion in target cells (five- against two-fold reduction in provirus formation), since lysine-to-arginine substitution precludes the acetylation of the target sites by any cellular HAT, while in GCN5 knockdown cells other HATs may, at least in part, substitute for the lack of GCN5 activity. The first candidate suited to compensate for the loss of GCN5 is, of course, p300, since we have shown that IN lysines 264, 266 and 273 are common substrates for the two HATs; however, a participation of PCAF in substituting for the lack of GCN5 cannot be excluded. In fact, GCN5 and PCAF are highly homologous HATs and an upregulation of PCAF expression in response to the loss of GCN5 has been documented in different cell lines (Kikuchi et al., 2005) (Nagy et al., 2010), suggesting that when GCN5 is inactivated, cells try to compensate the loss of its activity with that of its paralogue.

A recent debate: the role of IN acetylation during the replication cycle of HIV-1

The results of our *in vitro* integration assays demonstrate that GCN5-mediated acetylation of IN increases the catalytic activity of the protein, both in 3′-end processing and strand transfer reactions (see Figure 13D-E in the Results chapter). p300-catalyzed acetylation also enhances IN strand transfer activity, as reported in our previous work (Cereseto et al., 2005). This is not surprising, since IN lysines at positions 264, 266 and 273 are common substrates for the two HATs. Therefore, our previous and present studies suggest that one of the mechanisms through which acetylation of IN C-terminal lysines promotes HIV-1 integration is by enhancing the catalytic activity of the viral enzyme.

While it is clear that the acetylation of IN stimulates its catalytic activity *in vitro*, the role of this PTM during the HIV-1 replication cycle has been the subject of a recent debate. Our previous study showed that the replication level of a HIV-1_{BRU} clone

expressing a Flag-tagged triple-mutant IN (Flag-IN K264,266,273R) was severely impaired, and the replication deficiency was specifically due to a block at the integration step (Cereseto et al., 2005). A subsequent report by Topper and coworkers (Topper et al., 2007) presented contrasting results. The authors found that the virus expressing untagged triple-mutant IN was fully replication competent, exhibiting only a slight integration defect. In fact, as determined by a genetic integration assay based on the stable transduction of a dominant selectable marker, the integration frequency of the IN triple-mutant virus was about 60% of that of wild type, which did not preclude sustained viral replication in culture. Additional data on the topic came from a report by Apolonia and coworkers (Apolonia et al., 2007). Here, the integration efficiency of a HIV-derived lentiviral vector carrying lysine-to-arginine substitutions at IN positions 264, 266 and 273 was evaluated, in the context of a study aimed at the development of integration-deficient lentiviral vectors to be used as gene therapy tools. The IN K264,266,273R mutant vector exhibited a five-fold lower transgene expression than its wild type counterpart at fourteen days after transduction, indicating reduced viral DNA integration into the host cell genome. Quantification of the IN triple-mutant vector background integration by the same genetic assay used by Topper and coworkers revealed a fourteen-fold lower residual integration rate than the wild type vector. In the present study, we sought to clarify the role of IN acetylation during the HIV-1 replication cycle. To this aim, we performed single- and multiple-round infections with HIV-1 clones encoding IN either mutated at the lysine positions targeted in common by GCN5 and p300 (IN K264,266,273R), or carrying the additional lysine-to-arginine substitution at the site specifically modified by GCN5 *in vitro* (IN K258,264,266,273R). In multiple-round replication assays, the mutant HIV-1 clones exhibited delayed infectivity peaks and reduced virus production with respect to wild type (see Figure 16E-F in the Results chapter). The discrepancy of these findings with the data reported by Topper and coworkers (Topper et al., 2007) may be partially due to the different time-courses of analyses: although working in the same experimental conditions (10 ng of p24 antigen per 1×10^6 CEM T cells), detection of RT activity in the culture supernatants over a period of twenty-one days allowed us to monitor the peak of viral replication, while Topper and

coworkers terminated the replication curve before the highest point of infectivity was reached (at twelve days post infection). However, in our previous report (Cereseto et al., 2005), the introduction of lysine-to-arginine substitutions at IN positions 264, 266 and 273 in the virus resulted in a more severe replication defect than the one observed here. The different extent to which HIV-1 replication was affected by the mutations at IN acetylation sites in our previous and present studies may be likely due to the fact that the earlier finding was obtained by using viral clones containing a Flag-tag epitope fused downstream and in frame with the C-terminus of IN, as suggested in the report by Topper *et al* (Topper et al., 2007).

In single-round viral replication assays, consistent with the transgene expression data reported by Apolonia and coworkers (Apolonia et al., 2007), both the IN triple- and quadruple-mutant viruses exhibited a five-fold decrease in infectivity, which was paralleled by a five-fold reduction in the number of proviruses, as detected by RT-Q-PCR (see Figure 16 A and C in the Results chapter).

In our opinion, the results obtained in this study, together with those presented in previous reports, support the notion of the IN CTD acetylation by cellular HATs representing a mechanism through which the efficiency of HIV-1 integration is modulated, contributing to create the appropriate conditions for the virus to reach full infectivity, rather than serving as a simple “on/off” switch. The relevance of IN acetylation to HIV-1 infectivity has been recently highlighted by the finding that the cellular protein KAP1 specifically interacts with acetylated IN leading to HDAC1 recruitment and IN deacetylation (Allouch et al., 2011). Kap1 complex formation with IN ultimately determines integration impairment, thus suggesting a role for this host factor in an innate restriction mechanism in human cells that targets HIV-1 replication by specifically affecting the efficiency of the integration process. The development of a cellular defense mechanism targeting acetylated IN further reinforces the role of this PTM during HIV-1 infection.

CHAPTER 6

BIBLIOGRAPHY

Abbas, A.K., Lichtman, A.H., Pillai, S. (2010) *Cellular and Molecular Immunology Updated 6E*. Elsevier Inc.

Acharya, A., Xu, X.J., Husain-Ponnampalam, R.D., Hoffmann-Benning, S. and Kuo, M.H. (2005) Production of constitutively acetylated recombinant p53 from yeast and Escherichia coli by tethered catalysis. *Protein Expr Purif*, **41**, 417-425.

Adamson, C.S. and Freed, E.O. (2007) Human immunodeficiency virus type 1 assembly, release, and maturation. *Adv Pharmacol*, **55**, 347-387.

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.

Alfonso, P., Quetglas, J.I., Escribano, J.M. and Alonso, C. (2007) Protein pE120R of African swine fever virus is post-translationally acetylated as revealed by post-source decay MALDI mass spectrometry. *Virus Genes*, **35**, 81-85.

Alkhatib, G., Combadiere, C., Broder, C.C., Feng, Y., Kennedy, P.E., Murphy, P.M. and Berger, E.A. (1996) CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science*, **272**, 1955-1958.

Allouch, A. and Cereseto, A. (2011) Identification of cellular factors binding to acetylated HIV-1 integrase. *Amino Acids*, **41**, 1137-1145.

Allouch, A., Di Primio, C., Alpi, E., Lusic, M., Arosio, D., Giacca, M. and Cereseto, A. (2011) The TRIM family protein KAP1 inhibits HIV-1 integration. *Cell Host Microbe*, **9**, 484-495.

Anafi, M., Yang, Y.F., Barlev, N.A., Govindan, M.V., Berger, S.L., Butt, T.R. and Walfish, P.G. (2000) GCN5 and ADA adaptor proteins regulate triiodothyronine/GRIP1 and SRC-1 coactivator-dependent gene activation by the human thyroid hormone receptor. *Mol Endocrinol*, **14**, 718-732.

Ao, Z., Fowke, K.R., Cohen, E.A. and Yao, X. (2005) Contribution of the C-terminal tri-lysine regions of human immunodeficiency virus type 1 integrase for efficient reverse transcription and viral DNA nuclear import. *Retrovirology*, **2**, 62.

Ao, Z., Huang, G., Yao, H., Xu, Z., Labine, M., Cochrane, A.W. and Yao, X. (2007) Interaction of human immunodeficiency virus type 1 integrase with

cellular nuclear import receptor importin 7 and its impact on viral replication. *J Biol Chem*, **282**, 13456-13467.

Apolonia, L., Waddington, S.N., Fernandes, C., Ward, N.J., Bouma, G., Blundell, M.P., Thrasher, A.J., Collins, M.K. and Philpott, N.J. (2007) Stable gene transfer to muscle using non-integrating lentiviral vectors. *Mol Ther*, **15**, 1947-1954.

Appella, E. and Anderson, C.W. (2000) Signaling to p53: breaking the posttranslational modification code. *Pathol Biol (Paris)*, **48**, 227-245.

Arany, Z., Newsome, D., Oldread, E., Livingston, D.M. and Eckner, R. (1995) A family of transcriptional adaptor proteins targeted by the E1A oncoprotein. *Nature*, **374**, 81-84.

Ariumi, Y., Serhan, F., Turelli, P., Telenti, A. and Trono, D. (2006) The integrase interactor 1 (INI1) proteins facilitate Tat-mediated human immunodeficiency virus type 1 transcription. *Retrovirology*, **3**, 47.

Armon-Omer, A., Graessmann, A. and Loyter, A. (2004) A synthetic peptide bearing the HIV-1 integrase 161-173 amino acid residues mediates active nuclear import and binding to importin alpha: characterization of a functional nuclear localization signal. *J Mol Biol*, **336**, 1117-1128.

Balasubramanian, R., Pray-Grant, M.G., Selleck, W., Grant, P.A. and Tan, S. (2002) Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. *J Biol Chem*, **277**, 7989-7995.

Bannister, A.J. and Kouzarides, T. (1996) The CBP co-activator is a histone acetyltransferase. *Nature*, **384**, 641-643.

Bannister, A.J., Miska, E.A., Gorlich, D. and Kouzarides, T. (2000) Acetylation of importin-alpha nuclear import factors by CBP/p300. *Curr Biol*, **10**, 467-470.

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., Rozenbaum, W. and Montagnier, L. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, **220**, 868-871.

Beitzel, B. and Bushman, F. (2003) Construction and analysis of cells lacking the HMGA gene family. *Nucleic Acids Res*, **31**, 5025-5032.

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.

Bergel, M., Herrera, J.E., Thatcher, B.J., Prymakowska-Bosak, M., Vassilev, A., Nakatani, Y., Martin, B. and Bustin, M. (2000) Acetylation of novel sites in the nucleosomal binding domain of chromosomal protein HMG-14 by p300 alters its interaction with nucleosomes. *J Biol Chem*, **275**, 11514-11520.

Berger, E.A., Doms, R.W., Fenyo, E.M., Korber, B.T., Littman, D.R., Moore, J.P., Sattentau, Q.J., Schuitemaker, H., Sodroski, J. and Weiss, R.A. (1998) A new classification for HIV-1. *Nature*, **391**, 240.

Berger, S.L. (2007) The complex language of chromatin regulation during transcription. *Nature*, **447**, 407-412.

Berger, S.L., Pina, B., Silverman, N., Marcus, G.A., Agapite, J., Regier, J.L., Triezenberg, S.J. and Guarente, L. (1992) Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains. *Cell*, **70**, 251-265.

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.

Berkowitz, R., Fisher, J. and Goff, S.P. (1996) RNA packaging. *Curr Top Microbiol Immunol*, **214**, 177-218.

Bernassola, F., Salomoni, P., Oberst, A., Di Como, C.J., Pagano, M., Melino, G. and Pandolfi, P.P. (2004) Ubiquitin-dependent degradation of p73 is inhibited by PML. *J Exp Med*, **199**, 1545-1557.

Berry, C., Hannenhalli, S., Leipzig, J. and Bushman, F.D. (2006) Selection of target sites for mobile DNA integration in the human genome. *PLoS Comput Biol*, **2**, e157.

Bex, F., Murphy, K., Wattiez, R., Burny, A. and Gaynor, R.B. (1999) Phosphorylation of the human T-cell leukemia virus type 1 transactivator tax on adjacent serine residues is critical for tax activation. *J Virol*, **73**, 738-745.

Bieniasz, P.D. (2009) The cell biology of HIV-1 virion genesis. *Cell Host Microbe*, **5**, 550-558.

Bieniasz, P.D., Grdina, T.A., Bogerd, H.P. and Cullen, B.R. (1999) Recruitment of cyclin T1/P-TEFb to an HIV type 1 long terminal repeat promoter proximal RNA target is both necessary and sufficient for full activation of transcription. *Proc Natl Acad Sci U S A*, **96**, 7791-7796.

Bisgrove, D., Lewinski, M., Bushman, F. and Verdin, E. (2005) Molecular mechanisms of HIV-1 proviral latency. *Expert Rev Anti Infect Ther*, **3**, 805-814.

Blander, G., Zalle, N., Daniely, Y., Taplick, J., Gray, M.D. and Oren, M. (2002) DNA damage-induced translocation of the Werner helicase is regulated by acetylation. *J Biol Chem*, **277**, 50934-50940.

Boeke, J.D. and Devine, S.E. (1998) Yeast retrotransposons: finding a nice quiet neighborhood. *Cell*, **93**, 1087-1089.

Boese, A., Sommer, P., Gaussin, A., Reimann, A. and Nehrbass, U. (2004) Ini1/hSNF5 is dispensable for retrovirus-induced cytoplasmic accumulation of PML and does not interfere with integration. *FEBS Lett*, **578**, 291-296.

Bonaldi, T., Talamo, F., Scaffidi, P., Ferrera, D., Porto, A., Bachi, A., Rubartelli, A., Agresti, A. and Bianchi, M.E. (2003) Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion. *Embo J*, **22**, 5551-5560.

Borrow, J., Stanton, V.P., Jr., Andresen, J.M., Becher, R., Behm, F.G., Chaganti, R.S., Civin, C.I., Disteche, C., Dube, I., Frischauf, A.M., Horsman, D., Mitelman, F., Volinia, S., Watmore, A.E. and Housman, D.E. (1996) The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein. *Nat Genet*, **14**, 33-41.

Boudreault, A.A., Cronier, D., Selleck, W., Lacoste, N., Utley, R.T., Allard, S., Savard, J., Lane, W.S., Tan, S. and Cote, J. (2003) Yeast enhancer of polycomb defines global Esa1-dependent acetylation of chromatin. *Genes Dev*, **17**, 1415-1428.

Boyd, J.M., Subramanian, T., Schaeper, U., La Regina, M., Bayley, S. and Chinnadurai, G. (1993) A region in the C-terminus of adenovirus 2/5 E1a protein is required for association with a cellular phosphoprotein and

important for the negative modulation of T24-ras mediated transformation, tumorigenesis and metastasis. *Embo J*, **12**, 469-478.

Boyes, J., Byfield, P., Nakatani, Y. and Ogryzko, V. (1998) Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature*, **396**, 594-598.

Brand, M., Moggs, J.G., Oulad-Abdelghani, M., Lejeune, F., Dilworth, F.J., Stevenin, J., Almouzni, G. and Tora, L. (2001) UV-damaged DNA-binding protein in the TFTC complex links DNA damage recognition to nucleosome acetylation. *Embo J*, **20**, 3187-3196.

Brand, M., Yamamoto, K., Staub, A. and Tora, L. (1999) Identification of TATA-binding protein-free TAFII-containing complex subunits suggests a role in nucleosome acetylation and signal transduction. *J Biol Chem*, **274**, 18285-18289.

Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Xavier, R.J., Lieberman, J. and Elledge, S.J. (2008) Identification of host proteins required for HIV infection through a functional genomic screen. *Science*, **319**, 921-926.

Broder, S. and Gallo, R.C. (1984) A pathogenic retrovirus (HTLV-III) linked to AIDS. *N Engl J Med*, **311**, 1292-1297.

Brown, P.O., Bowerman, B., Varmus, H.E. and Bishop, J.M. (1989) Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. *Proc Natl Acad Sci U S A*, **86**, 2525-2529.

Brownell, J.E. and Allis, C.D. (1995) An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in Tetrahymena macronuclei. *Proc Natl Acad Sci U S A*, **92**, 6364-6368.

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y. and Allis, C.D. (1996) Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*, **84**, 843-851.

Bujacz, G., Alexandratos, J., Qing, Z.L., Clement-Mella, C. and Wlodawer, A. (1996) The catalytic domain of human immunodeficiency virus integrase: ordered active site in the F185H mutant. *FEBS Lett*, **398**, 175-178.

Bujacz, G., Alexandratos, J., Wlodawer, A., Merkel, G., Andrake, M., Katz, R.A. and Skalka, A.M. (1997) Binding of different divalent cations to the active site of avian sarcoma virus integrase and their effects on enzymatic activity. *J Biol Chem*, **272**, 18161-18168.

Bukrinsky, M., Sharova, N. and Stevenson, M. (1993a) Human immunodeficiency virus type 1 2-LTR circles reside in a nucleoprotein complex which is different from the preintegration complex. *J Virol*, **67**, 6863-6865.

Bukrinsky, M.I., Sharova, N., McDonald, T.L., Pushkarskaya, T., Tarpley, W.G. and Stevenson, M. (1993b) Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc Natl Acad Sci U S A*, **90**, 6125-6129.

Bulinski, J.C. (2007) Microtubule modification: acetylation speeds anterograde traffic flow. *Curr Biol*, **17**, R18-20.

Buratowski, S. (2003) The CTD code. *Nat Struct Biol*, **10**, 679-680.

Bushman, F.D. (1994) Tethering human immunodeficiency virus 1 integrase to a DNA site directs integration to nearby sequences. *Proc Natl Acad Sci U S A*, **91**, 9233-9237.

Bushman, F.D. (2001) *Lateral DNA transfer: Mechanisms and consequences*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Bushman, F.D. (2003) Targeting survival: integration site selection by retroviruses and LTR-retrotransposons. *Cell*, **115**, 135-138.

Bushman, F.D. and Craigie, R. (1992) Integration of human immunodeficiency virus DNA: adduct interference analysis of required DNA sites. *Proc Natl Acad Sci U S A*, **89**, 3458-3462.

Bushman, F.D., Engelman, A., Palmer, I., Wingfield, P. and Craigie, R. (1993) Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. *Proc Natl Acad Sci U S A*, **90**, 3428-3432.

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

Busschots, K., Voet, A., De Maeyer, M., Rain, J.C., Emiliani, S., Benarous, R., Desender, L., Debyser, Z. and Christ, F. (2007) Identification of the LEDGF/p75 binding site in HIV-1 integrase. *J Mol Biol*, **365**, 1480-1492.

Bustin, M. (1999) Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol Cell Biol*, **19**, 5237-5246.

Bustin, M. (2001a) Chromatin unfolding and activation by HMGN(*) chromosomal proteins. *Trends Biochem Sci*, **26**, 431-437.

Bustin, M. (2001b) Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends Biochem Sci*, **26**, 152-153.

Bustin, M., Lehn, D.A. and Landsman, D. (1990) Structural features of the HMG chromosomal proteins and their genes. *Biochim Biophys Acta*, **1049**, 231-243.

Cahn, P. and Sued, O. (2007) Raltegravir: a new antiretroviral class for salvage therapy. *Lancet*, **369**, 1235-1236.

Candau, R. and Berger, S.L. (1996) Structural and functional analysis of yeast putative adaptors. Evidence for an adaptor complex in vivo. *J Biol Chem*, **271**, 5237-5245.

Candau, R., Moore, P.A., Wang, L., Barlev, N., Ying, C.Y., Rosen, C.A. and Berger, S.L. (1996) Identification of human proteins functionally conserved with the yeast putative adaptors ADA2 and GCN5. *Mol Cell Biol*, **16**, 593-602.

Candau, R., Zhou, J.X., Allis, C.D. and Berger, S.L. (1997) Histone acetyltransferase activity and interaction with ADA2 are critical for GCN5 function in vivo. *Embo J*, **16**, 555-565.

Caron, C., Col, E. and Khochbin, S. (2003) The viral control of cellular acetylation signaling. *Bioessays*, **25**, 58-65.

Carrozza, M.J., Utley, R.T., Workman, J.L. and Cote, J. (2003) The diverse functions of histone acetyltransferase complexes. *Trends Genet*, **19**, 321-329.

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.

Cavusoglu, N., Brand, M., Tora, L. and Van Dorsselaer, A. (2003) Novel subunits of the TATA binding protein free TAFII-containing transcription complex identified by matrix-assisted laser desorption/ionization-time of flight mass spectrometry following one-dimensional gel electrophoresis. *Proteomics*, **3**, 217-223.

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.

Chang, D.D. and Sharp, P.A. (1989) Regulation by HIV Rev depends upon recognition of splice sites. *Cell*, **59**, 789-795.

Chen, D., Fong, Y. and Zhou, Q. (1999) Specific interaction of Tat with the human but not rodent P-TEFb complex mediates the species-specific Tat activation of HIV-1 transcription. *Proc Natl Acad Sci U S A*, **96**, 2728-2733.

Chen, H. and Engelman, A. (1998) The barrier-to-autointegration protein is a host factor for HIV type 1 integration. *Proc Natl Acad Sci U S A*, **95**, 15270-15274.

Chen, J.C., Krucinski, J., Miercke, L.J., Finer-Moore, J.S., Tang, A.H., Leavitt, A.D. and Stroud, R.M. (2000) Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. *Proc Natl Acad Sci U S A*, **97**, 8233-8238.

Chen, L., Fischle, W., Verdin, E. and Greene, W.C. (2001) Duration of nuclear NF-kappaB action regulated by reversible acetylation. *Science*, **293**, 1653-1657.

Cherepanov, P., Ambrosio, A.L., Rahman, S., Ellenberger, T. and Engelman, A. (2005) Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. *Proc Natl Acad Sci U S A*, **102**, 17308-17313.

Cherepanov, P., Devroe, E., Silver, P.A. and Engelman, A. (2004) Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. *J Biol Chem*, **279**, 48883-48892.

Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E. and Debyser, Z. (2003) HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J Biol Chem*, **278**, 372-381.

Cherepanov, P., Pluymers, W., Claeys, A., Proost, P., De Clercq, E. and Debyser, Z. (2000) High-level expression of active HIV-1 integrase from a synthetic gene in human cells. *Faseb J*, **14**, 1389-1399.

Chiu, T.K. and Davies, D.R. (2004) Structure and function of HIV-1 integrase. *Curr Top Med Chem*, **4**, 965-977.

Choe, H., Farzan, M., Sun, Y., Sullivan, N., Rollins, B., Ponath, P.D., Wu, L., Mackay, C.R., LaRosa, G., Newman, W., Gerard, N., Gerard, C. and Sodroski, J. (1996) The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell*, **85**, 1135-1148.

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

Christ, F., Voet, A., Marchand, A., Nicolet, S., Desimmie, B.A., Marchand, D., Bardiot, D., Van der Veken, N.J., Van Remoortel, B., Strelkov, S.V., De Maeyer, M., Chaltin, P. and Debyser, Z. (2010) Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat Chem Biol*, **6**, 442-448.

Chu, Z.L., Shin, Y.A., Yang, J.M., DiDonato, J.A. and Ballard, D.W. (1999) IKKgamma mediates the interaction of cellular IkappaB kinases with the tax transforming protein of human T cell leukemia virus type 1. *J Biol Chem*, **274**, 15297-15300.

Chuikov, S., Kurash, J.K., Wilson, J.R., Xiao, B., Justin, N., Ivanov, G.S., McKinney, K., Tempst, P., Prives, C., Gamblin, S.J., Barlev, N.A. and Reinberg, D. (2004) Regulation of p53 activity through lysine methylation. *Nature*, **432**, 353-360.

Ciuffi, A., Llano, M., Poeschla, E., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J.R. and Bushman, F. (2005) A role for LEDGF/p75 in targeting HIV DNA integration. *Nat Med*, **11**, 1287-1289.

Ciurciu, A., Komonyi, O., Pankotai, T. and Boros, I.M. (2006) The Drosophila histone acetyltransferase Gcn5 and transcriptional adaptor Ada2a are involved in nucleosomal histone H4 acetylation. *Mol Cell Biol*, **26**, 9413-9423.

Coffin, J.M., Hughes, S.H. and Varmus, H. (1997) *Retroviruses*. Cold Spring Harbor Laboratory Press, Plainview, N.Y.

Cohen, H.Y., Lavu, S., Bitterman, K.J., Hekking, B., Imahiyerobo, T.A., Miller, C., Frye, R., Ploegh, H., Kessler, B.M. and Sinclair, D.A. (2004) Acetylation of the C terminus of Ku70 by CBP and PCAF controls Bax-mediated apoptosis. *Mol Cell*, **13**, 627-638.

Cohen, T. and Yao, T.P. (2004) AcK-knowledge reversible acetylation. *Sci STKE*, **2004**, pe42.

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.

Cooper, D.A., Steigbigel, R.T., Gatell, J.M., Rockstroh, J.K., Katlama, C., Yeni, P., Lazzarin, A., Clotet, B., Kumar, P.N., Eron, J.E., Schechter, M., Markowitz, M., Loutfy, M.R., Lennox, J.L., Zhao, J., Chen, J., Ryan, D.M., Rhodes, R.R., Killar, J.A., Gilde, L.R., Strohmaier, K.M., Meibohm, A.R., Miller, M.D., Hazuda, D.J., Nessly, M.L., DiNubile, M.J., Isaacs, R.D., Teppler, H. and Nguyen, B.Y. (2008) Subgroup and resistance analyses of raltegravir for resistant HIV-1 infection. *N Engl J Med*, **359**, 355-365.

Corbeil, J., Sheeter, D., Genini, D., Rought, S., Leoni, L., Du, P., Ferguson, M., Masys, D.R., Welsh, J.B., Fink, J.L., Sasik, R., Huang, D., Drenkow, J., Richman, D.D. and Gingeras, T. (2001) Temporal gene regulation during HIV-1 infection of human CD4+ T cells. *Genome Res*, **11**, 1198-1204.

Corden, J.L. (2007) Transcription. Seven ups the code. *Science*, **318**, 1735-1736.

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

Cress, W.D. and Seto, E. (2000) Histone deacetylases, transcriptional control, and cancer. *J Cell Physiol*, **184**, 1-16.

Cribier, A., Segeral, E., Delelis, O., Parissi, V., Simon, A., Ruff, M., Benarous, R. and Emiliani, S. (2011) Mutations affecting interaction of integrase with TNPO3 do not prevent HIV-1 cDNA nuclear import. *Retrovirology*, **8**, 104.

Crowley, T.E., Kaine, E.M., Yoshida, M., Nandi, A. and Wolgemuth, D.J. (2002) Reproductive cycle regulation of nuclear import, euchromatic localization, and association with components of Pol II mediator of a mammalian double-bromodomain protein. *Mol Endocrinol*, **16**, 1727-1737.

Dalgleish, A.G., Beverley, P.C., Clapham, P.R., Crawford, D.H., Greaves, M.F. and Weiss, R.A. (1984) The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature*, **312**, 763-767.

Daniel, J.A., Torok, M.S., Sun, Z.W., Schieltz, D., Allis, C.D., Yates, J.R., 3rd and Grant, P.A. (2004) Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J Biol Chem*, **279**, 1867-1871.

De Rijck, J., Vandekerckhove, L., Gijsbers, R., Hombrouck, A., Hendrix, J., Vercammen, J., Engelborghs, Y., Christ, F. and Debyser, Z. (2006) Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. *J Virol*, **80**, 11498-11509.

Debyser, Z., Cherepanov, P., Pluymers, W. and De Clercq, E. (2001) Assays for the evaluation of HIV-1 integrase inhibitors. *Methods Mol Biol*, **160**, 139-155.

Debyser, Z. and Christ, F. (2010) On the cell biology of HIV integration from basic research to development of novel antiviral drugs. *Verh K Acad Geneeskd Belg*, **72**, 219-237.

Delelis, O., Malet, I., Na, L., Tchertanov, L., Calvez, V., Marcelin, A.G., Subra, F., Deprez, E. and Mouscadet, J.F. (2009) The G140S mutation in HIV integrases from raltegravir-resistant patients rescues catalytic defect due to the resistance Q148H mutation. *Nucleic Acids Res*, **37**, 1193-1201.

Demeny, M.A., Soutoglou, E., Nagy, Z., Scheer, E., Janoshazi, A., Richardot, M., Argentini, M., Kessler, P. and Tora, L. (2007) Identification of a small TAF complex and its role in the assembly of TAF-containing complexes. *PLoS One*, **2**, e316.

Demeulemeester, J., Tintori, C., Botta, M., Debyser, Z. and Christ, F. (2012) Development of an AlphaScreen-Based HIV-1 Integrase Dimerization Assay for Discovery of Novel Allosteric Inhibitors. *J Biomol Screen*.

Demirov, D.G. and Freed, E.O. (2004) Retrovirus budding. *Virus Res*, **106**, 87-102.

Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., Di Marzio, P., Marmon, S., Sutton, R.E., Hill, C.M., Davis, C.B., Peiper, S.C.,

Schall, T.J., Littman, D.R. and Landau, N.R. (1996) Identification of a major co-receptor for primary isolates of HIV-1. *Nature*, **381**, 661-666.

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.

Depienne, C., Roques, P., Creminon, C., Fritsch, L., Casseron, R., Dormont, D., Dargemont, C. and Benichou, S. (2000) Cellular distribution and karyophilic properties of matrix, integrase, and Vpr proteins from the human and simian immunodeficiency viruses. *Exp Cell Res*, **260**, 387-395.

Devroe, E., Engelman, A. and Silver, P.A. (2003) Intracellular transport of human immunodeficiency virus type 1 integrase. *J Cell Sci*, **116**, 4401-4408.

Dey, A., Chitsaz, F., Abbasi, A., Misteli, T. and Ozato, K. (2003) The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc Natl Acad Sci U S A*, **100**, 8758-8763.

Dey, A., Ellenberg, J., Farina, A., Coleman, A.E., Maruyama, T., Sciortino, S., Lippincott-Schwartz, J. and Ozato, K. (2000) A bromodomain protein, MCAP, associates with mitotic chromosomes and affects G(2)-to-M transition. *Mol Cell Biol*, **20**, 6537-6549.

Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K. and Zhou, M.M. (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature*, **399**, 491-496.

Di Fenza, A., Rocchia, W. and Tozzini, V. (2009) Complexes of HIV-1 integrase with HAT proteins: Multiscale models, dynamics, and hypotheses on allosteric sites of inhibition. *Proteins*.

Dingwall, C., Ernberg, I., Gait, M.J., Green, S.M., Heaphy, S., Karn, J., Lowe, A.D., Singh, M., Skinner, M.A. and Valerio, R. (1989) Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA in vitro. *Proc Natl Acad Sci U S A*, **86**, 6925-6929.

Dompierre, J.P., Godin, J.D., Charrin, B.C., Cordelieres, F.P., King, S.J., Humbert, S. and Saudou, F. (2007) Histone deacetylase 6 inhibition compensates for the transport deficit in Huntington's disease by increasing tubulin acetylation. *J Neurosci*, **27**, 3571-3583.

Dornan, D., Shimizu, H., Perkins, N.D. and Hupp, T.R. (2003) DNA-dependent acetylation of p53 by the transcription coactivator p300. *J Biol Chem*, **278**, 13431-13441.

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.

Durkin, S.S., Ward, M.D., Fryrear, K.A. and Semmes, O.J. (2006) Site-specific phosphorylation differentiates active from inactive forms of the human T-cell leukemia virus type 1 Tax oncoprotein. *J Biol Chem*, **281**, 31705-31712.

Durocher, Y., Perret, S. and Kamen, A. (2002) High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res*, **30**, E9.

Dyda, F., Hickman, A.B., Jenkins, T.M., Engelman, A., Craigie, R. and Davies, D.R. (1994) Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science*, **266**, 1981-1986.

Ebina, H., Aoki, J., Hatta, S., Yoshida, T. and Koyanagi, Y. (2004) Role of Nup98 in nuclear entry of human immunodeficiency virus type 1 cDNA. *Microbes Infect*, **6**, 715-724.

Eckert, D.M. and Kim, P.S. (2001) Mechanisms of viral membrane fusion and its inhibition. *Annu Rev Biochem*, **70**, 777-810.

Eckner, R., Ewen, M.E., Newsome, D., Gerdes, M., DeCaprio, J.A., Lawrence, J.B. and Livingston, D.M. (1994) Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes Dev*, **8**, 869-884.

Ehrlich, L.S., Agresta, B.E. and Carter, C.A. (1992) Assembly of recombinant human immunodeficiency virus type 1 capsid protein in vitro. *J Virol*, **66**, 4874-4883.

Eijkelenboom, A.P., Lutzke, R.A., Boelens, R., Plasterk, R.H., Kaptein, R. and Hard, K. (1995) The DNA-binding domain of HIV-1 integrase has an SH3-like fold. *Nat Struct Biol*, **2**, 807-810.

Ellison, V., Gerton, J., Vincent, K.A. and Brown, P.O. (1995) An essential interaction between distinct domains of HIV-1 integrase mediates assembly of the active multimer. *J Biol Chem*, **270**, 3320-3326.

Emerman, M., Vazeux, R. and Peden, K. (1989) The rev gene product of the human immunodeficiency virus affects envelope-specific RNA localization. *Cell*, **57**, 1155-1165.

Emiliani, S., Mousnier, A., Busschots, K., Maroun, M., Van Maele, B., Tempe, D., Vandekerckhove, L., Moisant, F., Ben-Slama, L., Witvrouw, M., Christ, F., Rain, J.C., Dargemont, C., Debyser, Z. and Benarous, R. (2005) Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication. *J Biol Chem*, **280**, 25517-25523.

Engelman, A. (1999) In vivo analysis of retroviral integrase structure and function. *Adv. Virus. Res.*, **52**, 411-426.

Engelman, A., Bushman, F.D. and Craigie, R. (1993) Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. *Embo J*, **12**, 3269-3275.

Engelman, A. and Craigie, R. (1992) Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J Virol*, **66**, 6361-6369.

Espeseth, A.S., Felock, P., Wolfe, A., Witmer, M., Grobler, J., Anthony, N., Egbertson, M., Melamed, J.Y., Young, S., Hamill, T., Cole, J.L. and Hazuda, D.J. (2000) HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase. *Proc Natl Acad Sci U S A*, **97**, 11244-11249.

Esposito, D. and Craigie, R. (1998) Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein-DNA interaction. *Embo J*, **17**, 5832-5843.

Facke, M., Janetzko, A., Shoeman, R.L. and Krausslich, H.G. (1993) A large deletion in the matrix domain of the human immunodeficiency virus gag gene redirects virus particle assembly from the plasma membrane to the endoplasmic reticulum. *J Virol*, **67**, 4972-4980.

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.

Farnet, C.M. and Haseltine, W.A. (1991) Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. *J Virol*, **65**, 1910-1915.

Fassati, A. and Goff, S.P. (2001) Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. *J Virol*, **75**, 3626-3635.

Feng, Y., Broder, C.C., Kennedy, P.E. and Berger, E.A. (1996) HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*, **272**, 872-877.

Ferris, A.L., Wu, X., Hughes, C.M., Stewart, C., Smith, S.J., Milne, T.A., Wang, G.G., Shun, M.C., Allis, C.D., Engelman, A. and Hughes, S.H. (2010) Lens epithelium-derived growth factor fusion proteins redirect HIV-1 DNA integration. *Proc Natl Acad Sci U S A*, **107**, 3135-3140.

Fields, B.N., Knipe, D.M., Howley, P.M., Griffin, D.E. and Ovid Technologies Inc. (2007) Fields' virology. Lippincott Williams & Wilkins, Philadelphia.

Flinn, E.M., Wallberg, A.E., Hermann, S., Grant, P.A., Workman, J.L. and Wright, A.P. (2002) Recruitment of Gcn5-containing complexes during c-Myc-dependent gene activation. Structure and function aspects. *J Biol Chem*, **277**, 23399-23406.

Foisner, R. (2001) Inner nuclear membrane proteins and the nuclear lamina. *J Cell Sci*, **114**, 3791-3792.

Fornerod, M., Ohno, M., Yoshida, M. and Mattaj, I.W. (1997) CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell*, **90**, 1051-1060.

Frank, S.R., Schroeder, M., Fernandez, P., Taubert, S. and Amati, B. (2001) Binding of c-Myc to chromatin mediates mitogen-induced acetylation of histone H4 and gene activation. *Genes Dev*, **15**, 2069-2082.

Freed, E.O. (2004) HIV-1 and the host cell: an intimate association. *Trends Microbiol*, **12**, 170-177.

Freed, E.O., Orenstein, J.M., Buckler-White, A.J. and Martin, M.A. (1994) Single amino acid changes in the human immunodeficiency virus type 1 matrix protein block virus particle production. *J Virol*, **68**, 5311-5320.

Fujii, K., Hurley, J.H. and Freed, E.O. (2007) Beyond Tsg101: the role of Alix in 'ESCRTing' HIV-1. *Nat Rev Microbiol*, **5**, 912-916.

Fujinaga, K., Cujec, T.P., Peng, J., Garriga, J., Price, D.H., Grana, X. and Peterlin, B.M. (1998) 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*, **72**, 7154-7159.

Fukuda, H., Sano, N., Muto, S. and Horikoshi, M. (2006) Simple histone acetylation plays a complex role in the regulation of gene expression. *Brief Funct Genomic Proteomic*, **5**, 190-208.

Fukuda, M., Asano, S., Nakamura, T., Adachi, M., Yoshida, M., Yanagida, M. and Nishida, E. (1997) CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature*, **390**, 308-311.

Gallay, P., Hope, T., Chin, D. and Trono, D. (1997) HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. *Proc Natl Acad Sci U S A*, **94**, 9825-9830.

Gallimore, P.H. and Turnell, A.S. (2001) Adenovirus E1A: remodelling the host cell, a life or death experience. *Oncogene*, **20**, 7824-7835.

Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., Kaplan, M., Haynes, B.F., Palker, T.J., Redfield, R., Oleske, J., Safai, B. and et al. (1984) Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science*, **224**, 500-503.

Gamper, A.M. and Roeder, R.G. (2008) Multivalent binding of p53 to the STAGA complex mediates coactivator recruitment after UV damage. *Mol Cell Biol*, **28**, 2517-2527.

Ganapathy, V., Daniels, T. and Casiano, C.A. (2003) LEDGF/p75: a novel nuclear autoantigen at the crossroads of cell survival and apoptosis. *Autoimmun Rev*, **2**, 290-297.

Ganser-Pornillos, B.K., Yeager, M. and Sundquist, W.I. (2008) The structural biology of HIV assembly. *Curr Opin Struct Biol*, **18**, 203-217.

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.

Gershey, E.L., Vidali, G. and Allfrey, V.G. (1968) Chemical studies of histone acetylation. The occurrence of epsilon-N-acetyllysine in the f2a1 histone. *J Biol Chem*, **243**, 5018-5022.

Gijsbers, R., Ronen, K., Vets, S., Malani, N., De Rijck, J., McNeely, M., Bushman, F.D. and Debyser, Z. (2010) LEDGF hybrids efficiently retarget lentiviral integration into heterochromatin. *Mol Ther*, **18**, 552-560.

Glozak, M.A., Sengupta, N., Zhang, X. and Seto, E. (2005) Acetylation and deacetylation of non-histone proteins. *Gene*, **363**, 15-23.

Glozak, M.A. and Seto, E. (2007) Histone deacetylases and cancer. *Oncogene*, **26**, 5420-5432.

Goldgur, Y., Craigie, R., Cohen, G.H., Fujiwara, T., Yoshinaga, T., Fujishita, T., Sugimoto, H., Endo, T., Murai, H. and Davies, D.R. (1999) Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: a platform for antiviral drug design. *Proc Natl Acad Sci U S A*, **96**, 13040-13043.

Gorlich, D. and Mattaj, I.W. (1996) Nucleocytoplasmic transport. *Science*, **271**, 1513-1518.

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

Grandgenett, D.P., Vora, A.C. and Schiff, R.D. (1978) A 32,000-dalton nucleic acid-binding protein from avian retravirus cores possesses DNA endonuclease activity. *Virology*, **89**, 119-132.

Grant, P.A., Duggan, L., Cote, J., Roberts, S.M., Brownell, J.E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C.D., Winston, F., Berger, S.L. and Workman, J.L. (1997) Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev*, **11**, 1640-1650.

Grobler, J.A., Stillmock, K., Hu, B., Witmer, M., Felock, P., Espeseth, A.S., Wolfe, A., Egbertson, M., Bourgeois, M., Melamed, J., Wai, J.S., Young, S., Vacca, J. and Hazuda, D.J. (2002) Diketo acid inhibitor mechanism and HIV-1 integrase: implications for metal binding in the active site of phosphotransferase enzymes. *Proc Natl Acad Sci U S A*, **99**, 6661-6666.

Gross, I., Hohenberg, H. and Krausslich, H.G. (1997) In vitro assembly properties of purified bacterially expressed capsid proteins of human immunodeficiency virus. *Eur J Biochem*, **249**, 592-600.

Gu, W. and Roeder, R.G. (1997) Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell*, **90**, 595-606.

Guelman, S., Kozuka, K., Mao, Y., Pham, V., Solloway, M.J., Wang, J., Wu, J., Lill, J.R. and Zha, J. (2009) The double-histone-acetyltransferase complex ATAC is essential for mammalian development. *Mol Cell Biol*, **29**, 1176-1188.

Guelman, S., Suganuma, T., Florens, L., Swanson, S.K., Kiesecker, C.L., Kusch, T., Anderson, S., Yates, J.R., 3rd, Washburn, M.P., Abmayr, S.M. and Workman, J.L. (2006) Host cell factor and an uncharacterized SANT domain protein are stable components of ATAC, a novel dAda2A/dGcn5-containing histone acetyltransferase complex in Drosophila. *Mol Cell Biol*, **26**, 871-882.

Guo, D., Hazbun, T.R., Xu, X.J., Ng, S.L., Fields, S. and Kuo, M.H. (2004) A tethered catalysis, two-hybrid system to identify protein-protein interactions requiring post-translational modifications. *Nat Biotechnol*, **22**, 888-892.

Hacein-Bey-Abina, S., von Kalle, C., Schmidt, M., Le Deist, F., Wulffraat, N., McIntyre, E., Radford, I., Villeval, J.L., Fraser, C.C., Cavazzana-Calvo, M. and Fischer, A. (2003a) A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N Engl J Med*, **348**, 255-256.

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E., Sorensen, R., Forster, A., Fraser, P., Cohen, J.I., de 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, L.E., Wissler, M., Prinz, C., Rabbitts, T.H., Le Deist, F., Fischer, A. and Cavazzana-Calvo, M. (2003b) LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science*, **302**, 415-419.

Haffar, O.K., Popov, S., Dubrovsky, L., Agostini, I., Tang, H., Pushkarsky, T., Nadler, S.G. and Bukrinsky, M. (2000) Two nuclear localization signals in the HIV-1 matrix protein regulate nuclear import of the HIV-1 pre-integration complex. *J Mol Biol*, **299**, 359-368.

Hallenberger, S., Bosch, V., Angliker, H., Shaw, E., Klenk, H.D. and Garten, W. (1992) Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. *Nature*, **360**, 358-361.

Hare, S., Di Nunzio, F., Labeja, A., Wang, J., Engelman, A. and Cherepanov, P. (2009) Structural basis for functional tetramerization of lentiviral integrase. *PLoS Pathog*, **5**, e1000515.

Hare, S., Gupta, S.S., Valkov, E., Engelman, A. and Cherepanov, P. (2010) Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature*.

Harhaj, E.W. and Sun, S.C. (1999) IKKgamma serves as a docking subunit of the IkappaB kinase (IKK) and mediates interaction of IKK with the human T-cell leukemia virus Tax protein. *J Biol Chem*, **274**, 22911-22914.

Hassan, A.H., Awad, S. and Prochasson, P. (2006) The Swi2/Snf2 bromodomain is required for the displacement of SAGA and the octamer transfer of SAGA-acetylated nucleosomes. *J Biol Chem*, **281**, 18126-18134.

Hassan, A.H., Neely, K.E., Vignali, M., Reese, J.C. and Workman, J.L. (2001) Promoter targeting of chromatin-modifying complexes. *Front Biosci*, **6**, D1054-1064.

Hassan, A.H., Prochasson, P., Neely, K.E., Galasinski, S.C., Chandy, M., Carrozza, M.J. and Workman, J.L. (2002) Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell*, **111**, 369-379.

Hazuda, D.J. (2010) Resistance to inhibitors of the human immunodeficiency virus type 1 integration. *Braz J Infect Dis*, **14**, 513-518.

Hazuda, D.J., Felock, P., Witmer, M., Wolfe, A., Stillmock, K., Grobler, J.A., Espeseth, A., Gabryelski, L., Schleif, W., Blau, C. and Miller, M.D. (2000) Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. *Science*, **287**, 646-650.

Hearps, A.C. and Jans, D.A. (2006) HIV-1 integrase is capable of targeting DNA to the nucleus via an importin alpha/beta-dependent mechanism. *Biochem J*, **398**, 475-484.

Helmlinger, D., Hardy, S., Abou-Sleymane, G., Eberlin, A., Bowman, A.B., Gansmuller, A., Picaud, S., Zoghbi, H.Y., Trottier, Y., Tora, L. and Devys, D. (2006) Glutamine-expanded ataxin-7 alters TFTC/STAGA recruitment

and chromatin structure leading to photoreceptor dysfunction. *PLoS Biol*, **4**, e67.

Helmlinger, D., Hardy, S., Sasorith, S., Klein, F., Robert, F., Weber, C., Miguet, L., Potier, N., Van-Dorsselaer, A., Wurtz, J.M., Mandel, J.L., Tora, L. and Devys, D. (2004) Ataxin-7 is a subunit of GCN5 histone acetyltransferase-containing complexes. *Hum Mol Genet*, **13**, 1257-1265.

Henderson, A., Bunce, M., Siddon, N., Reeves, R. and Tremethick, D.J. (2000) High-mobility-group protein I can modulate binding of transcription factors to the U5 region of the human immunodeficiency virus type 1 proviral promoter. *J Virol*, **74**, 10523-10534.

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.

Henderson, B.R. and Percipalle, P. (1997) Interactions between HIV Rev and nuclear import and export factors: the Rev nuclear localisation signal mediates specific binding to human importin-beta. *J Mol Biol*, **274**, 693-707.

Henry, K.W., Wyce, A., Lo, W.S., Duggan, L.J., Emre, N.C., Kao, C.F., Pillus, L., Shilatifard, A., Osley, M.A. and Berger, S.L. (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev*, **17**, 2648-2663.

Herrera, J.E., Sakaguchi, K., Bergel, M., Trieschmann, L., Nakatani, Y. and Bustin, M. (1999) Specific acetylation of chromosomal protein HMG-17 by PCAF alters its interaction with nucleosomes. *Mol Cell Biol*, **19**, 3466-3473.

Heuer, T.S. and Brown, P.O. (1997) Mapping features of HIV-1 integrase near selected sites on viral and target DNA molecules in an active enzyme-DNA complex by photo-cross-linking. *Biochemistry*, **36**, 10655-10665.

Heuer, T.S. and Brown, P.O. (1998) Photo-cross-linking studies suggest a model for the architecture of an active human immunodeficiency virus type 1 integrase-DNA complex. *Biochemistry*, **37**, 6667-6678.

Hindmarsh, P., Ridky, T., Reeves, R., Andrake, M., Skalka, A.M. and Leis, J. (1999) HMG protein family members stimulate human immunodeficiency virus type 1 and avian sarcoma virus concerted DNA integration in vitro. *J Virol*, **73**, 2994-3003.

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

Hombrouck, A., De Rijck, J., Hendrix, J., Vandekerckhove, L., Voet, A., De Maeyer, M., Witvrouw, M., Engelborghs, Y., Christ, F., Gijsbers, R. and Debyser, Z. (2007) Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV. *PLoS Pathog*, **3**, e47.

Horiuchi, J., Silverman, N., Marcus, G.A. and Guarente, L. (1995) ADA3, a putative transcriptional adaptor, consists of two separable domains and interacts with ADA2 and GCN5 in a trimeric complex. *Mol Cell Biol*, **15**, 1203-1209.

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.

Huang, M., Orenstein, J.M., Martin, M.A. and Freed, E.O. (1995) p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. *J Virol*, **69**, 6810-6818.

Huang, T.T., Kudo, N., Yoshida, M. and Miyamoto, S. (2000) A nuclear export signal in the N-terminal regulatory domain of IkappaBalpha controls cytoplasmic localization of inactive NF-kappaB/IkappaBalpha complexes. *Proc Natl Acad Sci U S A*, **97**, 1014-1019.

Hubbert, C., Guardiola, A., Shao, R., Kawaguchi, Y., Ito, A., Nixon, A., Yoshida, M., Wang, X.F. and Yao, T.P. (2002) HDAC6 is a microtubule-associated deacetylase. *Nature*, **417**, 455-458.

Hudson, B.P., Martinez-Yamout, M.A., Dyson, H.J. and Wright, P.E. (2000) Solution structure and acetyl-lysine binding activity of the GCN5 bromodomain. *J Mol Biol*, **304**, 355-370.

Iordanskiy, S., Berro, R., Altieri, M., Kashanchi, F. and Bukrinsky, M. (2006) Intracytoplasmic maturation of the human immunodeficiency virus type 1 reverse transcription complexes determines their capacity to integrate into chromatin. *Retrovirology*, **3**, 4.

Ivanov, D., Kwak, Y.T., Nee, E., Guo, J., Garcia-Martinez, L.F. and Gaynor, R.B. (1999) Cyclin T1 domains involved in complex formation with Tat and TAR RNA are critical for tat-activation. *J Mol Biol*, **288**, 41-56.

Jacobson, R.H., Ladurner, A.G., King, D.S. and Tjian, R. (2000) Structure and function of a human TAFII250 double bromodomain module. *Science*, **288**, 1422-1425.

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

Jaskolski, M., Alexandratos, J.N., Bujacz, G. and Wlodawer, A. (2009) Piecing together the structure of retroviral integrase, an important target in AIDS therapy. *Febs J*, **276**, 2926-2946.

Jeanmougin, F., Wurtz, J.M., Le Douarin, B., Chambon, P. and Losson, R. (1997) The bromodomain revisited. *Trends Biochem Sci*, **22**, 151-153.

Jenkins, T.M., Esposito, D., Engelman, A. and Craigie, R. (1997) Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking. *Embo J*, **16**, 6849-6859.

Jenkins, T.M., Hickman, A.B., Dyda, F., Ghirlando, R., Davies, D.R. and Craigie, R. (1995) Catalytic domain of human immunodeficiency virus type 1 integrase: identification of a soluble mutant by systematic replacement of hydrophobic residues. *Proc Natl Acad Sci U S A*, **92**, 6057-6061.

Jenkins, Y., McEntee, M., Weis, K. and Greene, W.C. (1998) Characterization of HIV-1 vpr nuclear import: analysis of signals and pathways. *J Cell Biol*, **143**, 875-885.

Jeong, J.W., Bae, M.K., Ahn, M.Y., Kim, S.H., Sohn, T.K., Bae, M.H., Yoo, M.A., Song, E.J., Lee, K.J. and Kim, K.W. (2002) Regulation and destabilization of HIF-1alpha by ARD1-mediated acetylation. *Cell*, **111**, 709-720.

Johnson, A.A., Santos, W., Pais, G.C., Marchand, C., Amin, R., Burke, T.R., Jr., Verdine, G. and Pommier, Y. (2006) Integration requires a specific interaction of the donor DNA terminal 5'-cytosine with glutamine 148 of the HIV-1 integrase flexible loop. *J Biol Chem*, **281**, 461-467.

Jones, K.S., Coleman, J., Merkel, G.W., Laue, T.M. and Skalka, A.M. (1992) Retroviral integrase functions as a multimer and can turn over catalytically. *J Biol Chem*, **267**, 16037-16040.

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.

Kahata, K., Hayashi, M., Asaka, M., Hellman, U., Kitagawa, H., Yanagisawa, J., Kato, S., Imamura, T. and Miyazono, K. (2004) Regulation of transforming growth factor-beta and bone morphogenetic protein signalling by transcriptional coactivator GCN5. *Genes Cells*, **9**, 143-151.

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.

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.

Kanno, T., Kanno, Y., Siegel, R.M., Jang, M.K., Lenardo, M.J. and Ozato, K. (2004) Selective recognition of acetylated histones by bromodomain proteins visualized in living cells. *Mol Cell*, **13**, 33-43.

Karageorgos, L., Li, P. and Burrell, C. (1993) Characterization of HIV replication complexes early after cell-to-cell infection. *AIDS Res Hum Retroviruses*, **9**, 817-823.

Karanam, B., Jiang, L., Wang, L., Kelleher, N.L. and Cole, P.A. (2006) Kinetic and mass spectrometric analysis of p300 histone acetyltransferase domain autoacetylation. *J Biol Chem*, **281**, 40292-40301.

Kataoka, N., Bachorik, J.L. and Dreyfuss, G. (1999) Transportin-SR, a nuclear import receptor for SR proteins. *J Cell Biol*, **145**, 1145-1152.

Katz, R.A., DiCandeloro, P., Kukolj, G. and Skalka, A.M. (2001) Role of DNA end distortion in catalysis by avian sarcoma virus integrase. *J Biol Chem*, **276**, 34213-34220.

Katz, R.A., Gravuer, K. and Skalka, A.M. (1998) A preferred target DNA structure for retroviral integrase in vitro. *J Biol Chem*, **273**, 24190-24195.

Katz, R.A., Merkel, G., Kulkosky, J., Leis, J. and Skalka, A.M. (1990) The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. *Cell*, **63**, 87-95.

Katz, R.A., Merkel, G. and Skalka, A.M. (1996) Targeting of retroviral integrase by fusion to a heterologous DNA binding domain: in vitro activities and incorporation of a fusion protein into viral particles. *Virology*, **217**, 178-190.

Kawaguchi, Y., Ito, A., Appella, E. and Yao, T.P. (2006) Charge modification at multiple C-terminal lysine residues regulates p53 oligomerization and its nucleus-cytoplasm trafficking. *J Biol Chem*, **281**, 1394-1400.

Kehn, K., Berro, R., de la Fuente, C., Strouss, K., Ghedin, E., Dadgar, S., Bottazzi, M.E., Pumfery, A. and Kashanchi, F. (2004) Mechanisms of HTLV-1 transformation. *Front Biosci*, **9**, 2347-2372.

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

Kikuchi, H., Takami, Y. and Nakayama, T. (2005) GCN5: a supervisor in all-inclusive control of vertebrate cell cycle progression through transcription regulation of various cell cycle-related genes. *Gene*, **347**, 83-97.

Kitamura, Y.I., Kitamura, T., Kruse, J.P., Raum, J.C., Stein, R., Gu, W. and Accili, D. (2005) FoxO1 protects against pancreatic beta cell failure through NeuroD and MafA induction. *Cell Metab*, **2**, 153-163.

Kleff, S., Andrulis, E.D., Anderson, C.W. and Sternglanz, R. (1995) Identification of a gene encoding a yeast histone H4 acetyltransferase. *J Biol Chem*, **270**, 24674-24677.

Knoepfler, P.S., Zhang, X.Y., Cheng, P.F., Gafken, P.R., McMahon, S.B. and Eisenman, R.N. (2006) Myc influences global chromatin structure. *Embo J*, **25**, 2723-2734.

Kobayashi, M., Nakahara, K., Seki, T., Miki, S., Kawauchi, S., Suyama, A., Wakasa-Morimoto, C., Kodama, M., Endoh, T., Oosugi, E., Matsushita, Y., Murai, H., Fujishita, T., Yoshinaga, T., Garvey, E., Foster, S., Underwood, M., Johns, B., Sato, A. and Fujiwara, T. (2008) Selection of diverse and clinically relevant integrase inhibitor-resistant human immunodeficiency virus type 1 mutants. *Antiviral Res*, **80**, 213-222.

Konig, R., Zhou, Y., Elleder, D., Diamond, T.L., Bonamy, G.M., Irelan, J.T., Chiang, C.Y., Tu, B.P., De Jesus, P.D., Lilley, C.E., Seidel, S., Opaluch, A.M., Caldwell, J.S., Weitzman, M.D., Kuhen, K.L., Bandyopadhyay, S., Ideker, T., Orth, A.P., Miraglia, L.J., Bushman, F.D., Young, J.A. and Chanda, S.K. (2008) Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell*, **135**, 49-60.

Kouzarides, T. (2000) Acetylation: a regulatory modification to rival phosphorylation? *Embo J*, **19**, 1176-1179.

Kovacs, J.J., Murphy, P.J., Gaillard, S., Zhao, X., Wu, J.T., Nicchitta, C.V., Yoshida, M., Toft, D.O., Pratt, W.B. and Yao, T.P. (2005) HDAC6 regulates Hsp90 acetylation and chaperone-dependent activation of glucocorticoid receptor. *Mol Cell*, **18**, 601-607.

Krebs, A.R., Karmodiya, K., Lindahl-Allen, M., Struhl, K. and Tora, L. (2011) SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers. *Mol Cell*, **44**, 410-423.

Krishnan, L., Matreyek, K.A., Oztop, I., Lee, K., Tipper, C.H., Li, X., Dar, M.J., Kewalramani, V.N. and Engelman, A. (2010) The requirement for cellular transportin 3 (TNPO3 or TRN-SR2) during infection maps to human immunodeficiency virus type 1 capsid and not integrase. *J Virol*, **84**, 397-406.

Kulkosky, J., Jones, K.S., Katz, R.A., Mack, J.P. and Skalka, A.M. (1992) Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. *Mol Cell Biol*, **12**, 2331-2338.

Kuo, M.H., Brownell, J.E., Sobel, R.E., Ranalli, T.A., Cook, R.G., Edmondson, D.G., Roth, S.Y. and Allis, C.D. (1996) Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature*, **383**, 269-272.

Kurash, J.K., Lei, H., Shen, Q., Marston, W.L., Granda, B.W., Fan, H., Wall, D., Li, E. and Gaudet, F. (2008) Methylation of p53 by Set7/9 mediates p53 acetylation and activity in vivo. *Mol Cell*, **29**, 392-400.

Kusch, T., Guelman, S., Abmayr, S.M. and Workman, J.L. (2003) Two Drosophila Ada2 homologues function in different multiprotein complexes. *Mol Cell Biol*, **23**, 3305-3319.

Ladurner, A.G., Inouye, C., Jain, R. and Tjian, R. (2003) Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. *Mol Cell*, **11**, 365-376.

Lamsoul, I., Lodewick, J., Lebrun, S., Brasseur, R., Burny, A., Gaynor, R.B. and Bex, F. (2005) Exclusive ubiquitination and sumoylation on overlapping

lysine residues mediate NF-kappaB activation by the human T-cell leukemia virus tax oncoprotein. *Mol Cell Biol*, **25**, 10391-10406.

Leavitt, A.D., Shiue, L. and Varmus, H.E. (1993) Site-directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions in vitro. *J Biol Chem*, **268**, 2113-2119.

Lebedeva, L.A., Nabirochkina, E.N., Kurshakova, M.M., Robert, F., Krasnov, A.N., Evgen'ev, M.B., Kadonaga, J.T., Georgieva, S.G. and Tora, L. (2005) Occupancy of the Drosophila hsp70 promoter by a subset of basal transcription factors diminishes upon transcriptional activation. *Proc Natl Acad Sci U S A*, **102**, 18087-18092.

Lee, D.Y., Hayes, J.J., Pruss, D. and Wolffe, A.P. (1993) A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell*, **72**, 73-84.

Lee, K.K. and Workman, J.L. (2007) Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol*, **8**, 284-295.

Lee, M.S. and Craigie, R. (1998) A previously unidentified host protein protects retroviral DNA from autointegration. *Proc Natl Acad Sci U S A*, **95**, 1528-1533.

Legube, G. and Trouche, D. (2003) Regulating histone acetyltransferases and deacetylases. *EMBO Rep*, **4**, 944-947.

Levin, A., Armon-Omer, A., Rosenbluh, J., Melamed-Book, N., Graessmann, A., Waigmann, E. and Loyter, A. (2009) Inhibition of HIV-1 integrase nuclear import and replication by a peptide bearing integrase putative nuclear localization signal. *Retrovirology*, **6**, 112.

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, A.G., Piluso, L.G., Cai, X., Gadd, B.J., Ladurner, A.G. and Liu, X. (2007) An acetylation switch in p53 mediates holo-TFIID recruitment. *Mol Cell*, **28**, 408-421.

Li, L., Yoder, K., Hansen, M.S., Olvera, J., Miller, M.D. and Bushman, F.D. (2000) Retroviral cDNA integration: stimulation by HMG I family proteins. *J Virol*, **74**, 10965-10974.

Li, M., Luo, J., Brooks, C.L. and Gu, W. (2002) Acetylation of p53 inhibits its ubiquitination by Mdm2. *J Biol Chem*, **277**, 50607-50611.

Li, M., Mizuuchi, M., Burke, T.R., Jr. and Craigie, R. (2006) Retroviral DNA integration: reaction pathway and critical intermediates. *Embo J*, **25**, 1295-1304.

Limon, A., Devroe, E., Lu, R., Ghory, H.Z., Silver, P.A. and Engelman, A. (2002) Nuclear localization of human immunodeficiency virus type 1 preintegration complexes (PICs): V165A and R166A are pleiotropic integrase mutants primarily defective for integration, not PIC nuclear import. *J Virol*, **76**, 10598-10607.

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.

Lin, F., Blake, D.L., Callebaut, I., Skerjanc, I.S., Holmer, L., McBurney, M.W., Paulin-Levasseur, M. and Worman, H.J. (2000) MAN1, an inner nuclear membrane protein that shares the LEM domain with lamina-associated polypeptide 2 and emerin. *J Biol Chem*, **275**, 4840-4847.

Liu, X., Tesfai, J., Evrard, Y.A., Dent, S.Y. and Martinez, E. (2003) c-Myc transformation domain recruits the human STAGA complex and requires TRRAP and GCN5 acetylase activity for transcription activation. *J Biol Chem*, **278**, 20405-20412.

Liu, X., Vorontchikhina, M., Wang, Y.L., Faiola, F. and Martinez, E. (2008) STAGA recruits Mediator to the MYC oncoprotein to stimulate transcription and cell proliferation. *Mol Cell Biol*, **28**, 108-121.

Llano, M., Delgado, S., Vanegas, M. and Poeschla, E.M. (2004a) Lens epithelium-derived growth factor/p75 prevents proteasomal degradation of HIV-1 integrase. *J Biol Chem*, **279**, 55570-55577.

Llano, M., Saenz, D.T., Meehan, A., Wongthida, P., Peretz, M., Walker, W.H., Teo, W. and Poeschla, E.M. (2006a) An essential role for LEDGF/p75 in HIV integration. *Science*, **314**, 461-464.

Llano, M., Vanegas, M., Fregoso, O., Saenz, D., Chung, S., Peretz, M. and Poeschla, E.M. (2004b) LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. *J Virol*, **78**, 9524-9537.

Llano, M., Vanegas, M., Hutchins, N., Thompson, D., Delgado, S. and Poeschla, E.M. (2006b) Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75. *J Mol Biol*, **360**, 760-773.

Lodewick, J., Lamsoul, I., Polania, A., Lebrun, S., Burny, A., Ratner, L. and Bex, F. (2009) Acetylation of the human T-cell leukemia virus type 1 Tax oncoprotein by p300 promotes activation of the NF-kappaB pathway. *Virology*, **386**, 68-78.

Lodi, P.J., Ernst, J.A., Kuszewski, J., Hickman, A.B., Engelman, A., Craigie, R., Clore, G.M. and Gronenborn, A.M. (1995) Solution structure of the DNA binding domain of HIV-1 integrase. *Biochemistry*, **34**, 9826-9833.

Lonard, D.M. and O'Malley B, W. (2007) Nuclear receptor coregulators: judges, juries, and executioners of cellular regulation. *Mol Cell*, **27**, 691-700.

Lu, R., Ghory, H.Z. and Engelman, A. (2005) Genetic analyses of conserved residues in the carboxyl-terminal domain of human immunodeficiency virus type 1 integrase. *J Virol*, **79**, 10356-10368.

Luduena, R.F. (1998) Multiple forms of tubulin: different gene products and covalent modifications. *Int Rev Cytol*, **178**, 207-275.

Lundblad, J.R., Kwok, R.P., Laurance, M.E., Harter, M.L. and Goodman, R.H. (1995) Adenoviral E1A-associated protein p300 as a functional homologue of the transcriptional co-activator CBP. *Nature*, **374**, 85-88.

MacRae, T.H. (1997) Tubulin post-translational modifications--enzymes and their mechanisms of action. *Eur J Biochem*, **244**, 265-278.

Madison, D.L., Yaciuk, P., Kwok, R.P. and Lundblad, J.R. (2002) Acetylation of the adenovirus-transforming protein E1A determines nuclear localization by disrupting association with importin-alpha. *J Biol Chem*, **277**, 38755-38763.

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.

Malek, S., Chen, Y., Huxford, T. and Ghosh, G. (2001) IkappaBbeta, but not IkappaBalpha, functions as a classical cytoplasmic inhibitor of NF-kappaB dimers by masking both NF-kappaB nuclear localization sequences in resting cells. *J Biol Chem*, **276**, 45225-45235.

Malet, I., Delelis, O., Soulie, C., Wirden, M., Tchertanov, L., Mottaz, P., Peytavin, G., Katlama, C., Mouscadet, J.F., Calvez, V. and Marcelin, A.G. (2009) Quasispecies variant dynamics during emergence of resistance to raltegravir in HIV-1-infected patients. *J Antimicrob Chemother*, **63**, 795-804.

Malet, I., Delelis, O., Valantin, M.A., Montes, B., Soulie, C., Wirden, M., Tchertanov, L., Peytavin, G., Reynes, J., Mouscadet, J.F., Katlama, C., Calvez, V. and Marcelin, A.G. (2008) Mutations associated with failure of raltegravir treatment affect integrase sensitivity to the inhibitor in vitro. *Antimicrob Agents Chemother*, **52**, 1351-1358.

Malim, M.H., Hauber, J., Le, S.Y., Maizel, J.V. and Cullen, B.R. (1989) The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature*, **338**, 254-257.

Manganaro, L., Lusic, M., Gutierrez, M.I., Cereseto, A., Del Sal, G. and Giacca, M. (2010) Concerted action of cellular JNK and Pin1 restricts HIV-1 genome integration to activated CD4+ T lymphocytes. *Nat Med*, **16**, 329-333.

Marchand, C., Johnson, A.A., Karki, R.G., Pais, G.C., Zhang, X., Cowansage, K., Patel, T.A., Nicklaus, M.C., Burke, T.R., Jr. and Pommier, Y. (2003) Metal-dependent inhibition of HIV-1 integrase by beta-diketo acids and resistance of the soluble double-mutant (F185K/C280S). *Mol Pharmacol*, **64**, 600-609.

Marchand, C., Neamati, N. and Pommier, Y. (2001) In vitro human immunodeficiency virus type 1 integrase assays. *Methods Enzymol*, **340**, 624-633.

Marcus, G.A., Silverman, N., Berger, S.L., Horiuchi, J. and Guarente, L. (1994) Functional similarity and physical association between GCN5 and ADA2: putative transcriptional adaptors. *Embo J*, **13**, 4807-4815.

Marinello, J., Marchand, C., Mott, B.T., Bain, A., Thomas, C.J. and Pommier, Y. (2008) Comparison of raltegravir and elvitegravir on HIV-1

integrase catalytic reactions and on a series of drug-resistant integrase mutants. *Biochemistry*, **47**, 9345-9354.

Maroun, M., Delelis, O., Coadou, G., Bader, T., Segeral, E., Mbemba, G., Petit, C., Sonigo, P., Rain, J.C., Mouscadet, J.F., Benarous, R. and Emiliani, S. (2006) Inhibition of early steps of HIV-1 replication by SNF5/Ini1. *J Biol Chem*, **281**, 22736-22743.

Marshall, H.M., Ronen, K., Berry, C., Llano, M., Sutherland, H., Saenz, D., Bickmore, W., Poeschla, E. and Bushman, F.D. (2007) Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. *PLoS One*, **2**, e1340.

Martinez, E., Kundu, T.K., Fu, J. and Roeder, R.G. (1998) A human SPT3-TAFII31-GCN5-L acetylase complex distinct from transcription factor IID. *J Biol Chem*, **273**, 23781-23785.

Martinez, E., Palhan, V.B., Tjernberg, A., Lymar, E.S., Gamper, A.M., Kundu, T.K., Chait, B.T. and Roeder, R.G. (2001) Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors in vivo. *Mol Cell Biol*, **21**, 6782-6795.

Martinez-Balbas, M.A., Bauer, U.M., Nielsen, S.J., Brehm, A. and Kouzarides, T. (2000) Regulation of E2F1 activity by acetylation. *Embo J*, **19**, 662-671.

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.

Marzio, G., Wagener, C., Gutierrez, M.I., Cartwright, P., Helin, K. and Giacca, M. (2000) E2F family members are differentially regulated by reversible acetylation. *J Biol Chem*, **275**, 10887-10892.

Matangkasombut, O. and Buratowski, S. (2003) Different sensitivities of bromodomain factors 1 and 2 to histone H4 acetylation. *Mol Cell*, **11**, 353-363.

Matsuyama, A., Shimazu, T., Sumida, Y., Saito, A., Yoshimatsu, Y., Seigneurin-Berny, D., Osada, H., Komatsu, Y., Nishino, N., Khochbin, S., Horinouchi, S. and Yoshida, M. (2002) In vivo destabilization of dynamic microtubules by HDAC6-mediated deacetylation. *Embo J*, **21**, 6820-6831.

Mattagajasingh, I., Kim, C.S., Naqvi, A., Yamamori, T., Hoffman, T.A., Jung, S.B., DeRicco, J., Kasuno, K. and Irani, K. (2007) SIRT1 promotes endothelium-dependent vascular relaxation by activating endothelial nitric oxide synthase. *Proc Natl Acad Sci U S A*, **104**, 14855-14860.

Mattsson, K., Kiss, C., Platt, G.M., Simpson, G.R., Kashuba, E., Klein, G., Schulz, T.F. and Szekely, L. (2002) Latent nuclear antigen of Kaposi's sarcoma herpesvirus/human herpesvirus-8 induces and relocates RING3 to nuclear heterochromatin regions. *J Gen Virol*, **83**, 179-188.

Maurice, T., Duclot, F., Meunier, J., Naert, G., Givalois, L., Meffre, J., Celerier, A., Jacquet, C., Copois, V., Mechti, N., Ozato, K. and Gongora, C. (2008) Altered memory capacities and response to stress in p300/CBP-associated factor (PCAF) histone acetylase knockout mice. *Neuropsychopharmacology*, **33**, 1584-1602.

McMahon, S.B., Van Buskirk, H.A., Dugan, K.A., Copeland, T.D. and Cole, M.D. (1998) The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. *Cell*, **94**, 363-374.

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.

Mitchell, R., Chiang, C.Y., Berry, C. and Bushman, F. (2003) Global analysis of cellular transcription following infection with an HIV-based vector. *Mol Ther*, **8**, 674-687.

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.

Mizzen, C.A., Yang, X.J., Kokubo, T., Brownell, J.E., Bannister, A.J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S.L., Kouzarides, T., Nakatani, Y. and Allis, C.D. (1996) The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell*, **87**, 1261-1270.

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

Mu, J.J., Tsay, Y.G., Juan, L.J., Fu, T.F., Huang, W.H., Chen, D.S. and Chen, P.J. (2004) The small delta antigen of hepatitis delta virus is an

acetylated protein and acetylation of lysine 72 may influence its cellular localization and viral RNA synthesis. *Virology*, **319**, 60-70.

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.

Mujtaba, S., He, Y., Zeng, L., Yan, S., Plotnikova, O., Sachchidanand, Sanchez, R., Zeleznik-Le, N.J., Ronai, Z. and Zhou, M.M. (2004) Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. *Mol Cell*, **13**, 251-263.

Mulder, L.C., Chakrabarti, L.A. and Muesing, M.A. (2002) Interaction of HIV-1 integrase with DNA repair protein hRad18. *J Biol Chem*, **277**, 27489-27493.

Mulder, L.C. and Muesing, M.A. (2000) Degradation of HIV-1 integrase by the N-end rule pathway. *J Biol Chem*, **275**, 29749-29753.

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.

Muller, S., Scaffidi, P., Degryse, B., Bonaldi, T., Ronfani, L., Agresti, A., Beltrame, M. and Bianchi, M.E. (2001) New EMBO members' review: the double life of HMGB1 chromatin protein: architectural factor and extracellular signal. *Embo J*, **20**, 4337-4340.

Munshi, N., Agalioti, T., Lomvardas, S., Merika, M., Chen, G. and Thanos, D. (2001) Coordination of a transcriptional switch by HMGI(Y) acetylation. *Science*, **293**, 1133-1136.

Munshi, N., Merika, M., Yie, J., Senger, K., Chen, G. and Thanos, D. (1998) Acetylation of HMG I(Y) by CBP turns off IFN beta expression by disrupting the enhanceosome. *Mol Cell*, **2**, 457-467.

Muratoglu, S., Georgieva, S., Papai, G., Scheer, E., Enunlu, I., Komonyi, O., Cserpan, I., Lebedeva, L., Nabirochkina, E., Udvardy, A., Tora, L. and Boros, I. (2003) Two different Drosophila ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes. *Mol Cell Biol*, **23**, 306-321.

Musacchio, A., Noble, M., Pauptit, R., Wierenga, R. and Saraste, M. (1992) Crystal structure of a Src-homology 3 (SH3) domain. *Nature*, **359**, 851-855.

Nagy, Z., Riss, A., Fujiyama, S., Krebs, A., Orpinell, M., Jansen, P., Cohen, A., Stunnenberg, H.G., Kato, S. and Tora, L. (2010) The metazoan ATAC and SAGA coactivator HAT complexes regulate different sets of inducible target genes. *Cell Mol Life Sci*, **67**, 611-628.

Nagy, Z. and Tora, L. (2007) Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene*, **26**, 5341-5357.

Nakai, H., Montini, E., Fuess, S., Storm, T.A., Grompe, M. and Kay, M.A. (2003) AAV serotype 2 vectors preferentially integrate into active genes in mice. *Nat Genet*, **34**, 297-302.

Nakatani, Y. (2001) Histone acetylases--versatile players. *Genes Cells*, **6**, 79-86.

Nakatani, Y. (2002) HIV-1 transcription: activation mediated by acetylation of Tat. *Structure*, **10**, 443-444.

Nasr, R., Chiari, E., El-Sabban, M., Mahieux, R., Kfoury, Y., Abdulhay, M., Yazbeck, V., Hermine, O., de The, H., Pique, C. and Bazarbachi, A. (2006) Tax ubiquitylation and sumoylation control critical cytoplasmic and nuclear steps of NF-kappaB activation. *Blood*, **107**, 4021-4029.

Nigg, E.A. (1997) Nucleocytoplasmic transport: signals, mechanisms and regulation. *Nature*, **386**, 779-787.

Nightingale, K.P., Wellinger, R.E., Sogo, J.M. and Becker, P.B. (1998) Histone acetylation facilitates RNA polymerase II transcription of the Drosophila hsp26 gene in chromatin. *Embo J*, **17**, 2865-2876.

Ocwieja, K.E., Brady, T.L., Ronen, K., Huegel, A., Roth, S.L., Schaller, T., James, L.C., Towers, G.J., Young, J.A., Chanda, S.K., Konig, R., Malani, N., Berry, C.C. and Bushman, F.D. (2011) HIV integration targeting: a pathway involving Transportin-3 and the nuclear pore protein RanBP2. *PLoS Pathog*, **7**, e1001313.

Ogryzko, V.V., Kotani, T., Zhang, X., Schiltz, R.L., Howard, T., Yang, X.J., Howard, B.H., Qin, J. and Nakatani, Y. (1998) Histone-like TAFs within the PCAF histone acetylase complex. *Cell*, **94**, 35-44.

Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H. and Nakatani, Y. (1996) The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell*, **87**, 953-959.

Ono, A. and Freed, E.O. (2005) Role of lipid rafts in virus replication. *Adv Virus Res*, **64**, 311-358.

Ono, A., Orenstein, J.M. and Freed, E.O. (2000) Role of the Gag matrix domain in targeting human immunodeficiency virus type 1 assembly. *J Virol*, **74**, 2855-2866.

Orpinell, M., Fournier, M., Riss, A., Nagy, Z., Krebs, A.R., Frontini, M. and Tora, L. (2010) The ATAC acetyl transferase complex controls mitotic progression by targeting non-histone substrates. *Embo J*, **29**, 2381-2394.

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.

Owen, D.J., Ornaghi, P., Yang, J.C., Lowe, N., Evans, P.R., Ballario, P., Neuhaus, D., Filetici, P. and Travers, A.A. (2000) The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *Embo J*, **19**, 6141-6149.

Pais, G.C., Zhang, X., Marchand, C., Neamati, N., Cowansage, K., Svarovskaia, E.S., Pathak, V.K., Tang, Y., Nicklaus, M., Pommier, Y. and Burke, T.R., Jr. (2002) Structure activity of 3-aryl-1,3-diketo-containing compounds as HIV-1 integrase inhibitors. *J Med Chem*, **45**, 3184-3194.

Palhan, V.B., Chen, S., Peng, G.H., Tjernberg, A., Gamper, A.M., Fan, Y., Chait, B.T., La Spada, A.R. and Roeder, R.G. (2005) Polyglutamine-expanded ataxin-7 inhibits STAGA histone acetyltransferase activity to produce retinal degeneration. *Proc Natl Acad Sci U S A*, **102**, 8472-8477.

Panet, A. and Cedar, H. (1977) Selective degradation of integrated murine leukemia proviral DNA by deoxyribonucleases. *Cell*, **11**, 933-940.

Parissi, V., Calmels, C., De Soultrait, V.R., Caumont, A., Fournier, M., Chaignepain, S. and Litvak, S. (2001) Functional interactions of human immunodeficiency virus type 1 integrase with human and yeast HSP60. *J Virol*, **75**, 11344-11353.

Parthun, M.R., Widom, J. and Gottschling, D.E. (1996) The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism. *Cell*, **87**, 85-94.

Patel, J.H., Du, Y., Ard, P.G., Phillips, C., Carella, B., Chen, C.J., Rakowski, C., Chatterjee, C., Lieberman, P.M., Lane, W.S., Blobel, G.A. and McMahon, S.B. (2004) The c-MYC oncoprotein is a substrate of the acetyltransferases hGCN5/PCAF and TIP60. *Mol Cell Biol*, **24**, 10826-10834.

Paulson, M., Press, C., Smith, E., Tanese, N. and Levy, D.E. (2002) IFN-Stimulated transcription through a TBP-free acetyltransferase complex escapes viral shutoff. *Nat Cell Biol*, **4**, 140-147.

Peloponese, J.M., Jr., Iha, H., Yedavalli, V.R., Miyazato, A., Li, Y., Haller, K., Benkirane, M. and Jeang, K.T. (2004) Ubiquitination of human T-cell leukemia virus type 1 tax modulates its activity. *J Virol*, **78**, 11686-11695.

Peng, C., Zhou, J., Liu, H.Y., Zhou, M., Wang, L.L., Zhang, Q.H., Yang, Y.X., Xiong, W., Shen, S.R., Li, X.L. and Li, G.Y. (2006) The transcriptional regulation role of BRD7 by binding to acetylated histone through bromodomain. *J Cell Biochem*, **97**, 882-892.

Peterlin, B.M. and Trono, D. (2003) Hide, shield and strike back: how HIV-infected cells avoid immune eradication. *Nat Rev Immunol*, **3**, 97-107.

Pina, B., Berger, S., Marcus, G.A., Silverman, N., Agapite, J. and Guarente, L. (1993) ADA3: a gene, identified by resistance to GAL4-VP16, with properties similar to and different from those of ADA2. *Mol Cell Biol*, **13**, 5981-5989.

Pivot-Pajot, C., Caron, C., Govin, J., Vion, A., Rousseaux, S. and Khochbin, S. (2003) Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. *Mol Cell Biol*, **23**, 5354-5365.

Polesskaya, A., Duquet, A., Naguibneva, I., Weise, C., Vervisch, A., Bengal, E., Hucho, F., Robin, P. and Harel-Bellan, A. (2000) CREB-binding protein/p300 activates MyoD by acetylation. *J Biol Chem*, **275**, 34359-34364.

Polesskaya, A., Naguibneva, I., Duquet, A., Bengal, E., Robin, P. and Harel-Bellan, A. (2001) Interaction between acetylated MyoD and the bromodomain of CBP and/or p300. *Mol Cell Biol*, **21**, 5312-5320.

Polevoda, B. and Sherman, F. (2002) The diversity of acetylated proteins. *Genome Biol*, **3**, reviews0006.

Pommier, Y., Johnson, A.A. and Marchand, C. (2005) Integrase inhibitors to treat HIV/AIDS. *Nat Rev Drug Discov*, **4**, 236-248.

Poulin, D.L., Kung, A.L. and DeCaprio, J.A. (2004) p53 targets simian virus 40 large T antigen for acetylation by CBP. *J Virol*, **78**, 8245-8253.

Pray-Grant, M.G., Schieltz, D., McMahon, S.J., Wood, J.M., Kennedy, E.L., Cook, R.G., Workman, J.L., Yates, J.R., 3rd and Grant, P.A. (2002) The novel SLIK histone acetyltransferase complex functions in the yeast retrograde response pathway. *Mol Cell Biol*, **22**, 8774-8786.

Pruss, D., Bushman, F.D. and Wolffe, A.P. (1994a) Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. *Proc Natl Acad Sci U S A*, **91**, 5913-5917.

Pruss, D., Reeves, R., Bushman, F.D. and Wolffe, A.P. (1994b) The influence of DNA and nucleosome structure on integration events directed by HIV integrase. *J Biol Chem*, **269**, 25031-25041.

Pryciak, P.M., Muller, H.P. and Varmus, H.E. (1992) Simian virus 40 minichromosomes as targets for retroviral integration in vivo. *Proc Natl Acad Sci U S A*, **89**, 9237-9241.

Pryciak, P.M. and Varmus, H.E. (1992) Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell*, **69**, 769-780.

Rain, J.C., Cribier, A., Gerard, A., Emiliani, S. and Benarous, R. (2009) Yeast two-hybrid detection of integrase-host factor interactions. *Methods*, **47**, 291-297.

Ran, Q. and Pereira-Smith, O.M. (2000) Identification of an alternatively spliced form of the Tat interactive protein (Tip60), Tip60(beta). *Gene*, **258**, 141-146.

Ratner, L., Portis, T., Robek, M., Harding, J. and Grossman, W. (2000) Studies of the immortalizing activity of HTLV type 1 Tax, using an infectious molecular clone and transgenic mice. *AIDS Res Hum Retroviruses*, **16**, 1647-1651.

Reed, N.A., Cai, D., Blasius, T.L., Jih, G.T., Meyhofer, E., Gaertig, J. and Verhey, K.J. (2006) Microtubule acetylation promotes kinesin-1 binding and transport. *Curr Biol*, **16**, 2166-2172.

Reifsnyder, C., Lowell, J., Clarke, A. and Pillus, L. (1996) Yeast SAS silencing genes and human genes associated with AML and HIV-1 Tat interactions are homologous with acetyltransferases. *Nat Genet*, **14**, 42-49.

Roberts, S.M. and Winston, F. (1996) SPT20/ADA5 encodes a novel protein functionally related to the TATA-binding protein and important for transcription in Saccharomyces cerevisiae. *Mol Cell Biol*, **16**, 3206-3213.

Roth, S.Y., Denu, J.M. and Allis, C.D. (2001) Histone acetyltransferases. *Annu Rev Biochem*, **70**, 81-120.

Rowland, S.J. and Dyke, K.G. (1990) Tn552, a novel transposable element from Staphylococcus aureus. *Mol Microbiol*, **4**, 961-975.

Ruiz-Garcia, A.B., Sendra, R., Pamblanco, M. and Tordera, V. (1997) Gcn5p is involved in the acetylation of histone H3 in nucleosomes. *FEBS Lett*, **403**, 186-190.

Rundlett, S.E., Carmen, A.A., Kobayashi, R., Bavykin, S., Turner, B.M. and Grunstein, M. (1996) HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. *Proc Natl Acad Sci U S A*, **93**, 14503-14508.

Sachchidanand, Resnick-Silverman, L., Yan, S., Mutjaba, S., Liu, W.J., Zeng, L., Manfredi, J.J. and Zhou, M.M. (2006) Target structure-based discovery of small molecules that block human p53 and CREB binding protein association. *Chem Biol*, **13**, 81-90.

Sampath, S.C., Marazzi, I., Yap, K.L., Krutchinsky, A.N., Mecklenbrauker, I., Viale, A., Rudensky, E., Zhou, M.M., Chait, B.T. and Tarakhovsky, A. (2007) Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. *Mol Cell*, **27**, 596-608.

Sanchez, R. and Zhou, M.M. (2009) The role of human bromodomains in chromatin biology and gene transcription. *Curr Opin Drug Discov Devel*, **12**, 659-665.

Sandeen, G., Wood, W.I. and Felsenfeld, G. (1980) The interaction of high mobility proteins HMG14 and 17 with nucleosomes. *Nucleic Acids Res*, **8**, 3757-3778.

Sandmeyer, S. (2003) Integration by design. *Proc Natl Acad Sci U S A*, **100**, 5586-5588.

Sartorelli, V., Puri, P.L., Hamamori, Y., Ogryzko, V., Chung, G., Nakatani, Y., Wang, J.Y. and Kedes, L. (1999) Acetylation of MyoD directed by PCAF is necessary for the execution of the muscle program. *Mol Cell*, **4**, 725-734.

Sattentau, Q.J. and Weiss, R.A. (1988) The CD4 antigen: physiological ligand and HIV receptor. *Cell*, **52**, 631-633.

Scaffidi, P., Misteli, T. and Bianchi, M.E. (2002) Release of chromatin protein HMGB1 by necrotic cells triggers inflammation. *Nature*, **418**, 191-195.

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.

Scottoline, B.P., Chow, S., Ellison, V. and Brown, P.O. (1997) Disruption of the terminal base pairs of retroviral DNA during integration. *Genes Dev*, **11**, 371-382.

Sendra, R., Tse, C. and Hansen, J.C. (2000) The yeast histone acetyltransferase A2 complex, but not free Gcn5p, binds stably to nucleosomal arrays. *J Biol Chem*, **275**, 24928-24934.

Shembade, N., Harhaj, N.S., Yamamoto, M., Akira, S. and Harhaj, E.W. (2007) The human T-cell leukemia virus type 1 Tax oncoprotein requires the ubiquitin-conjugating enzyme Ubc13 for NF-kappaB activation. *J Virol*, **81**, 13735-13742.

Shen, W., Xu, C., Huang, W., Zhang, J., Carlson, J.E., Tu, X., Wu, J. and Shi, Y. (2007) Solution structure of human Brg1 bromodomain and its specific binding to acetylated histone tails. *Biochemistry*, **46**, 2100-2110.

Sherman, M.P. and Greene, W.C. (2002) Slipping through the door: HIV entry into the nucleus. *Microbes Infect*, **4**, 67-73.

Shiama, N. (1997) The p300/CBP family: integrating signals with transcription factors and chromatin. *Trends Cell Biol*, **7**, 230-236.

Shimazu, T., Komatsu, Y., Nakayama, K.I., Fukazawa, H., Horinouchi, S. and Yoshida, M. (2006) Regulation of SV40 large T-antigen stability by reversible acetylation. *Oncogene*, **25**, 7391-7400.

Shimura, K., Kodama, E., Sakagami, Y., Matsuzaki, Y., Watanabe, W., Yamataka, K., Watanabe, Y., Ohata, Y., Doi, S., Sato, M., Kano, M., Ikeda, S. and Matsuoka, M. (2008) Broad antiretroviral activity and resistance profile of the novel human immunodeficiency virus integrase inhibitor elvitegravir (JTK-303/GS-9137). *J Virol*, **82**, 764-774.

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

Shun, M.C., Raghavendra, N.K., Vandegraaff, N., Daigle, J.E., Hughes, S., Kellam, P., Cherepanov, P. and Engelman, A. (2007b) LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev*, **21**, 1767-1778.

Sichtig, N., Sierra, S., Kaiser, R., Daumer, M., Reuter, S., Schulter, E., Altmann, A., Fatkenheuer, G., Dittmer, U., Pfister, H. and Esser, S. (2009) Evolution of raltegravir resistance during therapy. *J Antimicrob Chemother*, **64**, 25-32.

Silva, J.M., Li, M.Z., Chang, K., Ge, W., Golding, M.C., Rickles, R.J., Siolas, D., Hu, G., Paddison, P.J., Schlabach, M.R., Sheth, N., Bradshaw, J., Burchard, J., Kulkarni, A., Cavet, G., Sachidanandam, R., McCombie, W.R., Cleary, M.A., Elledge, S.J. and Hannon, G.J. (2005) Second-generation shRNA libraries covering the mouse and human genomes. *Nat Genet*, **37**, 1281-1288.

Silvers, R.M., Smith, J.A., Schowalter, M., Litwin, S., Liang, Z., Geary, K. and Daniel, R. (2010) Modification of integration site preferences of an HIV-1-based vector by expression of a novel synthetic protein. *Hum Gene Ther*, **21**, 337-349.

Smith, E.R., Belote, J.M., Schiltz, R.L., Yang, X.J., Moore, P.A., Berger, S.L., Nakatani, Y. and Allis, C.D. (1998) Cloning of Drosophila GCN5: conserved features among metazoan GCN5 family members. *Nucleic Acids Res*, **26**, 2948-2954.

Sorin, M., Yung, E., Wu, X. and Kalpana, G.V. (2006) HIV-1 replication in cell lines harboring INI1/hSNF5 mutations. *Retrovirology*, **3**, 56.

Spotts, G.D., Patel, S.V., Xiao, Q. and Hann, S.R. (1997) Identification of downstream-initiated c-Myc proteins which are dominant-negative inhibitors of transactivation by full-length c-Myc proteins. *Mol Cell Biol*, **17**, 1459-1468.

Stade, K., Ford, C.S., Guthrie, C. and Weis, K. (1997) Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell*, **90**, 1041-1050.

Starai, V.J., Celic, I., Cole, R.N., Boeke, J.D. and Escalante-Semerena, J.C. (2002) Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. *Science*, **298**, 2390-2392.

Sterner, D.E., Belotserkovskaya, R. and Berger, S.L. (2002) SALSA, a variant of yeast SAGA, contains truncated Spt7, which correlates with activated transcription. *Proc Natl Acad Sci U S A*, **99**, 11622-11627.

Sterner, D.E. and Berger, S.L. (2000) Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev*, **64**, 435-459.

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

Stimson, L., Wood, V., Khan, O., Fotheringham, S. and La Thangue, N.B. (2009) HDAC inhibitor-based therapies and haematological malignancy. *Ann Oncol*, **20**, 1293-1302.

Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications. *Nature*, **403**, 41-45.

Suganuma, T., Gutierrez, J.L., Li, B., Florens, L., Swanson, S.K., Washburn, M.P., Abmayr, S.M. and Workman, J.L. (2008) ATAC is a double histone acetyltransferase complex that stimulates nucleosome sliding. *Nat Struct Mol Biol*, **15**, 364-372.

Sullivan, N., Sun, Y., Sattentau, Q., Thali, M., Wu, D., Denisova, G., Gershoni, J., Robinson, J., Moore, J. and Sodroski, J. (1998) CD4-Induced conformational changes in the human immunodeficiency virus type 1 gp120 glycoprotein: consequences for virus entry and neutralization. *J Virol*, **72**, 4694-4703.

Szebeni, A., Mehrotra, B., Baumann, A., Adam, S.A., Wingfield, P.T. and Olson, M.O. (1997) Nucleolar protein B23 stimulates nuclear import of the HIV-1 Rev protein and NLS-conjugated albumin. *Biochemistry*, **36**, 3941-3949.

Takechi, S. and Nakayama, T. (1999) Sas3 is a histone acetyltransferase and requires a zinc finger motif. *Biochem Biophys Res Commun*, **266**, 405-410.

Tan, W., Dong, Z., Wilkinson, T.A., Barbas, C.F., 3rd and Chow, S.A. (2006) Human immunodeficiency virus type 1 incorporated with fusion proteins consisting of integrase and the designed polydactyl zinc finger protein E2C can bias integration of viral DNA into a predetermined chromosomal region in human cells. *J Virol*, **80**, 1939-1948.

Tan, W., Zhu, K., Segal, D.J., Barbas, C.F., 3rd and Chow, S.A. (2004) Fusion proteins consisting of human immunodeficiency virus type 1 integrase and the designed polydactyl zinc finger protein E2C direct integration of viral DNA into specific sites. *J Virol*, **78**, 1301-1313.

Tasaki, T., Mulder, L.C., Iwamatsu, A., Lee, M.J., Davydov, I.V., Varshavsky, A., Muesing, M. and Kwon, Y.T. (2005) A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-degrons. *Mol Cell Biol*, **25**, 7120-7136.

Taunton, J., Hassig, C.A. and Schreiber, S.L. (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science*, **272**, 408-411.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D. and Patel, D.J. (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol*, **14**, 1025-1040.

Terreni, M., Valentini, P., Liverani, V., Gutierrez, M.I., Di Primio, C., Di Fenza, A., Tozzini, V., Allouch, A., Albanese, A., Giacca, M. and Cereseto, A. (2010) GCN5-dependent acetylation of HIV-1 integrase enhances viral integration. *Retrovirology*, **7**, 18.

Thevenet, L., Mejean, C., Moniot, B., Bonneaud, N., Galeotti, N., Aldrian-Herrada, G., Poulat, F., Berta, P., Benkirane, M. and Boizet-Bonhoure, B. (2004) Regulation of human SRY subcellular distribution by its acetylation/deacetylation. *Embo J*, **23**, 3336-3345.

Thys, W., De Houwer, S., Demeulemeester, J., Taltynov, O., Vancraenenbroeck, R., Gerard, M., De Rijck, J., Gijsbers, R., Christ, F. and Debyser, Z. (2011) Interplay between HIV entry and transportin-SR2 dependency. *Retrovirology*, **8**, 7.

Timmermann, S., Lehrmann, H., Polesskaya, A. and Harel-Bellan, A. (2001) Histone acetylation and disease. *Cell Mol Life Sci*, **58**, 728-736.

Topper, M., Luo, Y., Zhadina, M., Mohammed, K., Smith, L. and Muesing, M.A. (2007) Posttranslational acetylation of the human immunodeficiency

virus type 1 integrase carboxyl-terminal domain is dispensable for viral replication. *J Virol*, **81**, 3012-3017.

Tse, C., Georgieva, E.I., Ruiz-Garcia, A.B., Sendra, R. and Hansen, J.C. (1998) Gcn5p, a transcription-related histone acetyltransferase, acetylates nucleosomes and folded nucleosomal arrays in the absence of other protein subunits. *J Biol Chem*, **273**, 32388-32392.

Tsuruno, C., Ohe, K., Kuramitsu, M., Kohma, T., Takahama, Y., Hamaguchi, Y., Hamaguchi, I. and Okuma, K. (2011) HMGA1a is involved in specific splice site regulation of human immunodeficiency virus type 1. *Biochem Biophys Res Commun*, **406**, 512-517.

Turelli, P., Doucas, V., Craig, E., Mangeat, B., Klages, N., Evans, R., Kalpana, G. and Trono, D. (2001) Cytoplasmic recruitment of INI1 and PML on incoming HIV preintegration complexes: interference with early steps of viral replication. *Mol Cell*, **7**, 1245-1254.

Turlure, F., Devroe, E., Silver, P.A. and Engelman, A. (2004) Human cell proteins and human immunodeficiency virus DNA integration. *Front Biosci*, **9**, 3187-3208.

Turlure, F., Maertens, G., Rahman, S., Cherepanov, P. and Engelman, A. (2006) A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. *Nucleic Acids Res*, **34**, 1653-1665.

Ugrinova, I., Pasheva, E.A., Armengaud, J. and Pashev, I.G. (2001) In vivo acetylation of HMG1 protein enhances its binding affinity to distorted DNA structures. *Biochemistry*, **40**, 14655-14660.

UNAIDS. (2010) Report on the global AIDS Epidemic.

Ura, K., Araki, M., Saeki, H., Masutani, C., Ito, T., Iwai, S., Mizukoshi, T., Kaneda, Y. and Hanaoka, F. (2001) ATP-dependent chromatin remodeling facilitates nucleotide excision repair of UV-induced DNA lesions in synthetic dinucleosomes. *Embo J*, **20**, 2004-2014.

van 't Wout, A.B., Lehrman, G.K., Mikheeva, S.A., O'Keeffe, G.C., Katze, M.G., Bumgarner, R.E., Geiss, G.K. and Mullins, J.I. (2003) Cellular gene expression upon human immunodeficiency virus type 1 infection of CD4(+)-T-cell lines. *J Virol*, **77**, 1392-1402.

van Gent, D.C., Groeneger, A.A. and Plasterk, R.H. (1992) Mutational analysis of the integrase protein of human immunodeficiency virus type 2. *Proc Natl Acad Sci U S A*, **89**, 9598-9602.

van Gent, D.C., Vink, C., Groeneger, A.A. and Plasterk, R.H. (1993) Complementation between HIV integrase proteins mutated in different domains. *Embo J*, **12**, 3261-3267.

Van Maele, B., Busschots, K., Vandekerckhove, L., Christ, F. and Debyser, Z. (2006) Cellular co-factors of HIV-1 integration. *Trends Biochem Sci*, **31**, 98-105.

Vandegraaff, N., Devroe, E., Turlure, F., Silver, P.A. and Engelman, A. (2006) Biochemical and genetic analyses of integrase-interacting proteins lens epithelium-derived growth factor (LEDGF)/p75 and hepatoma-derived growth factor related protein 2 (HRP2) in preintegration complex function and HIV-1 replication. *Virology*, **346**, 415-426.

Vandegraaff, N. and Engelman, A. (2007) Molecular mechanisms of HIV integration and therapeutic intervention. *Expert Rev Mol Med*, **9**, 1-19.

Vandekerckhove, L., Christ, F., Van Maele, B., De Rijck, J., Gijsbers, R., Van den Haute, C., Witvrouw, M. and Debyser, Z. (2006) Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. *J Virol*, **80**, 1886-1896.

Vanegas, M., Llano, M., Delgado, S., Thompson, D., Peretz, M. and Poeschla, E. (2005) Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering. *J Cell Sci*, **118**, 1733-1743.

Vettese-Dadey, M., Grant, P.A., Hebbes, T.R., Crane- Robinson, C., Allis, C.D. and Workman, J.L. (1996) Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *Embo J*, **15**, 2508-2518.

von Schwedler, U.K., Stemmler, T.L., Klishko, V.Y., Li, S., Albertine, K.H., Davis, D.R. and Sundquist, W.I. (1998) Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *Embo J*, **17**, 1555-1568.

Vries, R.G., Prudenziati, M., Zwartjes, C., Verlaan, M., Kalkhoven, E. and Zantema, A. (2001) A specific lysine in c-Jun is required for transcriptional repression by E1A and is acetylated by p300. *Embo J*, **20**, 6095-6103.

Walter, W., Clynes, D., Tang, Y., Marmorstein, R., Mellor, J. and Berger, S.L. (2008) 14-3-3 interaction with histone H3 involves a dual modification pattern of phosphoacetylation. *Mol Cell Biol*, **28**, 2840-2849.

Wang, J.Y., Ling, H., Yang, W. and Craigie, R. (2001) Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. *Embo J*, **20**, 7333-7343.

Wang, L., Mizzen, C., Ying, C., Candau, R., Barlev, N., Brownell, J., Allis, C.D. and Berger, S.L. (1997) Histone acetyltransferase activity is conserved between yeast and human GCN5 and is required for complementation of growth and transcriptional activation. *Mol Cell Biol*, **17**, 519-527.

Wang, W., Cote, J., Xue, Y., Zhou, S., Khavari, P.A., Biggar, S.R., Muchardt, C., Kalpana, G.V., Goff, S.P., Yaniv, M., Workman, J.L. and Crabtree, G.R. (1996) Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *Embo J*, **15**, 5370-5382.

Wang, W., Yang, X., Kawai, T., Lopez de Silanes, I., Mazan-Mamczarz, K., Chen, P., Chook, Y.M., Quensel, C., Kohler, M. and Gorospe, M. (2004) AMP-activated protein kinase-regulated phosphorylation and acetylation of importin alpha1: involvement in the nuclear import of RNA-binding protein HuR. *J Biol Chem*, **279**, 48376-48388.

Wang, Y.L., Faiola, F., Xu, M., Pan, S. and Martinez, E. (2008) Human ATAC Is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein. *J Biol Chem*, **283**, 33808-33815.

Weake, V.M., Lee, K.K., Guelman, S., Lin, C.H., Seidel, C., Abmayr, S.M. and Workman, J.L. (2008) SAGA-mediated H2B deubiquitination controls the development of neuronal connectivity in the Drosophila visual system. *Embo J*, **27**, 394-405.

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.

Westermann, S. and Weber, K. (2003) Post-translational modifications regulate microtubule function. *Nat Rev Mol Cell Biol*, **4**, 938-947.

Wieczorek, E., Brand, M., Jacq, X. and Tora, L. (1998) Function of TAF(II)-containing complex without TBP in transcription by RNA polymerase II. *Nature*, **393**, 187-191.

Winston, F. and Sudarsanam, P. (1998) The SAGA of Spt proteins and transcriptional analysis in yeast: past, present, and future. *Cold Spring Harb Symp Quant Biol*, **63**, 553-561.

Winter, S., Simboeck, E., Fischle, W., Zupkovitz, G., Dohnal, I., Mechtler, K., Ammerer, G. and Seiser, C. (2008) 14-3-3 proteins recognize a histone code at histone H3 and are required for transcriptional activation. *Embo J*, **27**, 88-99.

Woodward, C.L., Prakobwanakit, S., Mosessian, S. and Chow, S.A. (2009) Integrase interacts with nucleoporin NUP153 to mediate the nuclear import of human immunodeficiency virus type 1. *J Virol*, **83**, 6522-6533.

Wu, X., Li, Y., Crise, B. and Burgess, S.M. (2003) Transcription start regions in the human genome are favored targets for MLV integration. *Science*, **300**, 1749-1751.

Wu, X., Li, Y., Crise, B., Burgess, S.M. and Munroe, D.J. (2005) Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. *J Virol*, **79**, 5211-5214.

Wyatt, R. and Sodroski, J. (1998) The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science*, **280**, 1884-1888.

Wyce, A., Xiao, T., Whelan, K.A., Kosman, C., Walter, W., Eick, D., Hughes, T.R., Krogan, N.J., Strahl, B.D. and Berger, S.L. (2007) H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. *Mol Cell*, **27**, 275-288.

Xie, A.Y., Bermudez, V.P. and Folk, W.R. (2002) Stimulation of DNA replication from the polyomavirus origin by PCAF and GCN5 acetyltransferases: acetylation of large T antigen. *Mol Cell Biol*, **22**, 7907-7918.

Xu, W., Edmondson, D.G., Evrard, Y.A., Wakamiya, M., Behringer, R.R. and Roth, S.Y. (2000) Loss of Gcn5l2 leads to increased apoptosis and mesodermal defects during mouse development. *Nat Genet*, **26**, 229-232.

Xu, W., Edmondson, D.G. and Roth, S.Y. (1998) Mammalian GCN5 and P/CAF acetyltransferases have homologous amino-terminal domains important for recognition of nucleosomal substrates. *Mol Cell Biol*, **18**, 5659-5669.

Yamamoto, T. and Horikoshi, M. (1997) Novel substrate specificity of the histone acetyltransferase activity of HIV-1-Tat interactive protein Tip60. *J Biol Chem*, **272**, 30595-30598.

Yamamoto, Y. and Gaynor, R.B. (2004) IkappaB kinases: key regulators of the NF-kappaB pathway. *Trends Biochem Sci*, **29**, 72-79.

Yamashita, M. and Emerman, M. (2004) Capsid is a dominant determinant of retrovirus infectivity in nondividing cells. *J Virol*, **78**, 5670-5678.

Yamashita, M., Perez, O., Hope, T.J. and Emerman, M. (2007) Evidence for direct involvement of the capsid protein in HIV infection of nondividing cells. *PLoS Pathog*, **3**, 1502-1510.

Yamauchi, T., Yamauchi, J., Kuwata, T., Tamura, T., Yamashita, T., Bae, N., Westphal, H., Ozato, K. and Nakatani, Y. (2000) Distinct but overlapping roles of histone acetylase PCAF and of the closely related PCAF-B/GCN5 in mouse embryogenesis. *Proc Natl Acad Sci U S A*, **97**, 11303-11306.

Yanagisawa, J., Kitagawa, H., Yanagida, M., Wada, O., Ogawa, S., Nakagomi, M., Oishi, H., Yamamoto, Y., Nagasawa, H., McMahon, S.B., Cole, M.D., Tora, L., Takahashi, N. and Kato, S. (2002) Nuclear receptor function requires a TFTC-type histone acetyl transferase complex. *Mol Cell*, **9**, 553-562.

Yang, W.M., Inouye, C., Zeng, Y., Bearss, D. and Seto, E. (1996a) Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3. *Proc Natl Acad Sci U S A*, **93**, 12845-12850.

Yang, X.J. (2004a) Lysine acetylation and the bromodomain: a new partnership for signaling. *Bioessays*, **26**, 1076-1087.

Yang, X.J. (2004b) The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases. *Nucleic Acids Res*, **32**, 959-976.

Yang, X.J. (2005) Multisite protein modification and intramolecular signaling. *Oncogene*, **24**, 1653-1662.

Yang, X.J. and Gregoire, S. (2007) Metabolism, cytoskeleton and cellular signalling in the grip of protein Nepsilon - and O-acetylation. *EMBO Rep*, **8**, 556-562.

Yang, X.J., Ogryzko, V.V., Nishikawa, J., Howard, B.H. and Nakatani, Y. (1996b) A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature*, **382**, 319-324.

Yang, X.J. and Seto, E. (2007) HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene*, **26**, 5310-5318.

Yant, S.R., Wu, X., Huang, Y., Garrison, B., Burgess, S.M. and Kay, M.A. (2005) High-resolution genome-wide mapping of transposon integration in mammals. *Mol Cell Biol*, **25**, 2085-2094.

Yao, Y.L., Yang, W.M. and Seto, E. (2001) Regulation of transcription factor YY1 by acetylation and deacetylation. *Mol Cell Biol*, **21**, 5979-5991.

Yu, H., Rosen, M.K., Shin, T.B., Seidel-Dugan, C., Brugge, J.S. and Schreiber, S.L. (1992) Solution structure of the SH3 domain of Src and identification of its ligand-binding site. *Science*, **258**, 1665-1668.

Yu, Q., Minoda, Y., Yoshida, R., Yoshida, H., Iha, H., Kobayashi, T., Yoshimura, A. and Takaesu, G. (2008) HTLV-1 Tax-mediated TAK1 activation involves TAB2 adapter protein. *Biochem Biophys Res Commun*, **365**, 189-194.

Yuan, Z.L., Guan, Y.J., Chatterjee, D. and Chin, Y.E. (2005) Stat3 dimerization regulated by reversible acetylation of a single lysine residue. *Science*, **307**, 269-273.

Yung, E., Sorin, M., Pal, A., Craig, E., Morozov, A., Delattre, O., Kappes, J., Ott, D. and Kalpana, G.V. (2001) Inhibition of HIV-1 virion production by a transdominant mutant of integrase interactor 1. *Nat Med*, **7**, 920-926.

Yung, E., Sorin, M., Wang, E.J., Perumal, S., Ott, D. and Kalpana, G.V. (2004) Specificity of interaction of INI1/hSNF5 with retroviral integrases and its functional significance. *J Virol*, **78**, 2222-2231.

Zaitseva, L., Cherepanov, P., Leyens, L., Wilson, S.J., Rasaiyaah, J. and Fassati, A. (2009) HIV-1 exploits importin 7 to maximize nuclear import of its DNA genome. *Retrovirology*, **6**, 11.

Zamborlini, A., Coiffic, A., Beauclair, G., Delelis, O., Paris, J., Koh, Y., Magne, F., Giron, M.L., Tobaly-Tapiero, J., Deprez, E., Emiliani, S., Engelman, A., de The, H. and Saib, A. (2011) Impairment of human immunodeficiency virus type-1 integrase SUMOylation correlates with an early replication defect. *J Biol Chem*, **286**, 21013-21022.

Zeng, L., Zhang, Q., Gerona-Navarro, G., Moshkina, N. and Zhou, M.M. (2008) Structural basis of site-specific histone recognition by the bromodomains of human coactivators PCAF and CBP/p300. *Structure*, **16**, 643-652.

Zeng, Y., Wagner, E.J. and Cullen, B.R. (2002) Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell*, **9**, 1327-1333.

Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L. and Charneau, P. (2000) HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell*, **101**, 173-185.

Zhang, Q., Yao, H., Vo, N. and Goodman, R.H. (2000) Acetylation of adenovirus E1A regulates binding of the transcriptional corepressor CtBP. *Proc Natl Acad Sci U S A*, **97**, 14323-14328.

Zhang, R., Mehla, R. and Chauhan, A. (2010) Perturbation of host nuclear membrane component RanBP2 impairs the nuclear import of human immunodeficiency virus -1 preintegration complex (DNA). *PLoS One*, **5**, e15620.

Zhang, W. and Bieker, J.J. (1998) Acetylation and modulation of erythroid Kruppel-like factor (EKLF) activity by interaction with histone acetyltransferases. *Proc Natl Acad Sci U S A*, **95**, 9855-9860.

Zhang, W., Kadam, S., Emerson, B.M. and Bieker, J.J. (2001) Site-specific acetylation by p300 or CREB binding protein regulates erythroid Kruppel-like factor transcriptional activity via its interaction with the SWI-SNF complex. *Mol Cell Biol*, **21**, 2413-2422.

Zhang, X., Yuan, Z., Zhang, Y., Yong, S., Salas-Burgos, A., Koomen, J., Olashaw, N., Parsons, J.T., Yang, X.J., Dent, S.R., Yao, T.P., Lane, W.S.

and Seto, E. (2007) HDAC6 modulates cell motility by altering the acetylation level of cortactin. *Mol Cell*, **27**, 197-213.

Zhao, Y., Lang, G., Ito, S., Bonnet, J., Metzger, E., Sawatsubashi, S., Suzuki, E., Le Guezennec, X., Stunnenberg, H.G., Krasnov, A., Georgieva, S.G., Schule, R., Takeyama, K., Kato, S., Tora, L. and Devys, D. (2008) A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. *Mol Cell*, **29**, 92-101.

Zheng, R., Jenkins, T.M. and Craigie, R. (1996) Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proc Natl Acad Sci U S A*, **93**, 13659-13664.

Zheng, Y., Ao, Z., Wang, B., Jayappa, K.D. and Yao, X. (2011) Host protein Ku70 binds and protects HIV-1 integrase from proteasomal degradation and is required for HIV replication. *J Biol Chem*, **286**, 17722-17735.

Zhou, L., Sokolskaja, E., Jolly, C., James, W., Cowley, S.A. and Fassati, A. (2011) Transportin 3 promotes a nuclear maturation step required for efficient HIV-1 integration. *PLoS Pathog*, **7**, e1002194.

Zhu, Y., Dai, J., Fuerst, P.G. and Voytas, D.F. (2003) Controlling integration specificity of a yeast retrotransposon. *Proc Natl Acad Sci U S A*, **100**, 5891-5895.