ELSEVIER

Antiviral Research 36 (1997) 139–156

Antiviral Research

Mini-review

Molecular mechanisms in retrovirus DNA integration

Ernest Asante-Appiah, Anna Marie Skalka *

Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA

Received 9 July 1997; accepted 13 August 1997

Abstract

The integrase protein of retroviruses catalyzes the insertion of the viral DNA into the genomes of the cells that they infect. Integrase is necessary and sufficient for this recombination reaction in vitro; however, the enzyme's activity appears to be modulated in vivo by viral and cellular components included in the nucleoprotein pre-integration complex. In addition to integrase, *cis*-acting sequences at the ends of the viral DNA are important for integration. Solution of the structures of the isolated N- and C-terminal domains of HIV-1 integrase by nuclear magnetic resonance (NMR) and the available crystal structures of the catalytic core domains from human immunodeficiency virus type-1 (HIV-1) and avian sarcoma virus (ASV) integrases are providing a structural basis for understanding some aspects of the integration reaction. The role of the evolutionarily conserved acidic amino acids in the D,D(35)E motif as metal-coordinating residues that are critical for catalysis, has been confirmed by the metal–integrase (core domain) complexes of ASV integrase. The central role that integrase plays in the life cycle of the virus makes it an attractive target for the design of drugs against retroviral diseases such as AIDS. To this end, several compounds have been screened for inhibitory effects against HIV-1 integrase. These include DNA intercalators, peptides, RNA ligands, and small organic compounds such as bis-catechols, flavones, and hydroxylated arylamides. Although the published inhibitors are not very potent, they serve as valuable leads for the development of the next generation of tight-binding analogues that are more specific to integrase. In addition, new approaches are being developed, exemplified by intracellular immunization studies with conformation-sensitive inhibitory monoclonal antibodies against HIV-1 integrase. Increased knowledge of the mechanism of retroviral DNA integration should provide new strategies for the design of effective antivirals that inhibit integrase in the future. © 1997 Elsevier Science B.V.

Keywords: Retrovirus; Integrase; Integration

Abbreviations: AMV, avian myeloblastosis virus; ASV, avian sarcoma virus; EIAV, equine infectious anaemia virus; FIV, feline immunodeficiency virus; HFV, human foamy virus; HIV, human immunodeficiency virus; HMG, high-mobility group; HTLV, human T-cell leukemia virus; MoMLV, Moloney murine leukemia virus; SIV, simian immunodeficiency virus.

* Corresponding author. Tel.: +1 215 7282490; fax: +1 215 7282778.

PII S0166-3542(97)00046-6

1. Introduction

Infection of a cell by a retrovirus is established when the viral DNA (post reverse transcription) is successfully integrated into the host cell's genome. Integration of viral DNA allows the provirus access to the cellular machinery for viral gene expression. The integration event also ensures that the parental cell and all successive generations of daughter cells will carry the viral genes, thereby maintaining chronic infection. The retroviral enzyme, integrase, catalyzes the insertion of the viral DNA into the host chromosome. Significant effort has been directed in recent years to understanding the biochemical basis of this reaction. This effort has been spurred on by the continuing AIDS epidemic and the increasing evidence that targeting several viral proteins concurrently may be the best approach to blocking replication (Craig et al., 1993; Johnson et al., 1989; Connell et al., 1994). The heightened interest in integrase stems from its central role in the life cycle of the virus and the absence of a cellular homologue. Thus, potent inhibitors may be designed to inhibit its activity specifically. In this minireview, we discuss the current state of knowledge concerning integrase and its role in retroviral DNA integration. We emphasize recent advances in understanding the mechanism of integration, and highlight on-going and potential strategies for drug intervention.

2. The pre-integration complex

After gaining entry into the cytoplasm of a permissive cell, the RNA genome of a retrovirus is reverse transcribed. The resulting double stranded DNA, together with 50–100 molecules of integrase, is contained within a large nucleoprotein structure, called the pre-integration complex, which must then gain access to the host cell DNA to enable integration (Bowerman et al., 1989; Bukrinsky et al., 1992). In addition to the viral DNA and integrase (IN), pre-integration complexes isolated from HIV-1 infected cells have been reported to include the viral matrix protein (MA), reverse transcriptase (RT), nucleocapsid protein (NC), viral protein R (Vpr), and the cellular protein HMG I(Y) (Farnet and Haseltine, 1991; Bukrinsky et al., 1993a; Gallay et al., 1995b; Farnet and Bushman, 1997). There are indications that other cellular proteins may also be present in these complexes (Lee and Craigie, 1994). Thus, within the target cell, integrase is not the only protein associated with the viral nucleic acids and determining the significance of these associations is currently an active area of investigation.

3. The processing reaction

Analysis of viral DNA isolated from infected cells shows that preparation for DNA integration begins in the cytoplasm. As soon as viral DNA synthesis is completed, the linear duplex product is trimmed by (usually) two nucleotides at the 3′ ends of its long terminal repeats (LTRs) (Roth et al., 1989; Miller et al., 1997). Analysis of viral DNA ends from HIV-1 and MoMLV-infected cells has revealed that the processing reaction occurs within 5–6 h postinfection. This first step, called processing, has been recapitulated in vitro using purified integrase and DNA substrates that represent viral DNA ends (Katzman et al., 1989; Sherman and Fyfe, 1990). In this in vitro reaction, the appropriate 3′ terminal dinucleotide is specifically removed from oligonucleotides representing a single viral DNA end, to yield a product that is two nucleotides shorter in length (Fig. 1). The endonucleolytic cleavage always occurs at the 3′ of a conserved dinucleotide, CA, and results in exposure of a new hydroxyl group at the 3′ end of the viral DNA. Craigie and colleagues have determined the stereochemistry of the processing reaction with the aid of phosphorothioate oligonucleotide substrates. By substituting sulfur for oxygen at the susceptible phosphate, they demonstrated that the nucleophilic attack by lytic water occurs with inversion of chirality (Engelman et al., 1992). This is consistent with the in-line mechanism of a direct transesterification reaction. In vivo, processing of the two 3′ viral DNA ends is coordinated; mutations created in the conserved CA in the end of one LTR impairs integrase-mediated processing after the CA at the

other end (Murphy and Goff, 1992). Similar results have documented coordinated processing reactions in vitro (Kukolj and Skalka, 1995). The new 3′ hydroxyl group formed in the processing step is used as the attacking nucleophile in the subsequent joining step.

Pre-integration complexes may exploit more than one pathway to gain access to target host cell DNA in the nucleus. For MoMLV, entry is apparently gained only when the nuclear envelope breaks down during mitosis (Roe et al., 1993). Although this mechanism may account for nuclear entry of other retroviruses, the ability of lentiviruses, such as HIV, to infect non-dividing cells suggests an alternate or additional mechanism(s). Nuclear localization signals (NLS's) have now been identified on several viral proteins that are found in pre-integration complexes including IN (Kukolj et al., 1997), MA (Bukrinsky et al., 1993b; Gallay et al., 1995a), and Vpr (Heinzinger et al., 1994). At present there are conflicting reports in the literature (Freed and Martin, 1994; Freed et al., 1995, 1997; Von Schwedler et al., 1994; Gallay et al., 1995b; Bukrinskaya et al., 1996) concerning the contribution of particular viral proteins to nuclear entry, but an active nuclear transport for some retroviral pre-integration complexes seems likely (reviewed in Stevenson, 1996). Knowledge of the mechanism(s) of nuclear entry is important, as it may suggest new strategies for blocking integration.

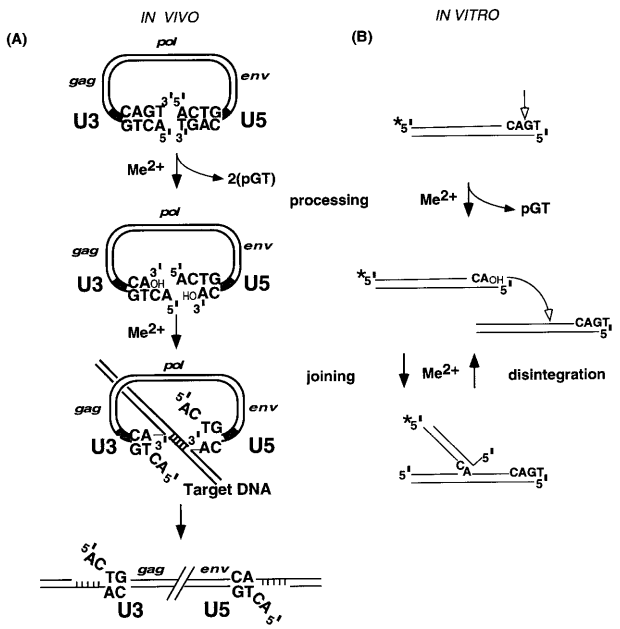


Fig. 1. The mechanism of retroviral DNA integration. (A) The terminal dinucleotide (pGT) is specifically removed from the viral DNA ends in vivo by integrase. This processing reaction exposes an OH group at the 3′ end of the invariant CA at the ends of both strands. The newly exposed OH's serve as the nucleophiles in the concerted cleavage-ligation joining step to insert the viral DNA into the target site in a 5 bp staggered fashion in the case of HIV-1 integrase. (B) The processing and joining reactions have been recapitulated in vitro with oligonucleotides representing the ends of the viral DNA. The oligonucleotide substrate DNA also serves as a surrogate target DNA in the in vitro reaction. A third reaction, disintegration, an apparent reversal of the joining step is catalyzed by integrase in vitro. In this reaction, a Y-shaped substrate representing the insertion of a viral DNA end into the target site can be resolved intramolecularly into the viral and target DNA components. The in vitro reactions in (B) are single-end events, representing 'half reactions' of the double-end events catalyzed by integrase in vivo (A). U3 and U5 refer to sequences that are unique to the 3′ and 5′ ends, respectively, of the viral RNA. They are found at the right (U3) and left (U5) ends of the DNA, respectively. Me refers to the metal cofactor, usually Mn^{2+} or Mg^{2+}.

4. The joining reaction

Once nuclear entry is gained, the processed viral DNA ends are inserted into the host cell's genome via a second catalytic event termed strand-transfer or joining (Fig. 1). As in the case of the processing reaction, the development of in vitro assays has enabled a more detailed understanding of the joining reaction (Fig. 1). Using oligonucleotide substrates, the joining of viral DNA ends to target DNA has also been shown to be catalyzed by the retroviral integrase in vitro (Katz et al., 1990; Craigie et al., 1990). In the simplest version of such an assay, the substrate oligonucleotides also serve as target DNA surrogates. Most products of this reaction are longer than the target molecules and can be distinguished from them by electrophoresis in denaturing polyacrylamide gels. The joining reaction also proceeds via direct transesterification and requires no compounds of high phosphate transfer potential (Engelman et al., 1992). The insertion of the viral

DNA into the target occurs in a relatively sequence-independent fashion; however, hotspots for joining have been observed in vitro (Kitamura et al., 1992; Fitzgerald et al., 1992; Shih et al., 1988). The choice of target site appears to be influenced by structural features in the DNA determined by chromatin organization and DNA-binding proteins, as well as by components of the nucleoprotein complex and the viral integrase itself (Pryciak et al., 1992; Pryciak and Varmus, 1992; Leavitt et al., 1992; Craigie, 1992). In general, bent or slightly unwound DNA appears to be a preferred target for joining (Müller and Varmus, 1994).

The joining reaction must also be coordinated in vivo. It comprises of a concerted cleavage and ligation reaction in which staggered phosphates in the duplex target DNA backbone are attacked by the two $3'_{OH}$ ends of the processed viral DNA. These viral DNA ends are inserted into opposite strands of the target DNA. The resulting single-stranded gaps in the target that flank the 5′ ends of the viral DNA (Fig. 1(a), bottom), reflect the distance between the joining sites. These gaps must be filled in, and the two-nucleotide 5′ overhang of viral DNA removed, to restore the continuity of the host DNA. The result is a short, direct duplication of the target sequence on either side of the integrated viral DNA whose length is characteristic of the virus (e.g. 6 bp for ASV; 5 bp for HIV-1; 4 bp for MoMLV; reviewed in Skalka, 1988; Varmus and Brown, 1989). In vivo, the gap repair machinery of the host cell is presumed to be responsible for these last steps. However, a role for RT and the disintegration activity of integrase (see below) has also been proposed (Chow et al., 1992; Roe et al., 1997). Concerted joining reactions catalyzed by purified integrases have recently been demonstrated in vitro (Vora et al., 1994; Goodarzi et al., 1995; Aiyar et al., 1996).

5. The disintegration reaction

A third catalytic activity of retroviral integrases has been described in vitro, termed disintegration. The reaction is an apparent reversal of the joining reaction in which a substrate representing a single or double-ended insertion of the viral DNA ends in a target DNA (Fig. 1b) is resolved intramolecularly into the viral and target components (Chow et al., 1992). A modification of this reaction, referred to as splicing, has been postulated to contribute to the gap repair following the joining reaction. However, no evidence of disintegration catalyzed by integrase has been demonstrated in vivo. The specificity of integrase for the disintegration reactions has been challenged because nonviral sequences presented in a similar configuration are also resolved by the enzyme (van den Ent et al., 1994). Regardless of its relevance in vivo, the disintegration reaction has been very useful in the biochemical characterization of several mutants, particularly those encoding substitutions in the core domains of integrase proteins. This is because the isolated core domains of integrases neither process nor join viral DNA ends to targets, but they do retain some disintegration activity. Integrase from human foamy (spume) virus, however, requires an intact N-terminal domain in order to catalyze disintegration (Pahl and Flügel, 1995).

6. The integrase protein

The integrases are 32–40 kDa proteins whose sequences are well conserved among retroviruses and retrotransposons. Integrase has been purified and characterized from several retroviruses including ASV, AMV, MoMLV, SIV, FIV, EIAV, HIV, HTLV, HFV and the visna virus. These proteins are generally quite insoluble upon purification and precipitate from solutions at concentrations approximating 1 mg/ml. The propensity to aggregate or precipitate has hampered detailed biochemical and biophysical characterization of the enzymes. The proteins from HIV and ASV, which have been characterized most extensively, exist in a dynamic equilibrium of monomers, dimers, tetramers, and higher-order oligomers (Jones et al., 1992; Hickman et al., 1994). A phenylalanine to lysine substitution at position 185 coupled with a cysteine to serine change at position 280 renders HIV-1 integrase considerably more soluble (>10 mg/ml) and predominantly

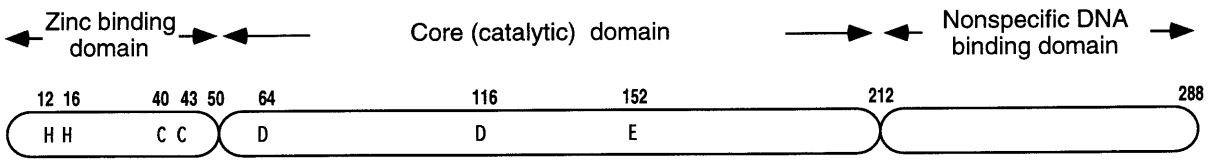


Fig. 2. A linear model of HIV-1 integrase showing the structural organization of the protein. The characteristic features or functions of the three domains are indicated. The horizontal arrows delimit the boundaries of the domains as characterized by biochemical and structural analyses. The residues that are evolutionary conserved among integrases are shown within the domains, with the HIV-1 integrase residue numbers indicated above.

dimeric (Jenkins et al., 1996). The in vitro catalytic activity of the soluble derivative is equivalent to that of the wild-type enzyme, however, virions harboring the F185K mutation fail to replicate. In contrast, a histidine substitution at position 185, which also improves solubility of the isolated protein, permits replication of virions (Engelman et al., 1997). Interestingly, the more soluble, wild type ASV integrase carries a histidine residue at the equivalent position; thus underscoring the relatedness of the retroviral integrases. Because the specific activities of integrases are quite low, the concentration of enzyme used in assays are generally in stoichiometric amounts with the substrate. The protein concentration-dependence of catalytic activity, and in vitro complementation studies with integrase derivatives, suggest that the functionally competent protein is a multimer, minimally a dimer (Jones et al., 1992; Engelman et al., 1993; van Gent et al., 1993).

7. Structural organization

Integrase polypeptides comprise three independently folding domains (Fig. 2). The N-terminal domain, of approximately 50 amino acids, is characterized by an HHCC motif (Johnson et al., 1986) that is reminiscent of the canonical CCHH zinc finger motif found in a variety of DNA binding proteins. Amino acid substitutions in the HHCC motif of integrase impair catalytic activity in vitro (Engelman and Craigie, 1992; Vincent et al., 1993; Kulkosky et al., 1992) and prevent the production of infectious virions (Roth et al., 1990; LaFemina et al., 1992). The HHCC motifs of MoMLV and HIV-1 integrase have been shown to bind Zn ions (McEuen et al., 1992; Burke et al., 1992); the binding of this metal appears to stabilize the N-terminus as an independently folded domain (Burke et al., 1992). Solution of the NMR structure of the N-terminal domain of HIV-1 integrase (residues 1–55) complexed with zinc ions has now substantiated the predicted metal-binding role of the HHCC motif (Cai et al., 1997). However, in contrast to the α–β structure observed in the canonical zinc-finger DNA binding proteins, the zinc binding domain of HIV-1 integrase is completely α-helical. A helix-turn-helix organization such as that typically found in some DNA binding domains is used, in this case, to stabilize the dimeric structure of the isolated N-terminal domain through protein–protein interactions. Recent studies with HIV-1 integrase suggest that the binding of Zn ions to the N-terminal domain promotes tetramerization of full length integrase in vitro and may also stimulate catalytic activity (Zheng et al., 1996; Lee et al., 1997). It is interesting to note that an N-terminally truncated derivative of ASV integrase (i.e. lacking the zinc binding motif) can catalyze both processing and joining reactions in vitro (Bushman and Wang, 1994; Katz et al., 1996). Because mutants encoding this apparently active N-terminal truncated form of ASV integrase do not replicate, it seems possible that the N-terminal domain may be necessary for protein-protein interactions in vivo. Further structural and biochemical analyses should help to clarify the role of this domain in viral DNA integration.

The core domain of integrase contains the evolutionarily conserved D,D(35)E motif, comprising three acidic residues, the last two separated by 35 amino acids in the primary sequence (Fayet et al.,

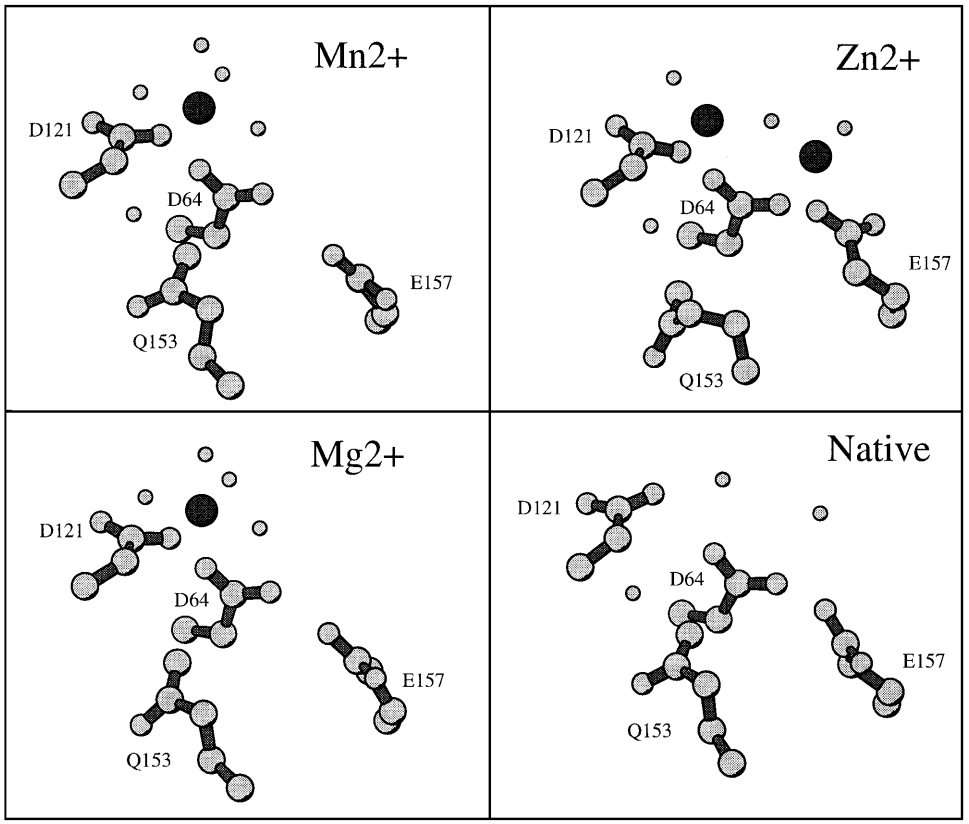


Fig. 3. Metal ion coordination at the active site of ASV integrase. Ball and stick models derived from crystal structures of the core domain of ASV integrase (apo–enzyme, bottom right) complexed with three divalent metal ions, Mn^{2+} (top left), Zn^{2+} (top right), Mg^{2+} (bottom left). Only the side-chain atoms are depicted. Large spheres represent metal ions; small spheres, water oxygens. The structures confirm that the conserved amino acids in the D,D(35)E motif (D64, D121, E157) coordinate the metal ion required for catalysis.

1990; Khan et al., 1991; Kulkosky et al., 1992). A similar constellation of acidic amino acids is also seen in the retrotransposon integrases and in the transposases of certain bacterial insertion sequences. Substitutions of any of these acidic amino acids in integrase leads to a significant decrease in all catalytic activities. The structural basis for the catalytic importance of the three amino acids has been revealed through solution of the crystal structures of metal complexes of the core domain of ASV integrase (residues 52–207) (Bujacz et al., 1995b; 1997). In these structures the acidic amino acids form binding pockets that coordinate the divalent metal ions required for catalysis (Fig. 3). The crystal structure of the isolated core domain of HIV-1 integrase was the first to be solved (Dyda et al., 1994). It revealed that the retroviral integrases belong to a superfamily of nucleases and recombinases including *E. coli* RNase H (Katayanagi et al., 1993), HIV RNase H (Davies et al., 1991), *E. coli* Ruv C (Ariyoshi et al., 1994), and bacteriophage Mu transposase (Rice and Mizuuchi, 1995) (reviewed in Andrake and Skalka, 1996; Rice et al., 1996; Yang and Steitz, 1995).

Despite the availability of structures of two integrase core domains, and their topological similarity to other recombinases, it is still not possible to deduce the exact manner in which DNA substrates are bound. A substrate (or inhibitor)-core domain complex has yet to be reported, and it seems likely that important structural elements

that help mold the active site are missing from the isolated core fragments. Thus, it is perhaps not surprising that the isolated core domains can neither process nor join viral DNA end sequences to target DNAs. Nevertheless, several studies indicate that determinants of specificity for recognition of the viral DNA ends reside primarily in the core domain of integrase (Kulkosky et al., 1995; Goulaouic and Chow, 1996; Pahl and Flügel, 1995; M. Katzman, personal communication). DNA-integrase crosslinking studies have suggested a critical role for the conserved lysine 159 in the HIV-1 integrase catalytic domain (Jenkins, Esposito and Craigie, personal communication).

The C-terminal domain of integrases comprises approximately 80 amino acids and it is the most evolutionarily diverged domain in these proteins. Deletion of this domain in integrases abolishes the ability to catalyze either the processing or joining reactions. Surprisingly, a C-terminally truncated derivative of FIV integrase has been reported to perform the processing and joining of viral DNA ends. The joining, however, occurs less efficiently than with the full-length protein (Shibagaki et al., 1997). Recent studies have demonstrated that structural elements in the C-terminal domain contribute to multimerization of integrases (Andrake and Skalka, 1995; Jenkins et al., 1996). The structure of an isolated peptide corresponding to the C-terminal domain of HIV-1 integrase (residues 220–270) has been solved by NMR (Lodi et al., 1995; Eijkelenboom et al., 1995). The structure is made up primarily of five β-strands that form two antiparallel β sheets. This fragment exhibits a fold similar to that of the Src homology 3 (SH3) domain and dimerizes in solution. It has been known for some time that the C-terminal domain of the retroviral enzymes possesses nonspecific DNA binding capability. A conserved residue, lysine 264, required for this property (Puras-Lutzke et al., 1994) was revealed to be located on a solvent exposed surface of this SH3-like dimer where it might be readily accessible to DNA (Eijkelenboom et al., 1995; Lodi et al., 1995). Despite these provocative observations, the significance of the C-terminal dimer structure and its relative spatial arrangement with respect to the catalytic core and the N-terminal domains in the full-length protein remain to be determined. Although details of the catalytic mechanism cannot be inferred directly from the isolated structures, recent studies on the role of divalent metal ions in the structural organization of HIV-1 integrase have provided some clues. They suggest that the C-terminal domain interacts more closely with the core domain in the presence of the metal cofactor (Asante-Appiah and Skalka, 1997). Such an interaction may bring the core and C-terminal DNA binding domains together in the catalytically competent enzyme.

8. The viral DNA substrate

In addition to integrase, certain *cis*-acting sequences in the viral DNA are required for integration (Panganiban and Temin, 1983; Colicelli and Goff, 1985). Deletions and base substitutions introduced into viral DNA have revealed important nucleotides within the long terminal repeats (LTRs) (LaFemina et al., 1991; Leavitt et al., 1992; Vink et al., 1991). An oligonucleotide duplex of about 15 bp of DNA is the minimum length that will support substrate cleavage in vitro (Katzman et al., 1989; Bushman and Craigie, 1991; Vink et al., 1991). Of these, nucleotides within the terminal 7 to 9 bp appear to be most important for specific recognition. Located within this terminal region is the conserved CA dinucleotide, usually at the 3rd and 4th position from the 3′ end of the viral sequence. Both conserved nucleotides are found near the 3′ ends of all retroviral DNA sequences and also in some transposable elements.

As noted above, the 3′-OH of the conserved CA dinucleotide, exposed during the processing reaction, is essential for the joining step. At the same time that processing reaction produces this new, recessed 3′ end, it leaves a two nucleotide overhang at the complementary 5′ end of substrates (Fig. 1). This 5′ end dinucleotide appears to guide the formation of a stable integrase–DNA complex in vitro (Ellison and Brown, 1994). Because the processing reaction takes place in the cytoplasm, the formation of such a stable intermediate may prevent the 'primed' substrate from being

diverted into potential nonproductive side reactions prior to entry of the pre-integration complex into the nucleus. Nucleotides upstream of the CA dinucleotide have also been found to influence the integration reactions. With the aid of chimeric proteins composed of complementary domains of integrase from HIV- 1 and visna viruses, recent studies have indicated that nucleotides at the 5th and 6th positions from the 3′ end influence specificity requirements of each enzyme for its cognate substrate (Katzman and Sudol, 1996). Nucleotides that are far removed from the viral DNA termini have also been shown to exert some influence on integration reactions in vitro (Hong et al., 1993).

The use of single-end oligonucleotide substrates has been invaluable in distinguishing and understanding the two major steps of integration. However, for the most part, only single ends are processed and joined in these assays. Thus, these 'half-reactions' do not represent the coordinated events that engage both viral DNA ends in vivo. The development and use of novel DNA substrates that show coordinated processing and joining in vitro is now providing a deeper insight into the molecular mechanism of integration (Vora et al., 1994; Kukolj and Skalka, 1995; Aiyar et al., 1996). For example, analysis of the crystal structures of the core domains of HIV-1 (Dyda et al., 1994) and ASV integrases (Bujacz et al., 1995b) raised new questions about how specifically spaced staggered cuts are made by the enzymes during the joining step. The isolated core domains of these integrases crystallize as tight dimers, with their active sites spaced 35 Å (HIV) and 38 Å (ASV) apart. These distances are a bit too far to accommodate the 5 and 6 bp staggered cut that separate the target phosphates during joining, if the target DNA is presumed to be in the B form. But they are not too far if one allows that substrate DNAs may be unwound by the enzyme prior to catalysis. That this is likely to be the case for processing was first demonstrated through the use of a novel oligonucleotide substrate in which viral DNA end sequences were joined together (synapsed) by a single strand tether, and with which coordinated processing was observed (Kukolj and Skalka, 1995). These studies revealed a requirement for torsional and rotational flexibility in the tether, and showed that the optimal separation between processing sites in these synapsed ends corresponded to 5 bp for HIV-1 integrase and 6 bp for ASV integrase. Fig. 4 shows a model that was proposed to accommodate these properties in the processing and, presumably symmetrical, joining steps. Results from recent studies that used viral DNA substrates with non-complementary termini support this model (Scottoline et al., 1997).

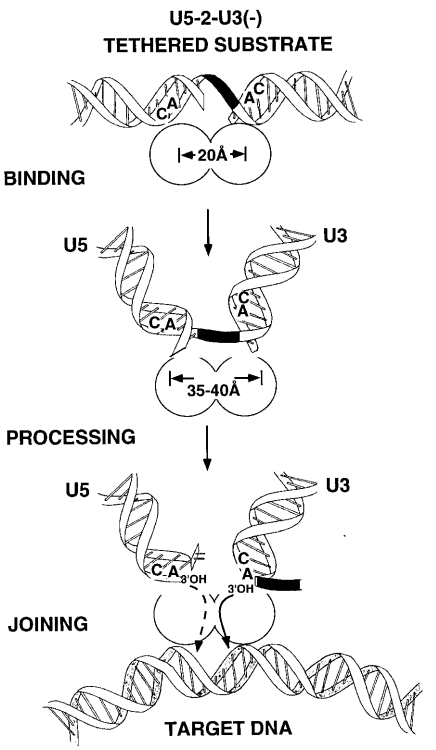


Fig. 4. A model for the processing and joining of synapsed-end substrates by retroviral integrases. The linear B-DNA form of the tethered retroviral DNA ends (U5 and U3) are separated by ~20 Å (shaded helical region). The flexibility of the tether allows repositioning of the processing sites to accommodate the separation of catalytic sites in an integrase dimer (circles). Because even nucleosomal targets may have less than the predicted (35–40 Å) spacing observed in the crystal structures of the core domains of ASV and HIV integrases following processing, the target DNA strands may be accommodated by the introduction of torsional relief after the first nucleophilic attack by one of the 3′ OH's of the processed strands at the target site. This would then facilitate the subsequent insertion of the other viral DNA processed end using the same or symmetrically paired catalytic sites in the functional integrase multimer. (Adapted from Kukolj and Skalka, 1995).

9. Cellular factors

Purified integrase protein is both necessary and sufficient for the catalysis of both processing and joining reactions in vitro. However, experiments with pre-integration complexes isolated from infected cells show a more efficient, concerted reaction suggesting that some components or features of the complex may enhance and modify the activity of integrase (Brown et al., 1987; Ellison et al., 1990; Lee and Coffin, 1990). The observation that certain nuclear proteins, such as the HMG class of DNA bending proteins, can enhance concerted integration of viral DNA ends in vitro suggested that such proteins might play a role as accessory factors in viral DNA integration in vivo (Aiyar et al., 1996). The cellular protein HMG I(Y), was subsequently found to be associated with pre-integration complexes isolated from HIV-1 infected cells (Farnet and Bushman, 1997). The integration competency of the HIV-1 pre-integration complexes is greatly reduced when HMG I(Y) is removed from the complex by salt extraction, and the subsequent addition of this cellular protein to the complex restores integrative efficiency. The involvement of HMG I(Y) in integration appears to depend on interaction with DNA and not the integrase (Farnet and Bushman, 1997; Hindmarsh, Skalka and Leis, unpublished observations).

The idea that cellular proteins contribute to HIV-1 DNA integration was first suggested by the isolation of a specific protein that binds to integrase, called Ini-1 (integrase interactor 1), that is a human homologue of the yeast SNF5 transcription factor (Kalpana et al., 1994). Ini-1 promotes the joining of viral DNA ends to a target DNA in vitro. However, the relevance of this interaction to the integration of viral DNA in vivo has not yet been established.

In the absence of sequence-specificity for joining, it is unclear how retroviruses avoid integrating into their own DNA. There are now hints from studies with MoMLV that a host cell protein (from NIH 3T3 cells) may play a role in preventing such auto-integration (Lee and Craigie, 1994). The absence of this protein in the nucleus, coupled with the availability of facilitating nuclear components such as the HMG's and Ini-1-like proteins, may contribute to the regulation of viral DNA integration.

10. The metal cofactor

The development of in vitro assays for integration revealed the requirement for a metal cofactor, Mn^{2+} or Mg^{2+}. Because of the low intracellular concentration of Mn^{2+} ($< 10^{-7}$ M), it is generally assumed that the more abundant Mg^{2+} (intracellular concentration $\sim 10^{-3}$ M) is the relevant cofactor in vivo. However, Mn^{2+} appears to be preferred in vitro; enzyme specific activities are at least 10-fold higher in Mn^{2+} than Mg^{2+} under similar conditions. This preference can be influenced by the length of the oligonucleotide substrate employed in the assay. Upon increasing the length of the substrate from the commonly used 19 or 21 bp to 35 bp, Lee et al. (1995) observed a switch of metal preference from Mn^{2+} to Mg^{2+}. In addition, under certain assay conditions, the activity of integrase appears not only to be higher but also more specific with Mg^{2+} than Mn^{2+} (Vora et al., 1994; Engelman and Craigie, 1995). Because integrase also produces nonspecific nicks in the presence of Mn^{2+} but not Mg^{2+}, these results reinforce the view that Mg^{2+} is the cofactor used for integration in vivo.

The mechanism by which metal ions support the integration reaction remains to be defined. The solution of crystal structures of metal–ASV integrase catalytic core domain complexes have provided some important clues regarding the role and number of metal ions employed in catalysis (Bujacz et al., 1997). The crystal structures of the Mn^{2+} and Mg^{2+} complexes of the ASV integrase core domain show a single bound metal ion coordinated by Asp64 and Aspl21 (site 1), the first two residues in the conserved D,D(35)E motif, with water molecules providing the additional ligands (Fig. 3). Although it does not support catalysis, a single Ca ion can bind in the same location. Two Cd or Zn ions are bound in the ASV integrase core domain complex. One ion is in site 1 and the second is bound between Asp64 and the last

residue of the conserved acidic triad, Glu157 (site 2); with water molecules completing the coordination (Bujacz et al., 1997) (Fig. 3). Unlike Cd, the Zn ion can support some enzymatic activity with both the full-length and core domain of ASV integrase. These structural models are consistent with a two-metal mechanism of catalysis by integrase (Bujacz et al., 1997). A mechanism similar to that proposed for the 3′–5′ exonuclease of *E. coli* DNA Pol 1 (Beese and Steitz, 1991) has been suggested in which one metal serves as an electrophilic catalyst while the other orients the substrate (Steitz et al., 1994; Yang and Steitz, 1995). The precise role of the metal ions, however, is yet to be delineated by biochemical studies. A structure of a metal–HIV-1 integrase complex is not currently available. However, the similar activities and topologies of ASV and HIV-1 core domains (Bujacz et al., 1995a; Dyda et al., 1994) together with results from site-directed mutagenesis studies (Engelman and Craigie, 1992; Kulkosky et al., 1992) predict that the corresponding residues of the D,D(35)E motif in HIV-1 integrase (Asp64, Aspl 16 and Glu152), will coordinate metal ions in a similar fashion.

Recent studies have shown that the specific activity of HIV-1 integrase increases when it is preincubated with metal ions (e.g. Mn^{2+} or Mg^{2+}; Ellison et al., 1995; Wolfe et al., 1996; Pemberton et al., 1996; Zheng et al., 1996; Lee et al., 1997; Asante-Appiah and Skalka, 1997). Addition of a metal cofactor to apo–HIV-1 integrase was shown to promote a conformational change in the protein that is required for optimal activity (Asante-Appiah and Skalka, 1997). Metal ions also facilitate the assembly of a stable metal–HIV integrase–DNA ternary complex (Engelman et al., 1994; Wolfe et al., 1996; Pemberton et al., 1996). The role of the metal cofactor has not been examined to the same degree with other retroviral integrases. However, ASV integrase appears to differ from HIV-1 integrase as it does not show a similar increase in enzymatic activity when pre-incubated with the metal cofactor, nor is there any evidence of a metal-induced conformational change in the ASV protein (Asante-Appiah and Skalka, 1997). The reason(s) for the differential effects of divalent metal ions on the conformation and activity of these related proteins is not known. As suggested by the structural analyses of its catalytic core–metal complexes, ASV integrase may have pre-formed metal binding sites and, therefore, may not require significant changes in protein conformation for activation (Bujacz et al., 1996). On the other hand, the crystal structure of the core domain of apo–HIV-1 integrase shows that the residues of the D,D(35)E motif are not properly oriented to bind metal ions (Bujacz et al., 1996). Thus, it seems likely that metal-induced HIV integrase conformational changes recruit the catalytic apparatus into the active configuration (Asante-Appiah and Skalka, 1997).

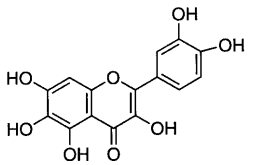
Because studies aimed at understanding the role of the metal cofactor have been performed primarily with in vitro integration reactions, their relevance to the situation in vivo can only be conjectural. However, recent studies with conformation-sensitive monoclonal antibodies against HIV-1 integrase suggest that the enzyme may undergo changes in vivo that are similar to the metal-induced changes observed in vitro (Levy-Mintz et al., 1996). One such monoclonal antibody (mAb33) recognizes HIV-1 integrase only in its non-activated conformation (Asante-Appiah and Skalka, 1997). When expressed intracellularly as a single-chain variable fragment derivative (sFv-33), it protects human SupT1 and peripheral blood mononuclear cells (PBMC's) against HIV-1 infection (Levy-Mintz et al., 1996). Interestingly, a derivative of sFv-33 which carries a nuclear localization signal (sFv-33-Nu) protects cells somewhat less well than the cytoplasmically-expressed version. Initial results suggest that sFv33 blocks an early step in viral replication, before integration occurs. Thus, it seems likely that at some time during the infectious cycle in vivo (either in the pre-integration complex or as part of the Gag-Pol precursor) HIV-1 integrase exists in a non-activated conformation that is recognized by the conformationally-sensitive antibody.

11. Integrase as a target for drug intervention

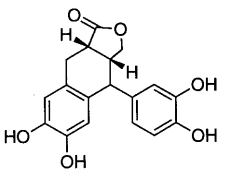
Unlike the other retroviral enzymes that have been targeted for inhibitory drug design (i.e. re-

verse transcriptase and protease), comparatively little has been achieved by way of promising therapies for AIDS intervention with integrase. The absence of a crystal structure for the full length protein has limited the application of computer modeling approaches to drug design. Although the structures of the isolated domains provide important starting points in understanding the structural basis of the mechanism of integration (Bujacz et al., 1995b; Dyda et al., 1994; Lodi et al., 1995; Eijkelenboom et al., 1995), we still do not know how the three domains are organized in the full length protein and within the functional multimer. As none of the isolated domains can perform the processing or joining of viral DNA ends, the residues that constitute the bona fide catalytic apparatus have yet to be revealed. This has also hampered the design of mechanism-based inhibitors or inactivators. The majority of compounds reported so far have been identified simply by screening for inhibition of integrase activity in vitro. Many do so by virtue of their ability to bind DNA; some also inhibit other DNA-metabolizing enzymes and, therefore, lack the desired specificity.

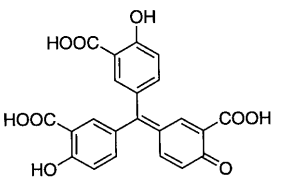
A variety of compounds have been tested as potential inhibitors against HIV-1 integrase in vitro; these include the bis-catechols (LaFemina et al., 1995), caffeic acid esters and analogues (Fesen et al., 1994; Robinson et al., 1996), aurintricarboxylic acids and cosalene analogues (Cushman and Sherman, 1992; Cushman et al., 1995), hydroxylated arylamides (Zhao et al., 1997), DNA intercalators (Fesen et al., 1993), topoisomerase inhibitors (Carteau et al., 1993), and several naturally occurring compounds including curcumin (Mazumder et al., 1995), lignans (Eich et al., 1996) flavones, and their analogues (Fesen et al., 1994). Structure–activity relationships of several organic compounds have revealed two main classes of inhibitors with respect to molecular configuration. The first type is characterized by multi-hydroxylation on an aromatic ring system. The structures of some examples of these phenolic compounds (the bis-catechol, β-conindendrol, aurintricarboxylic acid monomer, and the flavone quercetagenin) are shown in Fig. 5. The presence of at least two functional groups with hydrogen



QUERCETAGENIN



β–CONINDENDROL



AURINTRICARCARBOXYLIC ACID

Fig. 5. Chemical structures of some type 1 inhibitors of HIV-1 integrase.

bonding potential, and in ortho positions on the rings, appears to be important for inhibition. The second type of inhibitors is exemplified by caffeic acid phenethyl ester, curcumin, and L-chicoric acid (Fig. 6). The presence of at least one hydroxyl group on two aromatic rings separated by a long linker appears to be an important structural feature for this class of inhibitors. Energy minimization analysis on curcumin suggests that the linker region may allow the aromatic rings to fold back and stack on each other (Mazumder et al., 1995). The net effect may be to juxtapose the rings and hydroxyls in a spatial configuration similar to that in the class 1 type of inhibitors. It also seems possible that two non-overlapping binding loci for hydroxylated aromatic compounds exist at the active site of HIV-1 integrase. The observation that bis-catechols are generally more potent inhibitors compared to the monomeric analogues is consistent with such an

interpretation. Further work on enzyme-inhibitor interactions should shed more light on the binding modes of these compounds.

The bis-catechols, arylamides (e.g. the hydroxylated bis-naphthoylamides, Fig. 6) and dicaffeoylquinic acids are promising lead structures with IC_{50}s in the micromolar to submicromolar range. Unfortunately, with the exception of the dicaffeoylquinic acids (Robinson et al., 1996), these inhibitors have little or no inhibitory activity against HIV-1 replication when assayed in tissue culture. In vitro, the inhibitors appear to act by binding to the core domain of integrase which includes the active site of the enzyme. However, their kinetic mechanism of action cannot be ascertained directly by the available in vitro assays. As most of the inhibitors can also intercalate in DNA, it remains to be established that their inhibitory effects reflect specific binding to the enzyme. The presence of *ortho* bis-hydroxyl moieties on most of these compounds suggests that chelation of the essential metal cofactor may also contribute to their inhibitory effects. The propensity for these hydroxylated compounds to generate highly reactive cytotoxic species (e.g. quinones) upon oxidation raises concern regarding their utility in clinical settings.

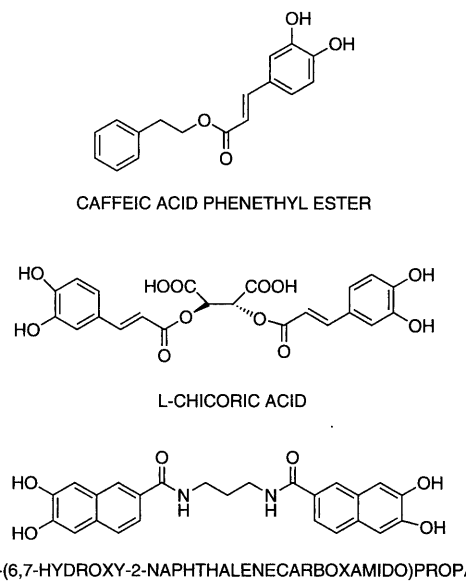


Fig. 6. Chemical structures of some type 2 inhibitors of HIV-1 integrase.

Although they are not very selective at the present time, it is encouraging that nucleotide analogues have also been reported to inhibit HIV-1 integrase (Lipford et al., 1994; Mazumder et al., 1994; 1996). As these inhibitors appear to block activity by binding to the catalytic core domain, they may serve as leads for the design of the next generation of more potent and specific analogues. The guanosine quartet-forming oligonucleotides described by Pommier and coworkers (Mazumder et al., 1996) are rather potent inhibitors of HIV-1 integrase with antiviral activity (Ojwang et al., 1995). Interestingly, these oligonucleotides appear to exert their effects by binding preferentially to the N-terminal domain of integrase. Further work should reveal the nature of this interaction and if it can be exploited for designing more potent simple compounds. In another approach, screening of synthetic combinatorial peptide libraries has identified a hexapeptide, HCKFWW, that inhibits HIV-1 integrase with an IC_{50} of 2 μM (Puras-Lutzke et al., 1995). Such peptides may subsequently be modified to evade cellular proteases and increase their bioavailability once their mechanism of inhibition has been established. A 30mer peptide spanning residues 147–175 that includes Glu152, the third member of the D,D(35)E motif, has been reported to inhibit HIV-1 integrase activity with an IC_{50} in the mM range (Scourgen et al., 1996). A K159P substitution in this peptide renders it inactive as an inhibitor. As this residue is important for DNA binding (Jenkins, Esposito and Craigie, personal communication), it seems possible that the peptide may inhibit integrase by sequestering the substrate oligonucleotide.

Libraries of potential RNA inhibitors have also been screened (Allen et al., 1995). Although RNA molecules that bind to integrase with dissociation constants in the nanomolar range have been identified, their ability to inhibit the enzyme's activity in a specific fashion remains to be demonstrated. In addition to using small organic compounds and polymers, alternate approaches to the development of antiviral agents against HIV-1 integrase may also be pursued. As we have already noted, single-chain variable fragments (sFv)

derived from inhibitory monoclonal antibodies have been shown to block viral infection upon their intracellular expression in human T1 cell lines and PBMC's (Levy-Mintz et al., 1996). SFvs have the potential to be exploited in an 'intracellular immunization,' gene therapy approach. In this regard, it will be of interest to know if the intracellularly expressed antibodies can block the production of infectious virions in chronically-infected cells.

12. Concluding remarks

The accumulating knowledge concerning the mechanism of viral DNA integration is beginning to provide the tools necessary for the design and development of inhibitors against integrase. Early efforts have focused on compounds that may compete for binding at the active site of this viral enzyme which we now know to be a member of a superfamily of nucleases and recombinases. Inhibitors that block other aspects of integrase function may be designed, as additional details of its mechanism of action and knowledge of the intracellular events required for integration become available. For example, the observation that HIV-1 integrase may exist in an alternate conformation prior to activation suggests the potential development of a class of inhibitors that bind to integrase and block the required conformational changes. The slow-binding and increased potency of CAPE (caffeic acid phenethyl ester) upon preincubation with the enzyme suggests that this inhibitor may recognize the activated form of the enzyme whereas quercetagenin, whose potency does not increase with preincubation, may recognize the non-activated or both conformations of the enzyme (Fesen et al., 1994). Other approaches to be explored include ligands that specifically target the C-terminus and those that may block dimerization or oligomerization. Interaction with cellular proteins that are important for the viral pre-integration complex to gain access to host DNA may also present opportunities for inhibition.

Although great strides have been made during the past few years in understanding the mechanism of integrase, the structural basis of catalysis still remains uncertain. In the absence of a detailed model of how integrase functions, structure–activity analyses with several lead compounds are providing some clues for the design of more potent inhibitors against this viral target. More information concerning the 3-dimensional structure of the full-length enzyme would certainly accelerate efforts in rational drug design. The success of the past few years should spur on further work in this and related areas. Understanding the molecular basis of retroviral DNA integration will surely provide additional avenues for drug discovery in the future.

Acknowledgements

This work was supported by National Institutes of Health grants CA-47486, CA-06927, CA-71515, a grant for infectious disease research from Bristol-Myers Squibb Foundation, and also by an appropriation from the Commonwealth of Pennsylvania.

References

Aiyar, A., Hindmarsh, P., Skalka, A.M., Leis, J., 1996. Concerted integration of linear retroviral DNA by the avian sarcoma virus integrase in vitro: dependence on both long terminal repeat termini. J. Virol. 70, 3571–3580.

Allen, P., Worland, S., Gold, L., 1995. Isolation of high-affinity RNA ligands to HIV-1 integrase from a random pool. Virology 209, 327–336.

Andrake, M., Skalka, A.M., 1995. Multimerization determinants reside in both the catalytic core and C-terminus of avian sarcoma virus integrase. J. Biol. Chem. 270, 29299–29306.

Andrake, M., Skalka, A.M., 1996. Retroviral integrases: putting the pieces together. J. Biol. Chem. 271, 19633–19636.

Asante-Appiah, E., Skalka, A.M., 1997. A metal-induced conformational change and activation in HIV-1 integrase. J. Biol. Chem. 272, 16196–16205.

Ariyoshi, M., Vassylyev, D.G., Iwasaki, H., Nakamura, H., Shinagawa, H., Morikawa, K., 1994. Atomic structure of Ruv C resolvase: a Holliday junction-specific endonuclease from *E. coli*. Cell 78, 1063–1072.

Beese, L.S., Steitz, T.A., 1991. Structural basis for the 3′–5′ exonuclease activity of *Escherichia coli* DNA polymerase I: a two-metal ion mechanism. EMBO J. 10, 25–33.

Bowerman, B., Brown, P.O., Bishop, J.M., Varmus, H.E., 1989. A nucleoprotein complex mediates the integration of retroviral DNA. Genes Dev. 3, 469–478.

Brown, P.O., Bowerman, B., Varmus, H.E., Bishop, J.M., 1987. Correct integration of retroviral DNA in vitro. Cell 49, 347–356.

Bujacz, G., Jaskólski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R.A., Skalka, A.M., 1995a. High resolution structure of the catalytic domain of avian sarcoma virus integrase. J. Mol. Biol. 253, 333–346.

Bujacz, G., Jaskólski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R.A., Skalka, A.M., 1995b. The catalytic domain of avian sarcoma virus integrase: confirmation of the active site residues in the presence of divalent cations. Structure 4, 89–96.

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

Bujacz, G., Jaskólski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Andrake, M.A., Katz, R.A., Skalka, A.M., 1997. Binding of different divalent cations to the active site of ASV integrase and their effects on enzymatic activity. J. Biol. Chem. 272, 18161–18168.

Bukrinskaya, A.G., Ghorpade, A., Heinzinger, N.K., Smithgall, T.E., Lewis, R.E., Stevenson, M., 1996. Phosphorylation-dependent human immunodeficiency virus type 1 infection and nuclear targeting of viral DNA. Proc. Natl. Acad. Sci. USA 93, 367–371.

Bukrinsky, M.I., Sharova, N., Dempsey, M.P., Stanwick, T.L., Bukrinskaya, A.G., Haggerty, S., Stevenson, M., 1992. Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. Proc. Natl. Acad. Sci. USA 89, 6580–6584.

Bukrinsky, M.I., Haggerty, S., Dempsey, M.P., Sharova, N., Adzhubei, A., Spitz, L., Lewis, P., Goldfarb, D., Emerman, M., Stevenson, M., 1993a. A nuclear localization signal with HIV-1 matrix protein that governs infection of non-dividing cells. Nature 365, 666–669.

Bukrinsky, M.I., Sharova, N., McDonald, T.L., Pusharskaya, T., Stevenson, M., 1993b. Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. Proc. Natl. Acad. Sci. USA 90, 6125–6129.

Burke, C.J., Sanyal, G., Bruner, M.W., Ryan, J.A., LaFemina, R.L., Robbins, H.L., Zeft, A.S., Middaugh, C.R., Cordingley, M.G., 1992. Structural implications of spectroscopic characterization of a putative zinc finger peptide from HIV-1 integrase. J. Biol. Chem. 267, 9639–9644.

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

Bushman, F.D., Wang, B., 1994. Rous sarcoma virus integrase: mapping functions for catalysis and substrate binding. J. Virol. 68, 2215–2223.

Cai, M., Zheng, R., Caffrey, R., Clove, M., Gronenborn, A.M., 1997. Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. Nature Struct. Biol. 4, 567–577.

Carteau, S., Mouscadet, J.F., Goulaouic, H., Subra, F., Auclair, C., 1993. Effect of topoisomerase inhibitors on the in vitro HIV DNA integration reaction. Biochem. Biophys. Res. Commun. 192, 1409–1414.

Chow, S.A., Vincent, K.A., Ellison, V., Brown, P.O., 1992. Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus type. Science 25, 723–726.

Colicelli, J., Goff, S.P., 1985. Mutants and pseudorevertants of Moloney murine leukemia virus with alterations at the integration site. Cell 42, 573–580.

Connell, E.V., Hsu, M.-C., Richman, D.D., 1994. Combinative interactions of a human immunodeficiency virus (HIV) tat antagonist with reverse transcriptase inhibitors and HIV protease inhibitor. Antimicrob. Agents Chemother. 38, 348–352.

Craig, J.C., Duncan, J.B., Whitakker, L., Roberts, N.A., 1993. Antiviral synergy between inhibitors of HIV proteinase and reverse transcriptase. Antivir. Chem. Chemother. 4, 161–166.

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

Craigie, R., 1992. Hotspots and warm spots: integration specificity of retroelements. Trends Genet. 8, 187–190.

Cushman, M., Sherman, P., 1992. Inhibition of HIV-1 integration protein by aurintricarboxylic acid monomers, monomer analogs, and polymer fractions. Biochem. Biophys. Res. Commun. 185, 85–90.

Cushman, M., Golebiewski, W.M., Pommier, Y., Mazumder, A., Reymen, D., De Clerk, E., Graham, L., Rice, W.G., 1995. Cosalene analogues with enhanced potencies as inhibitors of HIV-1 protease and integrase. J. Med. Chem. 38, 443–452.

Davies, J.F. II, Hostomska, Z., Hostomsky, Z., Jordan, S.R., Matthews, D.A., 1991. Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. Science 252, 88–95.

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

Eich, E., Pietz, H., Kaloga, M., Schultz, J., Fesen, M.R., Mazumder, A., Pommier, Y., 1996. (−)-Arctigenin as a lead structure for inhibitors of human immunodeficiency virus type 1 integrase. J. Med. Chem. 39, 86–95.

Eijkelenboom, A.P.A.M., Puras-Lutzke, R., Boelens, R., Plasterk, R.H.A., Kaptien, R., Hürd, K., 1995. The DNA binding domain of HIV-1 integrase has an SH3-like fold. Nat. Struct. Biol. 2, 807–810.

Ellison, V., Abrams, H., Roe, T., Lifson, J., Brown, P.O., 1990. Human immunodeficiency virus integration in a cell-free system. J. Virol. 64, 2711–2715.

Ellison, V., Brown, P.O., 1994. A stable complex between integrase and viral DNA ends mediates human immunodeficiency virus integration in vitro. Proc. Natl. Acad. Sci. USA 91, 7316–7320.

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

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

Engelman, A., Mizuuchi, K., Craigie, R., 1992. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell 67, 1211–1221.

Engelman, A., Bushman, F.D., Craigie, R., 1993. Identification of discreet functional domains of HIV-1 integrase and their organization within an active multimeric complex. EMBO J. 12, 3269–3275.

Engelman, A., Hickman, A.B., Craigie, R., 1994. The core and C-terminal domains of the integrase from human immunodeficiency virus type 1 each contribute to nonspecific DNA binding. J. Virol. 68, 5911–5917.

Engelman, A., Craigie, R., 1995. Efficient magnesium-dependent human immunodeficiency virus type 1 integrase activity. J. Virol. 69, 5908–5911.

Engelman, A., Ying, L., Chen, H., Farzan, M., Dyda, F., 1997. Structure-based mutagenesis of the catalytic domain of human immunodeficiency type 1 integrase. J. Virol. 71, 3507–3514.

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

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

Fayet, O., Ramond, P., Polard, P., Prére, M.F., Chandler, M., 1990. Functional similarities between retroviruses and the IS3 family of bacterial insertion sequences?. Mol. Microbiol. 4, 1771–1777.

Fesen, M.R., Kohn, K.W., Lefeurtre, F., Pommier, Y., 1993. Inhibitors of human immunodeficiency virus integrase. Proc. Nat. Acad. Sci. USA 90, 2399–2403.

Fesen, M.R., Pommier, Y., Leteurtre, F., Hiroguchi, S., Yung, J., Kohn, K., 1994. Inhibition of HIV-1 integrase by flavones, caffeic acid phenthyl ester (CAPE) and related compounds. Biochem. Pharm. 48, 595–608.

Fitzgerald, M.L., Vora, A.C., Zeh, W.G., Grandgenett, D.P., 1992. Concerted integration of viral DNA termini by purified avian myeloblastosis virus integrase. J. Virol. 66, 6257–6263.

Freed, E.O., Martin, M.A., 1994. HIV-1 infection of non-dividing cells. Nature 369, 107–108.

Freed, E.O., Englund, G., Martin, M.A., 1995. Role of the basic domain of human immunodeficiency virus type 1 matrix in macrophage infection. J. Virol. 69, 3949–3954.

Freed, E.O., Englund, G., Maldarelli, F., Martin, M.A., 1997. Phosphorylation of residue 131 of HIV-1 matrix is not required for macrophage infection. Cell 88, 171–174.

Gallay, P., Swingler, S., Aiken, C., Trono, D., 1995a. HIV-1 infection of nondividing cells: C-terminal tyrosine phosphorylation of the viral matrix protein is a key regulator. Cell 80, 379–388.

Gallay, P., Swingler, S., Song, J., Bushman, F., Trono, D., 1995b. HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase. Cell 83, 569–576.

Goodarzi, G., Im, G.-J., Brackmann, K., Grandgenett, D.P., 1995. Concerted integration of retrovirus-like DNA by human immunodeficiency virus type 1 integrase. J. Virol. 69, 6090–6097.

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

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

Hickman, A.B., Palmer, A., Engelman, R., Craigie, R., Wingfield, P., 1994. Biophysical and enzymatic properties of the catalytic domain of HIV-1 integrase. J. Biol. Chem. 46, 29279–29287.

Hong, T., Murphy, E., Groarke, J., Drlica, K., 1993. Human immunodeficiency virus type 1 DNA integration: Fine structure target analysis using synthetic oligonucleotides. J. Virol. 67, 1127–1131.

Jenkins, T.M., Engelman, A., Ghirlando, R., Craigie, R., 1996. A soluble active mutant of HIV-1 integrase: multimerization involves the core and C-terminal domains. J. Biol. Chem. 271, 7712–7718.

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

Johnson, V.A., Walker, B.D., Barhus, M.A., Paradis, T.J., Chou, T.-C., Hirsch, M.S., 1989. Synergistic inhibition of human immunodeficiency virus type 1 and type 2 replication in vitro by castanospermine and 3′-azido-3′-deoxythymidine. Antimicrob. Agents Chemother. 33, 53–57.

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

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

Katayanagi, K., Okumura, M., Morikawa, K., 1993. Crystal structure of *Escherichia coli* RNase H1 in complex with

Mg at 2.8 A resolution: proof of a single Mg-binding site. Prot. Struct. Func. Genet. 17, 337–346.

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

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

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

Katzman, M., Sudol, M., 1996. Influence of subterminal viral DNA nucleotides on differential susceptibility to cleavage by human immunodeficiency virus type 1 and visna virus integrases. J. Virol. 70, 9069–9073.

Kitamura, Y., Lee, Y.M., Coffin, J.M., 1992. Nonrandom integration of retroviral DNA in vitro: effect of CpG methylation. Proc. Natl. Acad. Sci. USA 89, 5532–5536.

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

Kukolj, G., Skalka, A.M., 1995. Enhanced and coordinated processing of viral DNA ends by retroviral integrates in vitro. Genes Dev. 9, 2556–2567.

Kukolj, G., Jones, K.S., Skalka, A.M., 1997. Subcellular localization of avian sarcoma virus and human immunodeficiency virus type 1 integrases. J. Virol. 71, 843–847.

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

Kulkosky, J., Katz, R.A., Merkel, G., Skalka, A.M., 1995. Activities and substrate specificity of the evolutionarily conserved central domain of retroviral integrase. Virology 206, 448–456.

LaFemina, R.L., Callahan, P.L., Cordingley, M.G., 1991. Substrate specificity of recombinant human immunodeficiency virus type 1 integrase protein. J. Virol. 65, 5624–5630.

LaFemina, R.L., Schneider, C.L., Robbins, H.L., Callahan, P.L., LeGrow, K., Roth, E., Schleif, W.A., Emini, E.A., 1992. Requirement of active human immunodeficiency virus type 1 integrase enzyme for proactive infection of human T-lymphoid cells. J. Virol. 66, 7414–7419.

LaFemina, R.L., Graham, P.L., LeGrow, K., Hastings, J.C., Wolffe, A., Young, S.D., Emini, E.A., Hazuda, D., 1995. Inhibition of human immunodeficiency virus by bis-catechols. Antimicrob. Agents Chemother. 39, 320–324.

Leavitt, A.D., Rose, R.B., Varmus, H.E., 1992. Both substrate and target oligonucleotide sequences affect in vitro integration mediated by human immunodeficiency virus type 1 integrase protein produced in *Saccharomyces cerevisiae*. J. Virol. 66, 2359–2368.

Lee, Y.M.H., Coffin, J.M., 1990. Efficient autointegration of avian retrovirus DNA in vitro. J. Virol. 64, 5624–5630.

Lee, M.S., Craigie, R., 1994. Protection of retroviral DNA from autointegration: involvement of a cellular factor. Proc. Natl. Acad. Sci. USA 64, 5958–5965.

Lee, S.P., Kim, H.G., Censullo, M.L., Han, M.K., 1995. Characterization of Mg^{2+}-dependent 3′-processing activity for human immunodeficiency virus type 1 integrase in vitro real time kinetic studies using fluorescence resonance energy transfer. Biochemistry 34, 10205–10214.

Lee, S.P., Xiao, J., Knutson, J.R., Lewis, M.S., Han, M.K., 1997. Zn^{2+} promotes the self-association of human immunodeficiency virus type-1 integrase in vitro. Biochemistry 36, 173–180.

Levy-Mintz, P., Duan, L., Zhang, H., Hu, B., Dornadula, G., Zhu, M., Kulkosky, J., Bizub-Bender, D., Skalka, A.M., Pomerantz, R., 1996. Intracellular expression of single-chain variable fragments to inhibit early stages of the viral life cycle by targeting human immunodeficiency virus type 1 integrase. J. Virol. 70, 8821–8832.

Lipford, J.R., Worland, S.T., Farnet, C.M., 1994. Nucleotide-binding by the HIV-1 integrase protein in vitro. J. AIDS 7, 1215–1223.

Lodi, P.J., Ernst, J.A., Kuszewski, J., Hickman, A.B., Engelman, A., Craigie, R., Clore, G.M., Gronenborn, A.M., 1995. Biochemistry 34, 9826–9833.

Mazumder, A., Cooney, D., Agbaria, R., Gupta, M., Pommier, Y., 1994. Inhibition of human immunodeficiency virus type 1 by 3′-azido-3′-deoxythymidylate. Proc. Natl. Acad. Sci. USA 91, 5771–5775.

Mazumder, A., Raghavan, K., Weinstein, J., Kohn, K.W., Pommier, Y., 1995. Inhibition of human immunodeficiency virus type 1 integrase by curcumin. Biochem. Pharm. 49, 1165–1170.

Mazumder, A., Neamati, N., Ojwang, J.O., Sunder, S., Rando, R.F., Pommier, Y., 1996. Inhibition of human immunodeficiency virus type 1 integrase by guanosine quartet structures. Biochemistry 43, 13762–13771.

McEuen, A.R., Edwards, B., Koepke, K.A., Ball, A.E., Wolstenholme, A.J., Danson, M.J., Hough, D.W., 1992. Zinc binding by retroviral integrase. Biochem. Biophys. Res. Commun. 189, 813–818.

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

Müller, H.-P., Varmus, H.E., 1994. DNA bending creates favored sites for retroviral integration: an explanation for preferred insertion sites in nucleosomes. EMBO J. 13, 4704–4714.

Murphy, J.E., Goff, S.P., 1992. A mutation at one end of Moloney murine leukemia virus DNA blocks cleavage of both ends by the viral integrase in vivo. J. Virol. 66, 5092–5095.

Ojwang, J.O., Buckheit, R.W., Pommier, Y., Mazumder, A., de Vresse, K., Esté, J.A., Reymen, D., Pallansch, L.A., Lackman-Smith, C., Wallace, T.L., deClercq, E., McGrath, M.S., Rando, R.F., 1995. T30177, an aligonucleotide stabilized by an intramolecular guanosine octet, is a potent inhibitor of laboratory strains and clinical isolates of human immunodeficiency virus type 1. Antimicrob. Agents Chemother. 39, 2426–2435.

Pahl, A., Flügel, R.M., 1995. Characterization of the human spuma retrovirus integrase by site-directed mutagenesis, by complementation, and by swapping the zinc finger domain of HIV-1. J. Biol. Chem. 270, 2957–2966.

Panganiban, A.T., Temin, H.M., 1983. The terminal nucleotides of retrovirus DNA are required for integration but not virus production. Nature 306, 155–160.

Pemberton, I.K., Buckle, M., Buc, H., 1996. The metal ion-induced cooperative binding of HIV-1 integrase to DNA exhibits preference for Mn(II) rather than Mg(II). J. Biol. Chem. 271, 1498–1506.

Pryciak, P.M., Sil, A., Varmus, H.E., 1992. Retroviral integration into minichromosomes in vitro. EMBO J. 11, 291–303.

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

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

Puras-Lutzke, R.A., Eppens, N.A., Weber, P.A., Houghten, R.A., Plasterk, R.H.A., 1995. Identification of a hexapeptide inhibitor of the human immunodeficiency virus integrase protein by using a combinatorial chemical library. Proc. Natl. Acad. Sci. USA 92, 11456–11460.

Rice, P., Mizuuchi, K., 1995. Structure of the bacteriophage Mu transposase core: a common structural motif for DNA transposition and retroviral integration. Cell 82, 209–220.

Rice, P., Craigie, R., Davies, D.R., 1996. Retroviral integrases and their cousins. Curr. Opin. Struct. Biol. 6, 76–83.

Robinson, W.E. Jr., Reinecke, M.G., Abdel-Malek, S., Chow, S.A., 1996. Inhibitors of HIV-1 replication that inhibit HIV integrase. Proc. Natl. Acad. Sci. USA 93, 6326–6331.

Roe, T., Reynolds, T.C., Yu, G., Brown, P.O., 1993. Integration of murine leukemia virus DNA depends on mitosis. EMBO J. 12, 2099–2108.

Roe, T., Chow, S.A., Brown, P.O., 1997. 3′-End processing and kinetics of 5′-end joining during retroviral integration in vivo. J. Virol. 71, 1334–1340.

Roth, M.J., Schwartzenberg, P.L., Goff, S.P., 1989. Structure of the termini of DNA intermediates in the integration of retroviral DNA: dependence of IN function and terminal DNA sequence. Cell 58, 47–54.

Roth, M.J., Schwartzenberg, P.L., Tanese, N., Goff, S.P., 1990. Analysis of mutations in the integration reactions of Maloney murine leukemia virus: effect on DNA binding and cutting. J. Virol. 64, 4709–4717.

Scottoline, B.P., Chow, S.A., Ellison, V., Brown, P.O., 1997. Disruption of the terminal base pairs of retroviral DNA during integration. Genes Dev. 11, 371–382.

Scourgen, F., Maroun, R.G., Frére, V., Bouziane, M., Anclair, C., Troalen, F., Fermandjian, S., 1996. A synthetic peptide from the human immunodeficiency virus type 1 integrase exhibits coiled-coil properties and interferes with the in vitro integration activity of the enzyme. Correlated biochemical and spectroscopic results. Eur. J. Biochem. 240, 765–773.

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

Shibagaki, Y., Holmes, M.L., Appa, R.S., Chow, S.A., 1997. Characterization of feline immunodeficiency virus integrase and analysis of functional domains. Virology 230, 1–10.

Shih, C.-C., Stoye, J.P., Coffin, J.M., 1988. Highly preferred targets for retrovirus integration. Cell 53, 531–537.

Skalka, A.M., 1988. Integrative recombination of retroviral DNA. In: Kucherlapati, R., Smith, R. (Eds.), Genetic recombination. American Society for Microbiology, Washington, DC, pp. 701–724.

Steitz, T.A., Smerdon, S.J., Jager, J., Joyce, C.M., 1994. A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. Science 266, 2022–2025.

Stevenson, M., 1996. Portals of entry: uncovering HIV nuclear transport pathways. Trends Cell Biol. 6, 9–15.

van den Ent, F.M., Vink, C., Plasterk, R.H., 1994. DNA substrate requirements for different activities of the human immunodeficiency virus type 1 integrase protein. J. Virol. 68, 7825–7832.

van Gent, D.C., Vink, C., Oude Groeneger, A.A.M., Plasterk, R.H.A., 1993. Complementation between HIV integrase proteins mutated in different domains. EMBO J. 12, 3261–3268.

Varmus, H.E., Brown, P.O., 1989. Retroviruses. In: Berg, D.E., Howe, M.M. (Eds.), Mobile DNA. American Society for Microbiology, Washington, DC, pp. 53–108.

Vincent, K.A., Ellison, V., Chow, S.A., Brown, P.O., 1993. Characterization of human immunodeficiency virus type 1 integrase expressed in *Escherichia coli* and analysis of variants with amino-terminal mutations. J. Virol. 67, 425–437.

Vink, C., van Gent, D.C., Elgersma, Y., Plasterk, R.H.A., 1991. Human immunodeficiency virus type 1 integrase protein requires a subterminal position of its viral DNA recognition sequence for efficient cleavage. J. Virol. 65, 4636–4644.

Von Schwedler, U., Kombluth, R.S., Trono, D., 1994. The nuclear localization signal of the matrix protein of human immunodeficiency virus type 1 allows the establishment of infection in macrophages and quiescient T lymphocytes. Proc. Natl. Acad. Sci USA 91, 6992–6996.

Vora, A.C., McCord, M., Fitzgerald, M.L., Inman, R.B., Grandgenett, D.P., 1994. Efficient concerted integration of retrovirus-like DNA in vitro by avian retrovirus integrase. Nucleic Acids Res. 22, 4454–4461.

Wolfe, A.L., Felock, P.J., Hastings, J.C., Blau, C.U., Hazuda, D.J., 1996. The role of manganese in promoting multimer-

ization and assembly of human immunodeficiency virus type 1 integrase as a catalytically active complex on immobilized long terminal repeat substrates. J. Virol. 70, 1424–1432.

Yang, W., Steitz, T.A., 1995. Recombining the structures of HIV integrase, Ruv C and RNase H. Structure 3, 88–95.

Zhao, H., Neamati, N., Mazumder, A., Sunder, S., Pommier, Y., Burke, T.R., 1997. Arylamide inhibitors of HIV-1 integrase. J. Med. Chem. 40, 1186–1194.

Zheng, R., Jenkins, T.M., Craigie, R., 1996. Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization and enhances activity. Proc. Natl. Acad. Sci. USA 93, 13644–13659.