Mol. Cells, Vol. 10, No. 1, pp. 96–101

Molecules and Cells

Minimal Core Domain of HIV-1 Integrase for Biological Activity

Do-Jin Kim, Sang Kwang Lee, You-Take Oh, and Cha-Gyun Shin*
Department of Biotechnology, Chung-Ang University, Kyunggido 456-756, Korea.

(Received on November 12, 1999)

The human immunodeficiency virus type-1 (HIV-1) integrase (IN) mediates insertion of viral DNA into human DNA, which is an essential step in the viral life cycle. In order to study minimal core domain in HIV-1 IN protein, we constructed nine deletion mutants by using PCR amplification. The constructs were expressed in *Escherichia coli*, and the proteins were subsequently purified and analyzed in terms of biological activity such as enzymatic and DNA-binding activities. The mutant INs with an N-terminal or C-terminal deletion showed strong disintegration activity though they failed to show endonucleolytic and strand transfer activities, indicating that the disintegration reaction does not require the fine structure of the HIV-1 IN protein. In the DNA-binding analysis using gel mobility shift assay and UV cross-linking method, it was found that both the central and C-terminal domains are essential for proper DNA-IN protein interaction although the central or C-terminal domain alone was able to be in close contact with DNA substrate. Therefore, our results suggest that the C-terminal domain act as a DNA-holding motive, which leads to proper interaction for enzymatic reaction between the IN protein and DNA.

Keywords: Deletion; Domain; Human Immunodeficiency Virus; Integrase.

Introduction

A critical step in the retroviral lifecycle is the integration of a double-stranded DNA copy of the viral RNA into the host genome. Many studies have shown that the integration is mediated by a specific viral enzyme, called integrase (IN) (Colicelli and Goff, 1988; Panganiban and Temin, 1983). Retroviral INs perform a sequence of DNA cleavage and joining reactions during the viral integration *in vivo*. In the first step, two nucleotides are cleaved on the 3' end of the linear viral DNA and a staggered cut is made in the host DNA by retroviral INs. Then the viral DNA with the recessive 3′ end is covalently joined to the 5′ protrusion of the host DNA. The integration procedures are completed by degradation of the two unpaired nucleotides at the 5′ end of the viral DNA, and by repair of the single strand gap between viral DNA and target DNA by cellular repair enzymes (Muesing *et al*., 1985).

The retroviral INs purified from the overexpressed bacteria have been successfully shown to have four distinctive, biochemical activities such as the endonucleolytic, integration, disintegration and DNA-binding activities *in vitro* (Brown *et al*., 1989; Bushman and Craigie, 1991; Chow *et al*., 1992; Katz, *et al*., 1990). The endonucleolytic activity catalyzes *in vitro* cleavage of two nucleotides on the 3′ ends of the short oligonucleotide DNA that mimics the ends of the viral DNA. The disintegration activity induces cleavage of the branched strand of the Y-shaped oligonucleotides *in vitro* (Chow *et al*., 1992). The integration activity mediates *in vitro* insertion of the pre-processed oligonucleotide substrate into the target DNA. The DNA binding activity induces binding of IN to oligonucleotides or DNA specificially or nonspecifically.

Retroviral IN proteins have three distinctive structural regions of N-teminal, central, and C-terminal domains. The N-terminal domain is known to constitute a zinc finger-like domain containing two pairs of histidine and cysteine residues (Donehower, 1988; Kulkosky *et al*., 1992). The N-terminal domain mediates dimerization of the enzyme that is essential for enzymatic function. Deletion mutants of the domain have showed loss of endonucleolytic and disintegration activities (Josson *et al*., 1996; Puras-Lutzke *et al*., 1994). The central domain is highly conserved and contains a single active site responsible for the enzymatic

* To whom correspondence should be addressed.
Tel: 82-334-670-3067; Fax: 82-334-675-0409
E-mail: cgshin@cau.ac.kr

Abbreviations: HIV-1, human immunodeficiency virus type-1; IN, integrase.

activities. This domain is characterized by the presence of three invariant amino acids essential for catalysis, two aspartic acids and one glutamic acid at the positions 64, 116, and 152, respectively (Engelman and Craigie, 1992; van Gent *et al*., 1992). The C-terminal region of the protein is the least conserved region among retroviral INs, and has been known to be the site of non-specific DNA-binding domain (Vink *et al*., 1993; Woerner and Marcus-Sekura, 1993). Although the N- and C-terminal regions of HIV-1 IN are known to be important for enzymatic activities, it is not yet known how each domain of HIV-1 IN is critically involved in the DNA binding reaction which is an essential step for enzymatic reactions. In this study, we investigate the enzymatic activities and DNA binding properties of the deletion mutants of HIV-1 IN, and characterize the minimal core region of HIV-1 IN protein associated with biological activities.

Materials and Methods

Materials NTA-resin and pQE9 vector were obtained from Qiagen. Radionucleotides of [γ-^{32}P]ATP and [α-^{32}P]ATP were purchased from Amersham at a specific activity of 3,000 Ci/mmol. Oligonucleotides used for preparation of the IN substrates and PCR primers were obtained from Bioneer Inc. Restriction enzymes, T4 DNA polynucleotide kinase, and T4 DNA ligase were obtained from New England Biolabs. Sequenase version 2.0 DNA polymerase was obtained from the United States Biochemical Corp. Other chemical agents were of molecular biology reagent grade from Sigma.

Plasmid construction Plasmid encoding the wild type IN have been described previously (Oh and Shin, 1996; Oh *et al*., 1997). Nine plasmids encoding deletion mutants (Δ1–Δ9) were each constructed by amplifying the IN coding region by PCR, digesting the amplifying DNA with *Bam*HI and *Hin*dIII, and ligating the cut DNA with *Bam*HI-*Hin*dIII-digested pQE9 vector DNA. The oligonucleotides used to constuct the mutant INs are shown in Table 1. All PCRs were performed with *pfu* DNA polymerase (Stratagene) to minimize error in DNA synthesis.

Protein production and purification The plasmid constructs were introduced into bacteria (*Escherichia coli*, XL1-blue) via electroporation (BTX, T820). HIV-1 IN proteins were purified from bacteria as described previously with the following modifications (Oh and Shin, 1996; Oh *et al*., 1997). The cell pellets obtained from the overexpressed culture were resuspended in buffer A (20 mM Tris-HCl, pH 8.0, 0.5 M NaCl, 2 mM β-mercaptoethanol (β-ME), 0.1 mM EDTA, 5 mM imidazole, 6 M guanidine-HCl). The suspension was centrifuged at 40,000 $\times$ *g* for 20 min at 4°C, and the supernatant was collected and loaded over a Ni-NTA column (Qiagen). The protein was eluted with 100 mM and 600 mM imidazoles in buffer A, respectively. The fractions were diluted with an equal volume of buffer containing 1 M NaCl, 20 mM Tris-HCl (pH 7.4), 2 mM β-ME and 2 mM EDTA. The protein was sequentially dialyzed for 1 h at 4°C against the three different solutions which are buffer I (20 mM Tris-HCl, pH 7.4, 2 M urea, 0.5 M NaCl, 5 mM EDTA, 10 mM β-ME, 10 mM CHAPS), buffer II (20 mM Tris-HCl, pH 7.4, 1 M NaCl, 1 mM EDTA, 1 mM DTT, 10 mM CHAPS, 10% glycerol), and buffer III (20 mM Tris-HCl, pH 7.4, 0.5 M NaCl, 1 mM EDTA, 1 mM DTT, 10 mM CHAPS, 10% glycerol). The dialysate was centrifuged at 19,000 $\times$ *g* for 10 min and the supernatant containing IN was saved in liquid nitrogen (Bushman *et al*., 1993; Taddeo *et al*., 1996).

***In vitro* enzymatic activities of HIV-1 IN** For *in vitro* assay of the disintegration activity, the disintegration substrate (Y-oligomer) was prepared by labeling the oligonucleotide T1 and annealed with two-fold amounts of the oligonucleotides T2, T3 and 6917. Disintegration reaction contained 0.1 pmol of oligonucleotide substrate and 20 pmol of IN protein in 15 mM Tris-HCl (pH 7.4), 10 mM Mn^{2+}, 100 mM NaCl, 2 mM β-ME, 2.5 mM CHAPS, 0.1 mM EDTA, 0.1 mM PMSF, 1% glycerol, and 10 mM imidazole in a total reaction volume of 10 μl. Reactions were incubated 33°C for 90 min and stopped by addition of 4 μl of 95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF. The reactions were heated to 90°C for 3 min and electrophoresed on a 20% denaturing polyacrylamide gel. The reaction products were visualized by autoradiography of the wet gel (Chow *et al*., 1992).

DNA binding activities DNA binding activities of the deletion mutants were investigated by two different methods — mobility shift assay and the UV-crosslinking method. DNA substrates for the binding reactions were constructed by using two oligonucleotides whose sequences resemble the end of U5-LTR of HIV-1 proviral DNA. Briefly, oligonucleotide 6916 (U5-LTR, +strand) was labeled at the 5′ end with ^{32}P-ATP and T4 polynucleotide kinase, and annealed with complementary oligonucleotide 6917 (U5-LTR, -strands).

For gel mobility shift assay, a standard reaction contained 0.1 pmol of duplex oligonucleotide substrate and 30 pmol of HIV-1 IN proteins in 25 mM Tris-HCl [pH 7.4], 5 mM NaCl, 0.5% Tween 20, 1% glycerol and 0.1 mM β-ME in a final volume of 10 μl. Reactions were incubated for 20 min at 30°C, and were stopped by addition of stop solution (0.25% bromophenol blue, 0.25% xylene FF and 30% glycerol). The samples were separated in 8% polyacrylamide gels in 0.5$\times$ TBE at room temperature.

For UV cross-linking analysis, reactions were performed under the same conditions as those when the gel mobility shift assay was done. Samples were pre-incubated for 15 min on ice prior to UV exposure and then irradiated on a chromato-vue transilluminator at 254 nm on ice (Engelman *et al*., 1994). Reactions were stopped by adding protein loading dye, and the samples were boiled for 5 min and subsequently loaded on sodium dodecyl sulfate (SDS)-16.5% polyacrylamide gel. UV-induced protein-DNA cross-linking products were visualized by autoradiography.

Results and Discussion

Expression and purification of HIV-1 in mutants HIV-1 IN, composed of 288 amino acids, is synthesized as a part of *Gag-Pol* precursor polyprotein, and activated by the proteolytic cleavage *in vivo* (Donehower and Varmus, 1984). The N-terminal domain containing about 50 amino

acids is known to be a dimeric region. The central domian that is composed of about 160 amino acids has an active site for enzymatic activities. The three amino acids (64 and 116 aspartic acids, 152 glutamic acid) in the domain are known to be catalytic residues for the activities. The C-terminal domain is composed of about 80 amino acids (Engelman and Craigie, 1992).

In order to construct the deletion mutants, the HIV-1 IN genes were amplified from the proviral DNA as shown in Fig. 1. The amplified IN genes were cloned into a prokaryotic expression vector, pQE9, which has an ATG codon for translation initiation and a hexahistidine tag sequence as vector-derived residues at the N-terminus of the IN gene. The hexahistidine tag facilitated purification of the IN proteins on the nickel-chelated NTA affinity chromatography. The IN proteins that contained 12 extra amino acids at N-terminus in the form of a fusion protein were eluted from the NTA column with 100 mM imidazole. It has been shown that presence of the 12 extra amino acids at N-terminus in HIV-1 IN did not interfere with enzymatic activities of HIV-1 IN (Oh and Shin, 1996). Judging from SDS-PAGE analysis, the IN proteins were purified to homogeneity of 95% or more (Fig. 2).

Enzymatic activities of the deletion INs Purified retroviral INs have been shown to have three enzymatic activities of endonucleolytic, strand transfer, and disintegration activities *in vitro* (Chow *et al.*, 1992; Engelman *et al.*, 1991; Sherman and Fyfe, 1990). The principles of the *in vitro* assays are based on the fact that retroviral INs make specific cleavage or joining on the

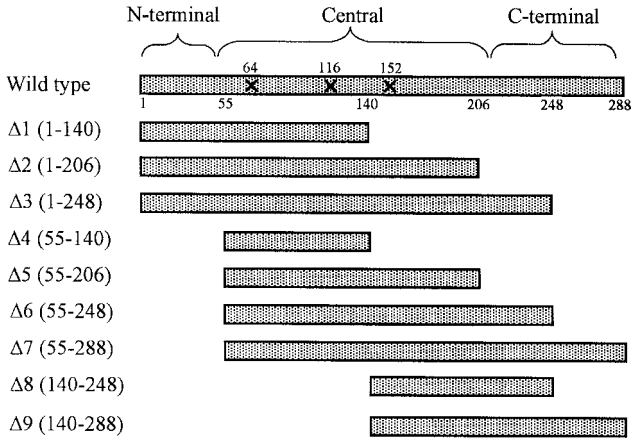

Fig. 1. Schematic presentation of the HIV-1 IN deletion mutants. The wild type protein with 288 amino acids is shown depicting the three functional domains. The amino acids known as critical residues for enzymatic functions are indicated at the positions of 64, 116 and 152. Nine deletion mutants were constructed by PCR amplification of the corresponding HIV-1 IN DNA regions.

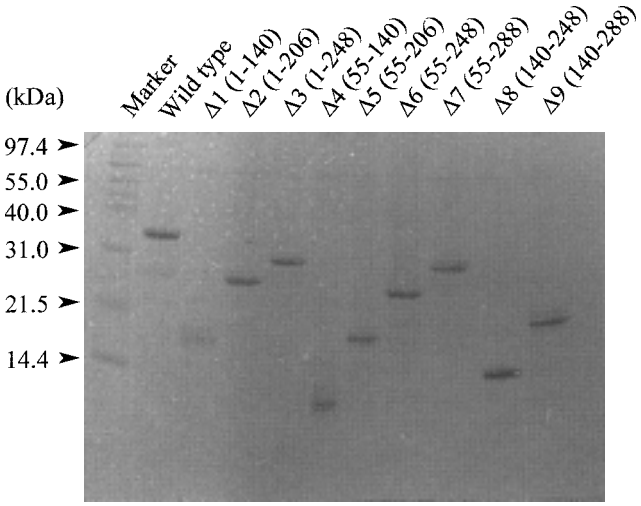

Fig. 2. SDS-PAGE analysis of the HIV-1 IN mutant proteins purified through NTA-chromatography. The purified proteins were separated by a 12.5% SDS-PAGE and visualized by Coomassie blue stain.

oligonucleotide substrates matching the 3′ ends of the U5 or U3 viral long terminal repeats (LTR) DNA sequences (Sherman and Fyfe, 1990). Several studies have shown that the endonucleolytic and strand transfer reactions require intact IN proteins and specific DNA substrates in their sequences and structure (Chow and Brown, 1994; Oh and Shin, 1999). However, the disintegration reaction is able to occur with less specificity. In our studies the nine deletion mutant INs did not show any endonucleolytic and strand transfer activities, reflecting that the reactions require intact INs (data not shown). But, strong disintegration activities were found in four deletion mutants; Δ2, Δ3, Δ6, and Δ7. Interestingly, the disintegration activities of Δ2 and Δ6 are far stronger than that of the wild type IN as their disintegration products are 400% and 340% of the wild type products, respectively (Fig. 3B). This result indicates that the intact structure of HIV-1 IN is not necessary to perform disintegration activity. The IN structure containing central domain and N-terminal or C-terminal domain seems to be better in performing disintegration reaction than the wild type IN.

DNA binding activities Several retroviral INs have been shown to bind DNA nonspecifically (Grandgenett *et al.*, 1978; Roth *et al.*, 1988; van Gent *et al.*, 1991). The DNA binding activity of HIV-1 IN has been mapped to the C-terminal region of the protein (Engelman *et al.*, 1994; Puras-Lutzke *et al.*, 1994; Vink *et al.*, 1993). In order to know the role of each domain of HIV-1 IN protein in DNA binding, the deletion mutants were tested for DNA-binding ability by two different methods (DNA mobility shift and protein-DNA UV cross-linking).

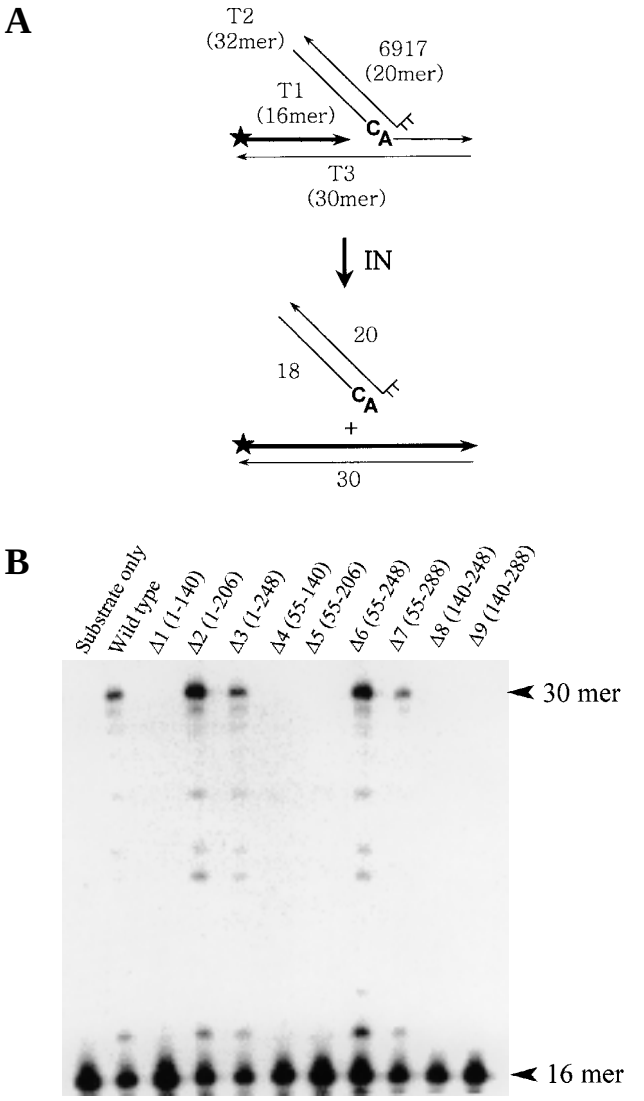

Fig. 3. Disintegration activities of the HIV-1 IN deletion mutants. **A.** Schematic illustration of *in vitro* disintegration reaction. **B.** Disintegration activities of IN proteins. To construct the DNA substrate, a 16mer-oligonucleotide (T1) was labeled at its 5′-end (★), and annealed with the three complementary oligonucleotides (T2, T3, and 6917). The reactions were carried out in 15 mM Tris·HCL (pH 7.4) and 10 μM $MnCl_2$ in the presence of 0.1 mM DNA substrate and 20 μM IN protein at 33°C for 90 min. The reaction products were analyzed in a 20% polyacrylamide gel and visualized by a phosphoimage analyzer (GS 525, BioRad).

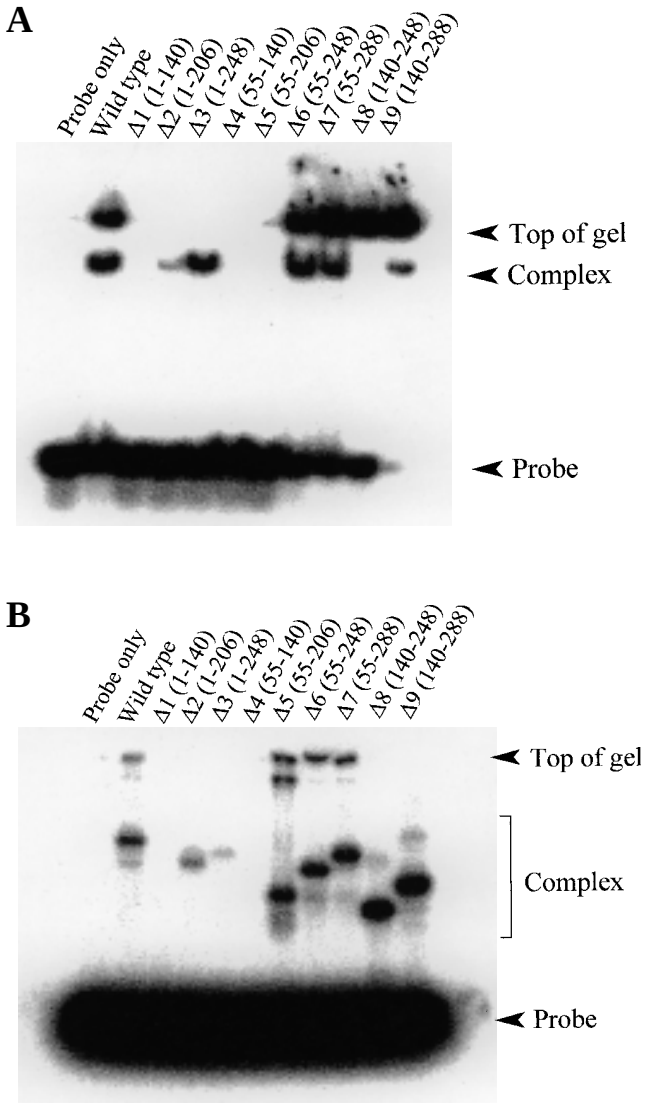

Fig. 4. DNA binding analysis of the HIV-1 IN deletion mutants. **A.** Mobility shift assay. A radiolabeled, double stranded oligonucleotide DNA, whose sequence resembles the end of U5-LTR of HIV-1 proviral DNA, was incubated with IN proteins in a non-denaturing condition and analyzed in 8% polyacrylamide gel. **B.** UV cross-linking analysis. A radiolabeled oligonucleotide DNA, as described above, was preincubated with IN proteins and then irradiated with UV-light. The reaction products were analyzed in a SDS-16.5% polyacrylamide gel.

In the gel mobility shift assay a radiolabeled oligonucleotide was incubated with IN protein for 20 min at 30°C and subsequently analyzed in non-denaturing gel electrophoresis (Fig. 4A). The IN proteins (Δ6, Δ7) containing central domain and C-terminal region showed strong DNA-binding. But, the IN proteins containing the central domain alone (Δ5) or deleted central domain with partial C-terminal region (Δ8) could not bind the DNA substrate, implying that structural integrity in the core region of IN protein is required for proper DNA-protein interaction.

In addition, in the UV cross-linking analysis, the IN proteins containing the central domain alone or central domain and C-terminal region were able to be cross-linked to the DNA substrate by UV irradiation. This result suggests that the amino acid residues present in the central or C-terminal region are in close contact with the DNA substrate (Fig. 4B; Δ5–Δ9). On the contrary, it is clear that

the N-terminal domain hardly contributes to DNA-binding of IN protein (Figs. 4A and 4B; Δ1–Δ3). It is likely that the N-terminal domain blocks UV-induced cross-linking of reactive residues present at the central or C-terminal domain as it exists in the external region of the tertiary structure and blocks UV-transmission. In particular, Δ5 and Δ8 showed UV-induced DNA cross-linking but failed to bind the DNA substrate in mobility shift analysis. It proposes that the C-terminal domain acts as a DNA-holding motive and its deletion leads to less fitting DNA-protein interaction even though there is still close contact between reactive residues and DNA substrate which can be detected in UV cross-linking.

HIV-1 IN has three functional domains. The three-dimensional structures for each domain have been reported using X-ray crystallography or nuclear magnetic resonance (Cai *et al*., 1997; Dyda *et al*., 1994; Lodi *et al.*, 1995). However, it is still difficult to understand the exact interaction between IN protein and DNA substrate as the three-dimensional structure for the intact protein has not been established yet. Therefore, biochemical studies, such as those reported here, may help to understand the relation and cooperation among the domains in performing biological activity, and also to suggest basic data to develop inhibitors against IN, targeted to functional domains.

Acknowledgments We are grateful to Se-Young Choi and Wan-Soo Kim for the excellent technical assistance in using a phosphoimage analyzer. This research was supported by the Chung-Ang University special research grant in 1997. In addition, this work was supported by a matching fund for research instrument (BTX, T820) of Chung-Ang University in 1998.

References

Brown, P. O., Bowerman, B., Varmus, H. E., and Bishop, J. M. (1989) Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. *Proc. Natl. Acad. Sci. USA* **86**, 2522–2529.

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

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

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

Chow, S. A. and Brown, P. O. (1994) Substrate features important for recognition and catalysis by human immunodeficiency virus type 1 integrase identified by using novel DNA substrates. *J. Virol*. **68**, 7869–7878.

Chow, S. A., Vincent, K. A., Ellison, V., and Brown, P. O. (1992) Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science* **255**, 723–726.

Colicelli, J. and Goff, S. P. (1988) Sequence and spacing requirements of a retrovirus integration sites. *J. Mol. Biol.* **199**, 47–59.

Donehower, L. (1988) Analysis of mutant Moloney leukemia viruses containing linker insertion mutations in the 3′ region of *pol*. *J. Virol.* **62**, 3958–3964.

Donehower, L. and Varmus, H. (1984) A mutant murine leukemia virus with a single missense codon in *pol* is defective in a function affecting integration. *Proc. Natl. Acad. Sci. USA* **81**, 6461–6465.

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

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

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

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

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

Josson, C. B., Donzella, G. A., Gaucan, E., Smith, C. M., and Roth, M. J. (1996) Functional domains of Moloney murine leukemia virus integrase defined by mutation and complementation analysis. *J. Virol*. **70**, 4585–4597.

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

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

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

Muesing, M. A., Smith, D. H., Cabradilla, C. D., Benton, C. V., Lasky, L. A., and Capton, D. J. (1985) Nucleic acid structure and expression of human AIDS/lymphadenopathy retrovirus. *Nature* **313**, 450–458.

Oh, J.-W. and Shin, C.-G. (1996) Purification and characterization of the human immunodeficiency virus type-1 integrase expressed in *Escherichia coli*. *Mol. Cells* **6**, 96-100.

Oh, J.-W., Oh, Y.-T., Kim, D.-J., and Shin, C.-G. (1997) Characterization of human immunodeficiency virus type-1 integrase mutants expressed in *Escherichia coli*. *Mol. Cells* **7**, 688–693.

Oh, Y.-T. and Shin, C.-G. (1999) Comparison of enzymatic activities of the HIV-1 and HFV integrases to their U5 LTR substrates. *Biochem. Mol. Bio. Int*. **47**, 621–629.

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

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

Roth, M. J., Tanese, N., and Goff, S. P. (1988) Gene product of Moloney murine leukimia virus required for proviral integration is a DNA-binding protein. *J. Mol. Biol*. **203**, 131–139.

Sherman, P. A. and 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.

Taddeo, B., Carlini, F., Verani, P., and Engelman, A. (1996) Reversion of a human immunodeficiency virus type 1 integrase mutant at a second site restores enzyme function and virus infectivity. *J. Virol.* **70**, 8277–8284.

van Gent, D. C., Elgersma, Y., Bolk, M. W., Vink, C., and Plasterk, R. H. A. (1991) DNA binding of the integrase proteins of human immunodeficiency viruses types 1 and 2. *Nucleic Acids Res*. **19**, 3821–3827.

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

Vink, C., Groenger, A. A. M. O., and Plastesk, R. H. A. (1993) Identification of the catalytic and DNA-binding region of the human immunodeficiency virus type 1 integrase protein. *Nucleic Acids Res*. **21**, 1414–1425.

Woerner, A. M. and Marcus-Sekura, C. J. (1993) Characterization of a DNA binding domain in the C-terminus of HIV-1 integrase by deletion mutagenesis. *Nucleic Acids Res*. **21**, 3507–3511.