

Retroviral DNA Integration

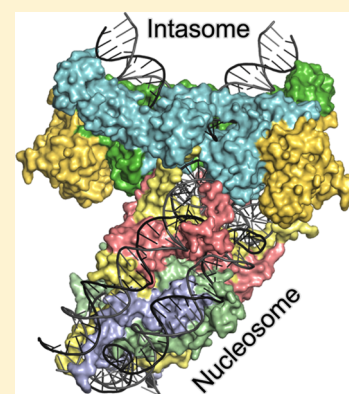
Paul Lesbats,[†] Alan N. Engelman,^{*,‡} and Peter Cherepanov^{*,†,§}

[†]Clare Hall Laboratories, The Francis Crick Institute, Blanche Lane, South Mimms, EN6 3LD, U.K.

[‡]Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute and Department of Medicine, Harvard Medical School, 450 Brookline Avenue, Boston, Massachusetts 02215 United States

[§]Imperial College London, St-Mary's Campus, Norfolk Place, London, W2 1PG, U.K.

ABSTRACT: The integration of a DNA copy of the viral RNA genome into host chromatin is the defining step of retroviral replication. This enzymatic process is catalyzed by the virus-encoded integrase protein, which is conserved among retroviruses and LTR-retrotransposons. Retroviral integration proceeds via two integrase activities: 3'-processing of the viral DNA ends, followed by the strand transfer of the processed ends into host cell chromosomal DNA. Herein we review the molecular mechanism of retroviral DNA integration, with an emphasis on reaction chemistries and architectures of the nucleoprotein complexes involved. We additionally discuss the latest advances on anti-integrase drug development for the treatment of AIDS and the utility of integrating retroviral vectors in gene therapy applications.



CONTENTS

1. Introduction	12730
2. IN and the Retroviral Life Cycle	12731
3. Enzymatic Steps in Retroviral DNA Integration	12731
3.1. Reactions Catalyzed by IN	12731
3.2. Postintegration DNA Repair	12733
4. Structure of Retroviral IN	12733
4.1. Domain Organization	12733
4.2. Architecture of the PFV Intasome	12735
4.3. Engagement of Target DNA by the PFV Intasome	12735
4.4. Mechanics of the IN Active Site	12736
4.5. ASLV and MMTV Intasome Structures	12738
5. Integration Site Selection	12738
5.1. Retroviral Integration Is Not Random with Respect to the Target Genome	12738
5.2. The Nexus between Lentiviruses and LEDGF/p75	12739
5.3. γ -Retroviruses and BET Proteins	12740
5.4. Integration Site Selection by Other Retroviruses and LTR Retrotransposons	12741
6. HIV-1 IN as a Target for Drug Development	12741
6.1. HIV-1 IN Strand Transfer Inhibitors (INSTIs)	12741
6.2. Viral Resistance to INSTIs	12743
6.3. Emerging Allosteric Inhibitors of HIV-1 IN	12743
7. Retroviral Integration as a Therapeutic Tool	12744
8. Perspectives	12745
Author Information	12745
Corresponding Authors	12745
Author Contributions	12745
Notes	12745

Biographies	12745
Acknowledgments	12746
Abbreviations Used	12746
References	12746

1. INTRODUCTION

More than 100 years ago, the Danes V. Ellerman and O. Bang and the American P. Rous passaged oncogenic variants of what is known today as the avian sarcoma-leukosis virus (ASLV).^{1,2} The importance of retroviruses in biology and medicine has greatly increased over the past 50 years with two major milestones: first, in 1981, the isolation of the human T-lymphotropic virus 1 (HTLV-1)³ and, soon after, the discovery of human immunodeficiency virus type 1 (HIV-1),^{4,5} which is responsible for one of the most dramatic pandemics in recent history. The flurry of high-octane research, initially driven by the suspected role of retroviruses in human cancer and later by the acquired immunodeficiency syndrome (AIDS) pandemic, yielded a plethora of discoveries and tools to bolster all disciplines of biology.⁶ It would be hard to imagine cancer biology without the concept of the oncogene or molecular biology without reverse transcriptase (RT).

Retroviridae is a large viral family comprising seven genera: α -through ϵ -retroviruses, lentivirus, and spumavirus (Table 1). HTLV-1 and HIV-1 (along with their respective types) belong to δ -retrovirus and lentivirus genera, respectively. Several other

Special Issue: Genome Modifying Mechanisms

Received: February 18, 2016

Published: May 20, 2016

Table 1. Classification of Retroviruses

subfamily	genus	species examples	duplication size (bp) ^a
orthoretrovirinae	α -retrovirus	avian sarcoma-leukosis virus (ASLV)	6
		mouse mammary tumor virus (MMTV)	6
	β -retrovirus	jaagsiekte sheep retrovirus (JSRV)	
		human endogenous retrovirus K (HERV-K) ^b	
		murine leukemia virus (MLV)	4 or 5
	γ -retrovirus	feline leukemia virus (FeLV)	
		reticuloendotheliosis virus strain A (RevA)	
	δ -retrovirus	human T-lymphotropic viruses 1 and 2 (HTLV-1,2)	6
		bovine leukemia virus (BLV)	
	ϵ -retrovirus	walleye dermal sarcoma virus (WDSV)	?
lentivirus	human immunodeficiency viruses 1 and 2 (HIV-1,2)		5
		maedi-visna virus (MVV)	
	equine infectious anemia virus (EIAV)		
	feline immunodeficiency virus (FIV)		
spumaretrovirinae	spumavirus	prototype foamy virus (PFV)	4
		bovine foamy virus (BFV)	
		feline foamy virus (FFV)	

^aSize of target DNA duplications flanking integrated proviruses.

^bExtinct as exogenous species.

retroviral species gained prominence as research models, for historical reasons or as animal pathogens. These include ASLV (an α -retrovirus), mouse mammary tumor virus (MMTV, a β -retrovirus), murine leukemia virus (MLV, a γ -retrovirus), simian immunodeficiency viruses (SIVs, lentiviruses highly related to HIV-1 and HIV-2), feline immunodeficiency virus (FIV, a lentivirus), and the prototype foamy virus (PFV, a spumavirus). Integration, which yields the establishment of the obligatory proviral state,⁷ is the one feature that distinguishes retroviruses from all other viral families. Herein, we present state-of-the-art interpretations of the structure of retroviral integrase (IN), the essential enzyme responsible for this process, as well as the role of IN in virus replication. Due to the conservation among IN proteins from different retroviral species, we will refer to them collectively as retroviral IN, except when discussing aspects that may be relevant to a particular retroviral genus or species.

2. IN AND THE RETROVIRAL LIFE CYCLE

Replication via formation of a stable DNA form makes retroviruses particularly amenable to reverse genetics. Accordingly, functions of retroviral gene products have been extensively probed through mutagenesis. In early studies, IN was identified as the protein product encoded within the 3' portion of the retroviral *pol* gene that was essential for efficient retroviral replication and integration.^{8–11} Reverse transcription of the diploid retroviral RNA genome results in the formation of a linear double-stranded viral DNA (vDNA) molecule carrying a copy of the long terminal repeat (LTR) sequence at either end.^{12–15} The vDNA molecule exists in the form of a preintegration complex (PIC)^{16,17} that is rather poorly

biophysically characterized due to the scarce level at which it forms, ca. one copy per cell, during acute virus infection. Nevertheless, PICs have been reported to contain a number of cellular and viral proteins, most notably IN.^{18–26} Once the PIC gains access to the nuclear compartment, the vDNA ends are inserted into a cellular chromosome. This step, initiated by the enzymatic action of IN and completed by the host cell DNA repair machinery, is a point of no return: the cell becomes a permanent carrier of the integrated viral genome, which is referred to as the provirus.

In addition to this well-established role, IN may play a range of less characterized functions in retroviral replication, as suggested by its unusually complex genetics (reviewed in ref 27). For instance, disruption of the IN coding portion of the HIV-1 *pol* gene can lead to production of viral particles with aberrant morphology and severe defects in reverse transcription.^{28–31} In fact, only a minority of HIV-1 IN mutants display defects solely at the integration step of the viral life cycle. Such mutants, which include amino acid substitutions within the IN active site, were collectively categorized as class I mutants.³² The hallmark of the associated phenotype is the predictable accumulation of nonintegrated forms of vDNA, including a circular form that contains two abutted copies of the LTR (2-LTR circles). Conversely, class II HIV-1 IN mutants disrupt viral replication at multiple steps while usually retaining at least partial IN enzymatic activity in vitro.^{33–36} The pleiotropic effects observed with class II HIV-1 IN mutants range from disrupted virion assembly to apparent nuclear import defects.^{30,33,34,37–41} Most notably, class II IN mutants typically show reduced levels of reverse transcription.²⁷ The abundance of HIV-1 IN mutations with pleiotropic phenotypes is a strong indication that the protein may play critical roles in the viral lifecycle outside of the integration step. Accordingly, HIV-1 IN was shown to interact with the viral RT and influence its activity in vitro.^{42–44} More recent work with allosteric IN inhibitors (described at length below) has highlighted a direct role for IN in HIV-1 particle maturation.^{45–47} Among the esoteric functions of HIV-1 IN, its proposed involvement in PIC nuclear import has been the subject of considerable and yet to be resolved debate.^{34,40,41,48–54}

3. ENZYMATIC STEPS IN RETROVIRAL DNA INTEGRATION

3.1. Reactions Catalyzed by IN

Biochemical studies of retroviral DNA integration began with partial purification of enzymatically competent PICs from acutely infected cells. Such preparations can catalyze vDNA integration into exogenous DNA in vitro, and the reaction products can be detected and quantified by Southern blotting or PCR.^{16,55–58} During infection, a considerable fraction of vDNA becomes circularized^{59–64} and, perhaps owing to the circular nature of the bacteriophage λ DNA substrate for integrative recombination,⁶⁵ initial studies proposed that retroviral integration proceeded through the 2-LTR circular DNA form.⁶⁶ However, subsequent experiments using native MLV PICs demonstrated that it is the linear vDNA that serves as the immediate precursor for integration.^{67,68} These landmark studies also established two activities associated with the PICs—3'-processing and strand transfer^{67–70}—and retroviral IN was shown to be sufficient to catalyze each of these reactions in vitro.^{71–77}

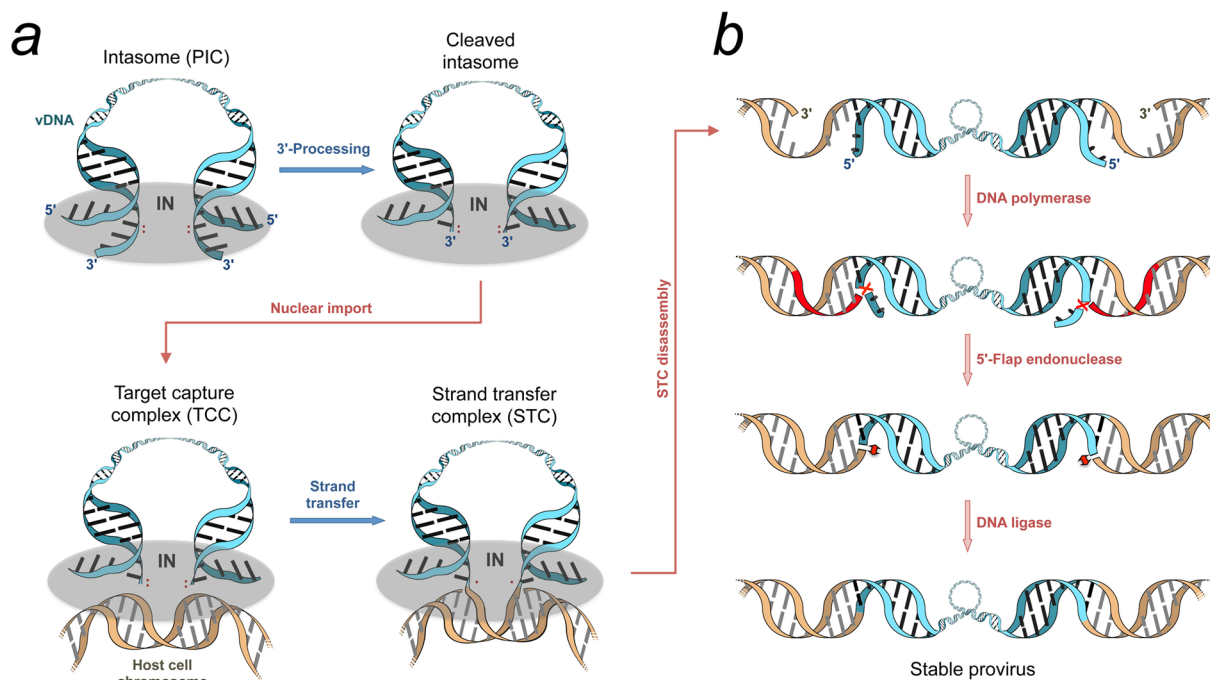


Figure 1. Retroviral DNA integration pathway: reactions catalyzed by IN (a, blue arrows) and host cell enzymes (b, red arrows). The intasome contains a multimer of IN (gray oval) assembled on vDNA ends. Following 3'-processing and nuclear entry, the cleaved intasome complex engages cellular chromosomal DNA, forming the TCC. Insertion of the 3' vDNA ends into host DNA results in formation of the STC with hemi-integrated vDNA. Formation of the stable provirus further requires disassembly of the STC and repair of the strand discontinuities flanking vDNA by the sequential actions of a DNA polymerase, 5'-flap endonuclease, and DNA ligase. Red dots represent magnesium ions in the intasome active sites.

Retroviral IN bears no similarities to its namesake from λ phage and is instead closely related to the DD(E/D) family of DNA transposases.⁷⁸ Crucially, the DNA cutting and strand transfer reactions catalyzed by IN and transposases proceed through phosphodiester transesterification, without formation of covalent protein–DNA intermediates.^{79–81} However, unlike transposases, retroviral IN requires a prelinearized DNA molecule—the product of reverse transcription—and cannot act on an integrated molecule that is flanked by continuous host DNA sequences. Therefore, while the active site of a prokaryotic cut-and-paste transposase must carry out four consecutive reactions,^{82,83} IN needs to accomplish only two, 3'-processing and strand transfer, equivalent to the first and the final steps in DNA transposition, respectively. These reactions are carried by a multimer of IN assembled on vDNA ends, referred to as the intasome (also known as the stable synaptic complex), at the business end of the PIC (Figure 1a).^{17,84,85} During 3'-processing IN hydrolyses a phosphodiester bond at either vDNA end, removing a di- or trinucleotide, liberating 3'-hydroxyl groups attached to invariant 5'-CA-3' dinucleotides. The intasome can then bind host chromosomal DNA, forming the target capture complex (TCC). Within the TCC, the enzyme utilizes vDNA 3'-hydroxyls as nucleophiles to cut host DNA in a staggered fashion, simultaneously joining both 3' vDNA ends to apposing strands of host DNA. The postcatalytic complex, referred to as the strand transfer complex (STC), is subject to disassembly, which is likely accomplished by host cell machinery.

The fact that 3'-processing of the vDNA can occur in the cell cytoplasm²¹ presents a potential problem for retroviruses. The high local concentration of vDNA in the confines of the PIC makes it a potential target for strand transfer, and products of viral self-integration—called autointegration—are readily de-

tectable in infected cells.^{86–88} It is obviously in the best interest of retroviruses to avoid autointegration, and various mechanisms have been identified. The barrier-to-autointegration factor (BAF), a small DNA-binding protein that possesses the unusual property to bridge and condense separate DNA molecules together,^{89,90} was identified via its ability to suppress the autointegration activity of MLV PICs in vitro.⁹¹ However, it has yet to be determined whether BAF provides this function during MLV infection, and the lentiviruses seem to utilize other protective measures. The SET complex, which harbors a variety of DNA metabolizing enzymes, can suppress autointegration during HIV-1 infection,⁹² while the viral capsid (CA) protein has been shown to play a role in regulating SIV autointegration.⁸⁸ Additional work should help to clarify whether there might be a universal mechanism or if, indeed, different viral species have evolved unique ways to protect themselves from suicidal integration as they move through the cell toward their preferred chromosomal DNA targets.

Recombinant IN proteins are generally proficient in supporting robust 3'-processing and strand transfer activities on short double-stranded oligonucleotide substrates that represent the U5 or U3 vDNA ends at the tips of the upstream and downstream LTRs, respectively.^{71–74,76,77,93} However, depending on the viral source and the method of enzyme preparation, unpaired insertions of vDNA ends into target DNA often account for the bulk of the strand transfer products formed. HIV-1 IN is particularly prone to such aberrant strand transfer events (referred to as half-site integration). Carefully optimized HIV-1 IN in vitro assays developed over the past 15 years greatly improved the efficiency of paired integration of vDNA ends (i.e., full-site or concerted integration).^{94–101} Conversely, some recombinant IN proteins, in particular PFV IN, are far more proficient at full-site strand transfer.^{102–104}

The underlying reasons for these differences between divergent IN proteins are not fully understood. The propensity of HIV-1 IN to self-associate into higher-order multimers in solution¹⁰⁵ could be a factor limiting its concerted integration activity in vitro. A third type of reaction that IN can catalyze in vitro is disintegration, representing a reversal of strand transfer.¹⁰⁶ Although this reaction is unlikely to be relevant in vivo, disintegration assays were instrumental to help map the different IN protein domains and the enzyme active site.^{107–115} Having evolved to catalyze a single reaction cycle, retroviral IN and transposases do not effectively dissociate from their reaction products, requiring active assistance from their host cells to disassemble the STC.^{116,117} Because they do not turn over, these enzymes are used in stoichiometric and often superstoichiometric ratios to their DNA substrates.

3.2. Postintegration DNA Repair

Having executed concerted strand transfer to join both 3' vDNA ends to host DNA, IN leaves a pair of single-stranded gaps and short 5' overhangs flanking the vDNA, which are repaired by cellular enzymes (Figure 1b; reviewed in ref 118). Notwithstanding their importance, the final steps of integration are among the least studied aspects of retroviral replication. Because disorderly repair of these discontinuities may lead to disruption of proviral and chromosomal DNA integrity, the handover of the retroviral hemi-integrand to the host DNA repair machinery is likely an exquisitely choreographed process. Disassembly of the thermodynamically stable STC complex must be the first step toward repair of the vDNA–chromosome junctions. Active degradation of IN/transposase, which is well-established in the bacteriophage Mu transposition system,^{116,117} may allow host cellular proteins to gain access to the sites of repair. Consistent with this model, HIV-1 IN is subject to ubiquitination and proteasome-dependent degradation when it is ectopically expressed in human cells.^{119–121} RAD18, an E3 ubiquitin ligase involved in DNA gap repair, was shown to interact with HIV-1 IN.¹²² However, a follow up study observed that knockdown of RAD18 in target cells did not reduce HIV-1 infectivity.¹²³ Subsequently, von Hippel–Lindau binding protein 1, a cellular subunit of the prefoldin chaperone, was implicated in proteasome-mediated HIV-1 IN degradation and, moreover, was shown to be required for efficient viral replication.¹²⁴

Following STC disassembly, three separate DNA repair enzymatic activities are required to complete the integration process by joining 5' vDNA ends to chromosomal DNA: a DNA polymerase, a 5' flap endonuclease, and a ligase (Figure 1b). Early studies used biochemical approaches to test candidate DNA cellular enzymes in repair of retroviral integration products in vitro.^{125–128} Synthetic model DNA containing a gap and 5' overhang served as a template for testing candidate proteins. In these experiments a cocktail of host DNA repair proteins was challenged to polymerize across the gap, to cleave the two-nucleotide flap, and to ligate the resulting product. Base excision repair (BER) pathway enzymes DNA polymerase β , flap endonuclease 1 (FEN1), and ligase I were shown sufficient to repair the substrate.¹²⁵ DNA polymerase δ also supported repair in the presence of FEN1 and ligase I and was stimulated by its cofactor PCNA. Alternatively, viral RT and IN have been proposed to mediate postintegration repair.¹⁰⁶ In this model, RT polymerizes across the gap sequence, while 5' flap removal and ligation of both strands are catalyzed by IN disintegration activity. However,

RT- and IN-dependent gap repair would require a complicated juggling of the vDNA ends between the two viral enzymes and, more problematically, would involve assembly of an unusual complex with 5' vDNA and freshly extended 3' chromosomal DNA ends in the IN active site. Even though RT is capable of polymerizing through the gap in vitro, the reaction displayed poor fidelity.¹²⁵ Moreover, addition of IN to the reaction did not afford completion of the repair process.^{125,127} Although HIV-1 IN interacts with FEN-1¹²⁹ and can stimulate its activity in vitro,¹²⁶ convincing evidence for a role of IN in repair of the hemi-integrand during retroviral infection is lacking. Cell-based studies have by contrast highlighted a role for the BER pathway of oxidative DNA damage in lentiviral DNA integration.^{130,131} Other studies have described that the alteration of the nonhomologous end-joining (NHEJ) pathway affects retroviral infection and survival of the infected cells.^{22,132–134} While some have shown a role for active IN and, by extension, the vDNA hemi-integrand, in NHEJ activation,^{132,134} a separate study concluded that the linear vDNA substrate was cytotoxic to cells.²²

As a result of integration across the major groove in target DNA and the subsequent gap repair, the proviral DNA is flanked by a short duplication of the target DNA sequence. The duplication size is virus-specific and ranges from 4 bp for MLV and PFV to 6 bp for HTLV-1, ASLV, and MMTV, while HIV-1 and other characterized lentiviruses generate 5-bp duplications (Table 1).^{63,135–146}

4. STRUCTURE OF RETROVIRAL IN

4.1. Domain Organization

IN proteins contain three conserved structural domains common to all retroviral genera: the N-terminal domain (NTD), the catalytic core domain (CCD), and the C-terminal domain (CTD) (Figure 2a). Initially identified by limited proteolysis of HIV-1 IN,¹⁰⁸ the domains were characterized in isolation using X-ray crystallography and/or nuclear magnetic resonance (NMR) spectroscopy (Figure 2a).^{147–153} The NTD folds into a compact three-helical bundle stabilized by coordination of a Zn^{2+} ion by side chains of His and Cys residues comprising the invariant HHCC motif.^{147,150} Often misidentified as a “zinc finger domain”, the NTD is most closely related to helix–turn–helix DNA binding domains.¹⁵⁴ In spumaviral and ϵ - and γ -retroviral INs, the NTD is expanded by a ~ 40 amino acid residue NTD extension domain (NED).^{85,155} The CTD is the least conserved of the three canonical IN domains and features a Src homology 3 (SH3)-like β -barrel fold.^{149,151} Although most structurally related to the Tudor family of chromatin binding domains, the IN CTD lacks the conserved hydrophobic cage used by Tudor domains to bind methylated Lys and Arg residues. During integration, the NTD and the CTD make important interactions with DNA substrates and play critical structural roles within the intasome assemblies.^{85,156,157} The flexible linkers connecting these domains to the CCD lack sequence conservation and vary in length between retroviral genera.^{155,158}

The CCD harbors the active site of the enzyme composed of three invariant carboxylates comprising the signature D,D-35-E motif.¹⁵⁹ In isolation, the CCD can catalyze disintegration, but neither of the biologically relevant activities of full-length IN.^{112–114} Determined over 2 decades ago, the crystal structure of the HIV-1 IN CCD revealed a protein fold shared by a superfamily of nucleotidyltransferases, which notably includes

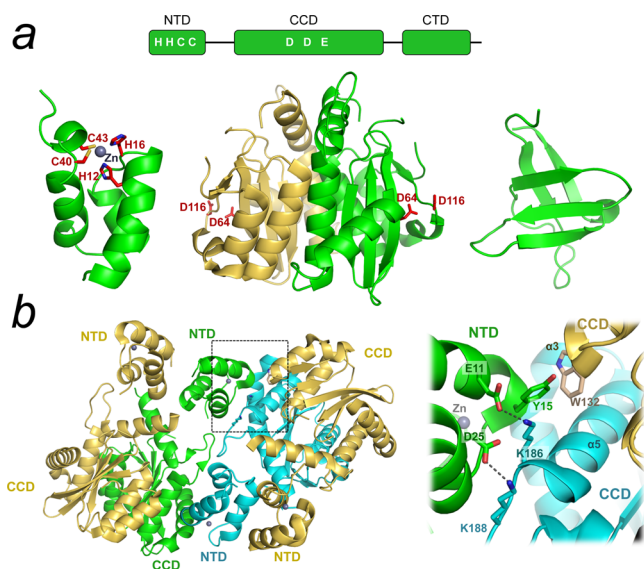


Figure 2. Domain organization of HIV-1 IN. (a) Schematic of the protein sequence with the NTD, CCD, and CTD shown as boxes (top). Structures of the individual domains determined in isolation (from left to right PDB IDs 1WJC, 1ITG, and 1IHV). Protein chains are shown as cartoons, with zinc-coordinating residues His12, His16, Cys40, and Cys43 and active site residues Asp64 and Asp116 shown as sticks with carbon atoms in red. (b) Structure of a two-domain HIV-1 IN construct (PDB ID 1K6Y). Details of the key NTD–CCD interface are shown as a blown up image to the right. Selected amino acid residues are indicated. Dashes represent hydrogen bonds; gray spheres are Zn^{2+} ions. All structural images in this review were prepared using PyMOL software (<http://www.pymol.org>).

bacterial and viral RNase H enzymes, Holliday junction resolvase RuvC, prokaryotic and eukaryotic DD(E/D) transposases, and RAG1 recombinase.^{148,160} This initial study and

the majority of subsequent IN crystal structures captured the CCD in a recurring dimeric form (Figure 2a). Given the wide separation of IN active sites in the CCD dimer, it became immediately clear that a higher-order IN multimer must be responsible for concerted integration of vDNA ends. Similarly, four copies of the mechanistically related Mu phage transposase assemble into its core synaptic complex, with only two of the four protomers contributing active sites for the catalysis of DNA cleavage and strand transfer.^{161–163} Of note, the HIV-1 virion packages as many as ~250 copies of IN,^{164,165} a large excess over the two active sites that are required to catalyze integration.

Over the last 20 years, considerable effort has been expended to establish the functional multimeric state of retroviral IN. The bulk of these studies has focused on HIV-1 IN, which was shown to form a variety of multimeric species.^{166–172} Unfortunately, it is not obvious how the tendency of IN to aggregate in solution relates to its active forms in complex with vDNA. Indeed, well-studied transposases tend to form functional multimers only when bound to their cognate DNA substrates,^{173–175} which is explained by the involvement of DNA in the synaptic interfaces.^{175–177} Due to perplexing difficulties in obtaining soluble intasome preparations, the functional state of HIV-1 IN has been particularly hard to address. The tetrameric form of HIV-1 IN has so far received the strongest experimental support,^{169,170} although evidence for octameric states was also reported.^{178,179} The breakthrough came from recent studies of nonlentiviral INs, which fortuitously allowed assembly of monodisperse preparations of the intasome complexes in vitro. Surprisingly, structural characterization of three such intasomes established that the functional multimeric states of IN proteins are not conserved among retroviral genera. Thus, whereas PFV IN is monomeric in solution and forms a tetramer within the intasome,^{85,103,180}

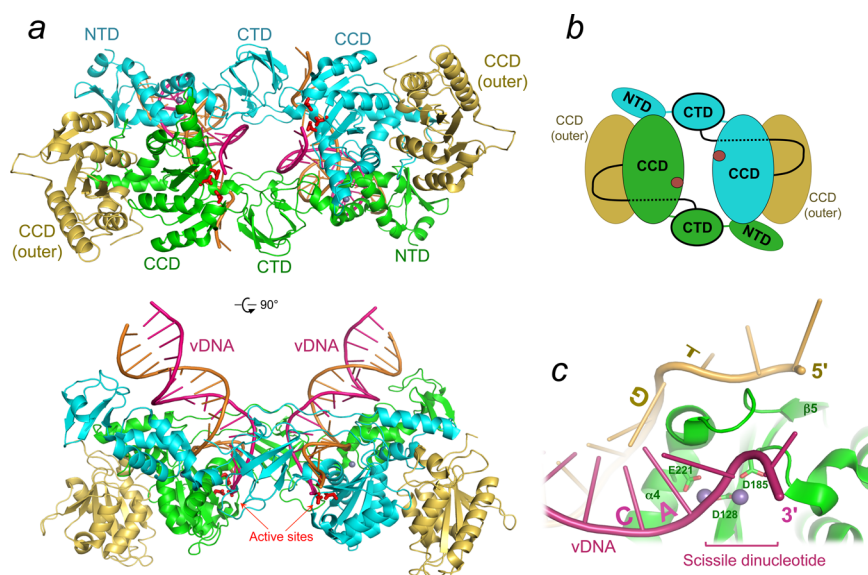


Figure 3. Structure of the PFV intasome. (a) Cleaved intasome complex (PDB ID 3L2Q) shown in two orthogonal views. Inner IN chains are colored green and cyan, and outer chains are yellow. Individual IN domains are indicated; invariant IN active site carboxylates are shown as red sticks; vDNA is shown as cartoons with transferred (i.e., the strand that IN acts upon) and nontransferred strands at each end shown in magenta and beige, respectively. (b) Schematic of the intasomal tetramer with IN domains shown as ovals and colored by chain; curved lines represent NTD–CCD (cyan and green) and CCD–CTD (black) linkers. Red circles represent inner IN chain active sites. (c) View on the IN active site prior to 3'-processing. Active site carboxylates and catalytic metal ions are respectively shown as sticks and gray spheres. The invariant vDNA CA nucleotides and their complements are letter-coded; the scissile 3' dinucleotide is indicated.

MMTV and ASLV INs are dimeric proteins that assemble into octamers on vDNA ends.^{155,181}

Initial hints for the structural organization of the different IN protomers within functional intasomes came from results of biochemical complementation studies.^{168,182–184} Mutant IN proteins that by themselves were defective for catalysis in vitro could recover significant levels of activity when studied as a mixture. In this way, IN NTD and CCD functions were shown to originate from different protomers within the multimeric complex. The first glimpse into the structural basis for IN multimerization beyond the canonical CCD dimerization was provided by the crystal structure of a two-domain construct spanning the NTD and the CCD of HIV-1 IN.¹⁸⁵ The structure revealed how pairs of IN dimers can interlock via formation of cross-dimer NTD–CCD interfaces (Figure 2b). Subsequently validated by site directed mutagenesis, the critical NTD–CCD interface, predicted from prior biochemical studies, became the second recurrent feature in structures of diverse INs and later intasomes.^{85,155,157,181,186} Further details on the partial structures of the preintosome era in retroviral IN structural biology can be found in recent reviews.^{158,187,188}

4.2. Architecture of the PFV Intasome

The first functional retroviral IN–DNA complex found to be amenable to structural characterization was the intasome from the spumavirus PFV.⁸⁵ Identified through comparative analysis of diverse orthologs, PFV IN was shown to be highly soluble and exceptionally active on short mimics of vDNA ends.^{103,189} The relatively simple architecture of the PFV intasome will provide a basis to describe the structures of the more complex α - and β -retroviral intasomes, which were characterized very recently.^{155,181} Although the U5 and U3 ends of retroviral LTRs are not identical, functional intasomes can generally be assembled with pairs of U5 sequences. Snapshots of such symmetrized PFV intasomes, visualized in the states prior to and after 3'-processing, as well as in complex with target host DNA prior to and after strand transfer, are now available.^{156,157,190}

The PFV intasome (Figure 3a) contains a tetramer of IN with a dimer-of-dimers architecture, composed of two structurally and functionally distinct IN subunits. The inner subunits (colored green and cyan in the figures) of each IN dimer are responsible for all interactions with vDNA, including provision of the active sites for catalysis. The outer IN subunits (shown in yellow) attach to the inner chains via the canonical CCD dimerization interface. The inner IN chains interact via forming intersubunit NTD–CCD contacts, and the two halves of the intasome are held together by the insertion of a pair of CTDs, which act as solid spacers between the CCD dimers (Figure 3a,b). The extended NTD–CCD and CCD–CTD linkers run parallel to each other, trussing the nucleoprotein assembly. The PFV intasome structures reported to date lack outer IN chain NEDs, NTDs, and CTDs, which were disordered in crystals and under cryo-electron microscopy (cryo-EM) conditions.^{85,191} Some clues about the average positions of the outer chain NTD and CTD could be gleaned from solution-state X-ray and neutron scattering.¹⁸⁰ However, these domains are dispensable for PFV intasome assembly and strand transfer activity in vitro, and their functions await clarification.¹⁹²

The vDNA ends enter the IN dimer–dimer interface, providing both 3' termini to the active sites of the inner IN subunits. The CCD, CTD, and NED of both inner IN chains

make intimate interactions with the vDNA backbone and bases, consistent with vDNA sequence specificity of retroviral IN.⁸⁵ Three bases of each 5' vDNA end are unpaired, tunneling through the CCD–CTD interface to protrude from the sides of the intasome structure (Figure 3c). Conceivably, such a protein–DNA complex can only form when the recognition sites are available on free vDNA ends, which makes the integration process fundamentally irreversible once post-integration repair is completed.

4.3. Engagement of Target DNA by the PFV Intasome

The PFV intasome assembled with bunt-ended vDNA ends readily undergoes 3'-processing in the presence of Mg^{2+} or Mn^{2+} ions.¹⁵⁶ Following dissociation of the cleaved 3'-dinucleotide, the saddle-shaped groove between the two halves of the intasome becomes available for host cell (target) DNA binding (Figure 4). The interaction primarily involves the

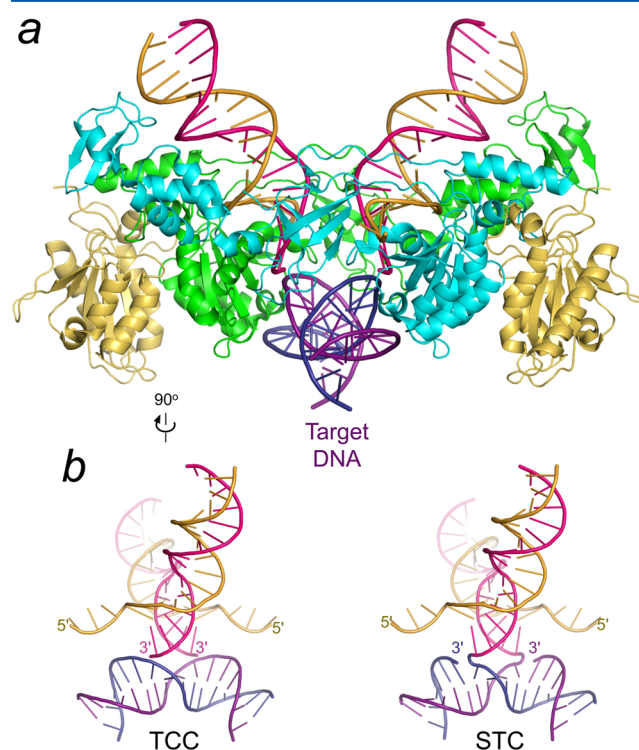


Figure 4. Target DNA capture by the PFV intasome. (a) Crystal structure of the PFV TCC (PDB ID 3OS1). (b) Conformation of vDNA and target DNA in the TCC (left) and STC (right, PDB ID 3OS0), with protein chains hidden for clarity.

target DNA backbone, which explains only weak preferences of IN for integration target site sequence.^{156,157,193} Target DNA binds to the PFV intasome in a sharply bent conformation, such that the base pair step at the center of the integration site unstacks with a 60° roll (Figure 4).¹⁵⁷ The associated expansion of the major groove to over 26 Å allows the widely spaced intasome active sites to align with a pair of phosphodiester separated by 4 bp on opposing strands in target DNA. The ability of DNA to form a sharp kink contributes to integration site selection at the level of nucleotide sequence, leading to a bias against rigid purine–pyrimidine base pair steps in the central positions of PFV integration sites.^{157,194} The energy of DNA deformation is offset by the interactions with the synaptic CTDs and the inner IN CCDs of the PFV intasome. Predictably, mutations affecting these contacts can lead to

increased bias for a bendable target DNA sequence.¹⁵⁷ Conversely, prebent and distorted DNA performs better as a target for retroviral integration.^{195,196}

The PFV intasome displays robust preference for chromatinized targets compared to naked DNA *in vitro*.^{191,197} These observations agree with early studies that, albeit using poorly defined IN–vDNA complexes, reported preferences of HIV-1 and MLV for nucleosomes.^{195,196,198–200} Moreover, nucleotide sequence periodicities observed in the immediate vicinity of retroviral integration sites are consistent with nucleosomes serving as targets for retroviral integration *in vivo*.^{194,201,202} Given the nonuniform availability of the major groove and the structure of the underlying histone octamer,²⁰³ it is not surprising that integration sites cluster in sharply defined hotspots along nucleosomal DNA. Recently, the PFV TCC containing the intasome and a nucleosome core particle was characterized by cryo-EM at 8 Å resolution (Figure 5a).¹⁹¹ The

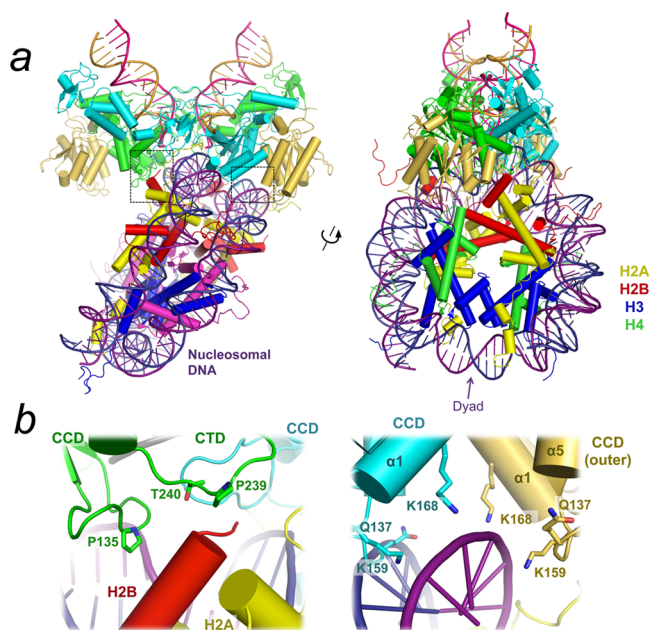


Figure 5. PFV intasome engaged with a mononucleosome. (a) Pseudoatomic model assembled by rigid body docking of the PFV intasome (PDB ID 3L2Q) and nucleosome (PDB ID 1KX5) structures into the cryo-EM map of the complex (EMDB ID 2992) and shown in two orientations. Individual histones are color coded as indicated. (b) Blow up of the regions boxed in the left side of panel a showing IN contacts with H2A and H2B histones (left) and with the second gyre of the nucleosomal DNA (right).

structure revealed that nucleosomal DNA engaged in the target DNA binding groove of the intasome is lifted from the surface of the histone octamer to assume the distorted conformation compatible with strand transfer. Outside of the target DNA binding groove, the intasome makes supporting contacts with one H2A–H2B heterodimer and the second gyre of the nucleosomal DNA (Figure 5b). These interactions presumably compensate for the energy of stretching and deforming histone-wrapped DNA and explain the strong preference of PFV to target superhelix ± 3.5 locations on nucleosomal DNA.

The strand transfer reactions occurring within the TCC that result in the joining of the 3' vDNA ends to the host DNA mark the end of IN catalytic function. Like 3'-processing, strand transfer does not lead to global conformational rearrangements

within the intasome structure.¹⁵⁶ It remains a mystery if and how IN signals completion of strand transfer for recruitment of STC disassembly and DNA repair machineries. Given the apparent preference of retroviruses to integrate into nucleosomes,^{191,194,197} chromatin remodellers appear to be good candidates for the job of STC disassembly.

4.4. Mechanics of the IN Active Site

Although the first CCD structures were determined in the 1990s, the functional organization of the IN active site was revealed only in the context of the functional PFV IN–DNA complexes. Dubbed “the active site loop” in the early literature,²⁰⁴ the IN residues connecting $\beta 5$ and $\alpha 4$ of the CCD, often including the essential Glu of the D,D-35-E motif, are disordered in most partial IN crystal structures. As expected, the region folds into a defined structure through interactions with the vDNA end (Figure 3c). Given a high degree of amino acid conservation within the IN active site, the lessons learned from the high-resolution PFV structures should be generally applicable. Accordingly, the PFV intasome was successfully used as a model to study active site inhibitors of HIV-1 IN.^{85,205,206}

The active sites of retroviral IN and the related DD(E/D) transposases contain three carboxylates, which serve to coordinate a pair of essential divalent metal cofactors (Figure 6). Although both Mg^{2+} and Mn^{2+} are capable of supporting IN activities, due to its greater abundance, the former metal ion is believed to be primarily utilized *in vivo*.^{207,208} The general mechanism of two- Mg^{2+} -ion catalysis at a phosphodiester bond, initially proposed in the early 1990s,^{209,210} has been corroborated by a growing body of experimental and theoretical work, including recent studies of *Bacillus halodurans* RNase H.^{208,211–213} The primary coordination spheres of the Mg^{2+} ions include essential active site carboxylates, the substrate phosphodiester, and the attacking nucleophile (a water molecule in the case of RNase H). The strong preference of Mg^{2+} for octahedral coordination enforces precise relative positioning of the reactants and aids in destabilizing the target phosphodiester group. During catalysis, the ions act as Lewis acids, with metal A assisting in deprotonation of the nucleophile and metal B neutralizing the negative charge developing on the phosphorane intermediate. The ability of Mn^{2+} to replace Mg^{2+} ion as a cofactor is explained by its similar size, pK_a , and coordination geometry. Flavors of the two- Mg^{2+} -ion mechanism are employed by a wide range of structurally and functionally diverse enzymes, which also includes protein kinases,^{214–216} DNA and RNA polymerases,^{214,217,218} and even ribozymes.²¹⁹

Due to the absolute dependence of the retroviral IN active site on its metal cofactors, it was possible to grow crystals of wild type (WT) PFV IN engaged with its DNA substrates. When exposed to Mg^{2+} or Mn^{2+} salts, such apo forms of the intasome and the TCC readily undergo 3'-processing and strand transfer “*in crystallo*”. Soaking the crystals in the presence of Mn^{2+} for very short periods of time (sufficient for diffusion of the salt through the crystal lattice but not catalysis) allowed freeze trapping the fully engaged configurations of the PFV nucleoprotein complexes in their ground states prior to 3'-processing and strand transfer.¹⁵⁶

For consistency with the RNase H literature, the ion cofactors coordinated to PFV IN residues Asp185 and Glu221 are designated metal A and metal B, respectively (Figure 6). The third carboxylate, Asp128, and a nonbridging

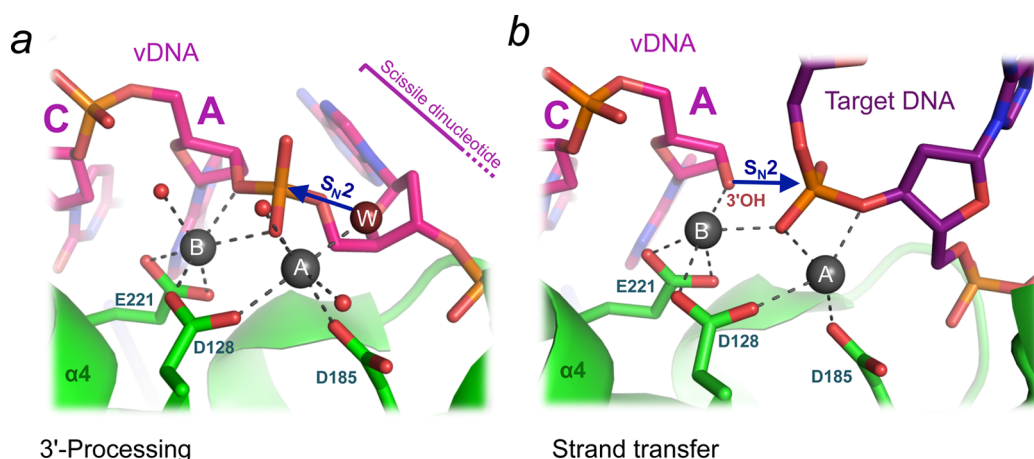


Figure 6. Configurations of the IN active site leading to 3'-processing (a, PDB ID 4E7I) and strand transfer (b, PDB ID 4E7K). Direction of nucleophilic attack by a water molecule (W, panel a) or 3'-hydroxyl of vDNA (panel b) is indicated with blue arrows marked S_N2 . Red spheres represent water molecules; lower TCC X-ray data resolution did not allow refinement of water molecules in panel b.

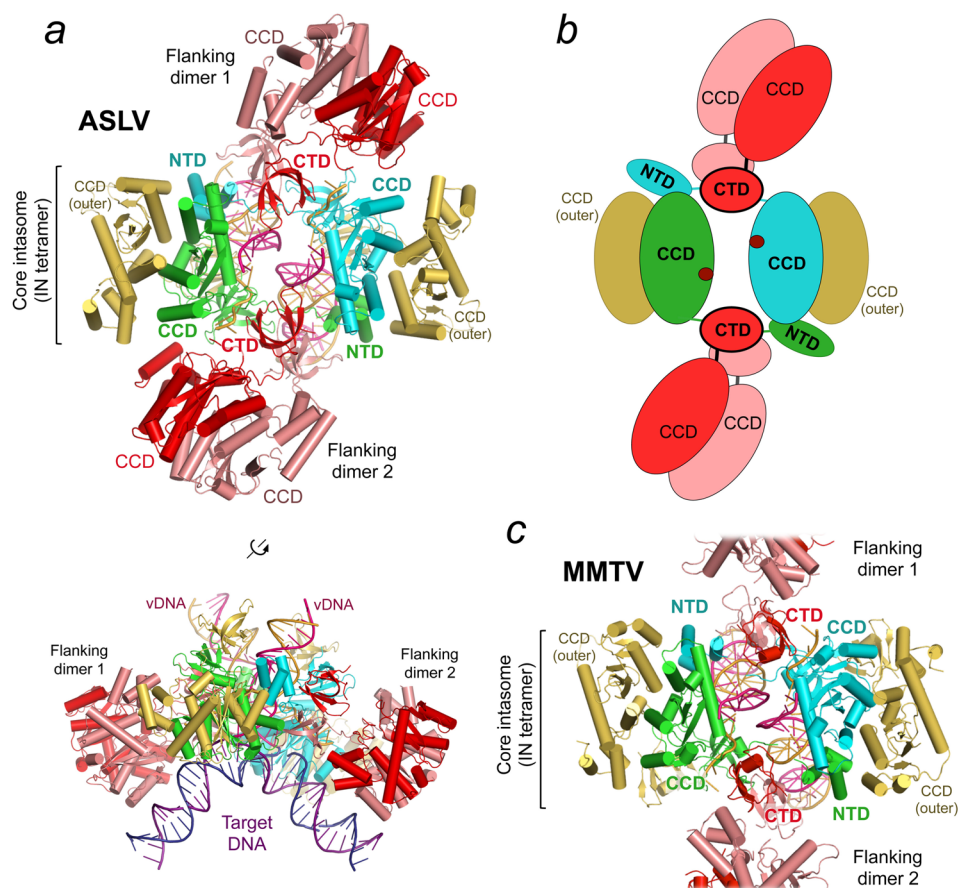


Figure 7. Architecture of the α and β -retroviral intasomes. (a) Crystal structure of the ASLV STC (PDB ID 5EJK) in two orientations. Target DNA is hidden in the top panel. The core IN tetramer containing inner and outer IN subunits and a pair of synaptic CTDs from the flanking IN chains is indicated. (b) Schematic of the intasomal IN octamer with IN domains shown as ovals and colored by chain. For clarity, CTDs and NTDs belonging to the outer IN subunits and NTDs from the flanking subunits present in the structures are not indicated in panel a or shown in panel c. (c) Pseudo-atomic model of MMTV intasome based on cryo-EM data (PDB ID 3JCA and EMDB ID 6441).

oxygen atom from the scissile phosphodiester are shared between both ions. In the pre-3'-processing state, the octahedral coordination sphere of metal A is nearly perfect and includes a water molecule positioned for an in-line nucleophilic attack at the phosphorus atom of the scissile phosphodiester (Figure 6a). Due to simultaneous bidentate

interactions with Glu221 and the phosphodiester, the environment of metal B cannot assume octahedral coordination. This departure from the preferred coordination geometry of metal B is thought to provide a destabilizing potential on the scissile phosphodiester.^{208,212} Following 3'-processing and dissociation of the cleaved dinucleotide, metal B remains coordinated to the

3' oxygen atom of the processed vDNA end. Consequently, it befalls metal B to activate the nucleophilic 3' hydroxyl of vDNA during strand transfer (Figure 6b). Thus, retroviral IN uses the inherent symmetry of the two-Mg²⁺-ion mechanism to carry out the consecutive reactions of hydrolysis and transesterification. Because the latter step does not change the number of high-energy bonds, strand transfer should in principle be reversible, at least for as long as the STC persists. However, such unproductive reversal of strand transfer is prevented by relocation of the newly formed phosphodiester out of the active site.¹⁵⁷ This reconfiguration is likely driven by conformational strain due to target DNA deformation, which is a conserved feature of retroviral intasomes and DD(E/D) transposases.^{157,173,177,220–222}

4.5. ASLV and MMTV Intasome Structures

In the PFV intasome, a pair of CTDs from the inner IN protomers are inserted in the dimer–dimer interface. The synaptic CTDs provide rigidity to the assembly and contribute to the host DNA binding platform. Crucially, this architecture depends on an extended polypeptide linker to track the linear distance of ~50 Å that separates the carboxyl terminal boundary of the inner chain CCD to the beginning of the CTD (depicted as a black curve in Figure 3b). Intriguingly, the length of the CCD–CTD linker is not conserved among retroviral IN proteins, ranging from 50 to 60 amino acid residues in γ - and ϵ -retroviruses and spumaviruses, to less than 10 residues in α - and β -retroviruses, whereas lentiviruses and δ -retroviruses possess intermediate-size linkers.^{155,158} Modeling suggested that the lentiviral IN CCD–CTD linker may stretch sufficiently to allow formation of the tetrameric intasomal architecture similar to that in PFV.^{205,223} However, this scenario is clearly not conceivable in the cases of α - and β -retroviral IN proteins.

Recently, the STC from ASLV and the intasome from MMTV were characterized by X-ray crystallography at 3.8 Å and cryo-EM at ~5 Å resolution, respectively.^{155,181} Unexpectedly, the structures revealed that the α - and β -retroviruses maintain a PFV-like core intasomal structure by employing additional IN dimers (referred to as flanking dimers) to source the pair of synaptic CTDs (Figure 7a,b). Strikingly, these intasomes contain homo-octamers of IN, each with four structurally and functionally distinct types of subunits. The intasomal core structure is further decorated by NTDs and CTDs belonging to the eight IN subunits. Locations of six CTDs, including the four provided by the flanking dimers, are conserved between the ASLV and MMTV intasome structures, and these make direct contacts with vDNA. The CCDs of the flanking IN dimers are considerably less defined in the cryo-EM structure, and their positions differ between the ASLV and MMTV intasomes (Figure 7a,c). Intriguingly, their locations are consistent with their potential roles in interactions with host DNA. Indeed, contacts between the flanking CCDs and the backbone of the target DNA are observed in the ASLV STC (Figure 7a, bottom panel).¹⁸¹ Results of biochemical complementation experiments moreover indicate that the flanking IN dimers are necessary for MMTV IN strand transfer activity.¹⁵⁵

5. INTEGRATION SITE SELECTION

5.1. Retroviral Integration Is Not Random with Respect to the Target Genome

Depending on the genomic location, the local chromatin environment of the provirus may be conducive to active viral expression or transcriptional silencing. Thus, the choice of integration site influences the level of ongoing viral replication and may contribute to the establishment of latent viral reservoirs.^{224–228} In fact, the propensity of HIV-1 to establish chronic and hitherto incurable infection is the direct consequence of its ability to establish latent reservoirs.^{225,229,230}

While the interaction of IN with vDNA ends is nucleotide-sequence- and structure-specific, the enzyme displays very little selectivity with regard to the host DNA. Alignments of retroviral integration sites revealed weak, virus-specific palindromic sequence consensi that do not extend farther than several bp from the integration site.^{193,194,231–233} These weak nucleotide sequence preferences are in part explained by the sparse interactions between IN and target DNA bases.^{157,194} However, far more interesting patterns emerge when the distributions of retroviral integration sites are scrutinized on the genomic scale.

Early research conducted in the 1970s and 1980s indicated that MLV integration may be associated with sites of DNase I hypersensitivity in the host cell genome.^{234–236} The availability of the draft human genome sequence^{237,238} and the relative ease of recovering vDNA–chromosomal junctions using PCR allowed Bushman and colleagues to address the distribution of HIV-1 integration sites in their landmark 2002 study.²³⁹ The field was subsequently bolstered by the advent of deep sequencing, which now permits recovery of hundreds of thousands of unique integration sites in a single infection experiment.^{201,240–242} It has emerged that retroviruses display distinct and contrasting preferences for various host cell genomic features (reviewed in ref 243). Thus, HIV-1 and other lentiviruses display strong preferences for transcription units with a sharp bias toward highly expressed and intron-rich genes.^{201,239,240,244} MLV and other γ -retroviruses strongly favor promoter regions and DNase I hypersensitive sites in general.^{194,235,236,241,242,245,246} In sharp contrast, the spumavirus PFV disfavors genes and loci of active transcription.^{191,247,248} It appears that the least selective retroviruses are from the α - and β -retrovirus genera, which show nearly random integration site distributions with respect to well-mapped genomic features.^{244,249,250}

Strong evidence implicated IN as a major determinant for integration site selection. Thus, implanting MLV IN into HIV-1 results in a chimeric virus with integration site preference which is closer to that of MLV.²⁴⁶ However, the same study found a subtler role for viral *gag* gene products, and more recent work has shown that amino acid substitutions in viral CA protein have considerable bearing on HIV-1 integration site distributions.^{246,251,252} A hallmark of lentiviruses is their ability to infect nondividing cells, with their PICs capable of traversing the nuclear envelope through the nuclear pore complex (NPC),^{253,254} while many other retroviruses need the nuclear envelope to break down during mitosis to access host chromatin.^{255–259} It seems plausible that CA may modulate HIV-1 PIC nuclear entry pathways, potentially via interactions with cleavage and polyadenylation specificity factor 6 (CPSF6) and/or nucleoporins (NUPs), which are the constitutive components of the NPC. HIV-1 CA has been shown to

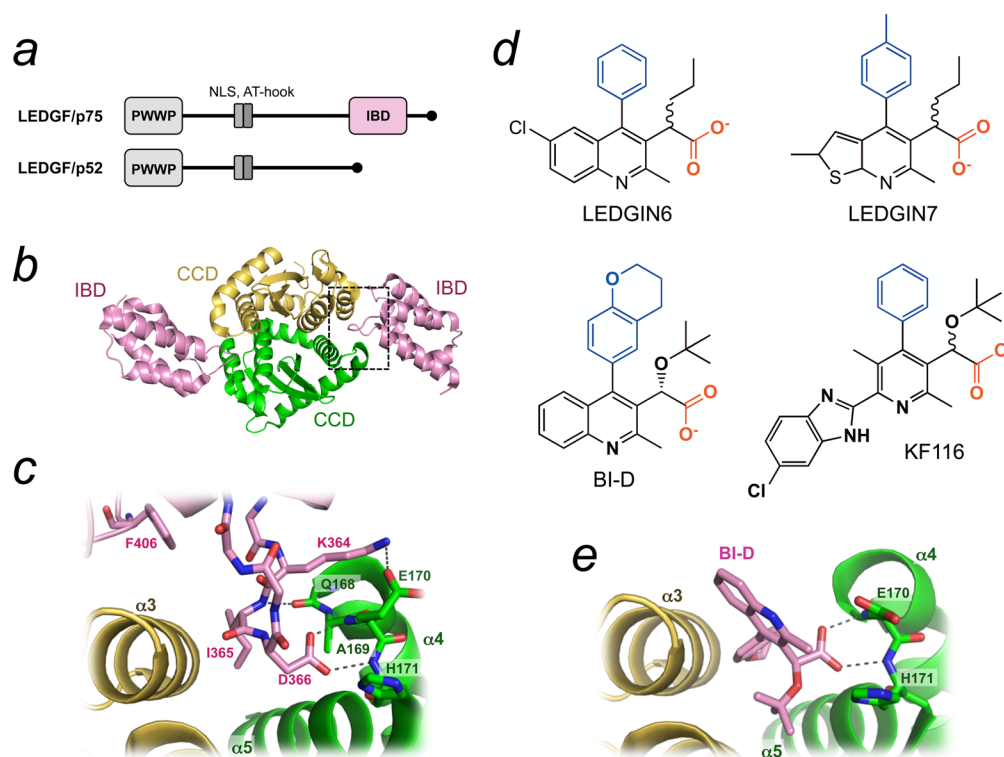


Figure 8. Interaction of LEDGF/p75 and ALLINIs with HIV-1 IN. (a) Schematic of LEDGF/p75 and p52 organization with NLS, AT-hooks, and structural domains (PWWP and IBD) shown as boxes. (b) Crystal structure of the HIV-1 IN CCD dimer in complex with LEDGF/p75 IBD (PDB ID 2B4J). (c) Details of the protein–protein interactions (rotated $\sim 90^\circ$ counterclockwise from the boxed region in panel b). (d) Chemical structures of selected ALLINIs; chemical groups mimicking LEDGF/p75 Asp366 and Ile365 are shown in red and blue, respectively. (e) Crystal structure of ALLINI BI-D in complex with HIV-1 IN CCD (PDB ID 4ID1), aligned with the panel c projection.

interact directly with CPSF6,^{260,261} NUP358,²⁶² and NUP153,²⁶³ and depletion of each of these factors significantly reduced the frequency of HIV-1 integration into gene-rich regions.^{251,262,264,265} The involvement of several NPC components in integration site selection is consistent with the observation that HIV-1 PICs preferentially target highly expressed genes in the nuclear periphery that are proximal to the nuclear pore.^{266,267} Conversely, HIV-1 integration is excluded from internal nuclear regions as well as from lamina-associated domains.²⁶⁷

5.2. The Nexus between Lentiviruses and LEDGF/p75

Whereas roles of retroviral structural proteins in PIC trafficking are only starting to transpire, the IN-dependent mechanisms of integration site selection are well-established. Lentiviruses and γ -retroviruses find their preferred genomic locations via recognition of specific chromatin-associated cellular proteins, which act as receptors or tethering factors for the PICs. Lens epithelium-derived growth factor (LEDGF) is a ubiquitous cellular chromatin-associated protein,²⁶⁸ initially described as transcriptional coactivator p75.^{269,270} Although a proposed extracellular function of the protein or specific roles in lens epithelium development were not corroborated, its misnomer has persisted in use. The protein was identified as a dominant HIV-1 IN binding partner in affinity-capture and yeast two-hybrid screening experiments^{171,271} and was later shown to interact with and stimulate the enzymatic activities of divergent lentiviral INs.^{102,272,273} LEDGF/p75 is composed of 530 amino acid residues and contains two small structured domains: an N-terminal PWWP domain and C-terminal IN binding domain (IBD) (Figure 8a).^{274,275} An alternative splice form, LEDGF/

p52,²⁶⁹ lacks the IBD (Figure 8a) and consequently neither interacts with IN nor has an effect on integration.²⁷⁶ The extended flexible regions of LEDGF/p75 harbor a classical importin α/β -dependent nuclear localization signal (NLS)²⁷⁷ and a pair of AT-hook motifs implicated in DNA binding.^{278,279} The PWWP domain belongs to the Tudor family and was shown to bind nucleosomes trimethylated on Lys36 of the histone 3 tail (H3K36me3), an epigenetic mark associated with transcription elongation and enriched within transcription units.^{280–282} Knockout of the gene encoding LEDGF does not affect cell proliferation in tissue culture but results in enhanced neonatal mortality and developmental abnormalities in mice.^{283,284} While its precise cellular functions remain to be elucidated, LEDGF/p75 reportedly interacts with a range of functionally diverse proteins, including the basal RNA polymerase cofactor PC2, mRNA splicing factors, multiple endocrine neoplasia type 1 protein product (menin), methyl CpG binding protein 2 (MeCP2), the end-resection protein CtIP, transcription factor JPO2, and the activating subunit of Cdc7 kinase.^{240,269,285–291} LEDGF/p75 was shown to recruit some of the binding partners to chromatin,^{276,285,290} suggesting that it may function as an adaptor protein in various chromatin-bound transactions.

Knockdown of LEDGF/p75 abolished the ability of ectopically expressed HIV-1 IN to bind chromatin, which provided the first hint about its role in lentiviral replication.^{24,276} However, the functional significance of the LEDGF/p75-IN interaction in vivo was initially unclear because the first attempts at LEDGF/p75 knockdown ostensibly failed to affect HIV-1 infectious titer and yielded only modest reductions in integration.^{24,292–294} Since infection of each new cell depends

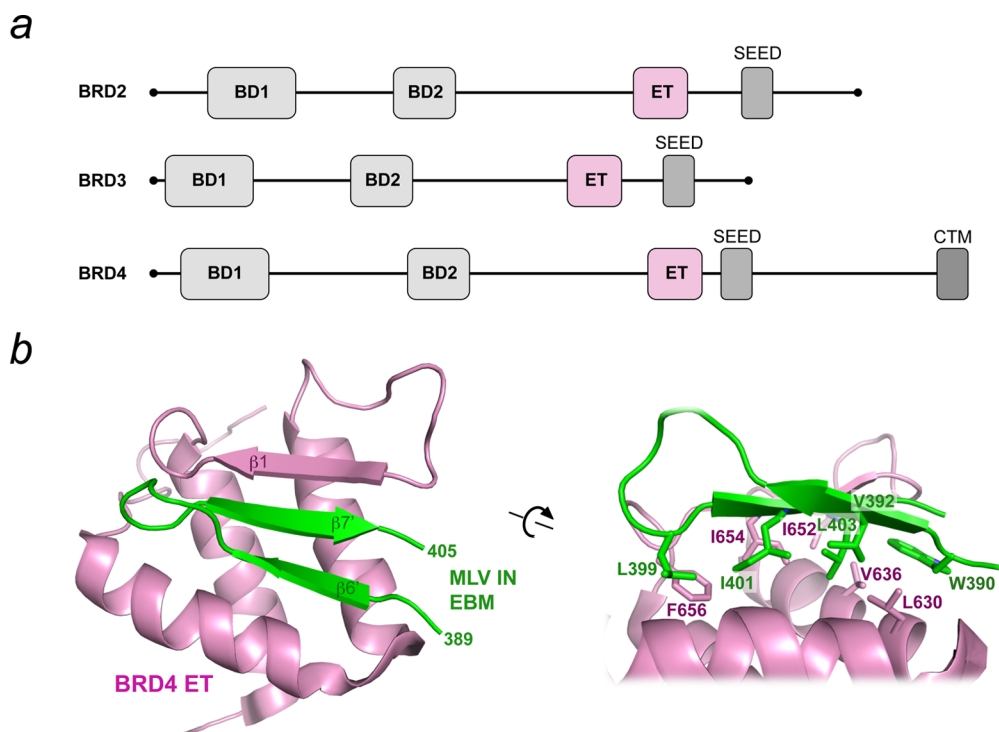


Figure 9. Interaction of BET proteins with MLV IN. (a) Schematic of domain compositions of BRD2, BRD3, and BRD4. (b) Solution structure of BRD4 ET domain in complex with the EBM of MLV IN (PDB ID 2N3K) shown in two orientations. BRD4 and IN residues participating in the hydrophobic core of the interface are shown as sticks and are indicated.

on a single-molecule integration event, even very limited levels of chromatin-associated LEDGF/p75 can be sufficient to support efficient viral replication.²⁹⁵ Clearer results came from infections conducted under conditions of intensified LEDGF/p75 knockdown and genetic knockout, which revealed considerable decreases of HIV-1 infectivity with the specific defect at the integration step.^{295–299}

The analysis of distribution of residual HIV-1 integration sites in LEDGF/p75-depleted cells demonstrated a dramatic loss of transcription unit targeting concomitant with a significant increase of integration near transcription start sites.^{296,299,300} The WT phenotype can be rescued upon restoration of LEDGF/p75 expression, confirming a role of the host factor in the targeting mechanism. Furthermore, by swapping the PWWP domain of LEDGF/p75 for alternative chromatin binding domains, it was possible to redirect HIV-1 integration toward chromatin regions bound by the heterologous tether.^{301–303} Collectively, the results support a model whereby LEDGF/p75 acts as a chromatin-bound tether that anchors the PIC by engaging its IN component in a direct protein–protein interaction. The results of a recent study that analyzed integration distribution patterns in cells knocked out for LEDGF/p75, CPSF6, or both factors clarified that the CA binding protein CPSF6 predominantly directs the HIV-1 PIC to actively transcribed euchromatin, where LEDGF/p75 determines positions of integration along gene bodies.²⁶⁵ LEDGF/p75 is additionally involved in HIV-1 latency by recruiting, post-integration, host factors IWS1 and SPT6 to the LTR, leading to the silencing of the provirus.²²⁶ Rapid silencing of the integrated viral genome is frequently observed during HIV-1 infection in proliferating CD4+ T cells.^{304–306} By interacting with LEDGF/p75, HIV-1 integration and transcriptional silencing may thus be coordinated to the establishment of long-lived viral reservoirs.²²⁶

The solution structure of the LEDGF/p75 IBD, determined using NMR, revealed a compact domain comprising a pair of HEAT repeatlike α -helical hairpins.³⁰⁷ On the virus side, the IN CCD is essential and minimally sufficient for the interaction with LEDGF/p75, whereas the NTD is required for high-affinity binding.²⁷⁶ X-ray crystallography was used to visualize atomic details of the virus–host interaction (Figure 8b,c).^{98,186,308} The hairpin loops at the tip of the elongated IBD structure contact the IN CCD dimer, with side chains of LEDGF/p75 residues Ile365 and Asp366 buried in a small pocket at the CCD dimerization interface, making hydrophobic interactions and a bifurcated hydrogen bond with the IN backbone, respectively. The protein–protein interaction is enhanced by intermolecular contacts involving Lys401, Lys402, and Arg405 on the basic side of the IBD and a conserved cluster of carboxylates on lentiviral IN NTDs.⁹⁸ Crucially, LEDGF/p75 contacts both the CCD and the NTD, which cooperate in higher-order IN multimerization and intasome assembly. It is not surprising, therefore, that the host factor enhances tetramerization and the strand transfer activity of HIV-1 IN *in vitro*.^{98,186,309} The pocket on the surface of the HIV-1 IN CCD involved in the interaction with the host factor has been targeted by small-molecule inhibitors, which in their binding mode to IN mimic the interactions made by LEDGF/p75 Ile365 and Asp366^{46,310–315} (Figure 8d,e; see section 6.3). Surprisingly, these small molecules potentially enhance IN multimerization,^{310,311,315–317} presumably without contacting the NTD.^{314,318}

5.3. γ -Retroviruses and BET Proteins

As LEDGF/p75 was shown to target lentiviral integration, it seemed possible that other retroviruses might similarly rely on genus-specific cellular factors to direct integration toward preferred genomic loci. Hence, it did not come as a surprise

when it was discovered that γ -retroviruses hijack cellular transcription factors as targeting factors.^{129,319–321} Highly related transcription factors BRD2, BRD3, and BRD4 belong to the bromodomain and extra-terminal (BET) family. These proteins play major roles in transcription regulation^{322,323} and had already been implicated in host–pathogen interactions.^{324–326} The characteristic features of BET members are two tandem bromodomains and a highly conserved extra-terminal (ET) domain within their N- and C-terminal regions, respectively (Figure 9a).²⁴³ Bromodomains belong to the well-studied group of chromatin readers with specificity for acetylated histone tails,³²⁷ while the ET domains were implicated in binding a range of cellular and viral proteins.³²³ In particular, BRD4 was shown to recruit P-TEFb to its target promoters to facilitate transcription elongation of cellular genes.³²⁸ Papilloma viruses tether their genomes to mitotic chromosomes via a direct interaction between the viral E2 protein and the C-terminal motif of BRD4, which allows stable segregation of vDNA copies between daughter cells.^{326,329} The latent nuclear antigen of Kaposi's sarcoma associated herpesvirus, essential for the viral episome maintenance and transcription, interacts with the ET domains of BET proteins.^{330,331}

Pull down experiments revealed a direct high-affinity interaction of γ -retroviral INs with the ET domains of BET proteins; moreover, the latter potently simulated IN strand transfer activity in vitro.^{194,319–321} Finally, the function of BET proteins in the context MLV replication was demonstrated using small-molecule inhibitors of the bromodomain–chromatin interaction as well as siRNA-mediated knockdown. The small molecules inhibited MLV but not HIV-1 integration in a dose-dependent manner.^{319–321} Notably, MLV integration sites correlate with binding sites of BET proteins as determined by chromatin immunoprecipitation studies.^{319,321,332} Treatment of cells with BET inhibitors or a siRNA cocktail specific to BRD2–4 mRNAs significantly reduced the preference of MLV to integrate near transcription start sites.^{319,320} As a complementary approach, LEDGF/p75-BRD4 hybrid proteins, containing the LEDGF/p75 PWWP and BRD4 ET domain, retargeted MLV integration toward active transcription units and away from transcription start sites,³²¹ confirming the major role of the BET proteins in γ -retroviral integration site selection.

Despite highly similar functional consequences, the structural bases for binding of lentiviral and γ -retroviral IN proteins to their cognate host factors are strikingly different. In contrast to lentiviruses, which engage the LEDGF/p75 IBD using a quaternary assembly of IN domains (minimally a CCD dimer), γ -retroviruses bind BET proteins using extended C-termini characteristic to the INs of this genus.^{321,333,334} The solution structure of the complex between the BRD4 ET domain and a conserved C-terminal ET binding motif (EBM) of MLV IN showed that the interaction involves the formation of an intermolecular three-stranded antiparallel β -sheet (Figure 9b).³³⁵ The folding of the interface only occurs upon binding of the two partners, as both the C-terminal tail of MLV IN and the BRD4 ET domain loop are unstructured on their own. The protein–protein interface contains a set of hydrophobic interactions involving buried side chains from both β 6 and β 7 of MLV IN and residues from helices α 1 and α 2 and the β 1 strand of the ET domain (Figure 9b). The interaction further depends on complementary electrostatics between the negatively charged amino acids from ET domain strand β 1

and the highly conserved positively charged residues of MLV C-terminal β 7. Mutation of the critical conserved amino acids showed a strong reduction of binding affinity as well as a shift of the integration pattern away from transcription start sites.^{320,334–336} Interestingly, BRD4 residues involved in the interaction with MLV IN were shown also to be important for binding its cognate cellular cofactors.³³⁵ Therefore, it seems likely that γ -retroviral IN evolved its C-terminal tail to mimic a cellular BET binding protein in order to optimize integration into transcriptionally active regions.

5.4. Integration Site Selection by Other Retroviruses and LTR Retrotransposons

The mechanisms of integration site selection employed by lentiviruses and γ -retroviruses present a remarkable case of convergent evolution, with both genera usurping cellular readers of the histone code to locate optimal target sites. Therefore, it is tempting to speculate that other retroviral genera may use similar strategies. It was recently reported that HTLV-1 and other δ -retroviral INs specifically interact with the B' protein phosphatase 2A (PP2A) regulatory subunits; moreover, recombinant B' proteins stimulated concerted integration activity of the δ -retroviral INs in vitro.³³⁷ Although not a classic chromatin binder, PP2A was implicated in dephosphorylating chromatin-resident targets.^{338,339} However, it remains to be determined whether PP2A is involved in directing δ -retroviral integration.

LTR retrotransposons, such as well-studied Ty and Tf elements from budding and fission yeasts, share most features of their replication cycle with retroviruses, although they complete it within the cell where they reside.³⁴⁰ The fundamental difference between a transposon and a virus is that the fate of the former wholly depends on the fitness of the host organism. Therefore, yeast LTR retrotransposons avoid instigating harmful insertional mutagenesis on the cell by precisely targeting new integration events to safe loci, and they achieve it by utilizing IN-binding host proteins. Thus, Ty5 retrotransposition is directed into transcriptionally silent regions of the *Saccharomyces cerevisiae* genome via the interaction between IN and the heterochromatin maintenance protein Sir4p.³⁴¹ Of note, the C-terminal peptide of Ty5 IN engages a patch on Sir4p, which is also recognized by the cellular interacting partner Esc1.^{342,343} Another *S. cerevisiae* retrotransposon, Ty1, engages the AC40 subunit of RNA polymerase III in a direct protein–protein interaction with IN, for specific integration upstream of RNA polymerase III transcribed genes.³⁴⁴

6. HIV-1 IN AS A TARGET FOR DRUG DEVELOPMENT

6.1. HIV-1 IN Strand Transfer Inhibitors (INSTIs)

Its essential role in viral replication and the lack of functional equivalents in human cells made IN an ideal target for anti-HIV/AIDS drug development. Intense interest and early screening efforts notwithstanding, the first class of small molecules capable of inhibiting HIV-1 replication by blocking IN was reported only in 2000.³⁴⁵ The key to identification of these molecules was the use of preassembled HIV-1 IN–vDNA complexes in screening assays.³⁴⁶ Empirical optimization of the original “diketo acid” pharmacophore led to the discovery of MK0518, now widely known as raltegravir (RAL),^{347,348} the first integrase strand transfer inhibitor (INSTI) to be approved for the treatment of AIDS in 2007. Two more antiretroviral drugs with an identical mode of action, elvitegravir (EVG) and

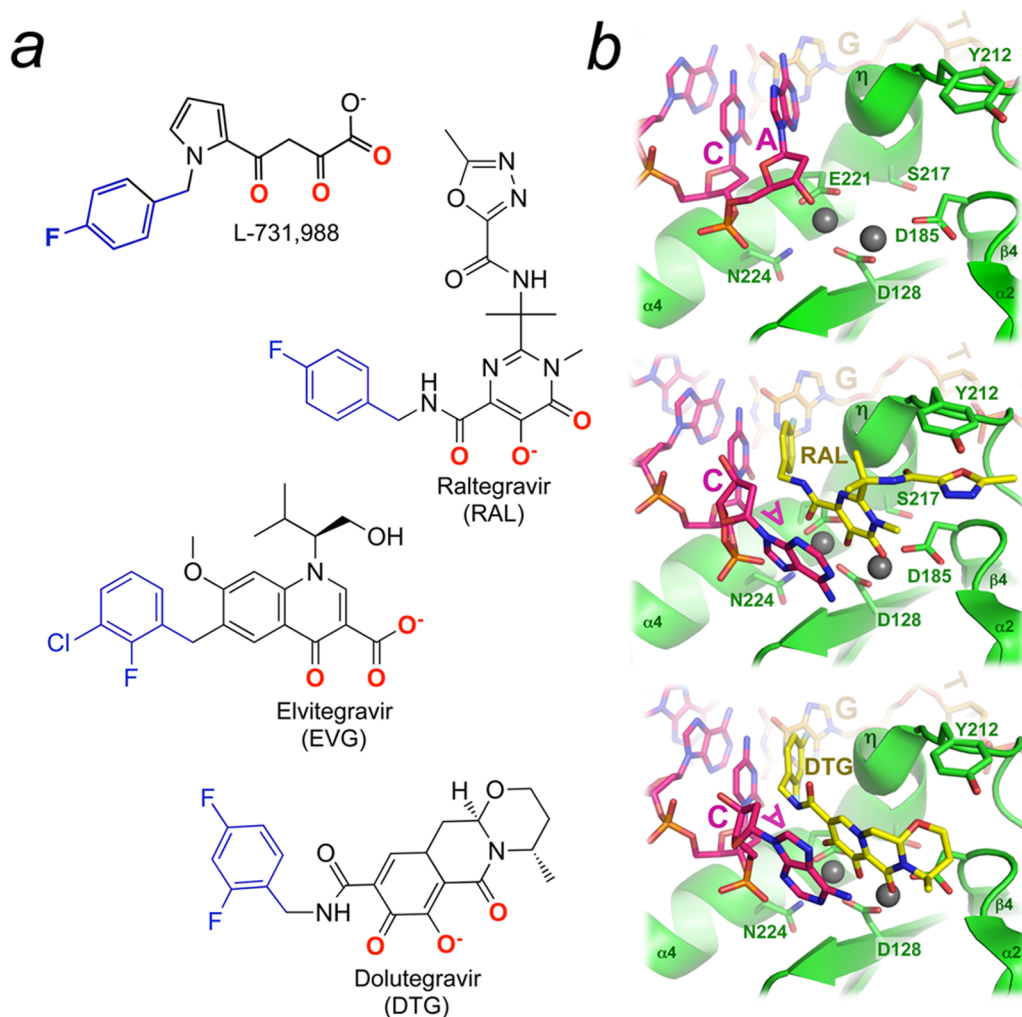


Figure 10. Inhibitors of HIV-1 IN strand transfer activity. (a) Chemical structures of a diketo acid (L-731,988) and the clinical INSTIs RAL, EVG, and DTG. Metal chelating atoms and halobenzyl moieties of the INSTIs are shown in red and blue, respectively. (b) Active site of the PFV intasome prior to (top) and after binding RAL (middle) or DTG (bottom). The 3_{10} helix that participates in the interactions with the INSTIs is indicated as η .

dolutegravir (DTG), have entered clinical use since then (Figure 10a).^{349,350}

Consistent with the method of their initial identification, INSTIs engage the active site of IN only when it is in complex with the vDNA end, competing with target DNA for binding to the intasome.⁸⁵ INSTIs specifically inhibit the strand transfer reaction, although they are capable of affecting 3'-processing at greatly elevated concentrations.^{345,351} These small molecules possess unusually tight binding to the HIV-1 intasome, with dissociative half-times measuring in hours (for EVG or RAL) or even days (for DTG).^{352,353} This property is likely very important for an inhibitor that blocks function of a long-lived complex, such as the PIC, which is geared for a one-off reaction event: to be effective, an INSTI must remain associated with the intasome until the cell destroys the PIC.

Despite their apparent chemical diversity, the INSTIs share two common functionalities: a Mg^{2+} chelating core, usually a triad of oxygen atoms attached to a rigid scaffold (colored red in Figure 10a), and a flexibly linked aromatic side chain, typically a halobenzyl group (shown in blue). Due to the conservation of the retroviral IN active site, these small molecules display broad-spectrum activity against diverse retroviruses.^{103,354,355} Accordingly, the structural basis for INSTI action could be characterized in the context of the

PFV intasome.^{85,205,206,356} Soaking the PFV intasome crystals with the different INSTIs invariably results in binding of the drug at the active site. In each studied case, the small molecules engaged the catalytic pair of Mg^{2+} ions in the IN active site (Figure 10b).⁸⁵ Here, the triad of metal-chelating heteroatoms of the small molecule closely imitates interactions made by the oxygen atoms of the scissile phosphodiester and the respective nucleophiles during 3'-processing and strand transfer.¹⁵⁶ The aromatic side chain of the INSTI assumes the position normally occupied by the base of the deoxyadenosine on the processed 3' vDNA end, intercalating between the base of the penultimate deoxycytidine and a short 3_{10} helix (designated η in Figure 10b) containing conserved PFV IN residues Pro214 and Gln215, which are equivalent to HIV-1 IN Pro145 and Gln146, respectively. INSTIs also make variable contacts with PFV IN Tyr212 (corresponding to HIV-1 IN Tyr143). In particular, the oxadiazole ring of RAL stacks with the phenolic side chain (Figure 10b, middle panel) and additionally makes a hydrogen bond to the backbone amide of the Tyr residue. Other compounds, such as EVG and DTG, make much less extensive van der Waals contacts with the Tyr side chain.

Crucially, by directly engaging the catalytic metals and displacing the 3' vDNA nucleotide, the INSTIs are incompatible with target DNA binding and strand transfer.

The requirement to displace the 3' adenosine from its natural position accounts for the slow kinetics of INSTIs binding to the intasome.³⁵⁷ A further energetic penalty associated with disengagement of a phosphodiester group from the Mg²⁺ ions in the intasomal active site explains the relative ineffectiveness of INSTIs to inhibit IN 3'-processing activity.¹⁵⁶

Adapted to operate on the bulky DNA substrate, the intasome harbors a voluminous active site, which is not ideal for the development of small-molecule inhibitors. X-ray structures of the PFV intasome prior to 3'-processing and strand transfer elucidated additional features of the DNA substrates that could potentially be mimicked in future INSTI design.¹⁵⁶ Furthermore, design of small molecules that more completely fill the substrate envelope within the intasome active site could lead to improved inhibitory properties.³⁵⁸

6.2. Viral Resistance to INSTIs

The clinical use of INSTIs has seen the emergence of HIV-1 variants with high-level resistance to RAL and EVG (for detailed reviews see refs 359, 360). Although not yet documented in INSTI naïve cohorts, resistance to DTG has been described in patients that initially failed RAL-based therapy.^{361–363} Due to their identical modes of action, INSTIs show substantially overlapping profiles of HIV-1 resistance mutations. The major genetic pathways leading to RAL resistance and virologic failure in patients are associated with substitutions of HIV-1 IN residues Tyr143 (typically to Cys or Arg), Asn155 (to His), and Gln148 (to His or Arg).³⁶⁴ Although, on their own, the primary substitutions result in modest levels of drug resistance, their effects are greatly amplified by secondary mutations. Most notably, combinations of Q148R/H with G140S/A in HIV-1 IN result in a loss of viral susceptibility to RAL and EVG, as well as substantial levels of resistance to DTG.^{361,362}

In the absence of HIV-1 intasome crystals, the PFV model was instrumental to shed some light on the mechanism of viral resistance to INSTIs.²⁰⁶ The observation that the oxadiazole ring of RAL interacts extensively with the aromatic side chain of Tyr212 readily explained why substitutions of HIV-1 IN residue Tyr143 cause viral resistance to RAL.⁸⁵ Because EVG and DTG make only weak contacts with the Tyr residue, their activities are largely unaffected by its substitutions.²⁰⁶ In contrast to Tyr212, PFV IN residues corresponding to HIV-1 Gln148 and Asn155 (Ser217 and Asn224, respectively) do not make direct interactions with the INSTIs. Nevertheless, akin to the effects of the analogous mutations in HIV-1 IN, S217H and N224H reduced susceptibility of PFV IN to RAL *in vitro*.^{206,356} Crystal structures revealed that the amino acid substitutions result in subtle but significant deformations of the intasomal active site. Expectedly, binding of the relatively rigid INSTIs requires the mutant active site to adopt a WT-like conformation. The energetic cost associated with the rehabilitation of the active site was proposed to be the reason for the apparent reduction in drug binding affinity.²⁰⁶ A Ser or an Ala residue at HIV-1 IN position 140 is predicted to make direct contacts with the side chain of His148, helping to explain the coevolution of the G140S and Q148H mutations.²⁰⁶ Ostensibly, a conformational adaptation to a large substrate, such as target DNA or chromatin, which make extensive interactions outside of the active site, will be offset to a lesser degree than small-molecule binding. The difference may provide the mutant viruses a sufficient selective advantage in the presence of the drug. This model accordingly predicts that INSTIs that make more

extensive interactions with immutable features of the IN active site will be less affected by mutations. Indeed, INSTIs such as DTG and MK2048, which display relatively long dissociative half-times from the WT HIV-1 intasome, are considerably less affected by the “shape-shifting” Q148H/R and N155H mutations.^{356,361,362} These small molecules, commonly referred to as second-generation INSTIs, tend to make van der Waals contacts to the main chain of the β 4- α 2 loop of the IN active site.^{206,356,358,365} Bulkier compounds, which occupy the substrate envelope of the intasomal active site more completely, tend to be more active against the classic RAL-resistant strains.³⁵⁸ Additionally, main chain IN amides involved in the interaction with the scissile phosphodiesters of the viral and target DNA substrates may provide useful immutable bonding points for the next-generation of INSTIs.¹⁵⁶

6.3. Emerging Allosteric Inhibitors of HIV-1 IN

Despite the great success of the combinatorial therapeutic approach against HIV/AIDS, new infections emerge from drug-resistant strains. In addition to developing new derivative compounds against current targets, new drugs that inhibit untapped steps of the HIV-1 replication cycle are of great importance. In the case of IN, the design of drugs that target positions different from the active site has the advantage of remaining theoretically potent against INSTI-resistant strains. Among the noncatalytic site inhibitors described so far, the most promising molecules target the LEDGF/p75 binding pocket at the HIV-1 IN CCD dimerization interface; these new molecules, which go by a variety of names (see refs 366, 367 for detailed reviews), will be referred to here as allosteric IN inhibitors (ALLINIs, Figure 8d).

Results of biochemical and cell-based infection assays provided initial evidence that the CCD–CCD interface might serve as a target for antiviral drug development. X-ray crystallography was used to screen for small-molecule binders of the HIV-1 IN CCD, and micromolar concentrations of one compound, 3,4-dihydroxyphenyltriphenylarsonium bromide, which engaged the CCD dimer interface at a region that was later confirmed as the LEDGF/p75 binding site, inhibited IN 3'-processing and strand transfer activities *in vitro*.³⁶⁸ Overexpression of LEDGF/p75 IBD-containing proteins that lacked chromatin-binding activity moreover inhibited HIV-1 replication at the integration step.^{295,369} Further highlighting the host factor-binding pocket for drug development, the combination of IBD protein overexpression with LEDGF/p75 knockdown could effectively cripple HIV-1 infection.³⁷⁰

The most advanced ALLINIs, derived from quinoline-based acetic acid, were independently discovered using two different approaches. Debyser and colleagues³¹² used the HIV-1 IN CCD-LEDGF/p75 IBD cocrystal structure to screen *in silico* for LEDGF/p75 binding site inhibitors, whereas another group discovered highly similar small molecules in a high-throughput screen for antagonists of IN 3'-processing activity.^{313,371} Optimized ALLINI compounds are highly potent, inhibiting HIV-1 replication with effective concentration 50% (EC₅₀) values in the low (~10–100) nanomolar range.^{310,313,315,317} This family of small molecules induces IN multimerization and inhibits IN catalysis^{310,311,316,317} and IN-LEDGF/p75 binding^{310–312,316,317} *in vitro*. Cocrystal structures of the inhibitors with the IN CCD dimer revealed that the molecules are bona fide allosteric inhibitors, as they engage the LEDGF/p75 binding site, which is distal from the enzyme active site.^{46,310–314,317} A crucial component of the ALLINI

pharmacophore is the carboxylic moiety (shown red in Figure 8d), which hydrogen bonds with the backbone amides of IN residues Glu170 and His171, mimicking the LEDGF/p75 Asp366 bidentate interaction with IN (Figure 8e). Another important feature is an aromatic side chain (blue in Figure 8d), which mimics hydrophobic interactions made by Ile365 of LEDGF/p75.

Initial experiments expectedly unveiled inhibition of integration during HIV-1 infection, with drug resistance mapping to the IN coding portion of the viral *pol* gene.^{310,312} However, follow up work revealed that the compounds are far more potent when they are present during the late phase of the HIV-1 lifecycle as compared to the acute phase of infection, when reverse transcription and integration occur.^{45,46,317,372} Consistent with the in vitro data, ALLINIs induce HIV-1 IN multimerization in the context of virus particles, which results in a catastrophic defect during virion maturation.^{45,46,372} The viral ribonucleoprotein (RNP) complex, composed mainly of viral RNA and nucleocapsid protein, is normally housed within a conical core composed of the CA protein. ALLINIs apparently uncouple the internal placement of the RNP within the core, yielding so-called “eccentric” virions with the RNP situated outside of the core, usually in association with the viral membrane.^{45–47,372} Virion protein, RNA, and cellular tRNA content are unaffected by ALLINI treatment,^{45,46,372,373} and the defective virions accordingly support normal levels of endogenous RT activity in vitro.^{45,373}

When present during the acute phase of HIV-1 infection, ALLINIs specifically inhibit integration without affecting the preceding reverse transcription step.^{45,46,310,312,315,317} Although drug-treated particles enter target cells normally,^{45,46} they are reportedly defective for reverse transcription, integration,^{45,46,315,372} and nuclear import of the PIC.³⁷² Careful comparisons of dose–response curves for inhibition of particle maturation, reverse transcription, and HIV-1 infection, however, suggest that the ability of the compounds to inhibit virion maturation accounts for their full antiviral activity.⁴⁷ In other words, drug-treated viruses are defective for reverse transcription because the misplaced RNP is unable to support DNA synthesis in the subsequently infected cell, as compared to a direct inhibition of reverse transcription by ALLINIs. Intriguingly, the range of replication defects ascribed to ALLINI-treated virions is reminiscent of the pleiotropic nature of class II IN mutations on HIV-1 replication (section 2 above). Moreover, class II IN mutant virus particles harbor an eccentric RNP that is indistinguishable from the one induced by ALLINI treatment.^{30,45,46,223} By extension, we suspect that the inability to encapsidate the RNP into the viral core underscores the majority of replication defects ascribed to class II HIV-1 IN mutant viruses. The ability to form eccentric RNPs has been shown to depend on the presence of viral RNA, indicating that IN may normally engage the viral genome, either directly or indirectly, to orchestrate RNP encapsidation into the viral core.⁴⁷

ALLINI potency is significantly increased when target cells are depleted for LEDGF/p75, suggesting that the host factor can compete with the compounds during the acute phase of infection.^{46,298,315,374} By contrast, LEDGF/p75 depletion or overexpression does not influence drug potency during particle assembly.^{45,46,297,315} As LEDGF/p75 associates constitutively with chromatin,^{171,268,276} it seems that the inability of LEDGF/p75 to compete for compound binding to IN during HIV-1 particle morphogenesis, which occurs at the cell periphery or

after the virus exits the cell, accounts for the unique pharmacology of this anti-IN drug class.

As might be expected from their known binding site at the IN CCD dimer interface, ALLINIs retain potency against INSTI-resistant HIV-1 strains.^{312,316,317,375} Unfortunately, ALLINIs seem to possess a relatively low genetic barrier to resistance, as several mutations mapping to the LEDGF/p75 binding cavity that greatly reduce drug potency have been described following ex vivo virus passage.^{310,312,315,316,375} Nevertheless, because these molecules perform well in concert with INSTIs,^{316,375} their clinical development is of great interest.

7. RETROVIRAL INTEGRATION AS A THERAPEUTIC TOOL

Due to the natural trait to stably integrate their genetic cargo into a cell chromosome, retroviruses have long been studied as tools for corrective gene therapy (see ref 376 for a current overview). Replication-defective vectors derived from MLV were used in pioneering studies to correct a variety of crippling diseases, including X-linked severe combined immunodeficiency (SCID-X1),³⁷⁷ Wiskott–Aldrich syndrome,³⁷⁸ and chronic granulomatous disease.^{379,380} However, a significant number of patients from these trials developed severe adverse effects due to clonal expansion of the treated cells, leading to leukemia.^{378,381–383} Detailed characterization revealed that many of these genotoxic events resulted from the integration of the MLV vector in the vicinity of a growth-promoting proto-oncogene, such as *LMO2*.^{378,382–384} Deregulation of proto-oncogene expression has long been known as a driver for retroviral genotoxicity,³⁸⁵ and the safety of retroviral vectors for human gene therapy applications has accordingly become a major priority in their development. An optimized vector would in theory efficiently express its transgene cargo over a long period of time without displaying adverse side effects on cellular gene expression or physiology.

The propensity for MLV to target integration to cellular promoters and enhancers,^{241,242,245} which was virtually unknown during the planning stages of the initial SCID-X1 trials, in hindsight likely made MLV an unfortunate choice for treatment. However, the field now has a much broader appreciation of how different types of retroviruses target potentially unsafe genomic features such as genes and enhancer regions.²⁴³ In this vein, vectors based on α -retroviruses,³⁸⁶ β -retroviruses,³⁸⁷ or spumaviruses,³⁸⁸ each of which targets genes and enhancers to lesser extents than lentiviruses and γ -retroviruses, respectively, might prove safer than MLV-based vectors. The identification of the mechanisms of integration targeting for the lentiviruses and γ -retroviruses has additionally opened up new approaches to vector design. As one example, MLV-derived vectors show greatly reduced propensity to integrate nearby transcriptional start sites in the presence of BET protein inhibitors,³¹⁹ raising the possibility of using such inhibitors with MLV vectors in the clinic. Several drawbacks however seem likely to limit such approaches. In addition to potential small-molecule toxicity, MLV retained a partial tendency to target promoter-proximal regions in the presence of BET inhibitors such as JQ-1.³¹⁹ Furthermore, MLV vector titer was reduced significantly by JQ-1 treatment due to inhibition of integration.^{319–321} An alternative strategy is to delete the IN C-terminal tail region that mediates BET protein binding.^{321,334} Such vectors greatly reduce promoter-proximal integration with only a modest effect on vector titer.^{334,389} It

will be instructive to determine the carcinogenic potential of such IN deletion constructs in animal models of MLV pathogenesis.

A separate approach to MLV vector modification looks extremely promising from initial clinical trials. The viral promoter and enhancer, which are situated within the U3 region of the LTR, can be deleted, resulting in so-called self-inactivating (SIN) vectors without severely affecting reverse transcription or integration.³⁹⁰ An SCID-X1 trial with an MLV SIN vector has failed to detect evidence for leukemia during an initial 1–3 year observational period, indicating that deletion of the viral enhancer might go a long way to improve MLV-based vector safety.³⁹¹

As fusion proteins between the LEDGF/p75 IBD^{301–303} or BET protein ET domain³²¹ and chromatin binding modules can effectively retarget integration, such protein hybrids could in theory be used to steer lentivirus or γ -retrovirus integration out of harm's way. For example, a LEDGF/p75-based fusion harboring the HP1 α heterochromatin protein yielded an overall integration pattern that was remarkably similar to random.³⁰¹ A key drawback of such approaches is the requirement to express the retargeting factor in the cells that will receive the retroviral vector. This approach accordingly seems to hold considerably less promise than the attempts to reduce or eliminate genotoxicity through direct vector modification.

Notwithstanding highly significant genotoxic integration targeting by the lentiviruses, there is little evidence to suggest an operational link between HIV-1 integration and carcinogenesis. A primary reason may be the highly cytopathic nature of infection: infected CD4-positive T cells, the primary targets of the virus, display an average half-life of only \sim 1.5 days.³⁹² However, cancer-related genes are targeted about 5-fold more frequently by HIV-1 PICs than expected.²⁴⁰ Integrations in the vicinity of growth-promoting genes moreover can help to drive the clonal expansion of cells that constitute the latent viral reservoir.^{227,228} Numerous HIV-1 gene products, including the surface envelope glycoprotein Gp120³⁹³ and viral protein R (Vpr),³⁹⁴ are acutely cytopathic, and such genes are deleted from HIV-based vectors. As is the case with MLV vectors, it is critical to monitor HIV-based gene therapy trials for sites of vector DNA integration by deep sequencing to catch potential longitudinal emergence of dominant cell clones in patients.

As is the case with MLV SIN vectors, results of preliminary clinical trials with HIV-based vectors look promising, with little to no evidence for the type of clonal dominance that was observed in initial MLV-based trials.^{395–398} The ability to safely integrate a corrective transgene in a long-lasting target cell can in theory be expected to go a long way—perhaps for a patient's lifetime—to correct certain debilitating diseases. The field accordingly cautiously awaits long-term follow up of ongoing retroviral-based gene therapy trials.

8. PERSPECTIVES

During the past 30 years the field of retroviral integration progressed from an epitome of experimental hardship and enigma to arguably the best-understood DNA recombination system. Combined efforts of academic groups and leading pharmaceutical companies resulted in the discovery of potent HIV-1 IN inhibitors, characterization of the first cellular cofactors of retroviral INs, and elucidation of the mechanistic and structural details of retroviral integration.

One clear vector for future research will be expansion of the intasome structure repertoire. In particular, elucidation of HIV-

1 or lentiviral intasome structure will be of great importance to help the rational design of next-generation INSTIs and improved ALLINIs. Furthermore, characterization of the HIV-1 intasome may help to identify new pockets that are potentially druggable by allosteric inhibitors. The development of resistance to the antiretroviral drug arsenal is a major health issue, making it important to find new therapeutic strategies. Many steps modulating retroviral integration are still poorly described and would necessitate further investigation. Among these is the mechanism by which retroviruses protect themselves against autointegration. One could imagine the development of small molecules that could be used to trigger suicidal autointegration before the PIC encounters host chromatin.

Another open question is the mechanism of retroviral STC disassembly. Since the available TCC and STC structures are very similar, how IN signals the completion of strand transfer to the cell to allow DNA repair remains unknown. Unravelling the mechanism and the cellular proteins involved in this process may well lead to the development of new drugs targeting these steps. A more detailed characterization of the unexpected role of IN in HIV-1 particle morphogenesis could also lead to new ways of inhibiting viral replication.

Studies on retroviral IN host factors have clarified the molecular mechanisms underlying integration site selection. The recent description of MLV IN targeting factors BRD2–4 together with the lentiviral IN cofactor LEDGF/p75 established the concept of bimodal tethering as a major mechanism for integration site selection. These discoveries opened new windows toward tuning the specificity of retroviral integration. Future studies will elucidate if the other retroviral families rely on similar pathways to select suitable chromatin environments and will hopefully give important insights into improving the safety of retroviral gene therapy vectors. Given the parallel development of gene-editing technologies such as CRISPR (clustered regularly interspaced short palindromic repeats)-Cas9 (reviewed in ref 399), there has been in recent years a resurgence of interest in human gene therapy. Yet, notwithstanding its lauded efficiency, gene editing using CRISPR-Cas9 involves generation of a double-strand DNA break at the target locus, a highly genotoxic chromosomal lesion in its own right. Therefore, it remains to be seen which methodology might evolve into a safer therapy approach in the future.

AUTHOR INFORMATION

Corresponding Authors

*A.N.E. e-mail: alan_engelman@dfci.harvard.edu.

*P.C. e-mail: peter.cherepanov@crick.ac.uk.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

Biographies

Paul Lesbats obtained his Ph.D. from the University of Bordeaux (Bordeaux, France) in 2011. During his doctorate under the supervision of Vincent Parissi he studied how HIV-1 IN assembles into the active complex and how its activity is modulated by chromatin compaction in vitro. In 2012 he joined as a postdoctoral fellow the laboratory of Peter Cherepanov at the Clare Hall Laboratories of the

Francis Crick Institute, where he is using structural biology to study the mechanisms of retroviral integration into chromatin and the role of Gag in this process.

Alan N. Engelman received his Ph.D. in 1990, studying the *v-abl* protein tyrosine kinase with Naomi Rosenberg at the Sackler School for Graduate Biomedical Sciences, Tufts University School of Medicine in Boston, Massachusetts. From 1990 to 1995 he conducted postdoctoral research on HIV-1 integration with Robert Craigie and Kiyoshi Mizuuchi at the National Institute of Diabetes and Digestive and Kidney Diseases in Bethesda, Maryland. In 1995 he joined the faculty of the Dana-Farber Cancer Institute in Boston, where he is now a professor of medicine in the Department of Cancer Immunology and Virology. Dr. Engelman's research focuses on the biochemistry, structural biology, and virology of retroviral DNA integration.

Peter Cherepanov obtained his Ph.D. from the University of Leuven (Leuven, Belgium) in 2000 in the laboratory of Zeger Debyser, where he identified LEDGF/p75 as an HIV-1 IN binding factor. He joined the laboratory of Alan Engelman at the Dana-Farber Cancer Institute in 2003 to study the structural aspects of the HIV-1 IN–LEDGF/p75 interaction. He established his group at Imperial College London in 2005, before moving to Clare Hall Laboratories of the Francis Crick Institute in 2011. His group is focusing on the structural biology of retroviral DNA integration, and their main contribution to the field was the elucidation of the PFV intasome, TCC, and STC structures.

ACKNOWLEDGMENTS

We thank H. Aihara, M. Kvaratskhelia, and M. Foster for sharing data prior to publication and V. Pie for critical reading of the manuscript. This work was in part funded by US National Institutes of Health grants R37 AI039394 and R01 AI052014 (to A.N.E) and P50 GM082251 (to P.C.). This work was additionally supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK, the UK Medical Research Council, and the Wellcome Trust.

ABBREVIATIONS USED

AIDS	acquired immunodeficiency syndrome
ALLINI	allosteric integrase inhibitor
ASLV	avian sarcoma-leukosis virus
BAF	barrier-to-autointegration factor
BER	base excision repair
BET	bromodomain and extraterminal domain (protein family)
CA	capsid
CCD	catalytic core domain
CPSF6	cleavage and polyadenylation specificity factor 6
CRISPR	clustered regularly interspaced short palindromic repeats
Cryo-EM	cryo-electron microscopy
CTD	carboxy-terminal domain
DTG	dolutegravir
EBM	ET binding motif
EC ₅₀	effective concentration 50%
ET	extra-terminal
EVG	elvitegravir
FEN-1	flap endonuclease 1
FIV	feline immunodeficiency virus
HIV	human immunodeficiency virus
HTLV	human T-lymphotropic virus
IBD	integrase binding domain
IN	integrase
INSTI	integrase strand transfer inhibitor

LEDGF	lens epithelium-derived growth factor
LTR	long terminal repeat
MeCP2	methyl CpG binding protein 2
MLV	murine leukemia virus
MMTV	mouse mammary tumor virus
NED	NTD extension domain
NHEJ	nonhomologous end joining
NLS	nuclear localization signal
NMR	nuclear magnetic resonance
NPC	nuclear pore complex
NUP	nucleoporin
NTD	amino-terminal domain
PFV	prototype foamy virus
PIC	preintegration complex
PP2A	protein phosphatase 2A
RAL	raltegravir
RNP	ribonucleoprotein
RT	reverse transcriptase
SCID-X1	X-linked severe combined immunodeficiency
SH3	Src homology 3
SIN	self-inactivating
SIV	simian immunodeficiency virus
STC	strand transfer complex
TCC	target capture complex
vDNA	viral DNA
Vpr	viral protein R
WT	wild type

REFERENCES

- (1) Rous, P. A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells. *J. Exp. Med.* **1911**, *13* (4), 397–411.
- (2) Ellermann, V.; Bang, O. Experimentelle Leukämie Bei Hühnern. *Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. Orig.* **1908**, *46*, 595–609.
- (3) Poiesz, B. J.; Ruscetti, F. W.; Gazdar, A. F.; Bunn, P. A.; Minna, J. D.; Gallo, R. C. Detection and Isolation of Type C Retrovirus Particles from Fresh and Cultured Lymphocytes of a Patient with Cutaneous T-Cell Lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* **1980**, *77* (12), 7415–7419.
- (4) Barre-Sinoussi, F.; Chermann, J. C.; Rey, F.; Nugeyre, M. T.; Chamaret, S.; Gruest, J.; Dauguet, C.; Axler-Blin, C.; Vezinet-Brun, F.; Rouzioux, C.; et al. Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). *Science* **1983**, *220* (4599), 868–871.
- (5) Popovic, M.; Sarngadharan, M. G.; Read, E.; Gallo, R. C. Detection, Isolation, and Continuous Production of Cytopathic Retroviruses (HTLV-III) from Patients with AIDS and Pre-AIDS. *Science* **1984**, *224* (4648), 497–500.
- (6) Coffin, J. M. The Discovery of HTLV-1, the First Pathogenic Human Retrovirus. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (51), 15525–15529.
- (7) Temin, H. M. The Participation of DNA in Rous Sarcoma Virus Production. *Virology* **1964**, *23*, 486–494.
- (8) Donehower, L. A.; Varmus, H. E. A Mutant Murine Leukemia Virus with a Single Missense Codon in *pol* Is Defective in a Function Affecting Integration. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81* (20), 6461–6465.
- (9) Panganiban, A. T.; Temin, H. M. The Retrovirus *pol* Gene Encodes a Product Required for DNA Integration: Identification of a Retrovirus *int* Locus. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 7885–7889.
- (10) Schwartzberg, P.; Colicelli, J.; Goff, S. P. Construction and Analysis of Deletion Mutations in the *pol* Gene of Moloney Murine Leukemia Virus: A New Viral Function Required for Productive Infection. *Cell* **1984**, *37*, 1043–1052.

- (11) Quinn, T. P.; Grandgenett, D. P. Genetic Evidence That the Avian Retrovirus DNA Endonuclease Domain of pol Is Necessary for Viral Integration. *J. Virol.* **1988**, *62* (7), 2307–2312.
- (12) Vande Woude, G. F.; Oskarsson, M.; Enquist, L. W.; Nomura, S.; Sullivan, M.; Fischinger, P. J. Cloning of Integrated Moloney Sarcoma Proviral DNA Sequences in Bacteriophage Lambda. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76* (9), 4464–4468.
- (13) Van Beveren, C.; Goddard, J. G.; Berns, A.; Verma, I. M. Structure of Moloney Murine Leukemia Viral DNA: Nucleotide Sequence of the 5' Long Terminal Repeat and Adjacent Cellular Sequences. *Proc. Natl. Acad. Sci. U. S. A.* **1980**, *77* (6), 3307–3311.
- (14) Shimotohno, K.; Mizutani, S.; Temin, H. M. Sequence of Retrovirus Provirus Resembles That of Bacterial Transposable Elements. *Nature* **1980**, *285* (5766), 550–554.
- (15) Ju, G.; Skalka, A. M. Nucleotide Sequence Analysis of the Long Terminal Repeat (LTR) of Avian Retroviruses: Structural Similarities with Transposable Elements. *Cell* **1980**, *22* (2), 379–386.
- (16) Bowerman, B.; Brown, P. O.; Bishop, J. M.; Varmus, H. E. A Nucleoprotein Complex Mediates the Integration of Retroviral DNA. *Genes Dev.* **1989**, *3* (4), 469–478.
- (17) Wei, S. Q.; Mizuuchi, K.; Craigie, R. A Large Nucleoprotein Assembly at the Ends of the Viral DNA Mediates Retroviral DNA Integration. *EMBO J.* **1997**, *16* (24), 7511–7520.
- (18) Farnet, C. M.; Haseltine, W. A. Determination of Viral Proteins Present in the Human Immunodeficiency Virus Type 1 Preintegration Complex. *J. Virol.* **1991**, *65* (4), 1910–1915.
- (19) Bukrinsky, M. I.; Sharova, N.; McDonald, T. L.; Pushkarskaya, T.; Tarpley, W. G.; Stevenson, M. Association of Integrase, Matrix, and Reverse Transcriptase Antigens of Human Immunodeficiency Virus Type 1 with Viral Nucleic Acids Following Acute Infection. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (13), 6125–6129.
- (20) Heininger, N. K.; Bukinsky, M. I.; Haggerty, S. A.; Ragland, A. M.; Kewalramani, V.; Lee, M. A.; Gendelman, H. E.; Ratner, L.; Stevenson, M.; Emerman, M. The Vpr Protein of Human Immunodeficiency Virus Type 1 Influences Nuclear Localization of Viral Nucleic Acids in Nondividing Host Cells. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (15), 7311–7315.
- (21) Miller, M. D.; Farnet, C. M.; Bushman, F. D. Human Immunodeficiency Virus Type 1 Preintegration Complexes: Studies of Organization and Composition. *J. Virol.* **1997**, *71* (7), 5382–5390.
- (22) Li, L.; Olvera, J. M.; Yoder, K. E.; Mitchell, R. S.; Butler, S. L.; Lieber, M.; Martin, S. L.; Bushman, F. D. Role of the Non-Homologous DNA End Joining Pathway in the Early Steps of Retroviral Infection. *EMBO J.* **2001**, *20* (12), 3272–3281.
- (23) Lin, C. W.; Engelman, A. The Barrier-to-Autointegration Factor Is a Component of Functional Human Immunodeficiency Virus Type 1 Preintegration Complexes. *J. Virol.* **2003**, *77* (8), 5030–5036.
- (24) Llano, M.; Vanegas, M.; Fregoso, O.; Saenz, D.; Chung, S.; Peretz, M.; Poeschla, E. M. LEDGF/p75 Determines Cellular Trafficking of Diverse Lentiviral but Not Murine Oncoretroviral Integrase Proteins and Is a Component of Functional Lentiviral Preintegration Complexes. *J. Virol.* **2004**, *78* (17), 9524–9537.
- (25) Raghavendra, N. K.; Shkriabai, N.; Graham, R.; Hess, S.; Kvaratskhelia, M.; Wu, L. Identification of Host Proteins Associated with HIV-1 Preintegration Complexes Isolated from Infected CD4+ Cells. *Retrovirology* **2010**, *7*, 66.
- (26) Schweitzer, C. J.; Jagdish, T.; Haverland, N.; Ciborowski, P.; Belshan, M. Proteomic Analysis of Early HIV-1 Nucleoprotein Complexes. *J. Proteome Res.* **2013**, *12* (2), 559–572.
- (27) Engelman, A. In *HIV-1 Integrase: Mechanism and Inhibitor Design*; Neamati, N., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, 2011; DOI:10.1002/9781118015377.ch6.
- (28) Shin, C. G.; Taddeo, B.; Haseltine, W. A.; Farnet, C. M. Genetic Analysis of the Human Immunodeficiency Virus Type 1 Integrase Protein. *J. Virol.* **1994**, *68* (3), 1633–1642.
- (29) Taddeo, B.; Haseltine, W. A.; Farnet, C. M. Integrase Mutants of Human Immunodeficiency Virus Type 1 with a Specific Defect in Integration. *J. Virol.* **1994**, *68* (12), 8401–8405.
- (30) Engelman, A.; Englund, G.; Orenstein, J. M.; Martin, M. A.; Craigie, R. Multiple Effects of Mutations in Human Immunodeficiency Virus Type 1 Integrase on Viral Replication. *J. Virol.* **1995**, *69* (5), 2729–2736.
- (31) Masuda, T.; Planelles, V.; Krogstad, P.; Chen, I. S. Genetic Analysis of Human Immunodeficiency Virus Type 1 Integrase and the U3 Att Site: Unusual Phenotype of Mutants in the Zinc Finger-Like Domain. *J. Virol.* **1995**, *69* (11), 6687–6696.
- (32) Engelman, A. In Vivo Analysis of Retroviral Integrase Structure and Function. *Adv. Virus Res.* **1999**, *52*, 411–426.
- (33) Leavitt, A. D.; Robles, G.; Alesandro, N.; Varmus, H. E. Human Immunodeficiency Virus Type 1 Integrase Mutants Retain In Vitro Integrase Activity yet Fail to Integrate Viral DNA Efficiently During Infection. *J. Virol.* **1996**, *70* (2), 721–728.
- (34) Lu, R.; Limon, A.; Devroe, E.; Silver, P. A.; Cherepanov, P.; Engelman, A. Class II Integrase Mutants with Changes in Putative Nuclear Localization Signals Are Primarily Blocked at a Postnuclear Entry Step of Human Immunodeficiency Virus Type 1 Replication. *J. Virol.* **2004**, *78* (23), 12735–12746.
- (35) Lu, R.; Vandegraaff, N.; Cherepanov, P.; Engelman, A. Lys-34, Dispensable for Integrase Catalysis, Is Required for Preintegration Complex Function and Human Immunodeficiency Virus Type 1 Replication. *J. Virol.* **2005**, *79* (19), 12584–12591.
- (36) Rahman, S.; Lu, R.; Vandegraaff, N.; Cherepanov, P.; Engelman, A. Structure-Based Mutagenesis of the Integrase-LEDGF/p75 Interface Uncouples a Strict Correlation between In Vitro Protein Binding and HIV-1 Fitness. *Virology* **2007**, *357*, 79–90.
- (37) Ansari-Lari, M. A.; Donehower, L. A.; Gibbs, R. A. Analysis of Human Immunodeficiency Virus Type 1 Integrase Mutants. *Virology* **1995**, *211* (1), 332–335.
- (38) Nakamura, T.; Masuda, T.; Goto, T.; Sano, K.; Nakai, M.; Harada, S. Lack of Infectivity of HIV-1 Integrase Zinc Finger-Like Domain Mutant with Morphologically Normal Maturation. *Biochem. Biophys. Res. Commun.* **1997**, *239* (3), 715–722.
- (39) Lu, R.; Ghory, H. Z.; Engelman, A. Genetic Analyses of Conserved Residues in the Carboxyl-Terminal Domain of Human Immunodeficiency Virus Type 1 Integrase. *J. Virol.* **2005**, *79* (16), 10356–10368.
- (40) Ikeda, T.; Nishitsuji, H.; Zhou, X.; Nara, N.; Ohashi, T.; Kannagi, M.; Masuda, T. Evaluation of the Functional Involvement of Human Immunodeficiency Virus Type 1 Integrase in Nuclear Import of Viral cDNA During Acute Infection. *J. Virol.* **2004**, *78* (21), 11563–11573.
- (41) Ao, Z.; Fowke, K. R.; Cohen, E. A.; Yao, X. Contribution of the C-Terminal Tri-Lysine Regions of Human Immunodeficiency Virus Type 1 Integrase for Efficient Reverse Transcription and Viral DNA Nuclear Import. *Retrovirology* **2005**, *2*, 62.
- (42) Hehl, E. A.; Joshi, P.; Kalpana, G. V.; Prasad, V. R. Interaction between Human Immunodeficiency Virus Type 1 Reverse Transcriptase and Integrase Proteins. *J. Virol.* **2004**, *78* (10), 5056–5067.
- (43) Tekeste, S. S.; Wilkinson, T. A.; Weiner, E. M.; Xu, X.; Miller, J. T.; Le Grice, S. F.; Clubb, R. T.; Chow, S. A. Interaction between Reverse Transcriptase and Integrase Is Required for Reverse Transcription During HIV-1 Replication. *J. Virol.* **2015**, *89* (23), 12058–12069.
- (44) Wu, X.; Liu, H.; Xiao, H.; Conway, J. A.; Hehl, E.; Kalpana, G. V.; Prasad, V.; Kappes, J. C. Human Immunodeficiency Virus Type 1 Integrase Protein Promotes Reverse Transcription through Specific Interactions with the Nucleoprotein Reverse Transcription Complex. *J. Virol.* **1999**, *73* (3), 2126–2135.
- (45) Balakrishnan, M.; Yant, S. R.; Tsai, L.; O'Sullivan, C.; Bam, R. A.; Tsai, A.; Niedziela-Majka, A.; Stray, K. M.; Sakowicz, R.; Cihlar, T. Non-Catalytic Site HIV-1 Integrase Inhibitors Disrupt Core Maturation and Induce a Reverse Transcription Block in Target Cells. *PLoS One* **2013**, *8* (9), e74163.
- (46) Jurado, K. A.; Wang, H.; Slaughter, A.; Feng, L.; Kessl, J. J.; Koh, Y.; Wang, W.; Ballandras-Colas, A.; Patel, P. A.; Fuchs, J. R.; et al. Allosteric Integrase Inhibitor Potency Is Determined through the

Inhibition of HIV-1 Particle Maturation. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (21), 8690–8695.

(47) Fontana, J.; Jurado, K. A.; Cheng, N.; Ly, N. L.; Fuchs, J. R.; Gorelick, R. J.; Engelman, A. N.; Steven, A. C. Distribution and Redistribution of HIV-1 Nucleocapsid Protein in Immature, Mature, and Integrase-Inhibited Virions: A Role for Integrase in Maturation. *J. Virol.* **2015**, *89* (19), 9765–9780.

(48) Gallay, P.; Hope, T.; Chin, D.; Trono, D. HIV-1 Infection of Nondividing Cells through the Recognition of Integrase by the Importin/Karyopherin Pathway. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 9825–9830.

(49) Bouyac-Bertoia, M.; Dvorin, J. D.; Fouchier, R. A.; Jenkins, Y.; Meyer, B. E.; Wu, L. I.; Emerman, M.; Malim, M. H. HIV-1 Infection Requires a Functional Integrase NLS. *Mol. Cell* **2001**, *7*, 1025–1035.

(50) Dvorin, J. D.; Bell, P.; Maul, G. G.; Yamashita, M.; Emerman, M.; Malim, M. H. Reassessment of the Roles of Integrase and the Central DNA Flap in Human Immunodeficiency Virus Type 1 Nuclear Import. *J. Virol.* **2002**, *76* (23), 12087–12096.

(51) Limón, A.; Devroe, E.; Lu, R.; Ghory, H. Z.; Silver, P. A.; Engelman, A. 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.* **2002**, *76* (21), 10598–10607.

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

(53) Krishnan, L.; Matreyek, K. A.; Oztop, I.; Lee, K.; Tipper, C. H.; Li, X.; Dar, M. J.; KewalRamani, V. N.; Engelman, A. The Requirement for Cellular Transportin 3 (TNPO3 or TRN-SR2) During Infection Maps to Human Immunodeficiency Virus Type 1 Capsid and Not Integrase. *J. Virol.* **2010**, *84* (1), 397–406.

(54) De Houwer, S.; Demeulemeester, J.; Thys, W.; Rocha, S.; Dirix, L.; Gijssbers, R.; Christ, F.; Debyser, Z. The HIV-1 Integrase Mutant R263A/K264A Is 2-Fold Defective for TRN-SR2 Binding and Viral Nuclear Import. *J. Biol. Chem.* **2014**, *289* (36), 25351–25361.

(55) Brown, P. O.; Bowerman, B.; Varmus, H. E.; Bishop, J. M. Correct Integration of Retroviral DNA in Vitro. *Cell* **1987**, *49* (3), 347–356.

(56) Farnet, C. M.; Haseltine, W. A. Integration of Human Immunodeficiency Virus Type 1 DNA in Vitro. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87* (11), 4164–4168.

(57) Ellison, V.; Abrams, H.; Roe, T.; Lifson, J.; Brown, P. Human Immunodeficiency Virus Integration in a Cell-Free System. *J. Virol.* **1990**, *64* (6), 2711–2715.

(58) Engelman, A.; Oztop, I.; Vandegraaff, N.; Raghavendra, N. K. Quantitative Analysis of HIV-1 Preintegration Complexes. *Methods* **2009**, *47*, 283–290.

(59) Gianni, A. M.; Smotkin, D.; Weinberg, R. A. Murine Leukemia Virus: Detection of Unintegrated Double-Stranded DNA Forms of the Provirus. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72* (2), 447–451.

(60) Shank, P. R.; Hughes, S. H.; Kung, H. J.; Majors, J. E.; Quintrell, N.; Guntaka, R. V.; Bishop, J. M.; Varmus, H. E. Mapping Unintegrated Avian Sarcoma Virus DNA: Termini of Linear DNA Bear 300 Nucleotides Present Once or Twice in Two Species of Circular DNA. *Cell* **1978**, *15* (4), 1383–1395.

(61) Shank, P. R.; Varmus, H. E. Virus-Specific DNA in the Cytoplasm of Avian Sarcoma Virus-Infected Cells Is a Precursor to Covalently Closed Circular Viral DNA in the Nucleus. *J. Virol.* **1978**, *25* (1), 104.

(62) Yoshimura, F. K.; Weinberg, R. A. Restriction Endonuclease Cleavage of Linear and Closed Circular Murine Leukemia Viral DNAs: Discovery of a Smaller Circular Form. *Cell* **1979**, *16* (2), 323–332.

(63) Shoemaker, C.; Goff, S.; Gilboa, E.; Paskind, M.; Mitra, S. W.; Baltimore, D. Structure of a Cloned Circular Moloney Murine Leukemia Virus DNA Molecule Containing an Inverted Segment: Implications for Retrovirus Integration. *Proc. Natl. Acad. Sci. U. S. A.* **1980**, *77* (7), 3932–3936.

(64) Pang, S.; Koyanagi, Y.; Miles, S.; Wiley, C.; Vinters, H. V.; Chen, I. S. High Levels of Unintegrated HIV-1 DNA in Brain Tissue of AIDS Dementia Patients. *Nature* **1990**, *343* (6253), 85–89.

(65) Mizuuchi, K.; Nash, H. A. Restriction Assay for Integrative Recombination of Bacteriophage Lambda DNA in Vitro: Requirement for Closed Circular DNA Substrate. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73*, 3524–3528.

(66) Panganiban, A. T.; Temin, H. M. Circles with Two Tandem LTRS Are Precursors to Integrated Retrovirus DNA. *Cell* **1984**, *36*, 673–679.

(67) Fujiwara, T.; Mizuuchi, K. Retroviral DNA Integration: Structure of an Integration Intermediate. *Cell* **1988**, *54* (4), 497–504.

(68) Brown, P. O.; Bowerman, B.; Varmus, H. E.; Bishop, J. M. Retroviral Integration: Structure of the Initial Covalent Product and Its Precursor, and a Role for the Viral in Protein. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86* (8), 2525–2529.

(69) Roth, M. J.; Schwartzberg, P. L.; Goff, S. P. Structure of the Termini of DNA Intermediates in the Integration of Retroviral DNA: Dependence on in Function and Terminal DNA Sequence. *Cell* **1989**, *58* (1), 47–54.

(70) Fujiwara, T.; Craigie, R. Integration of Mini-Retroviral DNA: A Cell-Free Reaction for Biochemical Analysis of Retroviral Integration. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86* (9), 3065–3069.

(71) Katzman, M.; Katz, R. A.; Skalka, A. M.; Leis, J. The Avian Retroviral Integration Protein Cleaves the Terminal Sequences of Linear Viral DNA at the in Vivo Sites of Integration. *J. Virol.* **1989**, *63* (12), 5319–5327.

(72) Vora, A. C.; Fitzgerald, M. L.; Grandgenett, D. P. Removal of 3'-OH-Terminal Nucleotides from Blunt-Ended Long Terminal Repeat Termini by the Avian Retrovirus Integration Protein. *J. Virol.* **1990**, *64* (11), 5656–5659.

(73) Katz, R. A.; Merkel, G.; Kulkosky, J.; Leis, J.; Skalka, A. M. The Avian Retroviral IN Protein Is Both Necessary and Sufficient for Integrative Recombination in Vitro. *Cell* **1990**, *63* (1), 87–95.

(74) Craigie, R.; Fujiwara, T.; Bushman, F. The IN Protein of Moloney Murine Leukemia Virus Processes the Viral DNA Ends and Accomplishes Their Integration in Vitro. *Cell* **1990**, *62*, 829–837.

(75) Bushman, F. D.; Fujiwara, T.; Craigie, R. Retroviral DNA Integration Directed by HIV Integration Protein in Vitro. *Science* **1990**, *249*, 1555–1558.

(76) Bushman, F. D.; Craigie, R. Activities of Human Immunodeficiency Virus (HIV) Integration Protein in Vitro: Specific Cleavage and Integration of HIV DNA. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 1339–1343.

(77) Sherman, P. A.; Fyfe, J. A. Human Immunodeficiency Virus Integration Protein Expressed in Escherichia Coli Possesses Selective DNA Cleaving Activity. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 5119–5123.

(78) Dyda, F.; Chandler, M.; Hickman, A. B. The Emerging Diversity of Transpososome Architectures. *Q. Rev. Biophys.* **2012**, *45* (4), 493–521.

(79) Mizuuchi, K.; Adzuma, K. Inversion of the Phosphate Chirality at the Target Site of Mu DNA Strand Transfer: Evidence for a One-Step Transesterification Mechanism. *Cell* **1991**, *66* (1), 129–140.

(80) Engelman, A.; Mizuuchi, K.; Craigie, R. HIV-1 DNA Integration: Mechanism of Viral DNA Cleavage and DNA Strand Transfer. *Cell* **1991**, *67* (6), 1211–1221.

(81) Kennedy, A. K.; Haniford, D. B.; Mizuuchi, K. Single Active Site Catalysis of the Successive Phosphoryl Transfer Steps by DNA Transposases: Insights from Phosphorothioate Stereoselectivity. *Cell* **2000**, *101* (3), 295–305.

(82) Reznikoff, W. S. Tn5 as a Model for Understanding DNA Transposition. *Mol. Microbiol.* **2003**, *47* (5), 1199–1206.

(83) Haniford, D. B. Transpososome Dynamics and Regulation in Tn10 Transposition. *Crit. Rev. Biochem. Mol. Biol.* **2006**, *41* (6), 407–424.

(84) Chen, H.; Wei, S.-Q.; Engelman, A. Multiple Integrase Functions Are Required to Form the Native Structure of the

Human Immunodeficiency Virus Type I Intasome. *J. Biol. Chem.* **1999**, *274* (24), 17358–17364.

(85) Hare, S.; Gupta, S. S.; Valkov, E.; Engelman, A.; Cherepanov, P. Retroviral Intasome Assembly and Inhibition of DNA Strand Transfer. *Nature* **2010**, *464* (7286), 232–236.

(86) Shoemaker, C.; Hoffman, J.; Goff, S. P.; Baltimore, D. Intramolecular Integration within Moloney Murine Leukemia Virus DNA. *J. Virol.* **1981**, *40*, 164–172.

(87) Li, Y.; Kappes, J. C.; Conway, J. A.; Price, R. W.; Shaw, G. M.; Hahn, B. H. 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.* **1991**, *65*, 3973–3985.

(88) Tipper, C.; Sodroski, J. Enhanced Autointegration in Hyperstable Simian Immunodeficiency Virus Capsid Mutants Blocked after Reverse Transcription. *J. Virol.* **2013**, *87* (7), 3628–3639.

(89) Zheng, R.; Ghirlando, R.; Lee, M. S.; Mizuuchi, K.; Krause, M.; Craigie, R. Barrier-to-Autointegration Factor (BAF) Bridges DNA in a Discrete, Higher-Order Nucleoprotein Complex. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (16), 8997–9002.

(90) Skoko, D.; Li, M.; Huang, Y.; Mizuuchi, M.; Cai, M.; Bradley, C. M.; Pease, P. J.; Xiao, B.; Marko, J. F.; Craigie, R.; et al. Barrier-to-Autointegration Factor (BAF) Condenses DNA by Looping. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (39), 16610–16615.

(91) Lee, M. S.; Craigie, R. A Previously Unidentified Host Protein Protects Retroviral DNA from Autointegration. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (4), 1528–1533.

(92) Yan, N.; Cherepanov, P.; Daigle, J. E.; Engelman, A.; Lieberman, J. The SET Complex Acts as a Barrier to Autointegration of HIV-1. *PLoS Pathog.* **2009**, *5*, e1000327.

(93) Pahl, A.; Flügel, R. M. Endonucleolytic Cleavages and DNA-Joining Activities of the Integration Protein of Human Foamy Virus. *J. Virol.* **1993**, *67*, 5426–5434.

(94) Sinha, S.; Pursley, M. H.; Grandgenett, D. P. Efficient Concerted Integration by Recombinant Human Immunodeficiency Virus Type 1 Integrase without Cellular or Viral Cofactors. *J. Virol.* **2002**, *76* (7), 3105–3113.

(95) Sinha, S.; Grandgenett, D. P. Recombinant Human Immunodeficiency Virus Type 1 Integrase Exhibits a Capacity for Full-Site Integration in Vitro That Is Comparable to That of Purified Preintegration Complexes from Virus-Infected Cells. *J. Virol.* **2005**, *79* (13), 8208–8216.

(96) Li, M.; Craigie, R. Processing of Viral DNA Ends Channels the HIV-1 Integration Reaction to Concerted Integration. *J. Biol. Chem.* **2005**, *280* (32), 29334–29339.

(97) Hindmarsh, P.; Ridky, T.; Reeves, R.; Andrade, M.; Skalka, A. M.; Leis, J. HMG Protein Family Members Stimulate Human Immunodeficiency Virus Type 1 and Avian Sarcoma Virus Concerted DNA Integration in Vitro. *J. Virol.* **1999**, *73* (4), 2994–3003.

(98) Hare, S.; Shun, M. C.; Gupta, S. S.; Valkov, E.; Engelman, A.; Cherepanov, P. A Novel Co-Crystal Structure Affords the Design of Gain-of-Function Lentiviral Integrase Mutants in the Presence of Modified PSIP1/LEDGF/p75. *PLoS Pathog.* **2009**, *5* (1), e1000259.

(99) Pandey, K. K.; Bera, S.; Grandgenett, D. P. The HIV-1 Integrase Monomer Induces a Specific Interaction with LTR DNA for Concerted Integration. *Biochemistry* **2011**, *50* (45), 9788–9796.

(100) Li, M.; Ivanov, V.; Mizuuchi, M.; Mizuuchi, K.; Craigie, R. DNA Requirements for Assembly and Stability of HIV-1 Intasomes. *Protein Sci.* **2012**, *21*, 249–257.

(101) Li, M.; Jurado, K. A.; Lin, S.; Engelman, A.; Craigie, R. Engineered Hyperactive Integrase for Concerted HIV-1 DNA Integration. *PLoS One* **2014**, *9*, e105078.

(102) Cherepanov, P. LEDGF/p75 Interacts with Divergent Lentiviral Integrases and Modulates Their Enzymatic Activity in Vitro. *Nucleic Acids Res.* **2007**, *35* (1), 113–124.

(103) Valkov, E.; Gupta, S. S.; Hare, S.; Helander, A.; Roversi, P.; McClure, M.; Cherepanov, P. Functional and Structural Characterization of the Integrase from the Prototype Foamy Virus. *Nucleic Acids Res.* **2009**, *37* (1), 243–255.

(104) Ballandras-Colas, A.; Naraharisetty, H.; Li, X.; Serrao, E.; Engelman, A. Biochemical Characterization of Novel Retroviral Integrase Proteins. *PLoS One* **2013**, *8* (10), e76638.

(105) Deprez, E.; Tauc, P.; Leh, H.; Mouscadet, J. F.; Auclair, C.; Brochon, J. C. Oligomeric States of the HIV-1 Integrase as Measured by Time-Resolved Fluorescence Anisotropy. *Biochemistry* **2000**, *39* (31), 9275–9284.

(106) Chow, S. A.; Vincent, K. A.; Ellison, V.; Brown, P. O. Reversal of Integration and DNA Splicing Mediated by Integrase of Human Immunodeficiency Virus. *Science* **1992**, *255* (5045), 723–726.

(107) van Gent, D. C.; Groeneger, A. A.; Plasterk, R. H. Mutational Analysis of the Integrase Protein of Human Immunodeficiency Virus Type 2. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (20), 9598–9602.

(108) Engelman, A.; Craigie, R. Identification of Conserved Amino Acid Residues Critical for Human Immunodeficiency Virus Type 1 Integrase Function in Vitro. *J. Virol.* **1992**, *66* (11), 6361–6369.

(109) Leavitt, A. D.; Shiue, L.; Varmus, H. E. Site-Directed Mutagenesis of HIV-1 Integrase Demonstrates Differential Effects on Integrase Functions in Vitro. *J. Biol. Chem.* **1993**, *268* (3), 2113–2119.

(110) Vink, C.; Oude Groeneger, A. M.; Plasterk, R. H. Identification of the Catalytic and DNA-Binding Region of the Human Immunodeficiency Virus Type 1 Integrase Protein. *Nucleic Acids Res.* **1993**, *21* (6), 1419–1425.

(111) Gerton, J. L.; Brown, P. O. The Core Domain of HIV-1 Integrase Recognizes Key Features of Its DNA Substrates. *J. Biol. Chem.* **1997**, *272* (41), 25809–25815.

(112) Bushman, F. D.; Engelman, A.; Palmer, I.; Wingfield, P.; Craigie, R. Domains of the Integrase Protein of Human Immunodeficiency Virus Type 1 Responsible for Polynucleotidyl Transfer and Zinc Binding. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 3428–3432.

(113) Bushman, F. D.; Wang, B. Rous Sarcoma Virus Integrase Protein: Mapping Functions for Catalysis and Substrate Binding. *J. Virol.* **1994**, *68*, 2215–2223.

(114) Kulkosky, J.; Katz, R. A.; Merkel, G.; Skalka, A. M. Activities and Substrate Specificity of the Evolutionarily Conserved Central Domain of Retroviral Integrase. *Virology* **1995**, *206*, 448–456.

(115) Donzella, G. A.; Jonsson, C. B.; Roth, M. J. Coordinated Disintegration Reactions Mediated by Moloney Murine Leukemia Virus Integrase. *J. Virol.* **1996**, *70* (6), 3909–3921.

(116) Levchenko, I.; Luo, L.; Baker, T. A. Disassembly of the Mu Transposase Tetramer by the ClpX Chaperone. *Genes Dev.* **1995**, *9* (19), 2399–2408.

(117) Burton, B. M.; Baker, T. A. Remodeling Protein Complexes: Insights from the AAA+ Unfoldase ClpX and Mu Transposase. *Protein Sci.* **2005**, *14*, 1945–1954.

(118) Skalka, A. M.; Katz, R. A. Retroviral DNA Integration and the DNA Damage Response. *Cell Death Differ.* **2005**, *12*, 971–978.

(119) Mulder, L. C.; Muesing, M. A. Degradation of HIV-1 Integrase by the N-End Rule Pathway. *J. Biol. Chem.* **2000**, *275* (38), 29749–29753.

(120) Devroe, E.; Engelman, A.; Silver, P. A. Intracellular Transport of Human Immunodeficiency Virus Type 1 Integrase. *J. Cell Sci.* **2003**, *116* (Pt 21), 4401–4408.

(121) Llano, M.; Delgado, S.; Vanegas, M.; Poeschla, E. M. Lens Epithelium-Derived Growth Factor/p75 Prevents Proteasomal Degradation of HIV-1 Integrase. *J. Biol. Chem.* **2004**, *279* (53), 55570–55577.

(122) Mulder, L. C.; Chakrabarti, L. A.; Muesing, M. A. Interaction of HIV-1 Integrase with DNA Repair Protein hRAD18. *J. Biol. Chem.* **2002**, *277* (30), 27489–27493.

(123) Lloyd, A. G.; Tateishi, S.; Bieniasz, P. D.; Muesing, M. A.; Yamaizumi, M.; Mulder, L. C. Effect of DNA Repair Protein Rad18 on Viral Infection. *PLoS Pathog.* **2006**, *2* (5), e40.

(124) Mousnier, A.; Kubat, N.; Massias-Simon, A.; Segeral, E.; Rain, J. C.; Benarous, R.; Emiliani, S.; Dargemont, C. Von Hippel Lindau Binding Protein 1-Mediated Degradation of Integrase Affects HIV-1

Gene Expression at a Postintegration Step. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (34), 13615–13620.

(125) Yoder, K. E.; Bushman, F. D. Repair of Gaps in Retroviral DNA Integration Intermediates. *J. Virol.* **2000**, *74* (23), 11191–11200.

(126) Faust, E. A.; Triller, H. Stimulation of Human Flap Endonuclease 1 by Human Immunodeficiency Virus Type 1 Integrase: Possible Role for Flap Endonuclease 1 in 5'-End Processing of Human Immunodeficiency Virus Type 1 Integration Intermediates. *J. Biomed. Sci.* **2002**, *9* (3), 273–287.

(127) Brin, E.; Yi, J.; Skalka, A. M.; Leis, J. Modeling the Late Steps in HIV-1 Retroviral Integrase-Catalyzed DNA Integration. *J. Biol. Chem.* **2000**, *275* (50), 39287–39295.

(128) Rumbaugh, J. A.; Fuentes, G. M.; Bambara, R. A. Processing of an HIV Replication Intermediate by the Human DNA Replication Enzyme FEN1. *J. Biol. Chem.* **1998**, *273* (44), 28740–28745.

(129) Studamire, B.; Goff, S. P. Host Proteins Interacting with the Moloney Murine Leukemia Virus Integrase: Multiple Transcriptional Regulators and Chromatin Binding Factors. *Retrovirology* **2008**, *5*, 48.

(130) Espeseth, A. S.; Fishel, R.; Hazuda, D.; Huang, Q.; Xu, M.; Yoder, K.; Zhou, H. siRNA Screening of a Targeted Library of DNA Repair Factors in HIV Infection Reveals a Role for Base Excision Repair in HIV Integration. *PLoS One* **2011**, *6*, e17612.

(131) Yoder, K. E.; Espeseth, A.; Wang, X. H.; Fang, Q.; Russo, M. T.; Lloyd, R. S.; Hazuda, D.; Sobol, R. W.; Fishel, R. The Base Excision Repair Pathway Is Required for Efficient Lentivirus Integration. *PLoS One* **2011**, *6*, e17862.

(132) Daniel, R.; Katz, R. A.; Merkel, G.; Hittle, J. C.; Yen, T. J.; Skalka, A. M. Wortmannin Potentiates Integrase-Mediated Killing of Lymphocytes and Reduces the Efficiency of Stable Transduction by Retroviruses. *Mol. Cell. Biol.* **2001**, *21* (4), 1164–1172.

(133) Daniel, R.; Katz, R. A.; Skalka, A. M. A Role for DNA-PK in Retroviral DNA Integration. *Science* **1999**, *284* (5414), 644–647.

(134) Cooper, A.; García, M.; Petrovas, C.; Yamamoto, T.; Koup, R. A.; Nabel, G. J. HIV-1 Causes CD4 Cell Death through DNA-Dependent Protein Kinase During Viral Integration. *Nature* **2013**, *498*, 376–379.

(135) Majors, J. E.; Varmus, H. E. Nucleotide Sequences at Host-Proviral Junctions for Mouse Mammary Tumour Virus. *Nature* **1981**, *289* (5795), 253–258.

(136) Dhar, R.; McClements, W. L.; Enquist, L. W.; Vande Woude, G. F. Nucleotide Sequences of Integrated Moloney Sarcoma Provirus Long Terminal Repeats and Their Host and Viral Junctions. *Proc. Natl. Acad. Sci. U. S. A.* **1980**, *77*, 3937–3941.

(137) Hishinuma, F.; DeBona, P. J.; Astrin, S.; Skalka, A. M. Nucleotide Sequence of Acceptor Site and Termini of Integrated Avian Endogenous Provirus Ev1: Integration Creates a 6 Bp Repeat of Host DNA. *Cell* **1981**, *23*, 155–164.

(138) Hughes, S. H.; Mutschler, A.; Bishop, J. M.; Varmus, H. E. A Rous Sarcoma Virus Provirus Is Flanked by Short Direct Repeats of a Cellular DNA Sequence Present in Only One Copy Prior to Integration. *Proc. Natl. Acad. Sci. U. S. A.* **1981**, *78*, 4299–5303.

(139) Chou, K. S.; Okayama, A.; Su, I. J.; Lee, T. H.; Essex, M. Preferred Nucleotide Sequence at the Integration Target Site of Human T-Cell Leukemia Virus Type I from Patients with Adult T-Cell Leukemia. *Int. J. Cancer* **1996**, *65*, 20–24.

(140) Chou, K. S.; Okayama, A.; Tachibana, N.; Lee, T. H.; Essex, M. Nucleotide Sequence Analysis of a Full-Length Human T-Cell Leukemia Virus Type I from Adult T-Cell Leukemia Cells: A Prematurely Terminated PX Open Reading Frame II. *Int. J. Cancer* **1995**, *60*, 701–706.

(141) Schweizer, M.; Fleps, U.; Jäckle, A.; Renne, R.; Turek, R.; Neumann-Haefelin, D. Simian Foamy Virus Type 3 (SFV-3) in Latently Infected Vero Cells: Reactivation by Demethylation of Proviral DNA. *Virology* **1993**, *192*, 663–666.

(142) Neves, M.; Périès, J.; Saib, A. Study of Human Foamy Virus Proviral Integration in Chronically Infected Murine Cells. *Res. Virol.* **1998**, *149*, 393–401.

(143) Vincent, K. A.; York-Higgins, D.; Quiroga, M.; Brown, P. O. Host Sequences Flanking the HIV Provirus. *Nucleic Acids Res.* **1990**, *18* (20), 6045–6047.

(144) Vink, C.; Groenink, M.; Elgersma, Y.; Fouchier, R. A.; Tersmette, M.; Plasterk, R. H. Analysis of the Junctions between Human Immunodeficiency Virus Type 1 Proviral DNA and Human DNA. *J. Virol.* **1990**, *64* (11), 5626–5627.

(145) Regier, D. A.; Desrosiers, R. C. The Complete Nucleotide Sequence of a Pathogenic Molecular Clone of Simian Immunodeficiency Virus. *AIDS Res. Hum. Retroviruses* **1990**, *6* (11), 1221–1231.

(146) Hacker, C. V.; Vink, C. A.; Wardell, T. W.; Lee, S.; Treasure, P.; Kingsman, S. M.; Mitrophanous, K. A.; Miskin, J. E. The Integration Profile of EIAV-Based Vectors. *Mol. Ther.* **2006**, *14* (4), 536–545.

(147) Cai, M.; Zheng, R.; Caffrey, M.; Craigie, R.; Clore, G. M.; Gronenborn, A. M. Solution Structure of the N-Terminal Zinc Binding Domain of HIV-1 Integrase [Published Erratum Appears in Nat Struct Biol 1997 Oct;4(10):839–40]. *Nat. Struct. Biol.* **1997**, *4* (7), 567–577.

(148) Dyda, F.; Hickman, A. B.; Jenkins, T. M.; Engelman, A.; Craigie, R.; Davies, D. R. Crystal Structure of the Catalytic Domain of HIV-1 Integrase: Similarity to Other Polynucleotidyl Transferases [See Comments]. *Science* **1994**, *266* (5193), 1981–1986.

(149) Eijkelenboom, A. P.; Puras Lutzke, R. A.; Boelens, R.; Plasterk, R. H.; Kaptein, R.; Hard, K. The DNA-Binding Domain of HIV-1 Integrase Has an SH3-Like Fold. *Nat. Struct. Biol.* **1995**, *2* (9), 807–810.

(150) Eijkelenboom, A. P.; van den Ent, F. M.; Vos, A.; Doreleijers, J. F.; Hard, K.; Tullius, T. D.; Plasterk, R. H.; Kaptein, R.; Boelens, R. The Solution Structure of the Amino-Terminal HHCC Domain of HIV-2 Integrase: A Three-Helix Bundle Stabilized by Zinc. *Curr. Biol.* **1997**, *7* (10), 739–746.

(151) Lodi, P. J.; Ernst, J. A.; Kuszewski, J.; Hickman, A. B.; Engelman, A.; Craigie, R.; Clore, G. M.; Gronenborn, A. M. Solution Structure of the DNA Binding Domain of HIV-1 Integrase. *Biochemistry* **1995**, *34* (31), 9826–9833.

(152) Bujacz, G.; Jaskolski, M.; Alexandratos, J.; Wlodawer, A.; Merkel, G.; Katz, R. A.; Skalka, A. M. High-Resolution Structure of the Catalytic Domain of Avian Sarcoma Virus Integrase. *J. Mol. Biol.* **1995**, *253* (2), 333–346.

(153) Aiyer, S.; Rossi, P.; Malani, N.; Schneider, W. M.; Chandar, A.; Bushman, F. D.; Montelione, G. T.; Roth, M. J. Structural and Sequencing Analysis of Local Target DNA Recognition by MLV Integrase. *Nucleic Acids Res.* **2015**, *43* (11), 5647–5663.

(154) Esposito, D.; Craigie, R. HIV Integrase Structure and Function. *Adv. Virus Res.* **1999**, *52*, 319–333.

(155) Ballandras-Colas, A.; Brown, M.; Cook, N. J.; Dewdney, T. G.; Demeler, B.; Cherepanov, P.; Lyumkis, D.; Engelman, A. N. Cryo-EM Reveals a Novel Octameric Integrase Structure for Betaretroviral Intasome Function. *Nature* **2016**, *530* (7590), 358–361.

(156) Hare, S.; Maertens, G. N.; Cherepanov, P. 3'-Processing and Strand Transfer Catalysed by Retroviral Integrase in Crystallo. *EMBO J.* **2012**, *31* (13), 3020–3028.

(157) Maertens, G. N.; Hare, S.; Cherepanov, P. The Mechanism of Retroviral Integration from X-Ray Structures of Its Key Intermediates. *Nature* **2010**, *468* (7321), 326–329.

(158) Li, X.; Krishnan, L.; Cherepanov, P.; Engelman, A. Structural Biology of Retroviral DNA Integration. *Virology* **2011**, *411*, 194–205.

(159) Kulkosky, J.; Jones, K. S.; Katz, R. A.; Mack, J. P.; Skalka, A. M. 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.* **1992**, *12* (5), 2331–2338.

(160) Nowotny, M. Retroviral Integrase Superfamily: The Structural Perspective. *EMBO Rep.* **2009**, *10* (2), 144–151.

(161) Kuo, C. F.; Zou, A. H.; Jayaram, M.; Getzoff, E.; Harshey, R. DNA-Protein Complexes During Attachment-Site Synapsis in Mu DNA Transposition. *EMBO J.* **1991**, *10* (6), 1585–1591.

(162) Harshey, R. M. The Mu Story: How a Maverick Phage Moved the Field Forward. *Mobile DNA* **2012**, *3* (1), 21.

- (163) Baker, T. A.; Mizuuchi, M.; Savilahti, H.; Mizuuchi, K. Division of Labor among Monomers within the Mu Transposase Tetramer. *Cell* **1993**, *74* (4), 723–733.
- (164) Jacks, T.; Power, M. D.; Masiarz, F. R.; Luciw, P. A.; Barr, P. J.; Varmus, H. E. Characterization of Ribosomal Frameshifting in HIV-1 Gag-Pol Expression. *Nature* **1988**, *331*, 280–283.
- (165) Briggs, J. A.; Simon, M. N.; Gross, I.; Kräusslich, H. G.; Fuller, S. D.; Vogt, V. M.; Johnson, M. C. The Stoichiometry of Gag Protein in HIV-1. *Nat. Struct. Mol. Biol.* **2004**, *11*, 672–675.
- (166) van Gent, D. C.; Elgersma, Y.; Bolk, M. W. J.; Vink, C.; Plasterk, R. H. A. DNA Binding Properties of the Integrase Proteins of Human Immunodeficiency Viruses Types 1 and 2. *Nucleic Acids Res.* **1991**, *19* (14), 3821–3827.
- (167) Vincent, K. A.; Ellison, V.; Chow, S. A.; Brown, P. O. Characterization of Human Immunodeficiency Virus Type 1 Integrase Expressed in *Escherichia Coli* and Analysis of Variants with Amino-Terminal Mutations. *J. Virol.* **1993**, *67* (1), 425–437.
- (168) Engelman, A.; Bushman, F. D.; Craigie, R. Identification of Discrete Functional Domains of HIV-1 Integrase and Their Organization within an Active Multimeric Complex. *EMBO J.* **1993**, *12*, 3269–3275.
- (169) Li, M.; Mizuuchi, M.; Burke, T. R., Jr.; Craigie, R. Retroviral DNA Integration: Reaction Pathway and Critical Intermediates. *EMBO J.* **2006**, *25* (6), 1295–1304.
- (170) Faure, A.; Calmels, C.; Desjobert, C.; Castroviejo, M.; Caumont-Sarcos, A.; Tarrago-Litvak, L.; Litvak, S.; Parissi, V. HIV-1 Integrase Crosslinked Oligomers Are Active in Vitro. *Nucleic Acids Res.* **2005**, *33* (3), 977–986.
- (171) Cherepanov, P.; Maertens, G.; Proost, P.; Devreese, B.; Van Beeumen, J.; Engelborghs, Y.; De Clercq, E.; Debyser, Z. HIV-1 Integrase Forms Stable Tetramers and Associates with LEDGF/p75 Protein in Human Cells. *J. Biol. Chem.* **2003**, *278* (1), 372–381.
- (172) Bera, S.; Pandey, K. K.; Vora, A. C.; Grandgenett, D. P. Molecular Interactions between HIV-1 Integrase and the Two Viral DNA Ends within the Synaptic Complex That Mediates Concerted Integration. *J. Mol. Biol.* **2009**, *389*, 183–198.
- (173) Hickman, A. B.; Ewis, H. E.; Li, X.; Knapp, J. A.; Laver, T.; Doss, A. L.; Tolun, G.; Steven, A. C.; Grishaev, A.; Bax, A.; et al. Structural Basis of Hat Transposon End Recognition by Hermes, an Octameric DNA Transposase from *Musca Domestica*. *Cell* **2014**, *158* (2), 353–367.
- (174) Baker, T. A.; Mizuuchi, K. DNA-Promoted Assembly of the Active Tetramer of the Mu Transposase. *Genes Dev.* **1992**, *6* (11), 2221–2232.
- (175) Bhasin, A.; Goryshin, I. Y.; Steiniger-White, M.; York, D.; Reznikoff, W. S. Characterization of a Tn5 Pre-Cleavage Synaptic Complex. *J. Mol. Biol.* **2000**, *302* (1), 49–63.
- (176) Richardson, J. M.; Colloms, S. D.; Finnegan, D. J.; Walkinshaw, M. D. Molecular Architecture of the Mos1 Paired-End Complex: The Structural Basis of DNA Transposition in a Eukaryote. *Cell* **2009**, *138* (6), 1096–1108.
- (177) Montano, S. P.; Pigli, Y. Z.; Rice, P. A. The Mu Transpososome Structure Sheds Light on DDE Recombinase Evolution. *Nature* **2012**, *491* (7424), 413–417.
- (178) Lee, S. P.; Xiao, J.; Knutson, J. R.; Lewis, M. S.; Han, M. K. Zn²⁺ Promotes the Self-Association of Human Immunodeficiency Virus Type-1 Integrase in Vitro. *Biochemistry* **1997**, *36* (1), 173–180.
- (179) Heuer, T. S.; Brown, P. O. Photo-Cross-Linking Studies Suggest a Model for the Architecture of an Active Human Immunodeficiency Virus Type 1 Integrase-DNA Complex. *Biochemistry* **1998**, *37* (19), 6667–6678.
- (180) Gupta, K.; Curtis, J. E.; Krueger, S.; Hwang, Y.; Cherepanov, P.; Bushman, F. D.; Van Duyn, G. D. Solution Conformations of Prototype Foamy Virus Integrase and Its Stable Synaptic Complex with U5 Viral DNA. *Structure* **2012**, *20* (11), 1918–1928.
- (181) Yin, Z.; Shi, K.; Banerjee, S.; Pandey, K. K.; Bera, S.; Grandgenett, D. P.; Aihara, H. Crystal Structure of the Rous Sarcoma Virus Intasome. *Nature* **2016**, *530* (7590), 362–366.
- (182) van Gent, D. C.; Vink, C.; Groeneger, A. A.; Plasterk, R. H. Complementation between HIV Integrase Proteins Mutated in Different Domains. *EMBO J.* **1993**, *12*, 3261–3267.
- (183) Pahl, A.; Flügel, R. M. Characterization of the Human Spuma Retrovirus Integrase by Site-Directed Mutagenesis, by Complementation Analysis, and by Swapping the Zinc Finger Domain of HIV-1. *J. Biol. Chem.* **1995**, *270* (7), 2957–2966.
- (184) Jonsson, C. B.; Donzella, G. A.; Gaucan, E.; Smith, C. M.; Roth, M. J. Functional Domains of Moloney Murine Leukemia Virus Integrase Defined by Mutation and Complementation Analysis. *J. Virol.* **1996**, *70* (7), 4585–4597.
- (185) Wang, J. Y.; Ling, H.; Yang, W.; Craigie, R. Structure of a Two-Domain Fragment of HIV-1 Integrase: Implications for Domain Organization in the Intact Protein. *EMBO J.* **2001**, *20* (24), 7333–7343.
- (186) Hare, S.; Di Nunzio, F.; Labeja, A.; Wang, J.; Engelman, A.; Cherepanov, P. Structural Basis for Functional Tetramerization of Lentiviral Integrase. *PLoS Pathog.* **2009**, *5* (7), e1000515.
- (187) Jaskolski, M.; Alexandratos, J. N.; Bujacz, G.; Wlodawer, A. Piecing Together the Structure of Retroviral Integrase, an Important Target in AIDS Therapy. *FEBS J.* **2009**, *276* (11), 2926–2946.
- (188) Engelman, A.; Cherepanov, P. Retroviral Integrase Structure and DNA Recombination Mechanism. *Microbiol. Spectr.* **2014**, *2*, 1–22.
- (189) Delelis, O.; Carayon, K.; Guiot, E.; Leh, H.; Tauc, P.; Brochon, J. C.; Mouscadet, J. F.; Deprez, E. Insight into the Integrase-DNA Recognition Mechanism. A Specific DNA-Binding Mode Revealed by an Enzymatically Labeled Integrase. *J. Biol. Chem.* **2008**, *283* (41), 27838–27849.
- (190) Yin, Z.; Lapkouski, M.; Yang, W.; Craigie, R. Assembly of Prototype Foamy Virus Strand Transfer Complexes on Product DNA Bypassing Catalysis of Integration. *Protein Sci.* **2012**, *21* (12), 1849–1857.
- (191) Maskell, D. P.; Renault, L.; Serrao, E.; Lesbats, P.; Matadeen, R.; Hare, S.; Lindemann, D.; Engelman, A. N.; Costa, A.; Cherepanov, P. Structural Basis for Retroviral Integration into Nucleosomes. *Nature* **2015**, *523*, 366–369.
- (192) Li, M.; Lin, S.; Craigie, R. Outer Domains of Integrase within Retroviral Intasomes Are Dispensable for Catalysis of DNA Integration. *Protein Sci.* **2016**, *25* (2), 472–478.
- (193) Holman, A. G.; Coffin, J. M. Symmetrical Base Preferences Surrounding HIV-1, Avian Sarcoma/Leukosis Virus, and Murine Leukemia Virus Integration Sites. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (17), 6103–6107.
- (194) Serrao, E.; Ballandras-Colas, A.; Cherepanov, P.; Maertens, G. N.; Engelman, A. N. Key Determinants of Target DNA Recognition by Retroviral Intasomes. *Retrovirology* **2015**, *12*, 39.
- (195) Pruss, D.; Bushman, F. D.; Wolffe, A. P. Human Immunodeficiency Virus Integrase Directs Integration to Sites of Severe DNA Distortion within the Nucleosome Core. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (13), 5913–5917.
- (196) Pruss, D.; Reeves, R.; Bushman, F. D.; Wolffe, A. P. The Influence of DNA and Nucleosome Structure on Integration Events Directed by HIV Integrase. *J. Biol. Chem.* **1994**, *269* (40), 25031–25041.
- (197) Benleulmi, M. S.; Matysiak, J.; Henriquez, D. R.; Vaillant, C.; Lesbats, P.; Calmels, C.; Naughtin, M.; Leon, O.; Skalka, A. M.; Ruff, M.; et al. Intasome Architecture and Chromatin Density Modulate Retroviral Integration into Nucleosome. *Retrovirology* **2015**, *12*, 13.
- (198) Pryciak, P. M.; Sil, A.; Varmus, H. E. Retroviral Integration into Minichromosomes in Vitro. *EMBO J.* **1992**, *11* (1), 291–303.
- (199) Pryciak, P. M.; Müller, H.-P.; Varmus, H. E. Simian Virus 40 minichromosomes as Targets for Retroviral Integration in Vivo. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (19), 9237–9241.
- (200) Pryciak, M. P.; Varmus, H. E. Nucleosomes, DNA-Binding Proteins, and DNA Sequence Modulate Retroviral Integration Target Site Selection. *Cell* **1992**, *69*, 769–780.
- (201) Wang, G. P.; Ciuffi, A.; Leipzig, J.; Berry, C. C.; Bushman, F. D. HIV Integration Site Selection: Analysis by Massively Parallel

Pyrosequencing Reveals Association with Epigenetic Modifications. *Genome Res.* **2007**, *17* (8), 1186–1194.

(202) Roth, S. L.; Malani, N.; Bushman, F. D. Gammaretroviral Integration into Nucleosomal Target DNA in Vivo. *J. Virol.* **2011**, *85* (14), 7393–7401.

(203) Luger, K.; Mader, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J. Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution. *Nature* **1997**, *389* (6648), 251–260.

(204) Greenwald, J.; Le, V.; Butler, S. L.; Bushman, F. D.; Choe, S. The Mobility of an HIV-1 Integrase Active Site Loop Is Correlated with Catalytic Activity. *Biochemistry* **1999**, *38* (28), 8892–8898.

(205) Krishnan, L.; Li, X.; Naraharisetty, H. L.; Hare, S.; Cherepanov, P.; Engelman, A. Structure-Based Modeling of the Functional HIV-1 Intasome and Its Inhibition. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (36), 15910–15915.

(206) Hare, S.; Vos, A. M.; Clayton, R. F.; Thuring, J. W.; Cummings, M. D.; Cherepanov, P. Molecular Mechanisms of Retroviral Integrase Inhibition and the Evolution of Viral Resistance. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (46), 20057–20062.

(207) Engelman, A.; Craigie, R. Efficient Magnesium-Dependent Human Immunodeficiency Virus Type 1 Integrase Activity. *J. Virol.* **1995**, *69* (9), 5908–5911.

(208) Yang, W.; Lee, J. Y.; Nowotny, M. Making and Breaking Nucleic Acids: Two-Mg²⁺-Ion Catalysis and Substrate Specificity. *Mol. Cell* **2006**, *22* (1), 5–13.

(209) Beese, L. S.; Steitz, T. A. Structural Basis for the 3'-5' Exonuclease Activity of Escherichia Coli DNA Polymerase I: A Two Metal Ion Mechanism. *EMBO J.* **1991**, *10* (1), 25–33.

(210) Steitz, T. A.; Steitz, J. A. A General Two-Metal-Ion Mechanism for Catalytic RNA. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (14), 6498–6502.

(211) Nowotny, M.; Gaidamakov, S. A.; Crouch, R. J.; Yang, W. Crystal Structures of RNase H Bound to an RNA/DNA Hybrid: Substrate Specificity and Metal-Dependent Catalysis. *Cell* **2005**, *121* (7), 1005–1016.

(212) Nowotny, M.; Yang, W. Stepwise Analyses of Metal Ions in RNase H Catalysis from Substrate Destabilization to Product Release. *EMBO J.* **2006**, *25* (9), 1924–1933.

(213) Rosta, E.; Nowotny, M.; Yang, W.; Hummer, G. Catalytic Mechanism of RNA Backbone Cleavage by Ribonuclease H from Quantum Mechanics/Molecular Mechanics Simulations. *J. Am. Chem. Soc.* **2011**, *133* (23), 8934–8941.

(214) Madhusudan; Akamine, P.; Xuong, N. H.; Taylor, S. S. Crystal Structure of a Transition State Mimic of the Catalytic Subunit of Camp-Dependent Protein Kinase. *Nat. Struct. Biol.* **2002**, *9* (4), 273–277.

(215) Bao, Z. Q.; Jacobsen, D. M.; Young, M. A. Briefly Bound to Activate: Transient Binding of a Second Catalytic Magnesium Activates the Structure and Dynamics of Cdk2 Kinase for Catalysis. *Structure* **2011**, *19* (5), 675–690.

(216) Bastidas, A. C.; Deal, M. S.; Steichen, J. M.; Guo, Y.; Wu, J.; Taylor, S. S. Phosphoryl Transfer by Protein Kinase A Is Captured in a Crystal Lattice. *J. Am. Chem. Soc.* **2013**, *135* (12), 4788–4798.

(217) Doublet, S.; Tabor, S.; Long, A. M.; Richardson, C. C.; Ellenberger, T. Crystal Structure of a Bacteriophage T7 DNA Replication Complex at 2.2 Å Resolution. *Nature* **1998**, *391* (6664), 251–258.

(218) Yin, Y. W.; Steitz, T. A. The Structural Mechanism of Translocation and Helicase Activity in T7 RNA Polymerase. *Cell* **2004**, *116* (3), 393–404.

(219) Stahley, M. R.; Strobel, S. A. Structural Evidence for a Two-Metal-Ion Mechanism of Group I Intron Splicing. *Science* **2005**, *309* (5740), 1587–1590.

(220) Gangadharan, S.; Mularoni, L.; Fain-Thornton, J.; Wheelan, S. J.; Craig, N. L. DNA Transposon Hermes Inserts into DNA in Nucleosome-Free Regions in Vivo. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (51), 21966–21972.

(221) Pribil, P. A.; Haniford, D. B. Target DNA Bending Is an Important Specificity Determinant in Target Site Selection in Tn10 Transposition. *J. Mol. Biol.* **2003**, *330* (2), 247–259.

(222) Hallet, B.; Rezsöházy, R.; Mahilton, J.; Delcour, J. IS231A Insertion Specificity: Consensus Sequence and DNA Bending at the Target Site. *Mol. Microbiol.* **1994**, *14* (1), 131–139.

(223) Johnson, B. C.; Métifiot, M.; Ferris, A.; Pommier, Y.; Hughes, S. H. A Homology Model of HIV-1 Integrase and Analysis of Mutations Designed to Test the Model. *J. Mol. Biol.* **2013**, *425*, 2133–2146.

(224) Jordan, A.; Defechereux, P.; Verdin, E. The Site of HIV-1 Integration in the Human Genome Determines Basal Transcriptional Activity and Response to Tat Transactivation. *EMBO J.* **2001**, *20* (7), 1726–1738.

(225) Jordan, A.; Bisgrove, D.; Verdin, E. HIV Reproducibly Establishes a Latent Infection after Acute Infection of T Cells in Vitro. *EMBO J.* **2003**, *22* (8), 1868–1877.

(226) Gerard, A.; Segéral, E.; Naughtin, M.; Abdouni, A.; Charmeteau, B.; Cheynier, R.; Rain, J. C.; Emiliani, S. The Integrase Cofactor LEDGF/p75 Associates with Iws1 and Spt6 for Post-integration Silencing of HIV-1 Gene Expression in Latently Infected Cells. *Cell Host Microbe* **2015**, *17* (1), 107–117.

(227) Maldarelli, F.; Wu, X.; Su, L.; Simonetti, F. R.; Shao, W.; Hill, S.; Spindler, J.; Ferris, A. L.; Mellors, J. W.; Kearney, M. F.; et al. Specific HIV Integration Sites Are Linked to Clonal Expansion and Persistence of Infected Cells. *Science* **2014**, *345* (6193), 179–183.

(228) Wagner, T. A.; McLaughlin, S.; Garg, K.; Cheung, C. Y.; Larsen, B. B.; Styrchak, S.; Huang, H. C.; Edlefsen, P. T.; Mullins, J. I.; Frenkel, L. M. Proliferation of Cells with HIV Integrated into Cancer Genes Contributes to Persistent Infection. *Science* **2014**, *345* (6196), 570–573.

(229) Wong, J. K.; Hezareh, M.; Gunthard, H. F.; Havlir, D. V.; Ignacio, C. C.; Spina, C. A.; Richman, D. D. Recovery of Replication-Competent HIV Despite Prolonged Suppression of Plasma Viremia. *Science* **1997**, *278* (5341), 1291–1295.

(230) Finzi, D.; Blankson, J.; Siliciano, J. D.; Margolick, J. B.; Chadwick, K.; Pierson, T.; Smith, K.; Lisiewicz, J.; Lori, F.; Flexner, C.; et al. Latent Infection of CD4+ T Cells Provides a Mechanism for Lifelong Persistence of HIV-1, Even in Patients on Effective Combination Therapy. *Nat. Med.* **1999**, *5* (5), 512–517.

(231) Wu, X.; Li, Y.; Crise, B.; Burgess, S. M.; Munroe, D. J. Weak Palindromic Consensus Sequences Are a Common Feature Found at the Integration Target Sites of Many Retroviruses. *J. Virol.* **2005**, *79* (8), 5211–5214.

(232) Stevens, S. W.; Griffith, J. D. Sequence Analysis of the Human DNA Flanking Sites of Human Immunodeficiency Virus Type 1 Integration. *J. Virol.* **1996**, *70* (9), 6459–6462.

(233) Stevens, S. W.; Griffith, J. D. Human Immunodeficiency Virus Type 1 May Preferentially Integrate into Chromatin Occupied by L1Hs Repetitive Elements. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (12), 5557–5561.

(234) Panet, A.; Cedar, H. Selective Degradation of Integrated Murine Leukemia Proviral DNA by Deoxyribonucleases. *Cell* **1977**, *11* (4), 933–940.

(235) Vijaya, S.; Steffen, D. L.; Robinson, H. L. Acceptor Sites for Retroviral Integrations Map near DNase I-Hypersensitive Sites in Chromatin. *J. Virol.* **1986**, *60* (2), 683–692.

(236) Rohdewohld, H.; Weiher, H.; Reik, W.; Jaenisch, R.; Breindl, M. Retrovirus Integration and Chromatin Structure: Moloney Murine Leukemia Proviral Integration Sites Map near DNase I-Hypersensitive Sites. *J. Virol.* **1987**, *61* (2), 336–343.

(237) Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; et al. The Sequence of the Human Genome. *Science* **2001**, *291* (5507), 1304–1351.

(238) Lander, E. S.; Linton, L. M.; Birren, B.; Nusbaum, C.; Zody, M. C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; et al. Initial Sequencing and Analysis of the Human Genome. *Nature* **2001**, *409* (6822), 860–921.

- (239) Schroder, A. R.; Shinn, P.; Chen, H.; Berry, C.; Ecker, J. R.; Bushman, F. HIV-1 Integration in the Human Genome Favors Active Genes and Local Hotspots. *Cell* **2002**, *110* (4), 521–529.
- (240) Singh, P. K.; Plumb, M. R.; Ferris, A. L.; Iben, J. R.; Wu, X.; Fadel, H. J.; Luke, B. T.; Esnault, C.; Poeschla, E. M.; Hughes, S. H.; et al. LEDGF/p75 Interacts with mRNA Splicing Factors and Targets HIV-1 Integration to Highly Spliced Genes. *Genes Dev.* **2015**, *29* (21), 2287–2297.
- (241) De Ravin, S. S.; Su, L.; Theobald, N.; Choi, U.; Macpherson, J. L.; Poidinger, M.; Symonds, G.; Pond, S. M.; Ferris, A. L.; Hughes, S. H.; et al. Enhancers Are Major Targets for Murine Leukemia Virus Vector Integration. *J. Virol.* **2014**, *88* (8), 4504–4513.
- (242) LaFave, M. C.; Varshney, G. K.; Gildea, D. E.; Wolfsberg, T. G.; Baxevanis, A. D.; Burgess, S. M. MLV Integration Site Selection Is Driven by Strong Enhancers and Active Promoters. *Nucleic Acids Res.* **2014**, *42* (7), 4257–4269.
- (243) Kvaratskhelia, M.; Sharma, A.; Larue, R. C.; Serrao, E.; Engelman, A. Molecular Mechanisms of Retroviral Integration Site Selection. *Nucleic Acids Res.* **2014**, *42* (16), 10209–10225.
- (244) Mitchell, R. S.; Beitzel, B. F.; Schroder, A. R.; Shinn, P.; Chen, H.; Berry, C. C.; Ecker, J. R.; Bushman, F. D. Retroviral DNA Integration: ASLV, HIV, and MLV Show Distinct Target Site Preferences. *PLoS Biol.* **2004**, *2* (8), e234.
- (245) Wu, X.; Li, Y.; Crise, B.; Burgess, S. M. Transcription Start Regions in the Human Genome Are Favored Targets for MLV Integration. *Science* **2003**, *300* (5626), 1749–1751.
- (246) Lewinski, M. K.; Yamashita, M.; Emerman, M.; Ciuffi, A.; Marshall, H.; Crawford, G.; Collins, F.; Shinn, P.; Leipzig, J.; Hannenhalli, S.; et al. Retroviral DNA Integration: Viral and Cellular Determinants of Target-Site Selection. *PLoS Pathog.* **2006**, *2* (6), e60.
- (247) Trobridge, G. D.; Miller, D. G.; Jacobs, M. A.; Allen, J. M.; Kiem, H.-P.; Kaul, R.; Russell, D. W. Foamy Virus Vector Integration Sites in Normal Human Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (5), 1498–1503.
- (248) Nowrouzi, A.; Dittrich, M.; Klanke, C.; Heinkelein, M.; Rammling, M.; Dandekar, T.; von Kalle, C.; Rethwilm, A. Genome-Wide Mapping of Foamy Virus Vector Integrations into a Human Cell Line. *J. Gen. Virol.* **2006**, *87*, 1339–1347.
- (249) Faschinger, A.; Rouault, F.; Sollner, J.; Lukas, A.; Salmons, B.; Gunzburg, W. H.; Indik, S. Mouse Mammary Tumor Virus Integration Site Selection in Human and Mouse Genomes. *J. Virol.* **2008**, *82* (3), 1360–1367.
- (250) Narezkina, A.; Taganov, K. D.; Litwin, S.; Stoyanova, R.; Hayashi, J.; Seeger, C.; Skalka, A. M.; Katz, R. A. Genome-Wide Analyses of Avian Sarcoma Virus Integration Sites. *J. Virol.* **2004**, *78* (21), 11656–11663.
- (251) Koh, Y.; Wu, X.; Ferris, A. L.; Matreyek, K. A.; Smith, S. J.; Lee, K.; KewalRamani, V. N.; Hughes, S. H.; Engelman, A. Differential Effects of Human Immunodeficiency Virus Type 1 Capsid and Cellular Factors Nucleoporin 153 and LEDGF/p75 on the Efficiency and Specificity of Viral DNA Integration. *J. Virol.* **2013**, *87* (1), 648–658.
- (252) Schaller, T.; Ocwieja, K. E.; Rasaiyaah, J.; Price, A. J.; Brady, T. L.; Roth, S. L.; Hue, S.; Fletcher, A. J.; Lee, K.; KewalRamani, V. N.; et al. HIV-1 Capsid-Cyclophilin Interactions Determine Nuclear Import Pathway, Integration Targeting and Replication Efficiency. *PLoS Pathog.* **2011**, *7* (12), e1002439.
- (253) Bukrinsky, M. I.; Sharova, N.; Dempsey, M. P.; Stanwick, T. L.; Bukrinskaya, A. G.; Haggerty, S.; Stevenson, M. Active Nuclear Import of Human Immunodeficiency Virus Type 1 Preintegration Complexes. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 6580–6584.
- (254) Lewis, P.; Hensel, M.; Emerman, M. Human Immunodeficiency Virus Infection of Cells Arrested in the Cell Cycle. *EMBO J.* **1992**, *11*, 3053–3058.
- (255) Roe, T.; Reynolds, T. C.; Yu, G.; Brown, P. O. Integration of Murine Leukemia Virus DNA Depends on Mitosis. *EMBO J.* **1993**, *12* (5), 2099–2108.
- (256) Lewis, P. F.; Emerman, M. Passage through Mitosis Is Required for Oncoretroviruses but Not for the Human Immunodeficiency Virus. *J. Virol.* **1994**, *68* (1), 510–516.
- (257) Hatzioannou, T.; Goff, S. P. Infection of Nondividing Cells by Rous Sarcoma Virus. *J. Virol.* **2001**, *75* (19), 9526–9531.
- (258) Katz, R. A.; Greger, J. G.; Darby, K.; Boimel, P.; Rall, G. F.; Skalka, A. M. Transduction of Interphase Cells by Avian Sarcoma Virus. *J. Virol.* **2002**, *76* (11), 5422–5434.
- (259) Mergia, A.; Chari, S.; Kolson, D. L.; Goodenow, M. M.; Ciccarone, T. The Efficiency of Simian Foamy Virus Vector Type-1 (SFV-1) in Nondividing Cells and in Human PBLs. *Virology* **2001**, *280*, 243–252.
- (260) Lee, K.; Ambrose, Z.; Martin, T. D.; Oztop, I.; Mulky, A.; Julias, J. G.; Vandegraaff, N.; Baumann, J. G.; Wang, R.; Yuen, W.; et al. Flexible Use of Nuclear Import Pathways by HIV-1. *Cell Host Microbe* **2010**, *7*, 221–233.
- (261) Price, A. J.; Fletcher, A. J.; Schaller, T.; Elliott, T.; Lee, K.; KewalRamani, V. N.; Chin, J. W.; Towers, G. J.; James, L. C. CPSF6 Defines a Conserved Capsid Interface That Modulates HIV-1 Replication. *PLoS Pathog.* **2012**, *8* (8), e1002896.
- (262) Ocwieja, K. E.; Brady, T. L.; Ronen, K.; Huegel, A.; Roth, S. L.; Schaller, T.; James, L. C.; Towers, G. J.; Young, J. A.; Chanda, S. K.; et al. HIV Integration Targeting: A Pathway Involving Transportin-3 and the Nuclear Pore Protein RanBP2. *PLoS Pathog.* **2011**, *7* (3), e1001313.
- (263) Matreyek, K. A.; Yucel, S. S.; Li, X.; Engelman, A. Nucleoporin NUP153 Phenylalanine-Glycine Motifs Engage a Common Binding Pocket within the HIV-1 Capsid Protein to Mediate Lentiviral Infectivity. *PLoS Pathog.* **2013**, *9* (10), e1003693.
- (264) Di Nunzio, F.; Fricke, T.; Miccio, A.; Valle-Casuso, J. C.; Perez, P.; Souque, P.; Rizzi, E.; Severgnini, M.; Mavilio, F.; Charneau, P.; et al. Nup153 and Nup98 Bind the HIV-1 Core and Contribute to the Early Steps of HIV-1 Replication. *Virology* **2013**, *440* (1), 8–18.
- (265) Sowd, G. A.; Serrao, E.; Wang, H.; Wang, W.; Fadel, H. J.; Poeschla, E. M.; Engelman, A. N. A Critical Role for Alternative Polyadenylation Factor CPSF6 in Targeting HIV-1 Integration to Transcriptionally Active Chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (8), E1054–1063.
- (266) Di Primio, C.; Quercioli, V.; Allouch, A.; Gijsbers, R.; Christ, F.; Debyser, Z.; Arosio, D.; Cereseto, A. Single-Cell Imaging of HIV-1 Provirus (SCIP). *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (14), 5636–5641.
- (267) Marini, B.; Kertesz-Farkas, A.; Ali, H.; Lucic, B.; Lisek, K.; Manganaro, L.; Pongor, S.; Luzzati, R.; Recchia, A.; Mavilio, F.; et al. Nuclear Architecture Dictates HIV-1 Integration Site Selection. *Nature* **2015**, *521* (7551), 227–231.
- (268) Nishizawa, Y.; Usukura, J.; Singh, D. P.; Chylack, L. T., Jr.; Shinohara, T. Spatial and Temporal Dynamics of Two Alternatively Spliced Regulatory Factors, Lens Epithelium-Derived Growth Factor (Ledgf/p75) and p52, in the Nucleus. *Cell Tissue Res.* **2001**, *305* (1), 107–114.
- (269) Ge, H.; Si, Y.; Roeder, R. G. Isolation of cDNAs Encoding Novel Transcription Coactivators p52 and p75 Reveals an Alternate Regulatory Mechanism of Transcriptional Activation. *EMBO J.* **1998**, *17* (22), 6723–6729.
- (270) Singh, D. P.; Kimura, A.; Chylack, L. T., Jr.; Shinohara, T. Lens Epithelium-Derived Growth Factor (LEDGF/p75) and p52 Are Derived from a Single Gene by Alternative Splicing. *Gene* **2000**, *242* (1–2), 265–273.
- (271) Emiliani, S.; Mousnier, A.; Busschots, K.; Maroun, M.; Van Maele, B.; Tempe, D.; Vandekerckhove, L.; Moisant, F.; Ben-Slama, L.; Witvrouw, M.; et al. Integrate Mutants Defective for Interaction with LEDGF/p75 Are Impaired in Chromosome Tethering and HIV-1 Replication. *J. Biol. Chem.* **2005**, *280* (27), 25517–25523.
- (272) Busschots, K.; Vercammen, J.; Emiliani, S.; Benarous, R.; Engelborghs, Y.; Christ, F.; Debyser, Z. The Interaction of LEDGF/p75 with Integrase Is Lentivirus-Specific and Promotes DNA Binding. *J. Biol. Chem.* **2005**, *280* (18), 17841–17847.

- (273) Pandey, K. K.; Sinha, S.; Grandgenett, D. P. Transcriptional Coactivator LEDGF/p75 Modulates Human Immunodeficiency Virus Type 1 Integrase-Mediated Concerted Integration. *J. Virol.* **2007**, *81* (8), 3969–3979.
- (274) Cherepanov, P.; Devroe, E.; Silver, P. A.; Engelman, A. Identification of an Evolutionarily Conserved Domain in Human Lens Epithelium-Derived Growth Factor/Transcriptional Co-Activator p75 (LEDGF/p75) That Binds HIV-1 Integrase. *J. Biol. Chem.* **2004**, *279* (47), 48883–48892.
- (275) Vanegas, M.; Llano, M.; Delgado, S.; Thompson, D.; Peretz, M.; Poeschla, E. Identification of the LEDGF/p75 HIV-1 Integrase-Interaction Domain and NLS Reveals NLS-Independent Chromatin Tethering. *J. Cell Sci.* **2005**, *118* (8), 1733–1743.
- (276) Maertens, G.; Cherepanov, P.; Pluyms, W.; Busschots, K.; De Clercq, E.; Debyser, Z.; Engelborghs, Y. LEDGF/p75 Is Essential for Nuclear and Chromosomal Targeting of HIV-1 Integrase in Human Cells. *J. Biol. Chem.* **2003**, *278* (35), 33528–33539.
- (277) Maertens, G.; Cherepanov, P.; Debyser, Z.; Engelborghs, Y.; Engelman, A. Identification and Characterization of a Functional Nuclear Localization Signal in the HIV-1 Integrase Interactor LEDGF/p75. *J. Biol. Chem.* **2004**, *279* (32), 33421–33429.
- (278) Turlure, F.; Maertens, G.; Rahman, S.; Cherepanov, P.; Engelman, A. A Tripartite DNA-Binding Element, Comprised of the Nuclear Localization Signal and Two at-Hook Motifs, Mediates the Association of LEDGF/p75 with Chromatin in Vivo. *Nucleic Acids Res.* **2006**, *34* (5), 1653–1665.
- (279) Llano, M.; Vanegas, M.; Hutchins, N.; Thompson, D.; Delgado, S.; Poeschla, E. M. Identification and Characterization of the Chromatin-Binding Domains of the HIV-1 Integrase Interactor LEDGF/p75. *J. Mol. Biol.* **2006**, *360* (4), 760–773.
- (280) Pradeepa, M. M.; Sutherland, H. G.; Ule, J.; Grimes, G. R.; Bickmore, W. A. Psp1/Ledgf p52 Binds Methylated Histone H3K36 and Splicing Factors and Contributes to the Regulation of Alternative Splicing. *PLoS Genet.* **2012**, *8* (5), e1002717.
- (281) Eidahl, J. O.; Crowe, B. L.; North, J. A.; McKee, C. J.; Shkriabai, N.; Feng, L.; Plumb, M.; Graham, R. L.; Gorelick, R. J.; Hess, S.; et al. Structural Basis for High-Affinity Binding of LEDGF PWWP to Mononucleosomes. *Nucleic Acids Res.* **2013**, *41* (6), 3924–3936.
- (282) van Nuland, R.; van Schaik, F. M.; Simonis, M.; van Heesch, S.; Cuppen, E.; Boelens, R.; Timmers, H. M.; van Ingen, H. Nucleosomal DNA Binding Drives the Recognition of H3K36-Methylated Nucleosomes by the PSIP1-PWWP Domain. *Epigenet. Chromatin* **2013**, *6* (1), 12.
- (283) Sutherland, H. G.; Newton, K.; Brownstein, D. G.; Holmes, M. C.; Kress, C.; Semple, C. A.; Bickmore, W. A. Disruption of Ledgf/Psp1 Results in Perinatal Mortality and Homeotic Skeletal Transformations. *Mol. Cell Biol.* **2006**, *26* (19), 7201–7210.
- (284) Wang, H.; Shun, M. C.; Dickson, A. K.; Engelman, A. N. Embryonic Lethality Due to Arrested Cardiac Development in Psp1/Hdgrp2 Double-Deficient Mice. *PLoS One* **2015**, *10*, e0137797.
- (285) Maertens, G. N.; Cherepanov, P.; Engelman, A. Transcriptional Co-Activator p75 Binds and Tethers the Myc-Interacting Protein JPO2 to Chromatin. *J. Cell Sci.* **2006**, *119* (12), 2563–2571.
- (286) Bartholomeeusen, K.; De Rijck, J.; Busschots, K.; Desender, L.; Gijbsbers, R.; Emiliani, S.; Benarous, R.; Debyser, Z.; Christ, F. Differential Interaction of HIV-1 Integrase and JPO2 with the C Terminus of LEDGF/p75. *J. Mol. Biol.* **2007**, *372*, 407–421.
- (287) Yokoyama, A.; Cleary, M. L. Menin Critically Links MLL Proteins with LEDGF on Cancer-Associated Target Genes. *Cancer Cell* **2008**, *14* (1), 36–46.
- (288) Leoh, L. S.; van Heertum, B.; De Rijck, J.; Philippova, M.; Rios-Colon, L.; Basu, A.; Martinez, S. R.; Tungteakkhun, S. S.; Philippov, V.; Christ, F.; et al. The Stress Oncoprotein LEDGF/p75 Interacts with the Methyl CpG Binding Protein MeCP2 and Influences Its Transcriptional Activity. *Mol. Cancer Res.* **2012**, *10* (3), 378–391.
- (289) Daugaard, M.; Baude, A.; Fugger, K.; Povlsen, L. K.; Beck, H.; Sorensen, C. S.; Petersen, N. H.; Sorensen, P. H.; Lukas, C.; Bartek, J.; et al. LEDGF (p75) Promotes DNA-End Resection and Homologous Recombination. *Nat. Struct. Mol. Biol.* **2012**, *19* (8), 803–810.
- (290) Hughes, S.; Jenkins, V.; Dar, M. J.; Engelman, A.; Cherepanov, P. Transcriptional Co-Activator LEDGF Interacts with Cdc7-Activator of S-Phase Kinase (ASK) and Stimulates Its Enzymatic Activity. *J. Biol. Chem.* **2010**, *285* (1), 541–554.
- (291) van Nuland, R.; Smits, A. H.; Pallaki, P.; Jansen, P. W.; Vermeulen, M.; Timmers, H. T. Quantitative Dissection and Stoichiometry Determination of the Human SET1/MLL Histone Methyltransferase Complexes. *Mol. Cell Biol.* **2013**, *33* (10), 2067–2077.
- (292) Vandekerckhove, L.; Christ, F.; Van Maele, B.; De Rijck, J.; Gijbsbers, R.; Van den Haute, C.; Witvrouw, M.; Debyser, Z. Transient and Stable Knockdown of the Integrase Cofactor LEDGF/p75 Reveals Its Role in the Replication Cycle of Human Immunodeficiency Virus. *J. Virol.* **2006**, *80* (4), 1886–1896.
- (293) Zielske, S. P.; Stevenson, M. Modest but Reproducible Inhibition of Human Immunodeficiency Virus Type 1 Infection in Macrophages Following LEDGFp75 Silencing. *J. Virol.* **2006**, *80* (14), 7275–7280.
- (294) Vandegraaff, N.; Devroe, E.; Turlure, F.; Silver, P. A.; Engelman, A. Biochemical and Genetic Analyses of Integrase-Interacting Proteins Lens Epithelium-Derived Growth Factor (LEDGF)/p75 and Hepatoma-Derived Growth Factor Related Protein 2 (HRP2) in Preintegration Complex Function and HIV-1 Replication. *Virology* **2006**, *346*, 415–426.
- (295) Llano, M.; Saenz, D. T.; Meehan, A.; Wongthida, P.; Peretz, M.; Walker, W. H.; Teo, W.; Poeschla, E. M. An Essential Role for LEDGF/p75 in HIV Integration. *Science* **2006**, *314* (5798), 461–464.
- (296) Marshall, H. M.; Ronen, K.; Berry, C.; Llano, M.; Sutherland, H.; Saenz, D.; Bickmore, W.; Poeschla, E.; Bushman, F. D. Role of PSIP1/LEDGF/p75 in Lentiviral Infectivity and Integration Targeting. *PLoS One* **2007**, *2* (12), e1340.
- (297) Fadel, H. J.; Morrison, J. H.; Saenz, D. T.; Fuchs, J. R.; Kvaratskhelia, M.; Ekker, S. C.; Poeschla, E. M. TALEN Knockout of the PSIP1 Gene in Human Cells: Analyses of HIV-1 Replication and Allosteric Integrase Inhibitor Mechanism. *J. Virol.* **2014**, *88* (17), 9704–9717.
- (298) Schrijvers, R.; De Rijck, J.; Demeulemeester, J.; Adachi, N.; Vets, S.; Ronen, K.; Christ, F.; Bushman, F. D.; Debyser, Z.; Gijbsbers, R. LEDGF/p75-Independent HIV-1 Replication Demonstrates a Role for HRP-2 and Remains Sensitive to Inhibition by LEDGINS. *PLoS Pathog.* **2012**, *8* (3), e1002558.
- (299) Shun, M. C.; Raghavendra, N. K.; Vandegraaff, N.; Daigle, J. E.; Hughes, S.; Kellam, P.; Cherepanov, P.; Engelman, A. LEDGF/p75 Functions Downstream from Preintegration Complex Formation to Effect Gene-Specific HIV-1 Integration. *Genes Dev.* **2007**, *21* (14), 1767–1778.
- (300) Ciuffi, A.; Llano, M.; Poeschla, E.; Hoffmann, C.; Leipzig, J.; Shinn, P.; Ecker, J. R.; Bushman, F. A Role for LEDGF/p75 in Targeting HIV DNA Integration. *Nat. Med.* **2005**, *11* (12), 1287–1289.
- (301) Ferris, A. L.; Wu, X.; Hughes, C. M.; Stewart, C.; Smith, S. J.; Milne, T. A.; Wang, G. G.; Shun, M. C.; Allis, C. D.; Engelman, A.; et al. Lens Epithelium-Derived Growth Factor Fusion Proteins Redirect HIV-1 DNA Integration. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (7), 3135–3140.
- (302) Gijbsbers, R.; Ronen, K.; Vets, S.; Malani, N.; De Rijck, J.; McNeely, M.; Bushman, F. D.; Debyser, Z. LEDGF Hybrids Efficiently Retarget Lentiviral Integration into Heterochromatin. *Mol. Ther.* **2010**, *18* (3), 552–560.
- (303) Silvers, R. M.; Smith, J. A.; Schowalter, M.; Litwin, S.; Liang, Z.; Geary, K.; Daniel, R. Modification of Integration Site Preferences of an HIV-1-Based Vector by Expression of a Novel Synthetic Protein. *Hum. Gene Ther.* **2010**, *21* (3), 337–349.
- (304) Dahabieh, M. S.; Ooms, M.; Simon, V.; Sadowski, I. A Doubly Fluorescent HIV-1 Reporter Shows That the Majority of Integrated HIV-1 Is Latent Shortly after Infection. *J. Virol.* **2013**, *87* (8), 4716–4727.

- (305) Jeeninga, R. E.; Westerhout, E. M.; van Gerven, M. L.; Berkhout, B. HIV-1 Latency in Actively Dividing Human T Cell Lines. *Retrovirology* **2008**, *5*, 37.
- (306) Duverger, A.; Jones, J.; May, J.; Bibollet-Ruche, F.; Wagner, F. A.; Cron, R. Q.; Kutsch, O. Determinants of the Establishment of Human Immunodeficiency Virus Type 1 Latency. *J. Virol.* **2009**, *83* (7), 3078–3093.
- (307) Cherepanov, P.; Sun, Z. Y.; Rahman, S.; Maertens, G.; Wagner, G.; Engelman, A. Solution Structure of the HIV-1 Integrase-Binding Domain in LEDGF/p75. *Nat. Struct. Mol. Biol.* **2005**, *12* (6), 526–532.
- (308) Cherepanov, P.; Ambrosio, A. L.; Rahman, S.; Ellenberger, T.; Engelman, A. Structural Basis for the Recognition between HIV-1 Integrase and Transcriptional Coactivator p75. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 17308.
- (309) McKee, C. J.; Kessler, J. J.; Shkriabai, N.; Dar, M. J.; Engelman, A.; Kvaratskhelia, M. Dynamic Modulation of HIV-1 Integrase Structure and Function by Cellular Lens Epithelium-Derived Growth Factor (LEDGF) Protein. *J. Biol. Chem.* **2008**, *283* (46), 31802–31812.
- (310) Tsiang, M.; Jones, G. S.; Niedziela-Majka, A.; Kan, E.; Lansdon, E. B.; Huang, W.; Hung, M.; Samuel, D.; Novikov, N.; Xu, Y.; et al. New Class of HIV-1 Integrase (IN) Inhibitors with a Dual Mode of Action. *J. Biol. Chem.* **2012**, *287* (25), 21189–21203.
- (311) Kessler, J. J.; Jena, N.; Koh, Y.; Taskent-Sezgin, H.; Slaughter, A.; Feng, L.; de Silva, S.; Wu, L.; Le Grice, S. F.; Engelman, A.; et al. Multimode, Cooperative Mechanism of Action of Allosteric HIV-1 Integrase Inhibitors. *J. Biol. Chem.* **2012**, *287* (20), 16801–16811.
- (312) Christ, F.; Voet, A.; Marchand, A.; Nicolet, S.; Desimmié, B. A.; Marchand, D.; Bardiou, D.; Van der Veken, N. J.; Van Remoortel, B.; Strelkov, S. V.; Debyser, Z.; et al. Rational Design of Small-Molecule Inhibitors of the LEDGF/p75-Integrase Interaction and HIV Replication. *Nat. Chem. Biol.* **2010**, *6* (6), 442–448.
- (313) Fader, L. D.; Malenfant, E.; Parisien, M.; Carson, R.; Bilodeau, F.; Landry, S.; Pesant, M.; Brochu, C.; Morin, S.; Chabot, C.; et al. Discovery of BI 224436, a Noncatalytic Site Integrase Inhibitor (NCINI) of HIV-1. *ACS Med. Chem. Lett.* **2014**, *5*, 422–427.
- (314) Gupta, K.; Brady, T.; Dyer, B. M.; Malani, N.; Hwang, Y.; Male, F.; Nolte, R. T.; Wang, L.; Velthuisen, E.; Jeffrey, J.; et al. Allosteric Inhibition of Human Immunodeficiency Virus Integrase: Late Block During Viral Replication and Abnormal Multimerization Involving Specific Protein Domains. *J. Biol. Chem.* **2014**, *289* (30), 20477–20488.
- (315) Sharma, A.; Slaughter, A.; Jena, N.; Feng, L.; Kessler, J. J.; Fadel, H. J.; Malani, N.; Male, F.; Wu, L.; Poeschla, E.; et al. A New Class of Multimerization Selective Inhibitors of HIV-1 Integrase. *PLoS Pathog.* **2014**, *10* (5), e1004171.
- (316) Christ, F.; Shaw, S.; Demeulemeester, J.; Desimmié, B. A.; Marchand, A.; Butler, S.; Smets, W.; Chaltin, P.; Westby, M.; Debyser, Z.; et al. Small-Molecule Inhibitors of the LEDGF/p75 Binding Site of Integrase Block HIV Replication and Modulate Integrase Multimerization. *Antimicrob. Agents Chemother.* **2012**, *56* (8), 4365–4374.
- (317) Le Rouzic, E.; Bonnard, D.; Chasset, S.; Bruneau, J. M.; Chevreuil, F.; Le Strat, F.; Nguyen, J.; Beauvoir, R.; Amadori, C.; Brias, J.; et al. Dual Inhibition of HIV-1 Replication by Integrase-LEDGF Allosteric Inhibitors Is Predominant at the Post-Integration Stage. *Retrovirology* **2013**, *10*, 144.
- (318) Shkriabai, N.; Dharmarajan, V.; Slaughter, A.; Kessler, J. J.; Larue, R. C.; Feng, L.; Fuchs, J. R.; Griffin, P. R.; Kvaratskhelia, M. A Critical Role of the C-Terminal Segment for Allosteric Inhibitor-Induced Aberrant Multimerization of HIV-1 Integrase. *J. Biol. Chem.* **2014**, *289* (38), 26430–26440.
- (319) Sharma, A.; Larue, R. C.; Plumb, M. R.; Malani, N.; Male, F.; Slaughter, A.; Kessler, J. J.; Shkriabai, N.; Coward, E.; Aiyer, S. S.; et al. BET Proteins Promote Efficient Murine Leukemia Virus Integration at Transcription Start Sites. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (29), 12036–12041.
- (320) Gupta, S. S.; Maetzig, T.; Maertens, G. N.; Sharif, A.; Rothe, M.; Weidner-Glunde, M.; Galla, M.; Schambach, A.; Cherepanov, P.; Schulz, T. F. Bromo- and Extraterminal Domain Chromatin Regulators Serve as Cofactors for Murine Leukemia Virus Integration. *J. Virol.* **2013**, *87* (23), 12721–12736.
- (321) De Rijck, J.; de Kogel, C.; Demeulemeester, J.; Vets, S.; El Ashkar, S.; Malani, N.; Bushman, F. D.; Landuyt, B.; Husson, S. J.; Busschots, K.; et al. The BET Family of Proteins Targets Moloney Murine Leukemia Virus Integration near Transcription Start Sites. *Cell Rep.* **2013**, *5* (4), 886–894.
- (322) Wu, S. Y.; Chiang, C. M. The Double Bromodomain-Containing Chromatin Adaptor Brd4 and Transcriptional Regulation. *J. Biol. Chem.* **2007**, *282* (18), 13141–13145.
- (323) Belkina, A. C.; Denis, G. V. BET Domain Co-Regulators in Obesity, Inflammation and Cancer. *Nat. Rev. Cancer* **2012**, *12* (7), 465–477.
- (324) Weidner-Glunde, M.; Ottinger, M.; Schulz, T. F. What Do Viruses BET On? *Front. Biosci., Landmark Ed.* **2010**, *15*, 537–549.
- (325) Zhu, J.; Gaiha, G. D.; John, S. P.; Pertel, T.; Chin, C. R.; Gao, G.; Qu, H.; Walker, B. D.; Elledge, S. J.; Brass, A. L. Reactivation of Latent HIV-1 by Inhibition of Brd4. *Cell Rep.* **2012**, *2* (4), 807–816.
- (326) You, J.; Croyle, J. L.; Nishimura, A.; Ozato, K.; Howley, P. M. Interaction of the Bovine Papillomavirus E2 Protein with Brd4 Tethers the Viral DNA to Host Mitotic Chromosomes. *Cell* **2004**, *117* (3), 349–360.
- (327) Dey, A.; Chitsaz, F.; Abbasi, A.; Misteli, T.; Ozato, K. The Double Bromodomain Protein Brd4 Binds to Acetylated Chromatin During Interphase and Mitosis. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (15), 8758–8763.
- (328) Jang, M. K.; Mochizuki, K.; Zhou, M.; Jeong, H. S.; Brady, J. N.; Ozato, K. The Bromodomain Protein Brd4 Is a Positive Regulatory Component of P-TEFb and Stimulates RNA Polymerase II-Dependent Transcription. *Mol. Cell* **2005**, *19* (4), 523–534.
- (329) Baxter, M. K.; McPhillips, M. G.; Ozato, K.; McBride, A. A. The Mitotic Chromosome Binding Activity of the Papillomavirus E2 Protein Correlates with Interaction with the Cellular Chromosomal Protein, Brd4. *J. Virol.* **2005**, *79* (8), 4806–4818.
- (330) You, J.; Srinivasan, V.; Denis, G. V.; Harrington, W. J., Jr.; Ballestas, M. E.; Kaye, K. M.; Howley, P. M. Kaposi's Sarcoma-Associated Herpesvirus Latency-Associated Nuclear Antigen Interacts with Bromodomain Protein Brd4 on Host Mitotic Chromosomes. *J. Virol.* **2006**, *80* (18), 8909–8919.
- (331) Ottinger, M.; Christalla, T.; Nathan, K.; Brinkmann, M. M.; Viejo-Borbolla, A.; Schulz, T. F. Kaposi's Sarcoma-Associated Herpesvirus LANA-1 Interacts with the Short Variant of BRD4 and Releases Cells from a BRD4- and BRD2/RING3-Induced G1 Cell Cycle Arrest. *J. Virol.* **2006**, *80* (21), 10772–10786.
- (332) LeRoy, G.; Chepelev, I.; DiMaggio, P. A.; Blanco, M. A.; Zee, B. M.; Zhao, K.; Garcia, B. A. Proteogenomic Characterization and Mapping of Nucleosomes Decoded by Brd and HP1 Proteins. *Genome Biol.* **2012**, *13* (8), R68.
- (333) Larue, R. C.; Plumb, M. R.; Crowe, B. L.; Shkriabai, N.; Sharma, A.; DiFiore, J.; Malani, N.; Aiyer, S. S.; Roth, M. J.; Bushman, F. D.; et al. Bimodal High-Affinity Association of Brd4 with Murine Leukemia Virus Integrase and Mononucleosomes. *Nucleic Acids Res.* **2014**, *42* (8), 4868–4881.
- (334) Aiyer, S.; Swapna, G. V.; Malani, N.; Aramini, J. M.; Schneider, W. M.; Plumb, M. R.; Ghanem, M.; Larue, R. C.; Sharma, A.; Studamire, B.; et al. Altering Murine Leukemia Virus Integration through Disruption of the Integrase and BET Protein Family Interaction. *Nucleic Acids Res.* **2014**, *42* (9), 5917–5928.
- (335) Crowe, B. L.; Larue, R. C.; Yuan, C.; Hess, S.; Kvaratskhelia, M.; Foster, M. P. Structure of the Brd4 ET Domain Bound to a C-Terminal Motif from Gamma-Retroviral Integrases Reveals a Conserved Mechanism of Interaction. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (8), 2086–2091.
- (336) El Ashkar, S.; De Rijck, J.; Demeulemeester, J.; Vets, S.; Madlala, P.; Cermakova, K.; Debyser, Z.; Gijssbers, R. BET-Independent MLV-Based Vectors Target Away from Promoters and Regulatory Elements. *Mol. Ther.–Nucleic Acids* **2014**, *3*, e179.

- (337) Maertens, G. N. B'-Protein Phosphatase 2A Is a Functional Binding Partner of Delta-Retroviral Integrase. *Nucleic Acids Res.* **2016**, *44* (1), 364–376.
- (338) Eichhorn, P. J.; Creighton, M. P.; Bernards, R. Protein Phosphatase 2A Regulatory Subunits and Cancer. *Biochim. Biophys. Acta, Rev. Cancer* **2009**, *1795* (1), 1–15.
- (339) Shanker, V.; Trincucci, G.; Heim, H. M.; Duong, H. T. Protein Phosphatase 2A Impairs IFN α -Induced Antiviral Activity against the Hepatitis C Virus through the Inhibition of STAT1 Tyrosine Phosphorylation. *J. Viral. Hepat.* **2013**, *20* (9), 612–621.
- (340) Curcio, M. J.; Lutz, S.; Lesage, P. The Ty1 LTR-Retrotransposon of Budding Yeast, *Saccharomyces Cerevisiae*. *Microbiol. Spectr.* **2015**, *3* (2), MDNA3-0053-2014.
- (341) Zhu, Y.; Dai, J.; Fuerst, P. G.; Voytas, D. F. Controlling Integration Specificity of a Yeast Retrotransposon. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (10), 5891–5895.
- (342) Brady, T. L.; Fuerst, P. G.; Dick, R. A.; Schmidt, C.; Voytas, D. F. Retrotransposon Target Site Selection by Imitation of a Cellular Protein. *Mol. Cell. Biol.* **2008**, *28* (4), 1230–1239.
- (343) Xie, W.; Gai, X.; Zhu, Y.; Zappulla, D. C.; Sternglanz, R.; Voytas, D. F. Targeting of the Yeast Ty5 Retrotransposon to Silent Chromatin Is Mediated by Interactions between Integrase and Sir4p. *Mol. Cell. Biol.* **2001**, *21* (19), 6606–6614.
- (344) Bridier-Nahmias, A.; Tchalian-Cosson, A.; Baller, J. A.; Menouni, R.; Fayol, H.; Flores, A.; Saib, A.; Werner, M.; Voytas, D. F.; Lesage, P. Retrotransposons. An RNA Polymerase III Subunit Determines Sites of Retrotransposon Integration. *Science* **2015**, *348* (6234), 585–588.
- (345) Hazuda, D. J.; Felock, P.; Witmer, M.; Wolfe, A.; Stillmock, K.; Grobler, J. A.; Espeseth, A.; Gabryelski, L.; Schleif, W.; Blau, C.; et al. Inhibitors of Strand Transfer That Prevent Integration and Inhibit HIV-1 Replication in Cells. *Science* **2000**, *287* (5453), 646–650.
- (346) Hazuda, D. J.; Hastings, J. C.; Wolfe, A. L.; Emini, E. A. A Novel Assay for the DNA Strand-Transfer Reaction of HIV-1 Integrase. *Nucleic Acids Res.* **1994**, *22* (6), 1121–1122.
- (347) Murray, J. M.; Emery, S.; Kelleher, A. D.; Law, M.; Chen, J.; Hazuda, D. J.; Nguyen, B. Y.; Tepler, H.; Cooper, D. A. Antiretroviral Therapy with the Integrase Inhibitor Raltegravir Alters Decay Kinetics of HIV, Significantly Reducing the Second Phase. *AIDS* **2007**, *21* (17), 2315–2321.
- (348) Summa, V.; Petrocchi, A.; Bonelli, F.; Crescenzi, B.; Donghi, M.; Ferrara, M.; Fiore, F.; Gardelli, C.; Gonzalez Paz, O.; Hazuda, D. J.; et al. Discovery of Raltegravir, a Potent, Selective Orally Bioavailable HIV-Integrase Inhibitor for the Treatment of HIV-AIDS Infection. *J. Med. Chem.* **2008**, *51* (18), 5843–5855.
- (349) Sato, M.; Motomura, T.; Aramaki, H.; Matsuda, T.; Yamashita, M.; Ito, Y.; Kawakami, H.; Matsuzaki, Y.; Watanabe, W.; Yamataka, K.; et al. Novel HIV-1 Integrase Inhibitors Derived from Quinolone Antibiotics. *J. Med. Chem.* **2006**, *49* (5), 1506–1508.
- (350) Johns, B. A.; Kawasuji, T.; Weatherhead, J. G.; Taishi, T.; Temelkoff, D. P.; Yoshida, H.; Akiyama, T.; Taoda, Y.; Murai, H.; Kiyama, R.; et al. Carbamoyl Pyridone HIV-1 Integrase Inhibitors 3. A Diastereomeric Approach to Chiral Nonracemic Tricyclic Ring Systems and the Discovery of Dolutegravir (S/GSK1349572) and (S/GSK1265744). *J. Med. Chem.* **2013**, *56* (14), 5901–5916.
- (351) Bar-Magen, T.; Sloan, R. D.; Faltenbacher, V. H.; Donahue, D. A.; Kuhl, B. D.; Oliveira, M.; Xu, H.; Wainberg, M. A. Comparative Biochemical Analysis of HIV-1 Subtype B and C Integrase Enzymes. *Retrovirology* **2009**, *6*, 103.
- (352) Hightower, K. E.; Wang, R.; Deanda, F.; Johns, B. A.; Weaver, K.; Shen, Y.; Tomberlin, G. H.; Carter, H. L., 3rd; Broderick, T.; Sigethy, S.; et al. Dolutegravir (S/GSK1349572) Exhibits Significantly Slower Dissociation Than Raltegravir and Elvitegravir from Wild-Type and Integrase Inhibitor-Resistant HIV-1 Integrase-DNA Complexes. *Antimicrob. Agents Chemother.* **2011**, *55* (10), 4552–4559.
- (353) Grobler, J.; Stillmock, K.; Miller, D. G.; Hazuda, D. Mechanism by Which the HIV Integrase Active-Site Mutation N155H Confers Resistance to Raltegravir. *Antiviral Ther* **2008**, *13*, A41.
- (354) Lewis, M. G.; Norelli, S.; Collins, M.; Barreca, M. L.; Iraci, N.; Chirullo, B.; Yalley-Ogunro, J.; Greenhouse, J.; Titti, F.; Garaci, E.; et al. Response of a Simian Immunodeficiency Virus (SIVmac251) to Raltegravir: A Basis for a New Treatment for Simian AIDS and an Animal Model for Studying Lentiviral Persistence During Antiretroviral Therapy. *Retrovirology* **2010**, *7*, 21.
- (355) Koh, Y.; Matreyek, K. A.; Engelman, A. Differential Sensitivities of Retroviruses to Integrase Strand Transfer Inhibitors. *J. Virol.* **2011**, *85* (7), 3677–3682.
- (356) Hare, S.; Smith, S. J.; Metifiot, M.; Jaxa-Chamiec, A.; Pommier, Y.; Hughes, S. H.; Cherepanov, P. Structural and Functional Analyses of the Second-Generation Integrase Strand Transfer Inhibitor Dolutegravir (S/GSK1349572). *Mol. Pharmacol.* **2011**, *80*, S65.
- (357) Langley, D. R.; Samanta, H. K.; Lin, Z.; Walker, M. A.; Krystal, M. R.; Dicker, I. B. The Terminal (Catalytic) Adenosine of the HIV LTR Controls the Kinetics of Binding and Dissociation of HIV Integrase Strand Transfer Inhibitors. *Biochemistry* **2008**, *47* (51), 13481–13488.
- (358) Zhao, X. Z.; Smith, S. J.; Maskell, D. P.; Metifiot, M.; Pye, V. E.; Fesen, K.; Marchand, C.; Pommier, Y.; Cherepanov, P.; Hughes, S. H.; et al. HIV-1 Integrase Strand Transfer Inhibitors with Reduced Susceptibility to Drug Resistant Mutant Integrases. *ACS Chem. Biol.* **2016**, *11*, 1074.
- (359) Grobler, J. A.; Hazuda, D. J. Resistance to HIV Integrase Strand Transfer Inhibitors: In Vitro Findings and Clinical Consequences. *Curr. Opin. Virol.* **2014**, *8*, 98–103.
- (360) Mesplede, T.; Wainberg, M. A. Resistance against Integrase Strand Transfer Inhibitors and Relevance to HIV Persistence. *Viruses* **2015**, *7* (7), 3703–3718.
- (361) Canducci, F.; Ceresola, E. R.; Boeri, E.; Spagnuolo, V.; Cossarini, F.; Castagna, A.; Lazzarin, A.; Clementi, M. Cross-Resistance Profile of the Novel Integrase Inhibitor Dolutegravir (S/GSK1349572) Using Clonal Viral Variants Selected in Patients Failing Raltegravir. *J. Infect. Dis.* **2011**, *204* (11), 1811–1815.
- (362) Goethals, O.; Vos, A.; Van Ginderen, M.; Geluykens, P.; Smits, V.; Schols, D.; Hertogs, K.; Clayton, R. Primary Mutations Selected in Vitro with Raltegravir Confer Large Fold Changes in Susceptibility to First-Generation Integrase Inhibitors, but Minor Fold Changes to Inhibitors with Second-Generation Resistance Profiles. *Virology* **2010**, *402* (2), 338–346.
- (363) Naeger, L. K.; Harrington, P.; Komatsu, T.; Deming, D. Effect of Dolutegravir Functional Monotherapy on HIV-1 Virological Response in Integrase Strand Transfer Inhibitor Resistant Patients. *Antiviral Ther.* **2016**, DOI: 10.3851/IMP3033.
- (364) Cooper, D. A.; Steigbigel, R. T.; Gatell, J. M.; Rockstroh, J. K.; Katlama, C.; Yeni, P.; Lazzarin, A.; Clotet, B.; Kumar, P. N.; Eron, J. E.; et al. Subgroup and Resistance Analyses of Raltegravir for Resistant HIV-1 Infection. *N. Engl. J. Med.* **2008**, *359* (4), 355–365.
- (365) Raheem, I. T.; Walji, A. M.; Klein, D.; Sanders, J. M.; Powell, D. A.; Abeywickrema, P.; Barbe, G.; Bennet, A.; Clas, S. D.; Dubost, D.; et al. Discovery of 2-Pyridinone Amins: A Prodrug Strategy to Advance a Second Generation of HIV-1 Integrase Strand Transfer Inhibitors. *J. Med. Chem.* **2015**, *58* (20), 8154–8165.
- (366) Jurado, K. A.; Engelman, A. Multimodal Mechanism of Action of Allosteric HIV-1 Integrase Inhibitors. *Expert Rev. Mol. Med.* **2013**, *15*, e14.
- (367) Engelman, A.; Kessler, J. J.; Kvaratskhelia, M. Allosteric Inhibition of HIV-1 Integrase Activity. *Curr. Opin. Chem. Biol.* **2013**, *17* (3), 339–345.
- (368) Molteni, V.; Greenwald, J.; Rhodes, D.; Hwang, Y.; Kwiatkowski, W.; Bushman, F. D.; Siegel, J. S.; Choe, S. Identification of a Small-Molecule Binding Site at the Dimer Interface of the HIV Integrase Catalytic Domain. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2001**, *57*, S36–S44.
- (369) De Rijck, J.; Vandekerckhove, L.; Gijsbers, R.; Hombrouck, A.; Hendrix, J.; Vercammen, J.; Engelborghs, Y.; Christ, F.; Debyser, Z. Overexpression of the Lens Epithelium-Derived Growth Factor/p75 Integrase Binding Domain Inhibits Human Immunodeficiency Virus Replication. *J. Virol.* **2006**, *80* (23), 11498–11509.

- (370) Meehan, A. M.; Saenz, D. T.; Morrison, J.; Hu, C.; Peretz, M.; Poeschla, E. M. LEDGF Dominant Interference Proteins Demonstrate Prenuclear Exposure of HIV-1 Integrase and Synergize with LEDGF Depletion to Destroy Viral Infectivity. *J. Virol.* **2011**, *85* (7), 3570–3583.
- (371) Tsantrizos, Y. S.; Boes, M.; Brochu, C.; Fenwick, C.; Malenfant, E.; Mason, S.; Pesant, M. Inhibitors of Human Immunodeficiency Virus Replication. Patent WO/2007/131350, 2007.
- (372) Desimmie, B. A.; Schrijvers, R.; Demeulemeester, J.; Borrenberghs, D.; Weydert, C.; Thys, W.; Vets, S.; Van Remoortel, B.; Hofkens, J.; De Rijck, J.; et al. LEDGINs Inhibit Late Stage HIV-1 Replication by Modulating Integrase Multimerization in the Virions. *Retrovirology* **2013**, *10*, 57.
- (373) van Bel, N.; van der Velden, Y.; Bonnard, D.; Le Rouzic, E.; Das, A. T.; Benarous, R.; Berkhout, B. The Allosteric HIV-1 Integrase Inhibitor BI-D Affects Virion Maturation but Does Not Influence Packaging of a Functional RNA Genome. *PLoS One* **2014**, *9*, e103552.
- (374) Wang, H.; Jurado, K. A.; Wu, X.; Shun, M. C.; Li, X.; Ferris, A. L.; Smith, S. J.; Patel, P. A.; Fuchs, J. R.; Cherepanov, P.; et al. HRP2 Determines the Efficiency and Specificity of HIV-1 Integration in LEDGF/p75 Knockout Cells but Does Not Contribute to the Antiviral Activity of a Potent LEDGF/p75-Binding Site Integrase Inhibitor. *Nucleic Acids Res.* **2012**, *40* (22), 11518–11530.
- (375) Fenwick, C.; Amad, M.; Bailey, M. D.; Bethell, R.; Bös, M.; Bonneau, P.; Cordingley, M.; Coulombe, R.; Duan, J.; Edwards, P.; et al. Preclinical Profile of BI 224436, a Novel HIV-1 Non-Catalytic-Site Integrase Inhibitor. *Antimicrob. Agents Chemother.* **2014**, *58*, 3233–3244.
- (376) Naldini, L. Gene Therapy Returns to Centre Stage. *Nature* **2015**, *526*, 351–360.
- (377) Cavazzana-Calvo, M.; Hacein-Bey, S.; de Saint Basile, G.; Gross, F.; Yvon, E.; Nussbaum, P.; Selz, F.; Hue, C.; Certain, S.; Casanova, J. L.; et al. Gene Therapy of Human Severe Combined Immunodeficiency (SCID)-X1 Disease. *Science* **2000**, *288* (5466), 669–672.
- (378) Braun, C. J.; Boztug, K.; Paruzynski, A.; Witzel, M.; Schwarzer, A.; Rothe, M.; Modlich, U.; Beier, R.; Göhring, G.; Steinemann, D.; et al. Gene Therapy for Wiskott-Aldrich Syndrome—Long-Term Efficacy and Genotoxicity. *Sci. Transl. Med.* **2014**, *6*, 227ra33.
- (379) Ott, M. G.; Schmidt, M.; Schwarzwaelder, K.; Stein, S.; Siler, U.; Koehl, U.; Glimm, H.; Kuhlcke, K.; Schilz, A.; Kunkel, H.; et al. Correction of X-Linked Chronic Granulomatous Disease by Gene Therapy, Augmented by Insertional Activation of MDS1-EVII, PRMD16 or SETBP1. *Nat. Med.* **2006**, *12* (4), 401–409.
- (380) Kohn, D. B. Update on Gene Therapy for Immunodeficiencies. *Clin. Immunol.* **2010**, *135* (2), 247–254.
- (381) Hacein-Bey-Abina, S.; von Kalle, C.; Schmidt, M.; Le Deist, F.; Wulffraat, N.; McIntyre, E.; Radford, I.; Villeval, J. L.; Fraser, C. C.; Cavazzana-Calvo, M.; et al. A Serious Adverse Event after Successful Gene Therapy for X-Linked Severe Combined Immunodeficiency. *N. Engl. J. Med.* **2003**, *348* (3), 255–256.
- (382) Hacein-Bey-Abina, S.; Garrigue, A.; Wang, G. P.; Soulier, J.; Lim, A.; Morillon, E.; Clappier, E.; Caccavelli, L.; Delabesse, E.; Beldjord, K.; et al. Insertional Oncogenesis in 4 Patients after Retrovirus-Mediated Gene Therapy of SCID-X1. *J. Clin. Invest.* **2008**, *118*, 3132–3142.
- (383) Howe, S. J.; Mansour, M. R.; Schwarzwaelder, K.; Bartholomae, C.; Hubank, M.; Kempinski, H.; Brugman, M. H.; Pike-Overzet, K.; Chatters, S. J.; de Ridder, D.; et al. Insertional Mutagenesis Combined with Acquired Somatic Mutations Causes Leukemogenesis Following Gene Therapy of SCID-X1 Patients. *J. Clin. Invest.* **2008**, *118*, 3143–3150.
- (384) Hacein-Bey-Abina, S.; Von Kalle, C.; Schmidt, M.; McCormack, M. P.; Wulffraat, N.; Leboulch, P.; Lim, A.; Osborne, C. S.; Pawliuk, R.; Morillon, E.; et al. LMO2-Associated Clonal T Cell Proliferation in Two Patients after Gene Therapy for SCID-X1. *Science* **2003**, *302* (5644), 415–419.
- (385) Kung, H. J.; Boerkoel, C.; Carter, T. H. Retroviral Mutagenesis of Cellular Oncogenes: A Review with Insights into the Mechanisms of Insertional Activation. *Curr. Top. Microbiol. Immunol.* **1991**, *171*, 1–25.
- (386) Suerth, J. D.; Labenski, V.; Schambach, A. Alpharetroviral Vectors: From a Cancer-Causing Agent to a Useful Tool for Human Gene Therapy. *Viruses* **2014**, *6* (12), 4811–4838.
- (387) Konstantoulas, C. J.; Indik, S. Mouse Mammary Tumor Virus-Based Vector Transduces Non-Dividing Cells, Enters the Nucleus Via a TNPO3-Independent Pathway and Integrates in a Less Biased Fashion Than Other Retroviruses. *Retrovirology* **2014**, *11*, 34.
- (388) Olszko, M. E.; Trobridge, G. D. Foamy Virus Vectors for HIV Gene Therapy. *Viruses* **2013**, *5*, 2585–2600.
- (389) Roth, M. J. Mutational Analysis of the Carboxyl Terminus of the Moloney Murine Leukemia Virus Integration Protein. *J. Virol.* **1991**, *65*, 2141–2145.
- (390) Yu, S. F.; von Rüden, T.; Kantoff, P. W.; Garber, C.; Seiberg, M.; Rüther, U.; Anderson, W. F.; Wagner, E. F.; Gilboa, E. Self-Inactivating Retroviral Vectors Designed for Transfer of Whole Genes into Mammalian Cells. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 3194–3198.
- (391) Hacein-Bey-Abina, S.; Pai, S.-Y.; Gaspar, H. B.; Armant, M.; Berry, C. C.; Blanche, S.; Bleesing, J.; Blondeau, J.; de Boer, H.; Buckland, K. F.; et al. A Modified Gammaretrovirus Vector for X-Linked Severe Combined Immunodeficiency. *N. Engl. J. Med.* **2014**, *371* (15), 1407–1417.
- (392) Perelson, A. S.; Neumann, A. U.; Markowitz, M.; Leonard, J. M.; Ho, D. D. HIV-1 Dynamics in Vivo: Virion Clearance Rate, Infected Cell Life-Span, and Viral Generation Time. *Science* **1996**, *271*, 1582–1586.
- (393) Cao, J.; Park, I. W.; Cooper, A.; Sodroski, J. Molecular Determinants of Acute Single-Cell Lysis by Human Immunodeficiency Virus Type 1. *J. Virol.* **1996**, *70* (3), 1340–1354.
- (394) Somasundaran, M.; Sharkey, M.; Brichacek, B.; Luzuriaga, K.; Emerman, M.; Sullivan, J. L.; Stevenson, M. Evidence for a Cytopathogenicity Determinant in HIV-1 Vpr. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (14), 9503–9508.
- (395) Cartier, N.; Hacein-Bey-Abina, S.; Bartholomae, C. C.; Veres, G.; Schmidt, M.; Kutschera, I.; Vidaud, M.; Abel, U.; Dal-Cortivo, L.; Caccavelli, L.; et al. Hematopoietic Stem Cell Gene Therapy with a Lentiviral Vector in X-Linked Adrenoleukodystrophy. *Science* **2009**, *326*, 818–823.
- (396) Aiuti, A.; Biasco, L.; Scaramuzza, S.; Ferrua, F.; Cicalese, M. P.; Baricordi, C.; Dionisio, F.; Calabria, A.; Giannelli, S.; Castiello, M. C.; et al. Lentiviral Hematopoietic Stem Cell Gene Therapy in Patients with Wiskott-Aldrich Syndrome. *Science* **2013**, *341* (6148), 1233151.
- (397) Biffi, A.; Montini, E.; Lorioli, L.; Cesani, M.; Fumagalli, F.; Plati, T.; Baldoli, C.; Martino, S.; Calabria, A.; Canale, S.; et al. Lentiviral Hematopoietic Stem Cell Gene Therapy Benefits Metachromatic Leukodystrophy. *Science* **2013**, *341*, 1233158.
- (398) Hacein-Bey-Abina, S.; Gaspar, H.; Blondeau, J.; et al. Outcomes Following Gene Therapy in Patients with Severe Wiskott-Aldrich Syndrome. *JAMA* **2015**, *313* (15), 1550–1563.
- (399) Doudna, J. A.; Charpentier, E. Genome Editing. The New Frontier of Genome Engineering with CRISPR-Cas9. *Science* **2014**, *346*, 1258096.