[Frontiers in Bioscience 9, 3187-3208, September 1, 2004]

HUMAN CELL PROTEINS AND HUMAN IMMUNODEFICIENCY VIRUS DNA INTEGRATION

Fanny Turlure [1], Eric Devroe [2], Pamela A. Silver [2] and Alan Engelman [1]

[1] Department of Pathology, Harvard Medical School, Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts, 02115, [2] Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Department of Cancer Cell Biology, Dana-Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts, 02115

TABLE OF CONTENTS

1. ABSTRACT

Integration, catalyzed by the viral integrase (IN) protein, is a crucial step in the life cycle of all retroviruses including human immunodeficiency virus type 1 (HIV-1). Although purified HIV-1 IN protein is sufficient to catalyze the DNA breakage and joining steps of integration in the absence of any other protein factor, a number of studies indicate that cellular proteins participate in the integration process in cells. These host cell proteins have been proposed to act through binding the pre-integrated viral cDNA substrate, by directly interacting with the IN protein, and/or by repairing the single-stranded DNA gaps that occur at viral/chromosomal DNA junctions during integration. In this paper we summarize the identification and potential roles of specific cell factors in HIV-1 integration. We also present experimental results of human cell proteins that coimmunoprecipitated with HIV-1 IN following its expression in HeLa cells and discuss these results in light of the previously-identified integration cofactors.

2. INTRODUCTION

Retroviruses are distinct amongst animal viruses in that they are the only viruses whose replication requires the stable integration of their genetic material into that of the host. This feature is made possible by the viral IN, a specialized DNA recombinase that is packaged into retroviral particles (reviewed in reference 1). Soon after entering a susceptible target cell, uncoated retroviral particles undergo reverse transcription in the context of high molecular weight nucleoprotein complexes known as reverse transcription complexes (2-4; reviewed in reference 5). Reverse transcription generates a linear, double-stranded cDNA containing a copy of the viral long terminal repeat (LTR) at each end. In order for a productive infection to occur, HIV-1 cDNA must become integrated into a host cell chromosome (6-13).

A number of studies have indicated that normal cellular proteins are likely to help the IN accomplish functional HIV-1 cDNA integration in infected cells (reviewed in reference 14). Since many of these proteins were identified using biochemical screens, their physiological relevant role(s) in integration requires confirmation through genetic tests. Here we review the identification and proposed roles of cellular proteins in HIV-1 cDNA integration and discuss experiments that

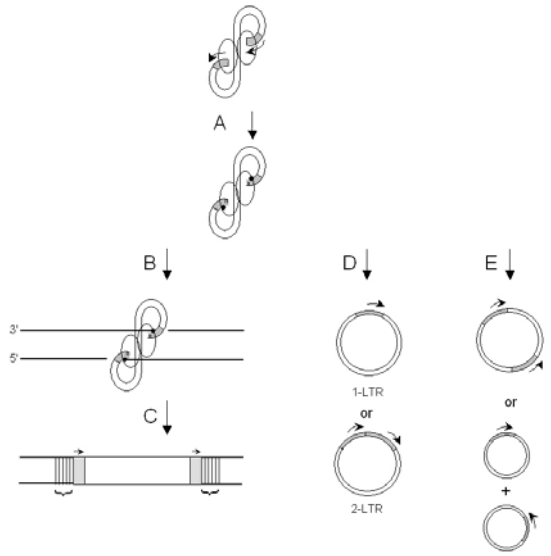

Figure 1. Mechanism of retroviral integration and alternative cDNA pathways. After completion of reverse transcription, multimers of IN (empty ovals) bind the *att* sites at the tips of the LTRs (light grey boxes; arrow tips represent the U5 ends of the LTRs). (A) During 3' processing, IN removes two nucleotides from the *att* sites, yielding recessed 3' ends (black balls). (B) During the subsequent DNA strand transfer step, IN uses the 3' ends to attack the two strands of host chromosomal DNA (thick lines). The resulting DNA product is a gapped structure with the 5' ends of viral DNA unjoined to target DNA. (C) Repair of the recombination intermediate yields the integrated provirus, which is flanked by a duplication (bracket) of the sequence of the double-stranded cut in B. (D) Host cell-mediated circularization. Homologous DNA recombination or NHEJ can lead to the formation of 1-LTR or 2-LTR circles, respectively. (E) IN-dependent autointegration. Depending on the strand of viral DNA that is used during autointegration, either a two LTR-containing circle with an inverted internal segment or 2 smaller circles each containing one LTR are formed (39).

could help clarify the function(s) of these proteins in the context of HIV-1 infection. Although similar reviews have been written (14-16), a number of recent findings including the results of the first cell-based genetic assay to investigate the roles of IN catalytic cofactors (17) and the identification of novel interacting proteins for HIV-1 IN (18, 19) warrant this timely update.

3. INTEGRATION MECHANISM

3.1. Reactions

Linear retroviral cDNA is the substrate for IN-mediated DNA recombination (20-23). Integration proceeds in three steps: 3' processing, DNA strand transfer, and gap repair (figure 1). 3' Processing and DNA strand transfer are distinct endonucleolytic reactions catalyzed by the IN. In contrast, gap repair is likely catalyzed by host cell enzymes (24, 25; reviewed in references 1 and 14).

IN performs its first endonucleolytic reaction soon after reverse transcription is completed. IN recognizes and acts on short sequences (12 to 20 bp) called attachment (*att*) sites that are located at the LTR termini (figure 1) (reviewed in reference 26). For HIV-1, a dinucleotide is hydrolyzed from each *att* site during 3' processing (figure 1A) (27-30). The resulting 3' hydroxyl groups are the nucleophiles that IN then uses to promote DNA strand transfer (29). The viral cDNA ends are used to cut the target DNA in a staggered fashion, which covalently links the viral 3' ends to the 5' phosphates of the cut (figure 1B). The resulting product is a gapped recombination intermediate with the 5' ends of the virus remaining unattached to the chromosome (21-23). DNA repair is then needed to complete the integration process (figure 1C). Since HIV-1 integration yields a 5 bp duplication flanking the integrated provirus (31, 32), HIV-1 IN affects a 5 bp staggered cut in chromosomal DNA during DNA strand transfer.

In addition to the productive integration pathway, some unintegrated cDNA travels down non-productive pathways that preclude virus infection. Covalently closed DNA circles containing either one or two copies of the LTR, referred to as 1-LTR and 2-LTR circles, respectively, are found in the nuclei of infected cells (20; reviewed in reference 33). Whereas 1-LTR circles are formed in part by the Rad50/Mre11/NBS1 complex implicated in cellular homologous DNA recombination (34), 2-LTR circles are formed by the non-homologous end joining (NHEJ) pathway of DNA recombination (35) (figure 1D). In addition to these host enzyme-mediated products, IN can generate non-productive DNA circles by a process called autointegration whereby retroviral cDNA prematurely integrates into itself instead of a host chromosome (figure 1E). Approximately 20% of Moloney murine leukemia virus (MoMLV) DNA circles were identified as products of autointegration (36), indicating that autointegration may occur at a relatively high frequency in MoMLV-infected cells. Preintegration complexes (PICs) are defined as high molecular weight nucleoprotein complexes isolated from cells that can catalyze the integration of endogenous viral cDNA into an exogenous recombinant target DNA *in vitro* (20-23, 37, 38). Although MoMLV PICs recovered from cells in isotonic salt-containing lysis buffer catalyzed normal intermolecular integration activity (21, 22, 39), autointegration was induced by treating the PICs with relatively high concentrations of salt and then purifying the complexes on the basis of size (39). This suggested that the high salt treatment removed a protein(s) that normally protected MoMLV from autointegration, and a subsequent study identified the barrier-to-autointegration factor (BAF) as a mouse cell protein that restored normal integration when it was added back to salt-treated MoMLV PICs (40). Autointegrated HIV-1 DNA can be identified in cell extracts (41) and PICs derived from avian cells infected with Rous sarcoma virus supported robust autointegration activity (23, 42), suggesting that autointegration is a normal part of retroviral infection and that all retroviruses have likely evolved a mechanism(s) to suppress this non-productive suicidal infection pathway.

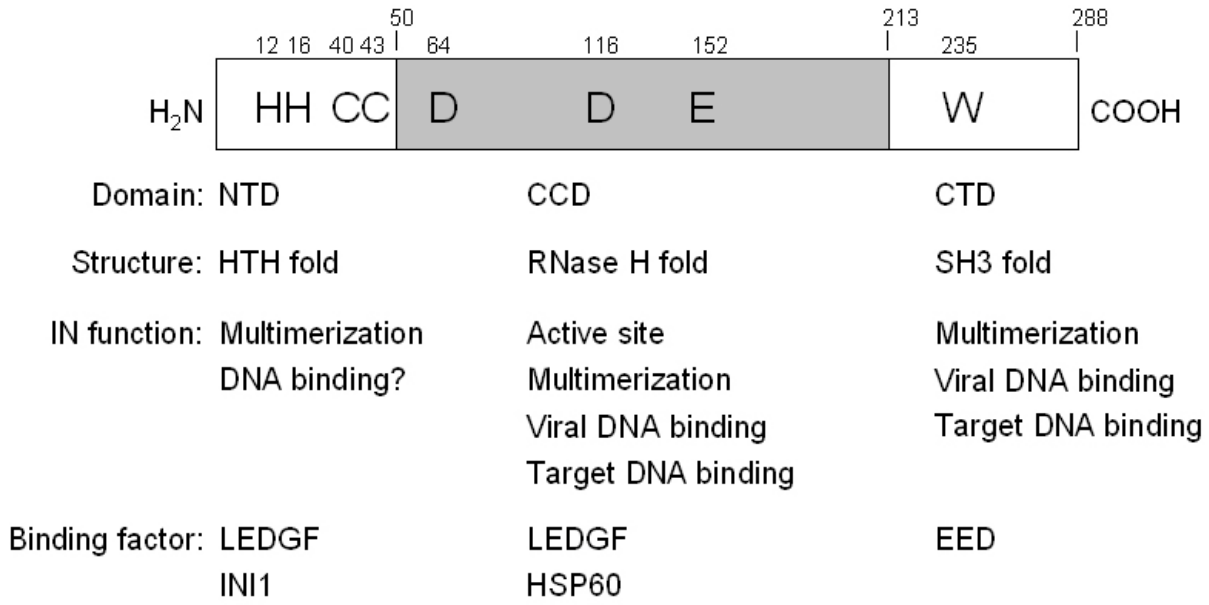

Figure 2. Structure/function diagram of HIV-1 IN. The three protein domains are indicated along with their structural folds, roles in IN function, and cellular binding partners. The positions of conserved His and Cys residues in the NTD and Asp and Glu residues in the CCD domain are indicated. Trp-235 in the CTD is conserved among retroviral IN proteins. HTH, helix-turn-helix.

3.2. IN and PICs

Results of numerous structural and functional analyses revealed that retroviral INs are comprised of three protein domains: the N-terminal domain (NTD), catalytic core domain (CCD), and C-terminal domain (CTD) (figure 2) (reviewed in references 43, 44). The NTD (HIV-1 residues 1 to 49) contains two His and two Cys residues that are conserved amongst retroviral and retrotransposon INs (45, 46). These residues coordinate zinc (47, 48) and contribute to functional IN multimerization (49, 50). The CCD (residues 50 to 212) harbors a triad of invariant residues (Asp-64, Asp-116, and Glu-152 for HIV-1) that comprise the enzyme active site (51-56). Other conserved CCD residues make important contacts with viral (57-61) and target (54, 61-63) DNA during integration. The less conserved CTD (HIV-1 residues 213 to 288) forms an SH3-like fold (64, 65) that also contributes to IN multimerization (66, 67) and makes sequence-specific (68) and non-specific (67, 69-71) contacts with DNA (figure 2).

Integration in cells occurs in the context of PICs (20-23, 37, 38) that are comprised of viral cDNA and viral (72-76) and cellular (35, 77-80) proteins. The viral cDNA is highly condensed (76) and results of micrococcal nuclease (72, 76), DNase I (81), and bacteriophage Mu-mediated PCR (MM-PCR) footprinting (82, 83) indicate that LTR termini are more tightly bound by proteins than are the internal regions of pre-integrative retroviral cDNA. An MM-PCR footprinting interference analysis defined the intasome as the functional nucleoprotein complex within PICs that catalyzes cDNA integration (82). Mutational analyses indicated that two functional cDNA ends (84-86) as well as multiple IN functions (83) are required for intasome formation in cells.

Purified retroviral IN is sufficient to catalyze the 3' processing and DNA strand transfer steps of integration in the absence of any other protein component (28-30, 87-91). Early reaction conditions with double-stranded oligonucleotide substrates revealed that in contrast to the concerted integration of both viral DNA ends (figure 1B) that generates the 5 bp sequence duplication in HIV-infected cells, the majority of DNA strand transfer products resulted from the integration of a single viral DNA end into just one strand of target DNA (89). Since HIV-1 PICs preferentially catalyzed concerted integration *in vitro* (38, 86), a potential role for PIC-associated host factors could be to increase the fidelity of HIV-1 IN's concerted integration activity. However more recent results have demonstrated that purified HIV-1 IN protein is not as inept at catalyzing concerted integration as once thought. Using optimized protein purification and integration reaction conditions, up to 35% of DNA strand transfer products resulted from concerted integration (91). Based on this, it is unclear if HIV-1 would require a host cell factor(s) to enable the high rate of concerted integration that is observed with PICs isolated from cells.

3.3. Target site selection

In the absence of any other protein, purified HIV-1 IN displays a slight preference for integration adjacent to purine residues (92). In cells, HIV-1 shows somewhat more target site specificity, as weak 5 bp GTA(A/T)(T/C) (93) and 5' GT(A/T)AC (94) integration consensus sites can be observed. More recent studies that utilized human genome-wide mapping approaches demonstrated that HIV-1 shows an approximate twofold preference for integration into genes (95, 96). Although MoMLV did not show a similar strong overall preference for genes, in this case the 5' ends of genes were preferentially targeted over downstream regions (96). These results can be compared and contrasted to the lifestyles of certain yeast retrotransposons. Retrotransposons are

similar to retroviruses in that reverse transcription generates a cDNA substrate for IN-catalyzed integration, and subsequent proviral gene expression generates polyproteins that coalesce into viruslike particles. The major distinction between retrotransposons and retroviruses is that the yeast elements lack an extracellular viral phase.

Yeast retrotransposons Ty1 and Ty3 integrate into specific regions upstream of actively transcribed RNA polymerase III genes (reviewed in references 14 and 97). The specificity for Ty3, which is quite exquisite as it occurs within one to two nucleotides of tRNA transcription start sites, is mediated through interactions with the TFIIIB and TFIIIC components of the yeast transcription machinery (reviewed in reference 98). Site-specific integration would seem to help maintain a gentle balance between intracellular retrotransposition and potentially lethal insertional mutagenesis in yeast. Retroviruses on the other hand need to spread from cell to cell and successfully multiply their overall systemic numbers. It therefore seems plausible that the unique patterns of HIV-1 and MoMLV integration observed in cells impart subtle yet important growth advantages in the face of active immune responses (99). Regardless of the precise reason, it seems reasonable to propose that HIV-1 and MoMLV interact with specific intracellular component(s) of the host to achieve their unique patterns of *in vivo* integration. Characterizing the mechanism of retroviral integration site targeting is an ongoing area of active research.

4. HOST PROTEINS IMPLICATED IN INTEGRATION

Cellular proteins that may function with HIV-1 to accomplish integration in cells can be grouped into two broad classes: those that act during IN catalysis (figure 1A and 1B), and those that act during the subsequent repair steps of DNA recombination (figure 1C). Each of these classes can be further subdivided into proteins that either do or do not directly interact with the IN protein. We will first discuss IN catalytic cofactors and then address those proteins proposed to participate in gap repair.

4.1. Potential IN catalytic cofactors

Cellular proteins that potentially aid the IN in accomplishing its catalytic activities in infected cells can be grouped into two classes: those that function via their innate ability to bind the cDNA substrate, versus those that additionally bind directly to HIV-1 IN. Whereas the high mobility group (HMG) protein A1 (HMGA1) (77, 100) and BAF (40, 101-103) bind DNA but do not appear to directly interact with IN, IN interactor (INI) 1 (104), lens epithelium-derived growth factor (LEDGF) (18), embryonic ectoderm development (EED) protein (19), and heat shock protein (HSP) 60 (105) have each been shown to directly interact with HIV-1 IN.

4.1.1. Factors that act through DNA binding

4.1.1.1. HMGA1

PICs isolated from HIV-1-infected cells can integrate their endogenous viral cDNA into an added target DNA *in vitro* (37, 38, 74-77, 80, 81, 83, 85, 86). Functional integration however can be disrupted by first treating HIV-1 PICs with relatively high concentrations of salt and then purifying the treated complexes by size (77, 106). Unlike MoMLV (39, 40), salt-treated HIV-1 PICs do not support detectable levels of autointegration activity and are instead catalytically inactive (77, 106). These observations suggested that high salt treatment either removed an important catalytic cofactor from HIV-1 PICs or more simply altered the IN such that it could no longer function. Since extracts of uninfected host cells restored integration activity to salt-treated HIV-1 PICs, it would seem that the salt removed a cofactor rather than inactivating the IN (77, 106). Additionally, since lysates of HIV-1 virions failed to restore PIC activity, the missing factor(s) was of cellular as compared to viral origin (77). HMGA1 was the first protein identified that reconstituted the integration activity of salt-treated HIV-1 PICs (77). Since HMGA1 also cofractionated with HIV-1 PICs *in vitro*, it seemed that HMGA1 was a component of PICs that supplied an important cofactor function for HIV-1 IN in the context of virus infection (77). The highly related HMGA2 protein similarly reconstituted the integration activity of salt-depleted HIV-1 PICs *in vitro* (107).

HMGA proteins are low molecular weight non-histone chromatin-associated proteins that preferentially bind to AT-rich regions located in the minor groove of DNA (108). The quasi specific nature of this DNA binding is mediated by highly conserved 11-residue domains known as AT-hooks that display affinity for a variety of altered DNA structures such as four-way junctions, supercoiled DNA, and nucleosomes (reviewed in reference 109). AT-hooks, like the intact HMGA1 protein, exhibit very little secondary structure but adopt a crescent shape that fits into the minor groove when DNA is bound (110). HMGA1 contains three AT-hooks, and these multiple DNA-binding motifs enable HMGA1 to induce local conformational changes in DNA such as bending, looping, unwinding or straightening out of bent regions. HMGA1 is suggested to participate in transcriptional regulation of genes whose promoters contain or are close to AT-rich regions. By binding to DNA and other proteins, HMGA1 is able to promote the formation of enhanceosomes, which in turn activate transcription (111).

In addition to its identification as a component of HIV-1 (77) and MoMLV (107) PICs, purified HMGA1 protein stimulated the DNA strand transfer activity of purified IN protein in *in vitro* integration assays (100, 112). Since HMGA1 does not physically interact with IN (77, 100), it would appear to act through DNA binding and it has been suggested that HMGA1 might bridge the two ends of viral cDNA together within the intasome. Potential binding sites for HMGA1 have been identified in the HIV-1 LTR (77, 113). A related end-bridging function for a host cell protein in integration likely functions in *E. coli*, where the bacterial Hu protein plays an important architectural role in transpososome formation and transposition (114). Alternatively, HMGA1 could help PIC-associated cDNA to fold by binding to multiple internal AT-rich regions, yielding an overall nucleoprotein structure proficient for integration (111, 112). A recent herculean effort in chicken

DT40 cell gene disruption yielded cell lines ablated for HMGA1, HMGA2, or both genes (17). These cell lines were tested for their ability to support retroviral integration using pseudotyped HIV-1 particles capable of single infectious cycles as well as replication-competent avian viruses. Somewhat surprisingly, integration of both types of virus appeared to proceed at normal rates with all derivative cell lines, indicating that neither HMGA1 nor HMGA2 plays an obligate role in retroviral cDNA integration under these conditions of virus infection (17).

4.1.1.2. BAF

BAF was identified through its ability to both suppress the autointegration activity of salt-depleted MoMLV PICs and restore their normal intermolecular integration activity (40). Unlike MoMLV, salt-treated HIV-1 PICs are catalytically inactive (77, 106). Purified BAF protein nonetheless restored intermolecular integration activity to salt-inactivated HIV-1 PICs (106). This observation enabled side-by-side comparisons of BAF and HMGA1 in functional PIC reconstitution assays.

BAF was significantly more active (about 500 fold) than HMGA1 in both HIV-1 (77, 106) and MoMLV (78, 79, 84) PIC reconstitution assays. Other results from *in vitro* reconstitution assays also highlighted a role for BAF in PIC structure and function. BAF restored the native intasome structures of salt-treated HIV-1 and MoMLV PICs under conditions wherein HMGA1 failed to function (84, 106). Salt-stripping partially impeded the migration of MoMLV PICs in sucrose gradients, and only BAF restored the normal pattern of MoMLV PIC sedimentation in sucrose (79). Like HMGA1 (77, 107), BAF is a component of MoMLV (79) and HIV-1 (80) PICs. However unlike HMGA1 (107, 112), BAF failed to detectably stimulate the activity of purified HIV-1 IN in *in vitro* integration assays (112, 115).

BAF is an 89 amino acid DNA-binding protein that is highly conserved in multicellular organisms from *C. elegans* to humans, showing 97% identity between human and murine homologs (116). This underscores the importance of BAF for cells and indicates that interactions with other proteins may have prevented its evolutionary divergence. Interactions with both DNA and nuclear proteins have been highlighted. BAF forms higher-order nucleoprotein complexes with double-stranded DNA at a ratio of approximately five BAF dimers to five DNA molecules (102). BAF also interacts with nuclear proteins that harbour a conserved LEM (LAP2 [lamin-associated polypeptide 2], emerin, MAN1) domain, both when DNA-bound (117) or free from DNA (118, 119). These inner nuclear membrane (INM) proteins bridge the INM to the nuclear lamina, providing a physical framework to the nucleus (reviewed in reference 120).

Knockdown of BAF in *C. elegans* embryos by RNA interference (RNAi) led to early embryonic lethality (under the 100 cell-stage) with almost 100% penetrance. Major defects were observed during mitosis and included abnormal chromosome segregation as well as trailing of chromatin between daughter cells (102). Interestingly, these anaphase chromatin bridges also resulted from double-knockdown of emerin and MAN1 in *C. elegans*, which could reflect the inability to exit mitosis (121). These results suggested that chromatin/BAF/emerin/MAN1 complexes play a crucial, non-redundant role in post-mitotic nuclear reassembly and chromatin decondensation (122). BAF could provide a physical link between recondensing chromosomes and emerin (123). Furthermore, since BAF binds to proteins that are part of the nuclear architecture as well as to DNA, it was postulated that BAF plays a role in linking chromatin to the INM, thus contributing to the stability of the interphase nucleus (117). Besides its hypothesized and extensively studied function in nuclear architecture and dynamics, BAF was recently shown to interact with and inhibit the transactivation function of the paired-like homeodomain transcription factor cone-rod homeobox (Crx) (124). Inhibition could arise from Crx sequestration in a subcellular compartment such that it fails to gain access to its DNA target, or from inducing conformational changes in BAF-bound Crx thus preventing it from binding to DNA. BAF has also been implicated in human disease, as it is the HB autoantigen associated with rare cases of rheumatoid arthritis and systemic lupus erythematosus (125).

Solution and X-ray crystallographic structures of BAF reveal a globular protein composed of five alpha helices surrounding a hydrophobic core with four Cys residues involved in hydrogen bonding and core stabilization (116, 126). BAF is a homodimer in solution with dimerization occurring through evolutionarily-conserved hydrophobic residues and a single salt bridge. BAF possesses a helix-hairpin-helix (HhH) motif between residues 20 and 35 (figure 3) that is frequently found in proteins such as DNA polymerases, ligases, glycosylases and helicases that bind DNA non-specifically (127). This region contains two anti-parallel alpha helices connected by a conserved GhG (Gly-hydrophobic residue-Gly) motif that adopts a hairpin conformation and interacts with DNA. BAF-DNA interactions (101, 102) are likely mediated via hydrogen bonds between positively charged residues in the HhH domain and phosphate groups on the DNA backbone (126). DNA binding appears to increase the affinity of BAF for interacting with LEM domain proteins, as residues essential for DNA binding (Lys-6, Gly-25, Gly-27, and Leu-46) were required for wild-type levels of emerin binding as well (122). The postulated emerin-binding site on BAF, comprised of Val-51, Lys-53, and Lys-54, is distinct from the HhH DNA binding domain (figure 3) (122). Thus, although DNA and emerin bind to different surfaces of the BAF dimer, DNA binding appears to induce a conformational change such that the affinity of BAF for emerin increases over that of DNA-free BAF (122).

How might this nuclear architectural factor function during the early phase of retroviral replication? After binding to newly synthesized cDNA and thereby helping to mediate PIC formation, BAF might compact the DNA to limit its availability for autointegration as well as promote association with target DNA when the PIC encounters a chromosome (40, 79, 106). Recent results demonstrate that BAF is incorporated into HIV-1 virions at

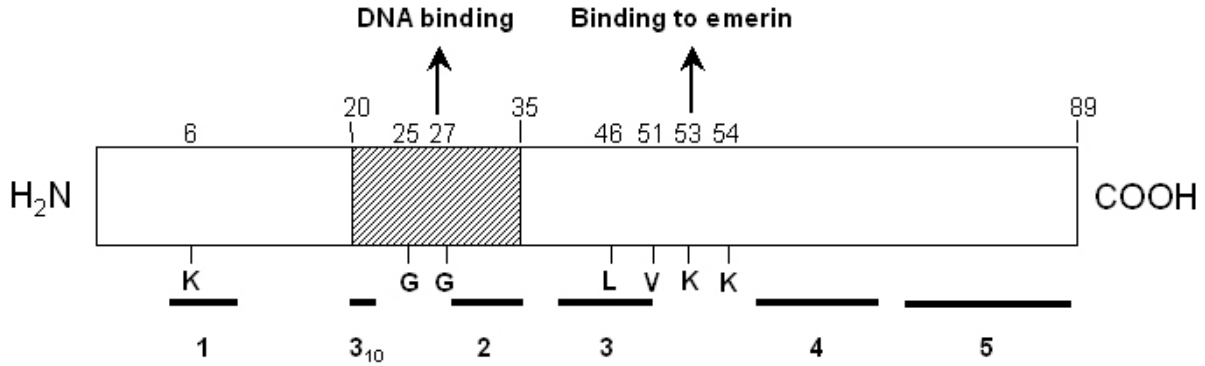

Figure 3. Structure/function features of human BAF. The HhH DNA-binding domain (residues 20 to 35) is indicated. Residues highlighted in text as important for emerin and DNA binding are also indicated. Also shown are the locations of the five alpha helices and one 3_{10} helix.

the fairly low stoichiometry of zero to three BAF dimers per virus (103). Although BAF failed to appreciably interact with HIV-1 IN, BAF bound the matrix (MA) domain of the Gag polyprotein precursor with an apparent affinity constant of 1.4 micromolar. BAF expression was unexpectedly low in purified resting T-cells until the cells were activated to divide (103). These findings suggest the intriguing model that BAF is incorporated into viral particles through its interaction with Gag/MA to ensure that sufficient quantities of the protein are available for PIC formation during the next round of infection. This scenario could be particularly important in resting human T-cells where the concentration of the factor appears exceptionally low. An important test of this model could include the construction of HIV-1 particles in cells that were depleted for the BAF protein through the use of RNAi, and subsequent infectivity measurements in resting versus activated primary T-cells.

4.1.2. IN binding proteins

Human proteins that directly interact with HIV-1 IN have been detected using yeast two-hybrid screens, coimmunoprecipitation via antibodies directed against recombinant IN protein, and/or *in vitro* interaction or "pull-down" assays. Of the resulting interacting factors, INI1, LEDGF/p75, EED, and human HSP60 (hHSP60) have been proposed to play catalytic cofactor roles during HIV-1 integration.

4.1.2.1. INI1

INI1, identified via yeast two-hybrid, was the first host protein discovered to interact with HIV-1 IN (104). The interaction, which was confirmed by glutathione S-transferase (GST) pull-down, mapped to the NTD of IN (figure 2). Recombinant INI1 protein stimulated the *in vitro* integration activity of purified IN ten- to twenty-fold, suggesting a catalytic cofactor role for INI1 in HIV-1 integration. INI1 is homologous to the yeast SNF5 protein (104), an essential component of the SWI/SNF chromatin remodeling complex. SWI/SNF complexes are ATP-dependent chromatin remodeling machines (reviewed in reference 128). Upon binding and subsequent ATP hydrolysis, free energy for interconversion between nucleosomal states is generated thus enabling a change of nucleosomal conformation. Interactions between DNA and histones are loosened during a transitional activated state that leads to transcriptional activation or repression of neighboring regions. INI1 has been reported to interact with other viral proteins such as the Epstein-Barr virus nuclear antigen 2 (129) and human papillomavirus E1 (130) as well as cellular factors c-MYC (131), ALL1 (132), and p53 (133). INI1 also acts as a tumor suppressor, as mutations in the *INI1* gene can lead to atypical teratoid and rhabdoid tumors, suggesting a role for the SWI/SNF complex in control of the cell cycle (134; see reference 135 for a complete review).

INI1 is normally found in the cell nucleus (136). However, shortly after HIV-1 entry INI1 is exported together with the promyelocytic (PML) protein to the cytoplasm of interphasic cells through the exportin-mediated pathway in response to a viral-induced signal (137). Relocalization occurred 30 min to 6 h post-infection during which time colocalization with incoming PICs was observed, which was then followed by INI1 and PML returning to the nucleus. Binding of INI1 to PICs via IN may mediate the subsequent recruitment of SWI/SNF components for what could be an effective way of targeting HIV-1 to SWI/SNF remodeled regions of the genome. In this vein it is noteworthy that SWI/SNF-mediated remodeling is more frequently associated with repression than activation events, at least in yeast (138). Extending this observation to humans, associated INI1 might preferentially target HIV-1 to repressed genes which is in contrast to the reported preferential integration into active gene regions (95). Alternatively, recruitment of PML to PICs could potentially promote the association of PML-interacting proteins such as histone acetyltransferase CBP/p300 or transcription factors (139) that could in turn trigger integration into transcribed regions of the genome.

INI1 contains three conserved regions including two direct imperfect repeats, repeat (Rpt) 1 and Rpt2, and a carboxyl-terminal coiled-coiled domain that is also referred to as homology region 3 (HR3) (figure 4) (140). INI1 harbors a DNA-binding region just upstream of Rpt1 (figure 4) (140). Overexpression of an internal S6 fragment of INI1, which corresponds to residues 183-294 and

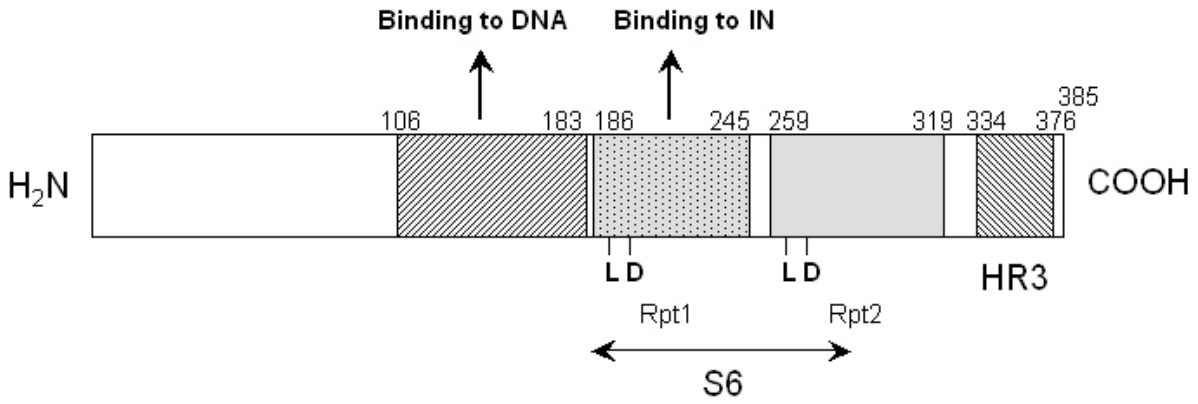

Figure 4. Structure/function features of INI1. The INI1 DNA binding domain (residues 106 to 183) is located immediately upstream of Rpt1. Imperfect repeats Rpt1 and Rpt2 are represented in grey areas, with the Rpt1 IN-binding region containing dots. The locations of HR3, the dominant-negative S6 fragment as well as highly-conserved Leu and Asp residues in Rpt1 and Rpt2 are also indicated.

contains IN-binding region Rpt1 (figure 4), dramatically reduced (10,000 to 100,000-fold) HIV-1 assembly and release from human 293T cells (136). IN is expressed in retroviral-infected cells as part of a multi-component Gag-Pol polyprotein precursor, and mature HIV-1 IN is liberated from Gag-Pol via the activity of the viral protease during or shortly after virus budding (reviewed in reference 141). Of note, particle inhibition required the interaction of the dominant-negative S6 fragment with IN in the context of Gag-Pol (136). Since INI1 was also identified as a component of HIV-1 virions (136), these results suggested that a physiologically-relevant interaction between INI1 and IN may occur during the late stage of HIV-1 replication when IN is still part of Gag-Pol as compared to the early phase of infection when IN is present as a mature PIC-associated polypeptide. Since INI1 was not packaged into the highly-related lentiviruses HIV-2 and simian immunodeficiency virus, the INI1-IN interaction is highly specific to HIV-1 (142). Consistent with this observation, over-expression of S6 in 293T cells shut-down HIV-1 assembly/release but not the egress of these related lentiviruses (142).

4.1.2.2. LEDGF/p75

LEDGF/p75 is a 530 residue DNA-binding protein implicated in cell differentiation and cellular response to environmental stress (reviewed in reference 143). Elevated levels of LEDGF/p75 were observed in mammalian cells following stress such as heat, exposure to toxic agents like ethanol, or oxidizing agents (144). Upon binding to stress-related elements or heat shock elements in DNA (145, 146), LEDGF/p75 activates the transcription of stress-related genes thereby triggering a survival response that protects cells from stress-induced damage (143).

LEDGF/p75 was identified as an IN-interacting protein via over-expressing FLAG-tagged HIV-1 IN in 293T cells and subsequent coimmunoprecipitation with anti-FLAG antibodies (18). When expressed in human cells on its own in the absence of other viral proteins, HIV-1 IN localizes to nuclei and is therefore karyophilic (18, 147-154). IN and LEDGF/p75 colocalized in cell nuclei during interphase and were both tightly bound to chromosomes (18, 154). Selective depletion of LEDGF/p75 via small-interfering RNA (siRNA) abolished IN nuclear localization, leading instead to a diffuse cytoplasmic pattern (154). The IN-LEDGF/p75 interaction was confirmed *in vitro* via a His_6 Tag pull-down assay using His_6-tagged IN and purified recombinant LEDGF/p75 (154). Purified LEDGF/p75 additionally stimulated the *in vitro* integration activity of recombinant HIV-1 IN protein (18). Taken together, these results suggest a role for LEDGF/p75 in HIV-1 nuclear localization, IN catalysis, and perhaps targeting PICs to particular chromosomal regions for integration. Defining the precise role(s) of LEDGF/p75 in integration and HIV-1 replication is an ongoing area of research. Many investigators are likely testing HIV-1 infectivity and integration in cells knocked-down for LEDGF/p75 protein using either the described siRNA (154) or an effective short-hairpin (sh) RNA-expressing retroviral vector (155).

LEDGF/p75 is evolutionarily conserved from *Xenopus* to humans, and several functional regions can be identified (figure 5). A functional nuclear localization signal (NLS) has been proposed to reside between residues 149 and 156 (143). Four potential DNA-binding regions, three helix-turn-helix (HTH) motifs and a basic leucine zipper, can also be identified (figure 5) (143). The N-terminal region of LEDGF/p75 harbors an approximate 80 residue PWWP domain, named for a conserved Pro-Trp-Trp-Pro motif (which is Pro-His-Trp-Pro in LEDGF/p75) found in a variety of chromatin-associated proteins (reviewed in reference 156). PWWP has been proposed to function as a DNA-binding (157) and/or protein-protein interaction domain (156). Delineating the precise role of the PWWP domain in the normal biology of LEDGF/p75 requires further experimentation.

The *LEDGF* gene also expresses a smaller splice variant that encodes for the 333 residue p52 protein (158). Interestingly, p52 failed to interact with HIV-1 IN in the *in vitro* pull-down assay and in live cells (154). Since the N-terminal 325 residues of LEDGF/p75 and p52 are identical,

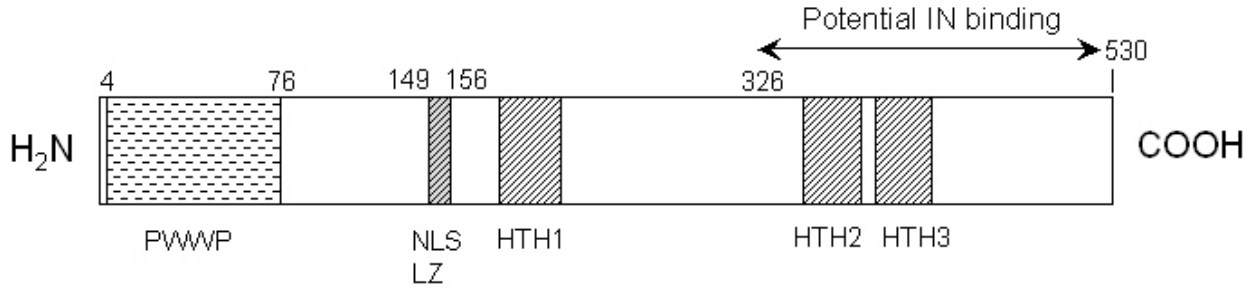

Figure 5. Structure/function features of LEDGF/p75. The N-terminal PWWP domain (approximate boundaries 4-76) is found in a variety of chromatin-associated proteins. The approximate locations of a potential NLS and a basic leucine zipper (LZ) and HTH DNA-binding domains are indicated. The potential IN-binding region in the C-terminal region of LEDGF/p75 is also indicated.

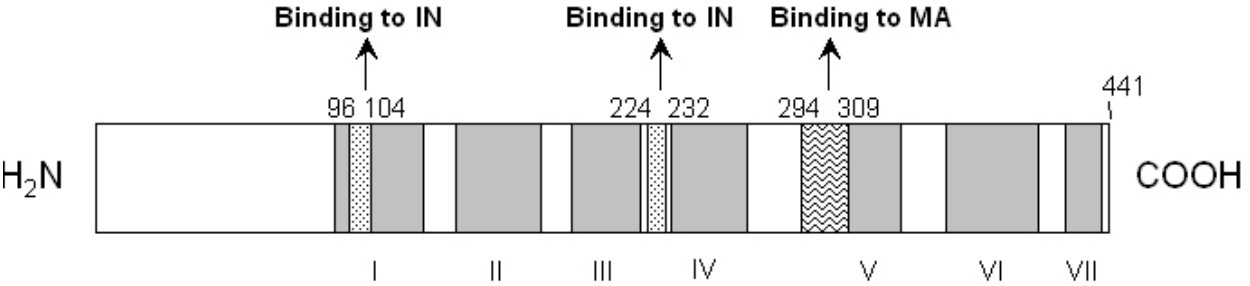

Figure 6. Structure/function features of human EED. The seven WD repeats are indicated by grey fill. The two putative IN binding sites are indicated by dots. The binding site to HIV-1 MA is indicated by wavy fill-in.

these results implicate the unique 205 residue C-terminal region of LEDGF/p75 as an IN binding region (figure 5). In terms of IN, analyses of fluorescent-tagged deletion mutants identified the CCD as the major interacting domain, with no appreciable involvement of the CTD (figure 2) (154). When expressed on their own, IN's NTD and CTD fragments displayed even cellular distribution and irregular nuclear accumulation, respectively, and did not co-localize with LEDGF/p75. Despite this finding, the NTD H12N point mutant protein, which displayed reduced affinity for zinc in an *in vitro* binding assay (159), failed to interact with LEDGF/p75 *in vitro* and displayed diffuse fluorescence in cells (154). Based on this it appears that an intact NTD likely stabilizes the overall IN-LEDGF/p75 interaction. Since mutations in the NTD zinc-binding residues of feline immunodeficiency virus IN also impaired the karyophilic properties of that protein (160), it is tempting to speculate that LEDGF/p75 may be a cellular binding partner for a variety of retroviral INs.

4.1.2.3. EED

EED (161) is a nuclear protein belonging to the WD-repeat superfamily (162) as well as the *Polycomb* group (PcG) of proteins (163). As for many PcG proteins (164), EED is part of a protein complex involved in transcriptional repression and gene silencing promoted by histone deacetylation (165). The most readily discernible structural features of EED are seven WD repeats (figure 6). WD repeats are conserved repeated units typically ending with Trp-Asp residues (162).

Prior to its discovery as an IN-binding protein, EED was identified as a cellular interactor for the HIV-1 MA protein (166). Bacteriophage panning, GST pull-down, and yeast two-hybrid assays then demonstrated that EED and IN interacted *in vitro* and in yeast (19). The EED-IN interaction mapped to the CTD of IN (figure 2) and to two distinct regions in the N-terminal half of EED (figure 6) (19). The EED C-terminal region also appears to contribute to the EED-IN interaction, perhaps by stabilizing the complex and/or effecting an overall conformational change that maximizes binding (19).

Purified EED protein stimulated the *in vitro* integration activity of recombinant HIV-1 IN between two- and thirty-fold (19). Interestingly, EED and IN colocalized within HIV-infected cells. The two proteins were observed together in cell nuclei and near nuclear pores with maximum colocalization occurring at 6 h post-infection (19). Triple colocalization of IN, EED, and MA was also observed in the nucleoplasm at 6 h post-infection, demonstrating the occurrence of these three proteins in what would seem to be the same multi-protein complex under physiologically-relevant conditions (19). EED also interacted with a histone deacetylase in the presence of transcription factor YY1 and the enhancer of Zeste 2 protein (167). These interactions suggest that YY1 might bridge EED and DNA, leading to silencing of targeted chromatin regions. It was proposed that recruitment of YY1-bound EED could abolish gene silencing, resulting in transcriptionally active regions (19) that could serve as preferential targets for HIV-1 integration (95). Confirmation of a role for EED in targeted HIV-1 integration requires further experimentation. EED was more recently shown to interact with the HIV-1 Nef protein (168). More experiments are required to address which of

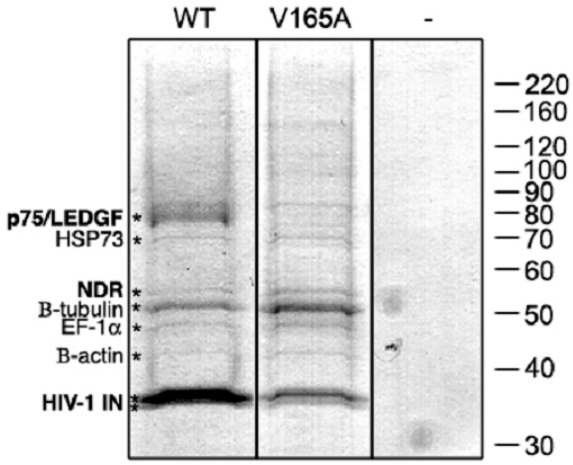

Figure 7. Identification and characterization of IN-interacting proteins. Lane 1, proteins recovered from HeLa cells transduced with the wild-type (WT) IN expression vector; lane 2, cells transduced with the V165A IN expression vector; lane 3, empty vector control. The identity of the eluted proteins is indicated to the left of the gel. Migration positions of molecular mass standards are shown to the right.

EED's interactions, those with IN, MA, or Nef, is most relevant in the context of HIV-1 infection.

4.1.2.4. HSP60

S. cerevisiae heat shock protein 60 (yHSP60) was identified as an interactor of HIV-1 IN via passing yeast cell extracts over an IN-affinity column (105). Three other yeast proteins, the neoglycogenesis PCK1 enzyme, the tubulin chaperone CCT4, and translation elongation factor (EF)-1-alpha were also identified as IN-binding proteins in this screen. HIV-1 IN also interacted with hHSP60, and the interaction was mapped to the IN CCD (figure 2) (105). Purified recombinant hHSP60 protein stimulated the *in vitro* 3' processing and DNA strand transfer activities of purified IN protein two- to four-fold (105). As a chaperone, HSP60 forms a complex with the co-chaperonine HSP10, which is able to bind misfolded proteins and refold them in an ATP-dependent manner (reviewed in reference 169). Since purified HIV-1 IN was a substrate for the HSP60-HSP10 complex, it was postulated that HSP60-HSP10 could activate and/or stabilize IN in HIV-infected cells.

HSP60 also binds hepatitis B virus DNA polymerase (170), and other chaperones were previously shown to associate with vesicular stomatitis virus (171), vaccinia virus (172), influenza virus (173), and adenovirus (174). Considering that the level of hHSP60 is elevated in virus-infected cells (175), HSP60 and, more generally, protein chaperones may represent a general cellular stress response to viral infection. Further experimentation is required to establish a physiologically-relevant role for hHSP60 during the early stages of HIV-1 infection.

4.1.2.5. A survey of human cell proteins that coimmunoprecipitate with HIV-1 IN

Mutations in IN can cause pleiotropic effects on the HIV-1 life cycle, which in addition to integration can impair virus assembly/release and/or reverse transcription (reviewed in reference 176). This observation suggests that IN may interact with a variety of host cell proteins during HIV-1 infection. To establish a comprehensive picture of human cell proteins that bind HIV-1 IN, we conducted the following large-scale coimmunoprecipitation experiment. Codon-optimized FLAG-tagged HIV-1 IN was expressed from the murine retroviral MSCVpuro vector (Clontech) as previously described (153). As controls we utilized the empty vector as well as one engineered to express the V165A mutant IN protein. Viruses containing the V165A change are defective in reverse transcription (R. Lu & A. Engelman, unpublished observations), integration (152), and possibly at a post-nuclear entry step (177). The V165A change was incorporated into the IN reading frame as previously described (153).

Suspension-adapted HeLa cells were transduced with the viral vectors, and extracts from approximately 6 L of cells were prepared by lysis in buffer containing 50 mM Tris pH 8.0, 500 mM KCl, 0.1% Triton X-100, 2.5% glycerol, and Complete Protease Inhibitor cocktail (Roche). IN-associated proteins were immunoprecipitated with anti-FLAG M2 affinity matrix (Sigma) and eluted with the FLAG peptide (Sigma). Eluted proteins were resolved by 4-20% Tris-Glycine SDS-PAGE (Novex) and stained with Simply BlueSafeStain (Invitrogen) (figure 7). Gel-resolved proteins were digested with trypsin, the mixtures fractionated on a Poros 50 R2 RP micro-tip, and the resulting peptide pools were analyzed by matrix-assisted laser-desorption / ionization reflectron time-of-flight (MALDI-reTOF) mass spectroscopy (MS) using a BRUKER UltraFlex TOF/TOF instrument (Bruker Daltonics) as previously described (178, 179). Selected experimental masses were taken to search a non-redundant protein database ('NR'; ~1.4 x 10^6 entries; National Center for Biotechnology Information) utilizing the PeptideSearch (Matthias Mann, Southern Denmark University) algorithm, with a mass accuracy restriction better than 40 ppm and maximum one missed cleavage site allowed per peptide. Mass spectrometric sequencing of selected peptides was done by MALDI-TOF/TOF (MS/MS) analysis on the same prepared samples using the UltraFlex instrument in 'LIFT' mode. Fragment ion spectra were taken to search NR using the MASCOT MS/MS Ion Search program (Matrix Science Ltd.). Identifications thus obtained were verified by comparing the computer generated fragment ion series of the predicted tryptic peptide with the experimental MS/MS data.

With one exception, every protein coimmunoprecipitating with the wild-type IN also purified with the V165A mutant (figure 7). Given their relative intracellular abundance, the significance of the beta-actin, beta-tubulin, EF-1-alpha, and HSP73 interactions is unknown. Indeed, several of these proteins were previously shown to interact with other HIV-1 proteins. Beta-actin interacts with HIV-1 nucleocapsid (180, 181), Nef (182), and reverse transcriptase (183). EF-1-alpha interacts with the Gag polyprotein (184) and as discussed in the previous section the yeast homolog of EF-1-alpha was previously identified as an IN-binding protein (105). Furthermore,

both EF-1-alpha and beta-actin are incorporated into HIV-1 virions (185, 186). With regard to beta-tubulin, a peptide-based yeast 2-hybrid screen identified microtubule-associated proteins as interactors of HIV-1 IN (187). Given the relative strength of the interaction between beta-tubulin and defective V165A IN (figure 7), the physiological relevance of interactions between HIV-1 IN and tubulin/microtubules is unclear and requires further experimentation.

The most abundant human protein coimmunoprecipitating with wild-type IN was LEDGF/p75 (figure 7). These results confirm those reported by Cherepanov *et al.* (18). Interestingly, we did not detect any peptides corresponding to LEDGF/p75 co-eluting with V165A IN, suggesting that alteration of Val-165 significantly reduced the affinity of LEDGF/p75 for IN. Directed reciprocal anti-LEDGF/p75 and anti-FLAG immunoprecipitation/Western blotting experiments confirmed the IN-LEDGF/p75 interaction and reinforced that V165A IN was defective for LEDGF/p75 binding (data not shown). Although Maertens *et al.* (154) previously reported that the H12N change in the NTD significantly reduced the binding of LEDGF/p75 to IN, ours is the first description of a CCD change that appears to significantly reduce the affinity of LEDGF/p75 for IN. A relatively low level of the serine-threonine kinase nuclear Dbf2-related (NDR) (see reference 188 for review) was also identified in this screen (figure 7). Although NDR appeared to specifically interact with both wild-type and V165A IN, subsequent experiments revealed low levels of NDR binding to the anti-FLAG M2 affinity matrix in the absence of IN expression (data not shown). Because of this, it is currently unclear if NDR is a specific interactor of HIV-1 IN.

4.2. Factors implicated in gap repair

The final step of the integration process is gap repair, wherein the DNA recombination intermediate formed by IN's DNA strand transfer activity is converted into the final integrated product (figure 1C). One can envision three different enzymatic activities required for this type of DNA repair: (i) a DNA polymerase to fill-in the single-strand gaps (which are 5 bases for HIV-1), (ii) a flap endonuclease to remove the non-templated dinucleotides from the 5' ends of the virus, and (iii) a DNA ligase to seal the single-strand nicks (figure 1C) (see reference 14 for further details). Experiments with oligonucleotide substrates that mimicked one side of the gapped HIV-1 structure demonstrated that a variety of DNA polymerases including polymerase beta, delta, or HIV-1 reverse transcriptase could fulfill the requirements of step (i), that flap endonuclease I effectively removed the non-templated 5' dinucleotide, and that different DNA ligases including ligase I, III, or IV could seal the nick (25, 189). Since several different polymerases and ligases functioned *in vitro*, it seems likely that different sets of repair enzymes operate in cells.

4.2.1. DNA repair enzymes and gap repair

Although the results of these *in vitro* assays identified proteins that fulfilled the roles of DNA repair, the identities of the proteins that carry out the repair process in HIV-infected cells remain unknown. However, cell-based genetic assays have implicated a variety of human proteins that may work via sensing the single-stranded breaks in the recombination intermediate as playing a role(s) in gap repair. Another potential gap repair protein, hRad18, interacts directly with HIV-1 IN (190).

4.2.1.1. DNA-PK and NHEJ

NHEJ is the main mechanism of double-stranded DNA break repair in mammalian cells. One of its key components is the DNA-dependent protein kinase (DNA-PK). DNA-PK is a multi-protein complex comprised of a large DNA-PK_{CS} catalytic subunit and the Ku70/Ku86 heterodimer that recognizes and binds double-stranded DNA ends. The other major proteins in NHEJ are XRCC4, DNA ligase IV, and Artemis (reviewed in reference 191).

A mutation in DNA-PK_{CS} that effectively disrupts NHEJ is responsible for the severe combined immunodeficiency (*scid*) phenotype in mice. *Scid* mouse pre-B cells underwent massive apoptosis approximately one day after infection with avian sarcoma virus, MoMLV, or HIV-1 viral vectors (24). Similar results were observed in cells deficient for Ku86 or XRCC4 (24). Since apoptosis required functional IN protein, Daniel *et al.* (24) concluded that cells lacking DNA-PK underwent apoptosis due to the inability to repair the integration intermediate, thus invoking a role for DNA-PK in the integration process. Considering that the Ku heterodimer was subsequently identified as a component of HIV-1 and MoMLV PICs (35), this would represent an attractive model. DNA-PK might play an active role in repair by recruiting additional cell factors, or a more passive role by binding the gapped structure and protecting it from degradation. However most investigators have not observed a strict correlation between retroviral infection and apoptosis of NHEJ-deficient cells. Notably, the pre-B *scid* cell lines employed by Daniel *et al.* were derived via transformation with Abelson murine leukemia virus (Ab-MLV), and Ab-MLV infection yielded similar numbers of transformants from *scid* and normal mouse bone marrow (192). Although a separate study confirmed *scid* mouse embryo fibroblasts (MEFs) and Ku-deficient cells underwent apoptosis as a result of infection, the requirement for a relatively high (>1) multiplicity of infection (moi) prompted the model that DNA-PK played a protective role against cellular toxicity induced by high moi as compared to a direct role in repairing the integration intermediate (193). A more recent study highlighted a limited connection between retroviral infection and apoptosis of NHEJ-defective cells (34). A survey of seven different cell lines deficient for DNA-PK, Ku80, or DNA ligase IV revealed that only Nalm-6 ligase IV mutant cells underwent apoptosis as a result of high moi infection and that these surviving cells contained proviruses at a level that was indistinguishable from control infections, indicating that a signal other than integration induced toxicity under these conditions. Consistent with this interpretation, Li *et al.* (35) previously noted that a catalytically-inactive IN mutant virus induced efficient apoptosis of Nalm-6 cells, leading to the model that pre-integrated linear viral cDNA supplied the apoptotic signal.

Due to the plethora of conflicting data on a connection between infection, integration, and apoptosis of NHEJ-deficient cells, it remains unclear if DNA-PK plays an active role in repairing the integration intermediate or if DNA-PK plays a protective role for the cell during the early stages of retroviral replication.

Daniel *et al.* (24) highlighted that residual integration occurred in NHEJ-deficient cells, raising the notion of an alternative post-integration repair pathway. DNA-PK_{CS} is a phosphatidylinositol-3 kinase related kinase, and other members of this family include the ataxia-telangiectasia mutated (ATM) and ATM- and Rad3-related (ATR) kinases (191). Wortmannin, a sterol-like fungal metabolite, sensitized normal mouse pre-B cells to apoptosis following high moi infection (194). ATM-deficient cells were particularly sensitive to the effects of wortmannin, leading to the suggestion that ATM activity was required for the residual levels of retroviral transduction observed in NHEJ-deficient cells (194). A follow-up study revealed that caffeine, which is an efficient inhibitor of DNA repair that likely acts through the ATM and ATR kinases, inhibited retroviral infection at sub-toxic concentrations (195). Since caffeine-treated cells supported near-normal levels of reverse transcription and 2-LTR circle formation but significantly less provirus formation, caffeine appeared to specifically inhibit integration. Although caffeine reduced infection of ATM-knockout cells, this effect was similar to reductions observed with ATM-expressing cells. In contrast, stable integration was significantly reduced in cells expressing a dominant-negative form of the ATR kinase. Based on this, Daniel *et al.* (195) proposed an important role for ATR kinase activity in integration gap repair. A potential limitation of these lines of experimentation is unanticipated pleiotropic effects of drugs like wortmannin and caffeine. Over-expressing dominant-negative mutants might also result in the perturbation of off-target cellular proteins/pathways (195).

4.2.1.2. PARP-1

Poly(ADP-ribose) polymerase 1 (PARP-1) is another protein-modifying enzyme implicated in DNA repair (see reference 196 for review). After binding to a DNA break, PARP-1 uses NAD+ to catalyze the attachment of poly(ADP-ribose) to a variety of chromatin-associated proteins, including itself. It then dissociates from the DNA, allowing access by other repair enzymes. Experimental approaches to the role of PARP-1 in retroviral integration have included the use of specific inhibitors of enzyme activity, antisense oligonucleotides that reduced enzyme concentration, and knockout mice ablated for both alleles of the *PARP-1* gene.

An initial study using three different competitive inhibitors including 3-methoxybenzamide (3-MB) noted two- to five-fold less stable transduction in the presence of drugs (197). Since an anti-PARP-1 antisense oligonucleotide resulted in parallel reductions of PARP-1 activity and stable transduction, it seemed that PARP-1 played a rate-limiting role in infection, which led to the suggestion that PARP-1 played an important role in gap repair (197). A subsequent study revealed that the infectivity of an HIV-1 vector was significantly impaired (approximately 23 fold) on MEFs derived from PARP-1-knockout mice as compared to matched PARP-1-expressing cells (198). However, not all investigators have observed infectivity defects in the absence of PARP-1 activity. Baekelandt *et al.* (193) observed that hamster cells infected in the presence of relatively high concentrations of 3-MB supported normal levels of virus transduction, discounting an essential role for PARP-1 in HIV-1 integration. Although Siva and Bushman (199) noted that the Ha *et al*. (198) vector infected PARP-$1^{-/-}$ MEFs two- to threefold less efficiently than control cells, knockout and control MEFs were infected equally by a separate HIV-1 vector (199). Based on these contradictory reports, one must posit that PARP-1 does not play an essential role in HIV-1 integration.

4.2.2. IN binding proteins

4.2.2.1. hRad18

The Rad18 protein is involved in the DNA postreplication repair/translesion pathway (reviewed in reference 200). Through binding to single-stranded breaks in DNA, Rad18 is thought to recruit Rad6, an E2 ubiquitin-conjugating enzyme (201) to the sites of DNA damage (202). The stable Rad6-Rad18 heterodimer (203, 204) has been hypothesized to destabilize the polymerase machinery, perhaps via ubiquitination of the stalled DNA polymerase complex, thus bypassing the lesion and allowing time for repair.

Human Rad18 (hRad18) was identified as an IN-interacting protein by coimmunoprecipitation of HIV-1 IN with FLAG-tagged hRad18 and colocalization of IN with hRad18 in cell nuclei (190). HIV-1 IN is relatively unstable when expressed on its own in mammalian cells, and stability can be increased via optimizing overall codon usage (205) or the identity of the N-terminal amino acid residue (206). This latter observation indicated that IN stability is dictated by the N-end rule. Cotransfection of hRad18 with either N-terminal Phe-IN or Arg-IN, which were both unstable (206), led to increased intracellular IN, suggesting that hRad18 might stabilize IN *in vivo* (190). Of note Maertens *et al.* (154) later described LEDGF/p75 as a cellular stabilizer of HIV-1 IN.

Rad18 possesses several functional domains, including an N-terminal RING finger that is often harbored by proteins carrying E3 ubiquitin ligase activity (207, 208), a zinc finger domain, and a putative SAP DNA-binding motif (209). The IN-binding region spans residues 65 to 226 and encompasses the zinc-finger domain (figure 8). Mulder *et al.* (190) proposed a potential role for hRad18 in HIV-1 integration that was based on studies of bacteriophage Mu transposition. MuA transposase protein is analogous to retroviral IN as MuA possesses two distinct endonucleolytic activities analogous to IN's 3' processing and DNA strand transfer activities that form a transpositional DNA recombination intermediate with single-stranded gaps (reviewed in reference 114). After DNA strand transfer, ClpX, a member of a conserved family of ATPases, unfolds MuA monomers, leading to

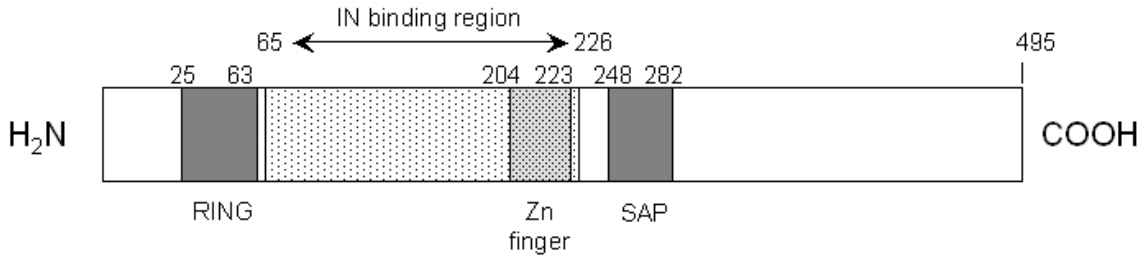

Figure 8. Structure/features of hRad18. The locations of the RING finger, zinc finger, and SAP domains are indicated as shaded boxes. The IN-binding area of the protein, which maps between residues 63 and 226, includes the zinc finger domain.

Table 1. Summary and potential roles of IN-binding proteins in HIV-1 integration

Protein	Method of identification	IN stimulation	IN binding region	Colocalization with IN	Putative function
INI1	Yeast two-hybrid	10-20 fold	NTD	?	Assembly, PIC Targeting
LEDGF/p75	Co-IP	Yes	NTD + CCD	Nuclear, associated with mitotic chromosomes	Catalytic, Nuclear import PIC Targeting
EED	Yeast two-hybrid	2-30 fold	CTD	Nuclear and cytoplasmic (next to nuclear pores)	Catalytic, Nuclear import PIC Targeting
HSP60	IN-affinity column	2-4 fold	CCD	?	Catalytic
Rad18	Co-IP	?	?	Nuclear in hRad18 typical structures	DNA repair

transpososome destabilization (210). Since MuA possesses overlapping binding sites for the phage-encoded MuB transposition activator and ClpX, release of MuB likely provides timed access to ClpX and thus a mechanism to regulate transpososome dissociation after DNA strand transfer (211). Such active transpososome disassembly plays an important role in the transition from integration to subsequent Mu DNA repair/replication (212, 213). Similarly, hRad18 might interact with HIV-1 IN soon after DNA strand transfer. The subsequently formed Rad6-Rad18 heterodimer could then destabilize PIC proteins that may remain tightly bound to the integration intermediate, allowing cellular DNA repair proteins to gain access to the gapped DNA (190).

5. SUMMARY AND PERSPECTIVE

This review detailed the potential roles of human cellular proteins in the catalytic and DNA repair steps of HIV-1 integration. Whereas HMGA1, BAF, INI1, LEDGF/p75, EED, and HSP60 have each been implicated as a potential catalytic cofactor for the IN (section 4.1), a variety of other proteins have been implicated as playing roles during the repair of the HIV-1 DNA recombination intermediate (section 4.2). However, solid genetic evidence for the obligate role of any of theses proteins in HIV-infected cells remains scant. Of all the potential catalytic cofactors, only HMGA1 has been studied via a genetic approach, in this case with knockout cell lines (17). Under these conditions neither HMGA1 nor the related HMGA2 protein played a discernable role in integration. With the relatively recent advent of siRNA technology (see reference 214 for review), it should be possible to selectively reduce the concentration of any one of these potential cofactors in cells and assay for resulting effect(s) on HIV-1 integration/infectivity. Indeed, the description of an effective synthetic siRNA (154) and shRNA-expression vector (155) that target human LEDGF/p75 suggests that such experiments are underway for at least a subset of the potential IN catalytic cofactors.

Many of the protein-modifying enzymes implicated in gap repair, in contrast, were identified through the use of cell-based genetic assays. However some of these results are controversial, as they have proven difficult to reproduce (see section 4.2.1). This suggests that certain experimental conditions like the identities of specific viral strains/vectors and/or cell types may have contributed to the results and if this is indeed the case, the roles of these repair-type proteins are likely not essential under all conditions of HIV-1 infection.

We have detailed five different human proteins that interact directly with HIV-1 IN, four of which could act as an IN catalytic cofactor and one as a gap repair factor (table 1). At least three other human proteins, karyopherin-alpha (215), importin 7 (216) and uracil deglycosylase 2 (UNG2) (217) have also been shown to interact with HIV-1 IN. Since karyopherin-alpha and importin 7 were implicated as playing roles in the translocation of HIV-1 PICs across intact nuclear pores and UNG2 was implicated in base excision repair of uracil-containing HIV-1 DNA, we considered these functions as outside the realm of the catalytic steps of DNA recombination and thus did not discuss these IN-interacting proteins in detail.

Finally, we identified HeLa cell proteins that coimmunoprecipitated with HIV-1 IN (figure 7). Our results highlighted the interaction between IN and LEDGF/p75, which corroborates recent findings from Debyser and colleagues (18, 154). Although we failed to detect any of the other previously-identified IN-interacting proteins outside of EF-1-alpha in this screen, this should not be taken as evidence for the lack of cofactor specificity.

For example, it is plausible that other IN-interacting proteins were simply not extracted from HeLa cells under conditions where both LEDGF/p75 and IN were efficiently recovered. Ongoing experiments in ours and others laboratories should help determine if LEDGF/p75 or any of the other human proteins discussed in detail in this review play an important role in allowing HIV-1 to accomplish its functional integration in human cells.

6. ACKNOWLEDGEMENTS

We thank P. Cherepanov and N. Vandegraaff for critically reviewing the manuscript, and H. Erdjument-Bromage and P. Tempst for performing mass spectroscopy. This work was supported by an HHMI Predoctoral Fellowship (E.D.), the Claudia Adams Barr Fund (P.S.), and NIH grants AI39394 and AI52014 (A.E.).

7. REFERENCES

1. Craigie, R: Retroviral DNA integration. In: Mobile DNA II. Eds: N. L. Craig, R. Craigie, M. Gellert, A. M. Lambowitz, ASM Press, Washington DC, 613-630 (2002)

2. Fassati, A. and S. P. Goff: Characterization of intracellular reverse transcription complexes of Moloney murine leukemia virus. *J Virol* 73, 8919-8925 (1999)

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

4. Nermut, M. V. and A. Fassati: Structural analyses of purified human immunodeficiency virus type 1 intracellular reverse transcription complexes. *J Virol* 77, 8196-8206 (2003)

5. Goff, S. P: Intracellular trafficking of retroviral genomes during the early phase of infection: viral exploitation of cellular pathways. *J Gene Med* 3, 517-528 (2001)

6. Stevenson, M., S. Haggerty, C. A. Lamonica, C. M. Meier, S.-K. Welch and A. Wasiak: Integration is not necessary for expression of human immunodeficiency virus type 1 protein products. *J Virol* 64, 2421-2425 (1990)

7. Lafemina, R. L., C. L. Schneider, H. L. Robbins, P. L. Callahan, K. LeGrow, E. Roth, W. A. Schleif and E. A. Emini: Requirement of active human immunodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. *J Virol* 66, 7414-7419 (1992)

8. Cannon, P. M., W. Wilson, E. Byles, S. M. Kingsman and A. J. Kingsman: Human immunodeficiency virus type 1 integrase: effect on viral replication of mutations at highly conserved residues. *J Virol* 68, 4768-4775 (1994)

9. Taddeo, B., W. A. Haseltine and C. M. Farnet: Integrase mutants of human immunodeficiency virus type 1 with a specific defect in integration. *J Virol* 68, 8401-8405 (1994)

10. Wiskerchen, M. and M. A. Muesing: Human immunodeficiency virus type integrase 1: effects of mutations on viral ability to integrate, direct gene expression from unintegrated viral DNA templates, and sustain propagation in primary cells. *J Virol* 69, 376-386 (1995)

11. Engelman, A., G. Englund, J. M. Orenstein, M. A. Martin and R. Craigie: Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication. *J Virol* 69, 2729-2736 (1995)

12. Ansari-Lari, M. L., L. A. Donehower and R. A. Gibbs: Analysis of human immunodeficiency virus type 1 integrase mutants. *Virology* 211, 332-335 (1995)

13. Nakajima, N., R. Lu and A. Engelman: Human immunodeficiency virus type 1 replication in the absence of integrase-mediated DNA recombination: Definition of permissive and nonpermissive T-cell lines. *J Virol* 75, 7944-7955 (2001)

14. Engelman, A: The roles of cellular factors in retroviral integration. *Curr Topics Microbiol Immunol* 281, 209-238 (2003)

15. Bushman, F. D: Host proteins in retroviral cDNA integration. *Adv Virus Res* 52, 301-317 (1999)

16. Hindmarsh, P. and J. Leis: Reconstitution of concerted DNA integration with purified components. *Adv Virus Res* 52, 397-410 (1999)

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

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

19. Violot, S., S. S. Hong, D. Rakotobe, C. Petit, B. Gay, K. Moreau, G. Billaud, S. Priet, J. Sire, O. Schwartz, J.-F. Mouscadet and P. Boulanger: The human Polycomb group EED protein interacts with the integrase of human immunodeficiency virus type 1. *J Virol* 77, 12507-12522 (2003)

20. Brown, P. O., B. Bowerman, H. E. Varmus and J. M. Bishop: Correct integration of retroviral DNA *in vitro*. *Cell* 49, 347-356 (1987)

21. Fujiwara, T. and K. Mizuuchi: Retroviral DNA integration: structure of an integration intermediate. *Cell* 54, 497-504 (1988)

22. Brown, P. O., B. Bowerman, H. E. Varmus and J. M.

Bishop: Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. *Proc Natl Acad Sci USA* 86, 2525-2529 (1989)

23. Lee, Y. M. L. and J. M. Coffin: Relationship of avian retrovirus DNA synthesis to integration *in vitro*. *Mol Cell Biol* 11, 1419-1430 (1991)

24. Daniel, R., R. A. Katz and A. M. Skalka: A role for DNA-PK in retroviral DNA integration. *Science* 284, 644-647 (1999)

25. Yoder, K. E. and F. D. Bushman: Repair of gaps in retroviral DNA integration intermediates. *J Virol* 74, 11191-11200 (2000)

26. Katzman, M. and R. A. Katz: Substrate recognition by retroviral integrases. *Adv Virus Res* 52, 371-395 (1999)

27. Pauza, C. D: Two bases are deleted from the termini of HIV-1 linear DNA during integrative recombination. *Virology* 179, 886-889 (1990)

28. Sherman, P. A. and J. A. Fyfe: Human immunodeficiency virus integration protein expressed in *Escherichia coli* possesses selective DNA cleaving activity. *Proc Natl Acad Sci USA* 87, 5119-5123 (1990)

29. Engelman, A., K. Mizuuchi and R. Craigie: HIV-1 DNA integration: Mechanism of viral DNA cleavage and DNA strand transfer. *Cell* 67, 1211-1221 (1991)

30. Vink, C., E. Yeheskiely, G. A. van der Marel, J. H. van Boom and R. H. A. Plasterk: Site-specific hydrolysis and alcoholysis of human immunodeficiency virus DNA termini by the viral integrase protein. *Nucleic Acids Res* 19, 6691-6698 (1991)

31. Vincent, K. A., D. York-Higgins, Q. Margarita and P. O. Brown: Host sequences flanking the HIV provirus. *Nucleic Acids Res* 18, 6045-6047 (1990)

32. Vink, C., M. Groenink, Y. Elgersma, R. A. M. Fouchier, M. Tersmette and R. H. A. Plasterk: Analysis of the junctions between human immunodeficiency virus type 1 proviral DNA and human DNA. *J Virol* 64, 5626-5627 (1990)

33. Goff, S. P: Retroviridae: The retroviruses and their replication. In: Fields Virology, 4th Edition. Eds: D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, S. E. Straus, Lippincott Williams and Wilkins, Philadelphia PA, 1871-1939 (2001)

34. Kilzer, J. M., T. Stracker, B. Beitzel, K. Meek, M. Weitzman and F. D. Bushman: Roles of host cell factors in circularization of retroviral DNA. *Virology* 314, 460-467 (2003)

35. Li, L., J. M. Olvera, K. E. Yoder, R. S. Mitchell, S. L. Butler, M. Lieber, S. L. Martin and F. D. Bushman: Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection. *EMBO J* 20, 3272-3281 (2001)

36. Shoemaker, C., J. Hoffman, S. P. Goff and D. Baltimore: Intramolecular integration within Moloney murine leukemia virus DNA. *J Virol* 40, 164-172 (1981)

37. Ellison, V., H. Abrams, T. Roe, J. Lifson and P. Brown: Human immunodeficiency virus integration in a cell-free system. *J Virol* 64, 2711-2715 (1990)

38. Farnet, C. M. and W. A. Haseltine: Integration of human immunodeficiency virus type 1 DNA *in vitro*. *Proc Natl Acad Sci USA* 87, 4164-4168 (1990)

39. Lee, M. S. and R. Craigie: Protection of retroviral DNA from autointegration: Involvement of a cellular factor. *Proc Natl Acad Sci USA* 91, 9823-9827 (1994)

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

41. Li, Y., J. C. Kappes, J. A. Conway, R. W. Price, G. M. Shaw and B. H. Hahn: Molecular characterization of human immunodeficiency virus type 1 cloned directly from uncultured human brain tissue: identification of replication-competent and -defective viral genomes. *J Virol* 65, 3973-3985 (1991)

42. Lee, Y. M. and J. M. Coffin: Efficient autointegration of avian retrovirus DNA *in vitro*. *J Virol* 64, 5958-5965 (1990)

43. Andrake, M. D. and A. M. Skalka: Retroviral integrase, putting the pieces together. *J Biol Chem* 271, 19633-19636 (1996)

44. Esposito, D. and R. Craigie: HIV integrase structure and function. *Adv Virus Res* 52, 319-333 (1999)

45. Johnson, M. S., M. A. McClure, D. F. Feng, J. Gray and R. F. Doolittle: Computer analysis of retroviral pol genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. *Proc Natl Acad Sci USA* 83, 7648-7652 (1986)

46. Khan, E., J. P. Mack, R. A. Katz, J. Kulkosky and A. M. Skalka: Retroviral integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Res* 19, 851-860 (1991)

47. Cai, M., R. Zheng, M. Caffrey, R. Craigie, G. M. Clore and A. M. Gronenborn: Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. *Nat Struct Biol* 4, 567-577 (1997)

48. Eijkelenboom, A. P., F. M. van den Ent, A. Vos, J. F. Doreleijers, K. Hard, T. D. Tullius, R. H. A. Plasterk, R. Kaptein and R. Boelens: The solution structure of the amino-terminal HHCC domain of HIV-2 integrase: a three-

helix bundle stabilized by zinc. *Curr Biol* 1, 739-746 (1997)

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

50. Lee, S. P., J. Xiao, J. R. Knutson, M. S. Lewis and M. K. Han: Zn2+ promotes the self-association of human immunodeficiency virus type-1 integrase *in vitro*. *Biochemistry* 36, 173-180 (1997)

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

52. Drelich, M., R. Wilhelm and J. Mous: Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein *in vitro*. *Virology* 188, 459-468 (1992)

53. Kulkosky, J., K. S. Jones, R. A. Katz, J. P. G. Mack and A. M. Skalka: 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 (1992)

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

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

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

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

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

59. Gerton, J. L. and P. O. Brown: The core domain of HIV-1 integrase recognizes key features of its DNA substrates. *J Biol Chem* 272, 25809-25815 (1997)

60. Gerton, J. L., S. Ohgi, M. Olsen, J. DeRisi and P. O. Brown: Effects of mutations in residues near the active site of human immunodeficiency virus type 1 integrase on specific enzyme-substrate interactions. *J Virol* 72, 5046-5055 (1998)

61. Heuer, T. S. and P. O. Brown: 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 (1997)

62. Appa, R. S., C.-G. Shin, P. Lee and S. A. Chow: Role of the non-specific DNA binding region and α helices within the core domain of retroviral integrase in selecting target DNA sites for integration. *J Biol Chem* 276, 45848-45855 (2001)

63. Harper, A. L., L. M. Skinner, M. Sudol and M. Katzman: Use of patient-derived human immunodeficiency virus type 1 integrases to identify a protein residue that affects target site selection. *J Virol* 75, 7756-7762 (2001)

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

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

66. Jenkins, T. M., A. Engelman, R. Ghirlando and R. Craigie: A soluble active mutant of HIV-1 integrase: involvement of both the core and C-terminal domains in multimerization. *J Biol Chem* 271, 7712-7718 (1996)

67. Puras Lutzke, R. A. and R. H. A. Plasterk: Structure-based mutational analysis of the C-terminal DNA binding domain of human immunodeficiency virus type 1 integrase: critical residues for protein oligomerization and DNA binding. *J Virol* 72, 4841-4848 (1998)

68. Gao, K., S. L. Butler and F. Bushman: Human immunodeficiency virus type 1 integrase: arrangement of protein domains in active cDNA complexes. *EMBO J* 20, 3565-3576 (2001)

69. Woerner, A. M., M. Klutch, J. G. Levin and C. J. Marcus-Sekura: Localization of DNA binding activity of HIV-1 integrase to the C-terminal half of the protein. *AIDS Res Human Retroviruses* 8, 297-304 (1992)

70. Puras Lutzke, R. A., C. Vink and R. H. A. Plasterk: Characterization of the minimal DNA-binding domain of the HIV integrase protein. *Nucleic Acids Res* 22, 4125-4131 (1994)

71. Engelman, A., A. B. Hickman and R. Craigie: The core and carboxyl terminal domains of the integrase protein of human immunodeficiency virus type 1 each contribute to non-specific DNA binding. *J Virol* 68, 5911-5917 (1994)

72. Bowerman, B., P. O. Brown, J. M. Bishop and H. E.

Varmus: A nucleoprotein complex mediates the integration of retroviral DNA. *Genes Dev* 3, 469-478 (1989)

73. Farnet, C. M. and W. A. Haseltine: Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. *J Virol* 65, 1910-1915 (1990)

74. Bukrinsky, M. I., N. Sharova, T. L. McDonald, T. Pushkarskaya, W. G. Tarpley and M. Stevenson: 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 USA* 90, 6125-6129 (1993)

75. Heinzinger, N. K., M. I. Bukinsky, S. A. Haggerty, A. M. Ragland, V. Kewalramani, M. A. Lee, H. E. Gendelman, L. Ratner, M. Stevenson and M. Emerman: The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. *Proc Natl Acad Sci USA* 91, 7311-7315 (1994)

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

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

78. Li, L., C. M. Farnet, W. F. Anderson and F. D. Bushman: Modulation of activity of Moloney murine leukemia virus preintegration complexes by host factors *in vitro*. *J Virol* 72, 2125-2131 (1998)

79. Suzuki, Y. and R. Craigie: Regulatory mechanisms by which barrier-to-autointegration factor blocks autointegration and stimulates intermolecular integration of Moloney murine leukemia virus preintegration complexes. *J Virol* 76, 12376-12380 (2002)

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

81. Khiytani, D. K. and N. J. Dimmock: Characterization of a human immunodeficiency virus type 1 pre-integration complex in which the majority of the cDNA is resistant to DNase I digestion. *J Gen Virol* 83, 2523-2532 (2002)

82. Wei, S. Q., K. Mizuuchi and R. Craigie: A large nucleoprotein assembly at the ends of the viral DNA mediates retroviral DNA integration. *EMBO J* 16, 7511-7520 (1997)

83. Chen, H., S. Q. Wei and A. Engelman: Multiple integrase functions are required to form the native structure of the human immunodeficiency virus type I intasome. *J Biol Chem* 274, 17358-17364 (1999)

84. Wei, S. Q., K. Mizuuchi and R. Craigie : Footprints on the viral DNA ends in Moloney murine leukemia virus preintegration complexes reflect a specific association with integrase. *Proc Natl Acad Sci USA* 95, 10535-10540 (1998)

85. Chen, H. and A. Engelman: Characterization of a replication-defective human immunodeficiency virus type 1 att site mutant that is blocked after the 3' processing step of retroviral integration. *J Virol* 74, 8188-8193 (2000)

86. Chen, H. and A. Engelman: Asymmetric processing of human immunodeficiency virus type 1 cDNA *in vivo*: implications for functional end coupling during the chemical steps of DNA transposition. *Mol Cell Biol* 21, 6758-6767 (2001)

87. Katzman, M., R. A. Katz, A. M. Skalka and J. Leis: The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the *in vivo* sites of integration. *J Virol* 63, 5319-5327 (1989)

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

89. Bushman, F. D. and R. Craigie: Activities of human immunodeficiency virus (HIV) integration protein *in vitro*: specific cleavage and integration of HIV DNA. *Proc Natl Acad Sci USA* 88, 1339-1343 (1991)

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

91. Sinha, S., M. H. Pursley and D. P. Grandgenett: Efficient concerted integration by recombinant human immunodeficiency virus type 1 integrase without cellular or viral cofactors. *J Virol* 76, 3105-3113 (2002)

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

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

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

95. Schroder, A. R., P. Shinn, H. Chen, C. Berry, J. R. Ecker and F. Bushman: HIV-1 integration in the human

genome favors active genes and local hotspots. *Cell* 110, 521-529 (2002)

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

97. Bushman, F. D: Integration site selection by lentiviruses: biology and possible control. *Curr Top Microbiol Immunol* 261, 165-177 (2002)

98. Sandmeyer, S: Integration by design. *Proc Natl Acad Sci USA* 100, 5586-5588 (2003)

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

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

101. Harris, D. and A. Engelman: Both the structure and DNA binding function of the barrier-to-autointegration factor contribute to reconstitution of HIV type 1 integration *in vitro*. *J Biol Chem* 275, 39671-39677 (2000)

102. Zheng, R., R. Ghirlando, M. S. Lee, K. Mizuuchi, M. Krause and R. Craigie: Barrier-to-autointegration (BAF) bridges DNA in a discrete, higher-order nucleoprotein complex. *Proc Natl Acad Sci USA* 97, 8997-9002 (2000)

103. Mansharamani, M., D. R. Graham, D. Monie, K. K. Lee, J. E. Hildreth, R. F. Siliciano and K. L. Wilson: Barrier-to-autointegration factor BAF binds p55 Gag and matrix and is a host component of human immunodeficiency virus type 1 virions. *J Virol* 77, 13084-13092 (2003)

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

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

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

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

108. Reeves, R. and M. S. Nissen: The A•T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J Biol Chem* 265, 8573-8582 (1990)

109. Reeves, R. and L. Beckerbauer: HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim Biophys Acta* 1519, 13-29 (2001)

110. Huth, J. R., C. A. Bewley, M. S. Nissen, J. N. S. Evans, R. Reeves, A. M. Gronenborn and G. M. Clore: The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat Struct Biol* 4, 657-665 (1997)

111. Thanos, D. and T. Maniatis: Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell* 83, 1091-1100 (1995)

112. Gao, K., R. J. Gorelick, D. G. Johnson and F. Bushman: Cofactors for human immunodeficiency virus type 1 cDNA integration *in vitro*. *J Virol* 77, 1598-1603 (2003)

113. Henderson, A., M. Bunce, N. Siddon, R. Reeves and D. J. Tremethick: 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 (2000)

114. Chaconas, G. and R. M Harshey: Transposition of phage Mu DNA. In: Mobile DNA II. Eds: N. L. Craig, R. Craigie, M. Gellert, A. M. Lambowitz, ASM Press, Washington DC, 384-456 (2002)

115. Carteau, S., R. J. Gorelick and F. D. Bushman: Coupled integration of human immunodeficiency virus type 1 cDNA ends by purified integrase *in vitro*: stimulation by the viral nucleocapsid protein. *J Virol* 73, 6670-6679 (1999)

116. Cai, M., Y. Huang, R. Zheng, S. Q. Wei, R. Ghirlando, M. S. Lee, R. Craigie, A. M. Gronenborn and G. M. Clore: Solution structure of the cellular factor BAF responsible for protecting retroviral DNA from autointegration. *Nat Struct Biol* 5, 903-909 (1998)

117. Shumaker, D. K., K. K. Lee, Y. C. Tanhehco, R. Craigie and K. L. Wilson: LAP2 binds to BAF•DNA complexes: requirement for the LEM domain and modulation by variable regions. *EMBO J* 20, 1754-1764 (2001)

118. Furukawa, K: LAP2 binding protein 1 (L2BP1/BAF) is a candidate mediator of LAP2-chromatin interaction. *J Cell Sci* 112, 2485-2492 (1999)

119. Lee, K. K., T. Haraguchi, R. S. Lee, T. Koujin, Y. Hiraoka and K. L. Wilson: Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF. *J Cell Sci* 114, 4567-4573 (2001)

120. Goldberg, M., A. Harel and Y. Gruenbaum: The

nuclear lamina: molecular organization and interaction with chromatin. *Crit Rev Eukaryot Gene Expr* 9, 285-293 (1999)

121. Liu, J., K. K. Lee, M. Segura-Totten, E. Neufeld, K. L. Wilson and Y. Gruenbaum: MAN1 and emerin have overlapping function(s) essential for chromosome segregation and cell division in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 100, 4598-4603 (2003)

122. Segura-Totten, M., A. K. Kowalski, R. Craigie and K. L. Wilson: Barrier-to-autointegration factor: major roles in chromatin decondensation and nuclear assembly. *J Cell Biol* 158, 475-485 (2002)

123. Haraguchi, T., T. Koujin, M. Segura-Totten, K. K. Lee, Y. Matsuoka, Y. Yoneda, K. L. Wilson and Y. Hiraoka: BAF is required for emerin assembly into the reforming nuclear envelope. *J Cell Sci* 114, 4575-4585 (2001)

124. Wang, X., S. Xu, C. Rivolta, L. Y. Li, G. H. Peng, P. K. Swain, C. H. Sung, A. Swaroop, E. L. Berson, T. P. Dryja and S. Chen: Barrier-to-autointegration factor interacts with the cone-rod homeobox and represses its transactivation function. *J Biol Chem* 277, 43288-43300 (2002)

125. Forné, I., M. Carrascal, L. Martinez-Lostao, J. Abian, J. L. Rodriguez-Sánchez and C. Juarez: Identification of the autoantigen HB as the barrier-to-autointegration factor. *J Biol Chem* 278, 50641-50644 (2003)

126. Umland, T. C., S. Q. Wei, R. Craigie and D. R. Davies: Structural basis of DNA bridging by barrier-to-autointegration factor. *Biochemistry* 39, 9130-9138 (2000)

127. Shao, X. and N. V. Grishin: Common fold in helix-hairpin-helix proteins. *Nucleic Acids Res* 28, 2463-2650 (2000)

128. Kingston, R. E. and G. J. Narlikar: ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev* 13, 2339-2352 (1999)

129. Wu, D. Y., G. V. Kalpana, S. P. Goff and W. H. Schubach: Epstein-Barr virus nuclear protein 2 (EBNA2) binds to a component of the human SNF-SWI complex, hSNF5/Ini1. *J Virol* 70, 6020-6028 (1996)

130. Lee, D., H. Sohn, G. V. Kalpana and J. Choe: Interaction of E1 and hSNF5 proteins stimulates replication of human papillomavirus DNA. *Nature* 399, 487-491 (1999)

131. Cheng, S. W., K. P. Davies, E. Yung, R. J. Beltran, J. Yu and G. V. Kalpana: c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function. *Nat Genet* 22, 102-105 (1999)

132. Rozenblatt-Rosen, O., T. Rozovskaia, D. Burakov, Y. Sedkov, S. Tillib, J. Blechman, T. Nakamura, C. M. Croce, A. Mazo and E. Canaani: The C-terminal SET domains of ALL-1 and TRITHORAX interact with the INI1 and SNR1 proteins, components of the SWI/SNF complex. *Proc Natl Acad Sci USA* 95, 4152-4157 (1998)

133. Lee, D., J. W. Kim, T. Seo, S. G. Hwang, E. J. Choi and J. Choe: SWI/SNF complex interacts with tumor suppressor p53 and is necessary for the activation of p53-mediated transcription. *J Biol Chem* 277, 22330-22337 (2002)

134. Versteege, I., N. Sevenet, J. Lange, M. F. Rousseau-Merck, P. Ambros, R. Handgretinger, A. Aurias and O. Delattre: Truncating mutations of hSNF5/INI1 in aggressive pediatric cancer. *Nature* 394, 203-206 (1998)

135. Muchardt C and M. Yaniv: When the SWI/SNF complex remodels… the cell cycle. *Oncogene* 20, 3067-3075 (2001)

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

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

138. Holstege, F. C., E. G. Jennings, J. J. Wyrick, T. I. Lee, C. J. Hengartner, M. R. Green, T. R. Golub, E. S. Lander and R. A. Young: Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95, 717-728 (1998)

139. Doucas, V., M. Tini, D. A. Egan and R. M. Evans: Modulation of CREB binding protein function by the promyelocytic (PML) oncoprotein suggests a role for nuclear bodies in hormone signaling. *Proc Natl Acad Sci USA* 96, 2627-2632 (1999)

140. Morozov, A., E. Jung and G. Kalpana: Structure-function analysis of integrase interactor INI1 reveals differential properties of two-repeat motifs present in the highly conserved region. *Proc Natl Acad Sci USA* 95, 1120-1125 (1998)

141. Cimarelli, A. and J. L. Darlix: Assembling the human immunodeficiency virus type 1. *Cell Mol Life Sci* 59, 1166-1184 (2002)

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

143. Shinohara, T., D. P. Singh and N. Fatma: LEGDF, a survival factor, activates stress-related genes. *Prog Retin Eye Res* 21, 341-358 (2002)

144. Singh, D. P., N. Ohguro, L. T. Chylack Jr and T. Shinohara: Lens epithelium-derived growth factor:

increased resistance to thermal and oxidative stresses. *Invest Ophthalmol Vis Sci* 40, 1444-1451(1999)

145. Singh, D. P., N. Fatma, A. Kimura, L. T. Chylack Jr. and T. Shinohara: LEDGF binds to heat shock (HSE) and stress-related elements (STRE) to activate the expression of stress-related genes. *Biochem Biophys Res Commun* 283, 943-955 (2001)

146. Fatma, N., D. P. Singh, T. Shinohara and L. T. Chylack Jr.: Transcriptional regulation of the antioxidant protein 2 gene, a thiol-specific antioxidant, by lens epithelium-derived growth factor to protect cells from oxidative stress. *J Biol Chem* 276, 48899-48907 (2001)

147. Pluymers, W., P. Cherepanov, D. Schols, E. De Clercq and Z. Debyser: Nuclear localization of human immunodeficiency virus type 1 integrase expressed as a fusion protein with green fluorescent protein. *Virology* 258, 327-332 (1999)

148. Petit, C., O. Schwartz and F. Mammano: Oligomerization within virions and subcellular localization of human immunodeficiency virus type 1 integrase. *J Virol* 73, 5079-5088 (1999)

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

150. Tsurutani, N., M. Kubo, Y. Maeda, T. Ohashi, N. Yamamoto, M. Kannagi and T. Masuda: Identification of critical amino acid residues in human immunodeficiency virus type 1 IN required for efficient proviral DNA formation at steps prior to integration in dividing and nondividing cells. *J Virol* 74, 4795-4806 (2000)

151. Bouyac-Bertoia, M., J. D. Dvorin, R. A. Fouchier, Y. Jenkins, B. E. Meyer, L. I. Wu, M. Emerman and M. H. Malim: HIV-1 infection requires a functional integrase NLS. *Mol Cell* 7, 1025-1035 (2001)

152. Limon, A., E. Devroe, R. Lu, H. Z. Ghory, P. A. Silver and A. Engelman: 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 (2002)

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

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

155. Devroe, E. and P. A. Silver: Retrovirus-delivered siRNA. *BMC Biotechnol* 2, 15 (2002)

156. Stec, I., S. B. Nagl, G. J. van Ommen and J. T. den Dunnen: The PWWP domain: a potential protein-protein domain in nuclear proteins influencing differentiation? *FEBS Lett* 473, 1-5 (2000)

157. Qiu, C., K. Sawada, X. Zhang and X. Cheng: The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. *Nat Struct Biol* 9, 217-224 (2002)

158. Ge, H., Y. Si and R. G. Roeder: Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. *EMBO J* 17, 6723-6729 (1998)

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

160. Woodward, C. L., Y. Wang, W. J. Dixon, H. Htun and S. A. Chow: Subcellular localization of feline immunodeficiency virus integrase and mapping of its karyophilic determinant. *J Virol* 77, 4516-4527 (2003)

161. Schumacher, A., O. Lichtarge, S. Schwartz and T. Magnuson: The murine Polycomb-group gene eed and its human orthologue: functional implications of evolutionary conservation. *Genomics* 54, 79-88 (1998)

162. Neer, E. J., C. J. Schmidt, R. Nambudripad and T. F. Smith: The ancient regulatory-protein family of WD-repeat proteins. *Nature* 371, 297-300 (1994)

163. Denisenko, O. N. and K. Bomsztyk: The product of the murine homolog of the Drosophila extra sex combs gene displays transcriptional repressor activity. *Mol Cell Biol* 17, 4707-4717 (1997)

164. Pirrotta, V: Polycombing the genome: PcG, trxG, and chromatin silencing. *Cell* 93, 333-336 (1998)

165. van der Vlag, J. and A. P. Otte: Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation. *Nat Genet* 23, 474-478 (1999)

166. Peytavi, R., S. S. Hong, B. Gay, A. D. d'Angeac, L. Selig, S. Benichou, R. Benarous and P. Boulanger: HEED, the product of the human homolog of the murine eed gene, binds to the matrix protein of HIV-1. *J Biol Chem* 274, 1635-1645 (1999)

167. Satijn, D. P., K. M. Hamer, J. den Blaauwen and A. P. Otte: The polycomb group protein EED interacts with YY1, and both proteins induce neural tissue in Xenopus embryos. *Mol Cell Biol* 21, 1360-1369 (2001)

168. Witte, V., B. Laffert, O. Rosorius, P. Lischka, K.

Blume, G. Galler, A. Stilper, D. Willbold, P. D'Aloja, M. Sixt, J. Kolanus, M. Ott, W. Kolanus, G. Schuler and A. S. Baur: HIV-1 Nef mimics an integrin receptor signal that recruits the Polycomb Group protein Eed to the plasma membrane. *Mol Cell* 13, 179-190 (2004)

169. Bukau, B. and A. L. Horwich: The Hsp70 and Hsp60 chaperone machines. *Cell* 92, 351-366 (1998)

170. Park, S. and G. Jung: Human hepatitis B virus polymerase interacts with the molecular chaperonin Hsp60. *J Virol* 75, 6962-6968 (2001)

171. Hammond, C. and A. Helenius: Folding of VSV G protein: sequential interaction with BiP and calnexin. *Science* 266, 456-458 (1994)

172. Jindal, S. and R. A. Young: Vaccinia virus infection induces a stress response that leads to association of Hsp70 with viral proteins. *J Virol* 66, 5357-5362 (1992)

173. Momose, F., T. Naito, K. Yano, S. Sugimoto, Y. Morikawa and K. Nagata: Identification of Hsp90 as a stimulatory host factor involved in influenza virus RNA synthesis. *J Biol Chem* 277, 45306-45314 (2002)

174. Macejak, D. G. and R. B. Luftig: Stabilization of actin filaments at early times after adenovirus infection and in heat-shocked cells. *Virus Res* 19, 31-45 (1991)

175. Morimoto, R. I., A. Tissieres and G. Georgopoulos: The biology of heat shock proteins and molecular chaperones. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1994)

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

177. Dvorin, J. D., P. Bell, G. G. Maul, M. Yamashita, M. Emerman and M. H. Malim: Reassessment of the roles of integrase and the central DNA flap in human immunodeficiency virus type 1 nuclear import. *J Virol* 76, 12087-12096 (2002)

178. Erdjument-Bromage, H., M. Lui, L. Lacomis, A. Grewal, R. S. Annan, D. E. McNulty, S. A. Carr and P. Tempst: Examination of micro-tip reversed-phase liquid chromatographic extraction of peptide pools for mass spectrometric analysis. *J Chromatogr A* 826, 167-181 (1998)

179. Sebastiaan Winkler, G., L. Lacomis, J. Philip, H. Erdjument-Bromage, J. Q. Svejstrup and P. Tempst: Isolation and mass spectrometry of transcription factor complexes. *Methods* 26, 260-269 (2002)

180. Liu, B., R. Dai, C.-J. Tian, L. Dawson, R. Gorelick and X.-F. Yu: Interaction of the human immunodeficiency virus type 1 nucleocapsid with actin. *J Virol* 73, 2901-2908 (1999)

181. Wilk, T., B. Gowen and S. D. Fuller: Actin associates with the nucleocapsid domain of the human immunodeficiency virus Gag polyprotein. *J Virol* 73, 1931-1940 (1999)

182. Fackler, O. T., N. Kienzle, E. Kremmer, A. Boese, B. Schramm, T. Klimkait, C. Kucherer, and N. Mueller-Lantzsch: Association of human immunodeficiency virus Nef protein with actin is myristoylation dependent and influences its subcellular localization. *Eur J Biochem* 247, 843-851 (1997)

183. Hottiger, M., K. Gramatikoff, O. Georgiev, C. Chaponnier, W. Schaffner and U. Hubscher: The large subunit of HIV-1 reverse transcriptase interacts with beta-actin. *Nucleic Acids Res* 23, 736-741 (1995)

184. Cimarelli, A. and J. Luban: Translation elongation factor 1-alpha interacts specifically with the human immunodeficiency virus type 1 Gag polyprotein. *J Virol* 73, 5388-5401 (1999)

185. Ott, D. E., L. V. Coren, B. P. Kane, L. K. Busch, D. G. Johnson, R. C. Sowder 2nd, E. N. Chertova, L. O. Arthur and L. E. Henderson: Cytoskeletal proteins inside human immunodeficiency virus type 1 virions. *J Virol* 70, 7734-7743 (1996)

186. Ott, D. E., L. V. Coren, D. G. Johnson, B. P. Kane, R. C. Sowder 2nd, Y. D. Kim, R. J. Fisher, X. Z. Zhou, K. P. Lu and L. E. Henderson: Actin-binding cellular proteins inside human immunodeficiency virus type 1. *Virology* 266, 42-51 (2000)

187. de Soultrait, V. R., A. Caumont, P. Durrens, C. Calmels, V. Parissi, P. Recordon, E. Bon, C. Desjobert, L. Tarrago-Litvak and M. Fournier: HIV-1 integrase interacts with yeast microtubule-associated proteins. *Biochim Biophys Acta* 1575, 40-48 (2002)

188. Tamaskovic, R., S. J. Bichsel and B. A. Hemmings: NDR family of AGC kinases--essential regulators of the cell cycle and morphogenesis. *FEBS Lett* 546, 73-80 (2003)

189. Brin, E., J. Yi, A. M. Skalka and J. Leis: Modeling the late steps in HIV-1 retroviral integrase-catalyzed DNA integration. *J Biol Chem* 275, 39287-39295 (2000)

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

191. Lees-Miller, S. P. and K. Meek: Repair of DNA double strand breaks by non-homologous end joining. *Biochimie* 85, 1161-1173 (2003)

192. Fulop, G. M., G. C. Bosma, M. J. Bosma and R. A. Phillips: Early B-cell precursors in scid mice: normal numbers of cells transformable with Abelson murine leukemia virus (A-MuLV). *Cell Immunol* 113, 192-201 (1988)

193. Baekelandt, V., A. Claeys, P. Cherepanov, E. De

Clercq, B. De Strooper, B. Nuttin and Z. Debyser: DNA-Dependent protein kinase is not required for efficient lentivirus integration. *J Virol* 74, 11278-11285 (2000)

194. Daniel, R., R. A. Katz, G. Merkel, J. C. Hittle , T. J. Yen and A. M. Skalka: Wortmannin potentiates integrase-mediated killing of lymphocytes and reduces the efficiency of stable transduction by retroviruses. *Mol Cell Biol* 21, 1164-1172 (2001)

195. Daniel, R., G. Kao, K. Taganov, J. G. Greger, O. Favorova, G. Merkel, T. J. Yen, R. A. Katz and A. M. Skalka: Evidence that the retroviral DNA integration process triggers an ATR-dependent DNA damage response. *Proc Natl Acad Sci USA* 100, 4778-4783 (2003)

196. D'Amours, D., S. Desnoyers, I. D'Silva and G. G. Poirier: Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *Biochem J* 342, 249-268 (1999)

197. Gaken, J. A., M. Tavassoli, S. U. Gan, S. Vallian, I. Giddings, D. C. Darling , J. Galea-Lauri, M. G. Thomas, H. Abedi, V. Schreiber, J. Menissier-de Murcia, M. K. Collins, S. Shall and F. Farzaneh: Efficient retroviral infection of mammalian cells is blocked by inhibition of poly(ADP-ribose) polymerase activity. *J Virol* 70, 3992-4000 (1996)

198. Ha, H. C., K. Juluri, Y. Zhou, S. Leung, M. Hermankova and S. H. Snyder: Poly(ADP-ribose) polymerase-1 is required for efficient HIV-1 integration. *Proc Natl Acad Sci USA* 98, 3364-3368 (2001)

199. Siva, A. C. and F. Bushman: Poly(ADP-ribose) polymerase 1 is not strictly required for infection of murine cells by retroviruses. *J Virol* 76, 11904-11910 (2002)

200. Broomfield, S., T. Hryciw and W. Xiao: DNA postreplication repair and mutagenesis in Saccharomyces cerevisiae. *Mutat Res* 486, 167-184 (2001)

201. Prakash, S., P. Sung and L. Prakash: DNA repair genes and proteins of Saccharomyces cerevisiae. *Annu Rev Genet* 27, 33-70 (1993)

202. Bailly, V., J. Lamb, P. Sung, S. Prakash and L. Prakash: Specific complex formation between yeast RAD6 and RAD18 proteins: a potential mechanism for targeting RAD6 ubiquitin-conjugating activity to DNA damage sites. *Genes Dev* 8, 811-820 (1994)

203. Bailly, V., S. Prakash and L. Prakash: Domains required for dimerization of yeast Rad6 ubiquitin-conjugating enzyme and Rad18 DNA binding protein. *Mol Cell Biol* 17, 4536-4543 (1997)

204. Bailly, V., J. Lamb, P. Sung, S. Prakash and L. Prakash: Yeast DNA repair proteins Rad6 and Rad18 form a heterodimer that has ubiquitin conjugating, DNA binding, and ATP hydrolytic activities. *J Biol Chem* 272, 23360-23365 (1997)

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

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

207. Lorick, K. L., J. P. Jensen, S. Fang, A. M. Ong, S. Hatakeyama and A. M. Weissman: RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc Natl Acad Sci USA* 96, 11364-11369 (1999)

208. Jackson, P. K., A. G. Eldridge, E. Freed, L. Furstenthal, J. Y. Hsu, B. K. Kaiser and J. D. Reimann: The lore of the RINGs: substrate recognition and catalysis by ubiquitin ligases. *Trends Cell Biol* 10, 429-439 (2000)

209. Aravind, L. and E. V. Koonin: SAP - a putative DNA-binding motif involved in chromosomal organization. *Trends Biochem Sci* 25, 112-114 (2000)

210. Burton, B. M., T. L. Williams and T. A. Baker: ClpX-mediated remodeling of mu transpososomes: selective unfolding of subunits destabilizes the entire complex. *Mol Cell* 8, 449-454 (2001)

211. Levchenko, I., M. Yamauch and T. A. Baker: ClpX and MuB interact with overlapping regions of Mu transposase: implications for control of the transposition pathway. *Genes Dev* 11, 1561-1572 (1997)

212. Nakai, H. and R. Kruklitis: Disassembly of the bacteriophage Mu transposase for the initiation of Mu DNA replication. *J Biol Chem* 270, 19591-19598 (1995)

213. Kruklitis, R., D. J. Welty and H. Nakai: ClpX protein of Escherichia coli activates bacteriophage Mu transposase in the strand transfer complex for initiation of Mu DNA synthesis. *EMBO J* 15, 935-944 (1996)

214. Shi, Y: Mammalian RNAi for the masses. *Trends Genet* 19, 9-12 (2003)

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

216. Fassati, A., D. Gorlich, I. Harrison, L. Zaytseva and J. M. Mingot: Nuclear import of HIV-1 intracellular reverse transcription complexes is mediated by importin 7. *EMBO J* 22, 3675-3685 (2003)

217. Priet, S., J. M. Navarro, N. Gros, G. Querat and J. Sire: Functional role of HIV-1 virion-associated uracil DNA glycosylase 2 in the correction of G:U mispairs to G:C pairs. *J Biol Chem* 278, 4566-4571 (2003)

Abbreviations: IN: integrase; HIV-1: human immunodeficiency virus type 1; LTR: long terminal repeat; *att*: attachment; NHEJ: non-homologous end joining;

MoMLV: Moloney murine leukemia virus; PIC: preintegration complex; BAF: barrier-to-autointegration factor; NTD: N-terminal domain; CCD: catalytic core domain; CTD: C-terminal domain; MM-PCR: Mu-mediated PCR; HMG: high mobility group; INI1: integrase interactor 1; LEDGF: lens epithelium-derived growth factor; EED: embryonic ectoderm development; HSP: heat shock protein; LEM: LAP2 (lamin-associated polypeptide 2), emerin, MAN1; INM: inner nuclear membrane; Crx: cone-rod homeobox; HhH: helix-hairpin-helix; MA: matrix; RNAi: RNA interference; GST: glutathione S-transferase; Rpt: repeat; HR3: homology region 3; siRNA: short-interfering RNA; shRNA: short-hairpin RNA; NLS: nuclear localization signal; HTH: helix-turn-helix; PcG: *Polycomb* group; EF: elongation factor; MALDI-reTOF: matrix-assisted laser-desorption / ionization reflectron time-of-flight; MS: mass spectroscopy; NDR: nuclear Dbf2-related; DNA-PK: DNA-dependent protein kinase: *scid*: severe combined immunodeficiency; Ab-MLV: Abelson murine leukemia virus; MEF: mouse embryo fibroblast; ATM: ataxia-telangiectasia mutated; ATR: ATM- and Rad3-related; PARP-1: poly(ADP-ribose) polymerase 1; 3-MB: 3-methoxybenzamide; UNG2: uracil deglycosylase 2

Key Words: HIV-1, Integration, Integrase, Integration cofactor, DNA recombination, Review

Send correspondence to: Dr Alan Engelman, Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney St., Boston, MA 02115, Tel: 617-632-4361, Fax: 617-632-3113, E-mail: alan_engelman@dfci.harvard.edu