Central Core Domain of Retroviral Integrase Is Responsible for Target Site Selection*

(Received for publication, October 25, 1996, and in revised form, January 13, 1997)

Yoshio Shibagaki‡ and Samson A. Chow§

From the Department of Molecular and Medical Pharmacology and Molecular Biology Institute, UCLA School of Medicine, Los Angeles, California 90095

Integration of retroviral DNA can occur into many sites on target DNA with a wide variation in preference. One factor known to affect target site selection is integrase, the viral protein required for the integration reaction. In this study, assays that measure the distribution and frequency of retroviral DNA integration showed that purified integrases of human immunodeficiency virus type 1 (HIV-1) and feline immunodeficiency virus (FIV) had different patterns of target site usage. The integrase domain involved in target site selection was mapped by analyzing the integration pattern of chimeric proteins formed between HIV-1 and FIV integrases and of deletion variants of the two wild-type integrases. The results indicate that the domain responsible for target site selection resides in the central core region of integrase.

Integration of a double-stranded DNA copy of the viral RNA genome into a chromosome of the host cell can be divided into three steps: 3'-end processing, 3'-end joining, and 5'-end joining. Purified integrase alone can catalyze the first two steps of the reaction (1–7). 3'-End processing involves the removal of two nucleotides from the 3'-end of each strand of linear viral DNA so that the viral 3'-ends termate with the conserved CA dinucleotide. 3'-End joining or strand transfer is a concerted cleavage-ligation reaction during which integrase makes a staggered cut in the target DNA and ligates the recessed 3'-ends of the viral DNA to the 5'-ends of the target DNA at the cleavage site. Integrase also catalyzes a reversal of the 3'-end joining reaction, termed disintegration, in which a substrate mimicking one end of viral DNA joined to target DNA is resolved to its viral and target DNA parts (8).

Mutational analyses of several retroviral integrases have identified at least three distinct functional domains (9–11): (i) an HHCC domain in the N terminus, named for its zinc finger-like motif HX_{3–7}HX_{23–32}CX_{2}C, (ii) a central domain located within a protease-resistant core containing the DD35E motif (DX_{80–85}DX_{83–90}E) that is highly conserved in retroviruses and retrovirus-like elements (12), and (iii) a C terminus domain that represents the least conserved region among retroviral integrases. The HHCC domain is critical for forming a stable complex with viral DNA, and the domain may be involved in multimerization of integrase (13, 14). The central core domain is important for catalysis (9, 15–17) and dimerization (18, 19). The C-terminal domain is capable of binding DNA nonspecifically and is believed to contain the site for binding target DNA (20–25).

One characteristic feature of retroviral DNA integration is that many sites on target DNA can be used for integration. However, the frequency and distribution of integration are not random, and there are hot spots and cold spots of integration on target DNA (Refs. 26 and 27 and references therein). The mechanism for target site selection is not well understood. The transcriptional status of DNA, methylation, association of DNA with histones or other DNA-binding proteins, and DNA bending are known to affect target site specificity (28–32). In vitro studies showed that integrases from different retroviruses, such as human immunodeficiency virus type 1 (HIV-1), murine leukemia virus, feline immunodeficiency virus (FIV), and visna virus, produce different integration patterns (as defined by frequency and distribution), indicating that integrase is a major factor in determining target site preference (30, 33). However, the integrase domain involved in target site selection has not been defined. Similarities in the integration patterns among wild-type and various deletion derivatives of HIV-1 integrase suggest that the domain involved in target site selection resides in the central core region of integrase (34). Analysis of the patterns of nonspecific alcoholysis obtained from chimeras between HIV-1 and visna virus integrases also implicates the core region in selecting target sites (33). It is not clear, however, whether the determinant for nonspecific alcoholysis is equivalent to that for target site selection.

In the present study, we prepared chimeric proteins between HIV-1 and FIV integrase to identify domains responsible for target site selection. FIV is a lentivirus with a complex genomic organization resembling that of primate lentiviruses (35). The amino acid sequence of FIV integrase is 37% identical to that of HIV integrase, and mutational analysis of FIV integrase showed that its enzymatic activities and domain organization are similar to that of HIV-1 integrase (Fig. 1). Since the integrases of HIV-1 and FIV produce different integration patterns and thus may possess different target site determinants, analysis of integration patterns of chimeras between HIV-1 and FIV integrases represents a useful strategy for mapping integrase domains involved in target site selection. Comparison of the patterns of 3'-end joining of the various chimeras indicates that the central core domains of both HIV-1 and FIV integrases contain the determinants for target site selection.

* This work was supported by National Institutes of Health Grant R01 CA68859. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Recipient of the UCLA-DOE Innovation Postdoctoral Fellowship.

§ To whom correspondence should be addressed. Tel.: 310-825-9600; Fax: 310-825-6267; E-mail: schow@pharm.medsch.ucla.edu.

1 The abbreviations used are: HIV-1, human immunodeficiency virus type 1; CHAPS, 3-(3-cholamidopropyl)-dimethylammonio)-1-propanesulfonic acid; DTT, dithiothreitol; FIV, feline immunodeficiency virus; PCR, polymerase chain reaction.

2 Y. Shibagaki, M. L. Holmes, R. S. Appa, and S. A. Chow (1997) Virology, in press.
The C terminus, although it binds DNA nonspecifically, is not involved in target site selection.

**EXPERIMENTAL PROCEDURES**

**Enzymes and Reagents—**Wild-type and deletion mutants of HIV-1 and FIV integrases were purified as described previously (34). A rabbit antipeptide antiserum to FIV integrase was obtained from Dr. John H. Elder at the Scripps Research Institute. T4 polynucleotide kinase, T4 DNA ligase, and restriction endonucleases were purchased from New England Biolabs; AmpliTaq DNA polymerase was from Perkin-Elmer; modified T7 DNA polymerase (Sequenase version 2.0) and exonuclease-free Klenow fragment of *E. coli* DNA polymerase I were from U.S. Biochemical Corp. Deoxyribonucleotides were purchased from Pharmacia Biotech Inc. [γ-32P]ATP and [α-32P]dTTP were obtained from Amersham at a specific activity of 6,000 C/mmol. Oligonucleotides were purchased from Operon Technologies, Inc., and were purified by electrophoresis through a 15% denaturing polyacrylamide gel before use.

**Cloning and Expression of HIV/FIV Chimeric Integrases—**HIV/FIV chimeric integrases were prepared by constructing plasmids encoding the chimeric integrase genes. The chimeras H/H/F and F/F/H, which contained a swap of the C terminus, were prepared by exchanging the Aval site (underlined) and the AvaII site (boldface) in the H-U5V2 fragment between pT7-7/H-INF, which encodes HIV-1 integrase, and pT7-7/F-INF, which encodes FIV integrase (34). The chimeras containing a swap at the N terminus, H/F/F and H/F/H, were prepared by performing two sequential polymerase chain reactions (PCRs). For H/F/F, in the first PCR, primer pT7-7/H-INF was used as the template and the F3 primer that contains an NdeI site (underlined) and anneals to the N terminus of HIV-1 integrase. The 3’-primer was H38F (5’-TTCAATTGCTCCACTGTTCTCTCTATTAGTGAAG-3’) and it contains DNA sequences derived from HIV-1 (roman type) and FIV (italic type) integrase. The PCR product was purified by a QIAquick spin column (Qiagen) and was used as the 5’-primer in a second PCR, which included pT7-7/F-INF as the template and F3’ (5’-GGTCGGATCCATGTGATTCTGCATA-3’) and H38F as the 3’-primer. The DNA sequence from the middle of the unique NdeI site of pT7-7 (34). The resulting plasmids were cleaved with NdeI, and a double-stranded oligonucleotide (5’TATGGTTCTCCGGCTGCAG-3’) was single-stranded. Purification of Chimeric Integrase—The amino acid residues of human integrase were determined by Edman degradation (39). The 3’-end joining activity was measured using a modified assay that uses separate oligonucleotides as the donor and target substrates (38, 40). The donor substrate, prepared by annealing the H- and F-U5V2 strands with their complementary oligonucleotides H- and F-U5V2, respectively. The substrate was singly labeled at the 3’-end of the H- and F-U5V2-2 strands using exonuclease-free Klenow fragment of *E. coli* DNA polymerase I, dGTP and [α-32P]dTTP (39).

**Assays for Integrate Activity—**The 3’-end processing, 3’-end joining, and disintegration activities of the chimeric proteases were assayed as described previously (37, 38). The following oligonucleotides (Oligo Technologies, Inc.) were used as DNA substrates (boldface letters denote the invariant CAVTG dinucleotide pair): H-U5V1-2, 5’-ATGGTGAAAATCTCTAGACCA; H-U5V2, 5’-ACTGCTAGAAGATTTCCCTACAT; H-U5V1L-2, 5’-CCGGCGCCAGAAGATTTCCCTACAT; H-U5V2L, 5’-ACTGCTAGAAGATTTCCCTACAT. The DNA sequences of all of the PCR-amplified DNA fragments were determined by the dideoxynucleotide chain termination method.

**The DNA constructs were transformed into E. coli BL21 (DE3).** The cells were grown at 35 °C in 4 liters of LB medium containing 50 μg of ampicillin/ml. An optical density of 0.8, isoprropyl-1-thio-β-D-galacto-pyranoside was added to 0.3 ml for induction expression, and the culture was grown for an additional 5 h. After harvesting, the cell pellet was frozen at −80 °C.

**Purification of Chimeric Integrase—**The frozen bacterial pellet was resuspended in 100 ml of a buffer containing 20 mM HEPS, pH 7.5, 0.2 mM EDTA, 1 mM NaCl, 10% glycerol, 5 mM 2-mercaptoethanol, 1 mM phenylmethylsulfonfyl fluoride, and 0.05% Nonidet P-40. The cell suspension was sonicated and centrifuged at 100,000 × g for 1 h at 4 °C. The supernatant, after dialysis against buffer A (20 mMHEPS, pH 7.5, 1 mM NaCl, 10% glycerol, 5 mM 2-mercaptoethanol, 0.1% Nonidet P-40), was incubated on ice for 2 h with 2.5 ml of Ni2+–nitrilotriacetic acid-agarose resin (Qiagen). The wash was passed five times with 20 ml of buffer A and then packed on a column. The protein was eluted in a total volume of 20 or 30 ml with a linear gradient from buffer A plus 50 mM imidazole to buffer A plus 500 mM imidazole. The fractions containing the protein were pooled and dialyzed against buffer C (20 mM HEPS, pH 7.5, 0.1 mM EDTA, 0.5 mM NaCl, 20% glycerol, 1 mM dithiothreitol (DTT), and 10 mM CHAPS).

**The His tag (underlined) was significant for a protein.** The substrate for nonspecific alcoholysis, prepared by annealing T5 and T6 and was singly labeled at the 3’-end of the H- and F-U5V2-2 strands using exonuclease-free Klenow fragment of *E. coli* DNA polymerase I, dGTP and [α-32P]dTTP (39). The 3’-end joining activity was measured using a modified assay that uses separate oligonucleotides as the donor and target substrates (38, 40). The donor substrate, prepared by annealing the H- and F-U5V2 strands with their complementary oligonucleotides H- and F-U5V2, respectively. The substrate was singly labeled at the 3’-end of the H- and F-U5V2-2 strands using exonuclease-free Klenow fragment of *E. coli* DNA polymerase I, dGTP and [α-32P]dTTP (39).

**The HIV-1 disintegration substrate (Y-oligomer) was prepared by annealing the labeled T1 strand with oligonucleotides T3, H-V1/T2, and F-V1/T2, then packed on a column.** The 3’-end joining activity was measured using a modified assay that uses separate oligonucleotides as the donor and target substrates (38, 40). The donor substrate, prepared by annealing the H- and F-U5V2-2 strands with their complementary oligonucleotides H- and F-U5V2, respectively. The substrate was singly labeled at the 3’-end of the T5 strand using exonuclease-free Klenow fragment of *E. coli* DNA polymerase I and [α-32P]dTTP (39).

**The chimeric integrases were purified by electrophoresis through a 15% denaturing polyacrylamide gel.** The substrates used to assay the 3’-end processing activity of wild-type and chimeric integrases were double-stranded oligonucleotides containing sequences derived from the U5 end of the HIV-1 (H-U5V1-2/H-U5V2) or FIV (F-U5V1-2/F-U5V2) long terminal repeat terminal repeat. The DNA sequences were prepared by annealing the H- and F-U5V2-2 strands with their complementary oligonucleotides H- and F-U5V2, respectively. The substrate was singly labeled at the 3’-end of the H- and F-U5V1-2 strands using exonuclease-free Klenow fragment of *E. coli* DNA polymerase I, dGTP and [α-32P]dTTP (39).

**The 3’-end joining activity was measured using a modified assay that uses separate oligonucleotides as the donor and target substrates (38, 40).** The donor substrate, prepared by annealing the H- and F-U5V2-2 strands with their complementary oligonucleotides H- and F-U5V2, respectively. The substrate was singly labeled at the T5 strand using exonuclease-free Klenow fragment of *E. coli* DNA polymerase I and [α-32P]dTTP (39).

**The 3’-end joining activity was measured using a modified assay that uses separate oligonucleotides as the donor and target substrates (38, 40).** The donor substrate, prepared by annealing the H- and F-U5V2-2 strands with their complementary oligonucleotides H- and F-U5V2, respectively. The substrate was singly labeled at the T5 strand using exonuclease-free Klenow fragment of *E. coli* DNA polymerase I and [α-32P]dTTP (39).
Integrase Domain Responsible for Target Site Selection

RESULTS

Experimental Plan—Since the patterns of preferred target sites (determined by frequency and distribution of integration events) are different between HIV-1 and FIV integrases (see Figs. 4 and 5), two approaches to identify the integrase domain responsible for target site selection is through analysis of HIV/FIV chimeric integrases. If the selection of target sites by the wild-type integrase is mediated by a discrete domain, then exchanging that domain between a pair of chimeric proteins should produce a corresponding exchange of integration patterns. We chose to form chimeras between HIV-1 and FIV integrases because the two proteins are related by sequence and size, and they share similar reaction conditions for optimal activities in vitro (35, 37, 41).

Based on the amino acid alignment (Fig. 1) and the domain organization of HIV-1 and FIV integrases (9–11), the protein was divided into three domains: N terminus, central core, and C terminus. Peptides derived from HIV-1 integrase (H) are depicted by hatched boxes, and peptides derived from FIV integrase (F) are depicted by open boxes. The numbers in parentheses correspond to the amino acid residues of the indicated wild-type integrase (WT) that are included in each chimeric protein, B, SDS-polyacrylamide gel electrophoresis of HIV/FIV chimeric integrases. Purified wild-type and HIV/FIV chimeric integrases were separated by 13% SDS-polyacrylamide gel electrophoresis, and the gel was stained with Coomassie Blue. Lanes 1 and 10 contain molecular weight standards (Life Technologies, Inc.) with the masses (in kilodaltons) indicated on the right. The nomenclature of the chimeric proteins is as described under ‘Results.’

Fig. 1. Alignment of amino acid sequences of HIV-1 and FIV integrases. Identical amino acids between the two integrases are noted by vertical lines. The highly conserved HHCC and DD(35)E motifs are highlighted by asterisks. Dashes represent gaps introduced by the alignment. The striped, open, and stippled boxes represent the N terminus, central core, and C terminal domains, respectively, that were exchanged between the proteins.

Fig. 2. Structures and SDS-polyacrylamide gel electrophoresis of HIV/FIV chimeric integrases. A, primary structures. The integrases of HIV-1 and FIV are divided into three domains: N terminus, central core, and C terminus. Peptides derived from HIV-1 integrase (H) are depicted by hatched boxes, and peptides derived from FIV integrase (F) are depicted by open boxes. The numbers in parentheses correspond to the amino acid residues of the indicated wild-type integrase (WT) that are included in each chimeric protein, B, SDS-polyacrylamide gel electrophoresis of HIV/FIV chimeric integrases. Purified wild-type and HIV/FIV chimeric integrases were separated by 13% SDS-polyacrylamide gel electrophoresis, and the gel was stained with Coomassie Blue. Lanes 1 and 10 contain molecular weight standards (Life Technologies, Inc.) with the masses (in kilodaltons) indicated on the right. The nomenclature of the chimeric proteins is as described under ‘Results.’
**FIG. 3. In vitro activities of chimeric integrases.**

**A**, 3′-end processing activity. HIV-1 U5 substrate (H-U5V1/H-U5V2; 5 nM) singly labeled at the 3′-end of the processed strand was incubated with the wild-type (WT) or the various chimeric integrases. The open arrowhead indicates the position of the labeled substrate (21-mer). The filled arrowheads correspond to the two major processing products, the dinucleotide (GT*) and the glycerol-dinucleotide adduct (Glycerol-GT*).

**B**, 3′-end joining (strand transfer) activity. Unlabeled, preprocessed HIV-1 U5 DNA (H-U5V1-2/H-U5V2; 15 nM) was preincubated with the wild-type or chimeric integrase for 10 min at room temperature in the standard reaction buffer. The reaction was started by adding 5 nM of the 3′-end-labeled nonviral DNA substrate (T5/T6) and incubating at 37 °C for 60 min. In lanes 3, 4, 5, 8, and 9, the autoradiogram was exposed twice as long as the other lanes so that the weak joining products could be seen. The asterisk corresponds to a contaminant present in the unreacted, labeled substrate and is not a joining product. The open arrowhead indicates the position of the labeled substrate. The 3′-end joining products are marked by IP.

**C**, disintegration activity. The reaction was carried out with 5 nM of the Y-oligomer substrate. The open arrowhead indicates the position of the 5′-end-labeled T1 strand (16-mer) of the Y-oligomer substrate. The filled arrowhead indicates the position of the disintegration product (30-mer). The bands between the substrate and the disintegration product represent products resulting from reintegration of released viral DNA ends into the disintegration product (8, 37), nonspecific alcoholysis of the disintegration product, or both (data not shown). In all three assays, the concentrations of the wild-type or the chimeric integrase used were 200 (H-IN wild type, F-IN wild type, F/F/H), 260 (F/H/F), 300 (H/F/H, H/F/F), 460 (F/H/H), and 580 nM (H/H/F). The concentrations chosen were determined in preliminary experiments and represent the optimal concentration of the particular wild-type or chimeric integrase.
Integrase Domain Responsible for Target Site Selection

TABLE I

| Integrase derivative | 3′-End processing | 3′-End joining | Disintegration |
|----------------------|-------------------|---------------|-------------|
| H-IN wild type       | ++               | ++            | ++         |
| F-IN wild type       | +++              | ++            | ++         |
| F/H/H                | +/−              | +/−           | ++         |
| F/H/F                | −/−              | +/−           | ++         |
| H/H/F                | +/−              | −/−           | ++         |
| F-IN53–281           | +/−              | −/−           | ++         |
| F-IN1–235            | +/−              | +/−           | +/−        |
| F-IN53–235           | −/−              | −/−           | −/−        |

Activities were determined using oligonucleotide-based assays as described under “Experimental Procedures.” The results are expressed as the percentage of the substrate converted to the product under standard reaction conditions. +, 20% of conversion; +/−, less than 2% of conversion; −, no conversion. The 3′-end joining activity of two chimeric proteins, H/H/F and F/H/F, and two deletion mutants, F-IN53–281 and F-IN53–235, although undetectable with the oligonucleotide-based assay, was detected by the PCR-based assay (see Fig. 5).

Pattern of Target Site Usage Was Independent of the Sequence of Donor DNA—Previous studies have shown that selection of target sites by HIV-1 integrase is dependent on the donor DNA and independent of the donor DNA (30, 38, 40).

Whether the same is true with FIV integrase has not been investigated. The earlier finding that HIV-1 and FIV integrases have different preferred target sites also needed to be confirmed. The target site usage of wild-type HIV-1 and FIV integrases was determined using an oligonucleotide-based assay and a PCR-based 3′-end joining assay (Fig. 4).

In the oligonucleotide-based assay (38, 40), separate oligonucleotide and several chimeric proteins that contained the core region of FIV integrase, such as H/H/F, F/F/H, and H/H/F, showed prominent products of about 68 kDa in size (Fig. 2B) that reacted positively with an antipeptide antiserum to FIV integrase by Western blot analysis (data not shown). Based on their sizes and reactivity to anti-integrase antibody, the slowly migrating products were presumably dimers of the respective proteins.

The activities of all six chimeric proteins were analyzed in vitro using oligonucleotide-based assays (Fig. 3), and the results are summarized in Table I. All chimeric proteins could mediate disintegration, indicating that their catalytic sites were functional. The ability of the proteins to carry out 3′-end processing and 3′-end joining was less than that of the wild-type HIV or FIV integrase. No processing or joining activity was detected with H/H/F or F/F/H using the oligonucleotide-based assay. A weak 3′-end joining activity for both chimeric proteins could be detected using the more sensitive PCR-based assay when a preprocessed viral DNA end was the donor DNA (see Fig. 5).

Fig. 4. Differences in distribution of integration sites between HIV-1 and FIV integrases. A, oligonucleotide-based assay. Fifteen nanomolar preprocessed HIV-1 U5 (lanes 1, 3, and 5) or FIV U5 (lanes 2, 4, and 6) was preincubated in the absence of integrase (lanes 1 and 2) or in the presence of 200 nm of HIV-1 integrase (lanes 3 and 4) or FIV integrase (lanes 5 and 6) for 10 min at room temperature in the standard reaction buffer. This was followed by the addition of 5 nm 3′-end labeled target DNA (T5/T6) and incubation at 37 °C for 60 min. The symbols are the same as in Fig. 3B. B, PCR-based assay. Preprocessed HIV-1 U5 (lanes 2, 4, and 6) or FIV U5 (lanes 3, 5, and 7) was preincubated with HIV-1 integrase (lanes 4 and 5) or FIV integrase (lanes 6 and 7) as described above. The reaction was started by adding 50 μg/ml of pBluescript SK II + plasmid DNA and incubating at 37 °C for 60 min. The reaction products were amplified by PCR as described under “Experimental Procedures.” Oligonucleotide BS + was used as the 3′-primer, and H-U5V1–2 (lanes 2, 4, and 6) or F-U5V1–2 (lanes 3, 5, and 7) was used as the 5′-primer. Lanes 1 and 2 contain DNA size markers with lengths in nucleotides indicated on the right. F, HIV-1; F, FIV; IN, integrase.
cleotides were used as donor and target molecules. The target oligonucleotide contains arbitrary DNA sequences and was singly labeled at the 3' end of the T5 strand. The donor oligonucleotide, which contains sequences derived from the U5 end of HIV-1 or FIV long terminal repeat, was not radiolabeled. Therefore, the slowly migrating products were generated from 3' end joining of the donor DNA to the labeled target DNA (Fig. 4A). With HIV-1 integrase (Fig. 4A, lanes 3 and 4; see also Fig. 7), the most preferred position for integration was nucleotide 13 (counting from the 5' end), and other preferred sites were 14, 16, 17, 18, and 19. With FIV integrase, the most preferred position was nucleotide 17, and other preferred sites were 13, 16, 18, and 19 (Fig. 4A, lanes 5 and 6). The integration pattern was unchanged as the concentration of both integrases varied from 25 to 500 nM (data not shown).

The result from the oligonucleotide-based assay confirmed the previous observation that HIV-1 and FIV integrases have different target site preferences. However, the selection of target sites by HIV-1 or FIV integrase was identical regardless of whether the donor DNA was derived from HIV-1 U5 (Fig. 4A, lanes 3 and 5) or FIV U5 (Fig. 4A, lanes 4 and 6). Therefore, using an identical target DNA, the integration pattern depended on the source of integrase and was independent of the donor DNA.

In the PCR-based 3'-end joining assay, the donor DNA was identical to the one used in the oligonucleotide assay, and the target DNA was a Bluescript plasmid. After the reaction, the integration site in the recombinant product was amplified by PCR using two primers. One primer anneals to the plasmid, and the other primer, which was labeled at the 5' end, anneals to the donor DNA. The PCR products were analyzed on a denaturing polyacrylamide gel (30, 34). Each band on the gel corresponds to an integration event at a given phosphodiester bond (Fig. 4B). The frequency of integration at a particular site is proportional to the intensity of the band, and its position can be determined with use of DNA size markers or a sequencing ladder.

Similar observations to those of the oligonucleotide-based assay were made with the PCR-based assay; HIV-1 and FIV integrases had different integration patterns, and the integration pattern was independent of the donor DNA (Fig. 4B). The integration patterns of some chimeric proteins was also analyzed using the PCR-based assay and again was found to be independent of the source of donor DNA (data not shown). In all subsequent experiments, the integration pattern was determined using HIV-1 U5 as the donor DNA.
minus, F/H/H and H/F/F (Fig. 3, lanes 4 and 8; Fig. 5, lanes 5 and 11), were similar to those of the wild-type HIV-1 (Fig. 3, lane 2; Fig. 5, lane 3) and FIV integrase (Fig. 3, lane 6; Fig. 5, lane 9), respectively. Likewise, the integration patterns of chimeric proteins containing a swap at the C terminus, H/H/F and F/F/H (Fig. 3, lanes 3 and 7; Fig. 5, lanes 4 and 10), were similar to those of the wild-type HIV-1 and FIV integrases, respectively. Swapping of the central core domain, however, produced a reciprocal exchange of integration patterns between the two resultant chimeric proteins, F/H/F (Fig. 5, lane 6) and H/F/H (Fig. 3, lane 9; Fig. 5, lane 12). The result indicates that the domain responsible for target selection is not in the C or N terminus and instead resides in the core region of integrase.

The role of the core domain in target site selection was further supported by the results obtained from various deletion mutants. Like some of the chimeric integrases, the 3′-end joining activity of the deletion mutants could only be detected by the PCR-based assay (Fig. 5). HIV-1 integrase containing a deletion in the C terminus (H-IN1–234) had an integration pattern (Fig. 5, lane 8) similar to that of the full-length HIV-1 integrase, whereas FIV integrase containing a deletion in the N terminus (F-IN3–235) or C terminus (F-IN53–281) had an integration pattern (Fig. 5, lanes 13 and 14) similar to that of the full-length FIV integrase. Most strikingly, similar integration patterns were obtained when the target site usage was analyzed with only the core domain of HIV-1 (H-IN50–234; Fig. 5, lane 9) or FIV integrase (F-IN53–235; Fig. 5, lane 15). We conclude that both the N and C termini are not essential in determining target site usage.

Patterns of Nonspecific Alcoholysis of Wild-type and Chimeric Integrases—In addition to having sequence- and site-specific 3′-end processing activity, the integrases of HIV-1, Rous sarcoma virus, and visna virus exhibit a nonspecific endonuclease activity (42). The nonspecific alcoholysis of different integrases shows different preferences for target sites. The nonspecific cleavage pattern has been used previously to map the domain responsible for target site selection (33). Consistent with previous observations (42), the preferred sites for nonspecific alcoholysis of FIV integrase were different from those of HIV-1 integrase (Fig. 6, lanes 3 and 7; see also Fig. 7). We therefore examined the patterns of nonspecific alcoholysis of chimeric integrases of HIV-1 and FIV as another means of mapping the domain for target site selection (Fig. 6). For proteins with a core domain derived from FIV integrase (Fig. 6, lanes 8–10), the nonspecific alcoholysis patterns were similar to that of the wild-type FIV integrase (Fig. 6, lane 7). Of the three chimeric proteins that contained the core domain of HIV-1 integrase, only F/H/H (Fig. 6, lane 5) showed a nonspecific alcoholysis pattern that was identical to the wild-type HIV-1 integrase (Fig. 6, lane 3). The nonspecific alcoholysis patterns of H/H/F and F/F/H (Fig. 6, lanes 4 and 6) were identical to each other but were different from that of wild-type integrases of HIV-1 or FIV. The result suggests that domains other than the core can influence the selection of nonspecific alcoholysis sites. Furthermore, the core domain of HIV-1 integrase or FIV integrase had a very weak to undetectable level of nonspecific cleavage activity (Fig. 6, lanes 11 and 12), suggesting that nonspecific alcoholysis requires domains in addition to the core.

**DISCUSSION**

In vivo and in vitro studies showed that integration of retroviral DNA occurs into many sites on target DNA (Ref. 26 and references therein). The process, however, is not entirely random; integration into some sites occurs at a frequency several hundred times random (27). One factor known to affect target site selection is integrase, a viral enzyme that catalyzes the 3′-end processing and 3′-end joining steps of the integration reaction. Even when tested under identical reaction conditions and using an identical DNA substrate as the integration target,
integrase from different retroviruses have different target site preferences (30, 33). To identify the domain responsible for target site selection, the integrases of HIV-1 and FIV were divided into three domains, N terminus, central core, and C terminus, and a total of six chimeras were prepared by exchanging each of the three domains between the two proteins.

Analysis of the integration patterns of the various chimeras between HIV-1 and FIV integrases mapped the domain responsible for selection of target sites to the central core region of integratease. Using both oligonucleotide- and PCR-based assays for determining the distribution and frequency of integration events, we found that a reciprocal exchange of the core domain resulted in a corresponding exchange in the integration pattern of the resultant chimeras. In contrast, an exchange of the N or the C terminus did not alter the integration patterns of the chimeras from their respective wild-type integratease. The chimeric result is further corroborated by analysis of deletion mutants of HIV-1 and FIV integrases. HIV-1 integratease containing a deletion at the C terminus or FIV integratease containing a deletion at the N or C terminus showed that the deletion did not appreciably change the integration pattern. Moreover, similar integration patterns to those of the wild-type integrases could be obtained using only the core domain of HIV-1 or FIV integratease. Taken together, the results indicate that the central core domain of integratease is responsible for target site selection.

Close examination of the integration patterns obtained by the PCR-based assay revealed that subtle differences existed between the patterns of the chimeras and those of the wild-type integrases. The differences were reproducible and manifested mainly as a change in integration frequencies in a small fraction of integration sites. The difference was more apparent in the patterns obtained from the core swap chimeras and the deletion mutants. The subtle difference in integration patterns may be caused by slight changes in the core structure induced by the presence of exogenous terminal domains or the absence of terminal domains. Alternatively, the subtle difference implies that the terminal domains may contain secondary determinants for specifying target site usage.

All of the HIV/FIV chimeric proteins tested had poorer activities than the wild-type HIV-1 and FIV integrases. Overall, chimeric integrases containing a core domain derived from FIV integrase had higher activities than those containing an HIV-1 integrase core domain. Since the core domains of HIV-1 and FIV integrases alone had similar activities (data not shown), we speculate that the core domain of FIV integratease can be better complemented by the exogenous N- and C-terminal domains than that of HIV-1 integrase. The best example is the chimeric protein F/F/H, which possessed activities ranging from 50 to 80% of the wild-type integrases.

The finding that the central core domain controls the target site selection is consistent with the previous data showing that various N and C terminus deletion variants of HIV-1 integrase retain a similar integration pattern to that of the full-length integrase (34). Mutational analysis of human immunodeficiency virus type 2 integrase also showed that target site preference is altered by single amino acid substitutions of the asparagine at position 120 within the central core domain (43). We are not aware of any mutations in the N or C terminus of integrase that can lead to a change in target site preference.

Two regions of HIV-1 integrase have been shown to bind DNA: The core domain, which requires the presence of a divalent metal ion for DNA binding (23), and the C terminus, which does not require a divalent metal ion (22–25). Besides HIV-1 integrase, the C terminus of several integrases, including FIV, human immunodeficiency virus type 2, and avian sarcoma-leukosis virus, is capable of binding DNA nonspecifically (20, 22, 24). Because it binds DNA and is required for efficient 3′-end joining, the C terminus is generally regarded as the target DNA-binding domain. The present finding that exchanging or deleting the C terminus of integrase did not alter the integration pattern indicates that the C terminus, although capable of binding target DNA, is not involved in target site selection. Whether the metal ion-dependent, DNA-binding region in the core domain (23) is equivalent to the region for target site selection awaits further investigation.

In addition to the site-specific 3′-end processing activity, integrase possesses a nonspecific nuclease activity termed nonspecific alcoholysis (42, 44). As in the joining reaction, the cleavage site on the DNA substrate is not entirely random, and different integrase exhibit different preferred cleavage sites (33). By analyzing the pattern of nonspecific alcoholysis of chimeric proteins between HIV-1 and visna virus integrases and assuming nonspecific alcoholysis reflects target site selection during integration, the domain responsible for determining target site usage was mapped to the central core region (33). However, the integration pattern of the HIV-1/visna chimeric integrase was not examined. In the present study, the patterns of nonspecific alcoholysis of HIV-1/FIV chimeric integrases were determined and correlated with the patterns of 3′-end joining. For the most part, nonspecific alcoholysis and 3′-end joining activities parallel each other, but several discrepancies exist. First, the patterns of nonspecific alcoholysis of H/H/F and F/H/F were different from those of wild-type HIV-1 integrase and F/H/H, although their core domains were all derived from HIV-1 integrase. Second, with both HIV-1 and FIV integrases, the preferred DNA sites for nonspecific alcoholysis were different from those for 3′-end joining (Fig. 7). Third, the core domain alone of both HIV-1 and FIV integrases had no significant nonspecific alcoholysis activity. These results lead us to conclude that the determinants for the two activities are not identical and that the pattern of nonspecific alcoholysis is not a reliable marker for identifying the minimal domain responsible for target site selection.

Since integration of retroviral DNA is essential for the subsequent expression of viral genes and production of progenies (45–47), selection of target sites may have a significant effect on the fate of the infecting retrovirus. Analysis of the preferred target sites may provide information on the interaction between integratease and the sequence and structure of target DNA. Future studies will be focused on identifying the minimal peptides or amino acids involved in target site selection and examining target site usage of infectious viruses bearing integratease with altered target specificity.

Acknowledgments—We thank Hélène Goulaouic for helpful discussions, Jocelyn Atienza for technical assistance, and TaiYun Roe and Janice Chow for comments on the manuscript.

REFERENCES
1. Craigie, R., Fujisawa, T., and Bushman, F. (1990) Cell 62, 829–837
2. Katz, R. A., Merkel, G., Kulkosky, J., Leis, J., and Skalka, A. M. (1990) Cell 63, 47–65
3. Vink, C., van Gent, D. C., and Plasterk, R. H. (1990) J. Virol. 64, 5219–5222
4. Bushman, F. D., and Craigie, R. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 1339–1343
5. Pahl, A., and Flugel, R. M. (1993) J. Virol. 67, 5426–5434
6. Bushman, F. D., and Wang, B. (1994) J. Virol. 68, 2215–2223
7. Katzman, M., and Sudol, M. (1994) J. Virol. 68, 3558–3569
8. Chow, S. A., Vincent, K. A., Ellison, V., and Brown, P. O. (1992) Science 255, 725–729
9. Engelkamp, A., and Craigie, R. (1992) J. Virol. 66, 6361–6369
10. Engelkamp, A., Bushman, F. D., and Craigie, R. (1993) EMBO J. 12, 3269–3275
11. van Gent, D., Vink, C., Oude Groenegte, A. A. M., and Plasterk, R. H. A. (1993) EMBO J. 12, 3261–3267
12. Polard, P., and Chandler, M. (1995) Mol. Microbiol. 15, 13–23
13. Ellison, V., and Brown, P. O. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 7316–7320
14. Ellison, V., Gerton, J., Vincent, K. A., and Brown, P. O. (1995) J. Biol. Chem. 270, 3320–3326
15. Kulkosky, J., Jones, K. S., Katz, R. A., Mack, J. P. G., and Skalka, A. M. (1992) *Mol. Cell. Biol.* 12, 2331–2338
16. Bushman, F. D., Engelman, A., Palmer, I., Wingfield, P., and Craigie, R. (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 3428–3432
17. Leavitt, A. D., Shiue, L., and Varmus, H. E. (1993) *J. Biol. Chem.* 268, 2113–2119
18. Kalpana, G. V., and Goff, S. P. (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 10593–10597
19. Hickman, A. B., Palmer, I., Engelman, A., Craigie, R., and Wingfield, P. (1994) *J. Biol. Chem.* 269, 29279–29287
20. Kahn, E., Mack, J. P. G., Katz, R. A., Kulkosky, J., and Skalka, A. M. (1991) *Nucleic Acids Res.* 19, 851–860
21. Woerner, A. M., Klutch, M., Levin, J. G., and Markus-Sekura, C. J. (1992) *AIDS Res. Hum. Retroviruses* 8, 2433–2437
22. Vink, C., Oude Groeneger, A. A. M., and Plasterk, R. H. A. (1993) *Nucleic Acids Res.* 21, 1419–1425
23. Engelman, A., Hickman, A. B., and Craigie, R. (1994) *J. Virol.* 68, 5911–5917
24. Puras Lutzke, R. A., Vink, C., and Plasterk, R. H. A. (1994) *Nucleic Acids Res.* 22, 4125–4131
25. Lodi, P. J., Ernst, J. A., Kuszewski, J., Hickman, A. B., Engelman, A., Craigie, R., Clore, G. M., and Gronenborn, A. M. (1995) *Biochemistry* 34, 9826–9833
26. Craigie, R. (1992) *Trends Genet.* 8, 187–190
27. Withers-Ward, E. S., Kitamura, Y., Barnes, J. P., and Coffin, J. (1994) *Genes & Dev.* 8, 1473–1487
28. Rohdevedhold, H., Weiher, H., Reik, W., Jaenisch, R., and Breindl, M. (1987) *J. Virol.* 61, 336–343
29. Kitamura, Y., Lee, Y. M. H., and Coffin, J. M. (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89, 5532–5536
30. Pryciak, P. M., and Varmus, H. E. (1992) *Cell* 69, 769–780
31. Pruss, D., Reeves, R., Bushman, F. D., and Wolfe, A. P. (1994) *J. Biol. Chem.* 269, 25031–25041
32. Muller, H. P., and Varmus, H. E. (1994) *EMBO J.* 13, 4704–4714
33. Katzman, M., and Sudol, M. (1995) *J. Virol.* 69, 5687–5696
34. Goulaouic, H., and Chow, S. A. (1996) *J. Virol.* 70, 37–46
35. Tahbott, R. L., Sparger, E. E., Lovelace, K. M., Fitch, W. M., Pedersen, N. C., Lucw, P. A., and Elder, J. H. (1989) *Proc. Natl. Acad. Sci. U. S. A.* 86, 5743–5747
36. Tabor, S., and Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. U. S. A.* 82, 1074–1078
37. Vincent, K. A., Ellison, V., Chow, S. A., and Brown, P. O. (1993) *J. Virol.* 67, 425–437
38. Chow, S. A., and Brown, P. O. (1994) *J. Virol.* 68, 3896–3907
39. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
40. Leavitt, A. D., Rose, R. B., and Varmus, H. E. (1992) *J. Virol.* 66, 2359–2368
41. Vink, C., van der Linden, R. H., and Plasterk, R. H. A. (1994) *J. Virol.* 68, 1468–1474
42. Katzman, M., and Sudol, M. (1996) *J. Virol.* 70, 2598–2604
43. van Gent, D. C., Oude Groeneger, A. A. M., and Plasterk, R. H. A. (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89, 9598–9602
44. Engelman, A., and Craigie, R. (1995) *J. Virol.* 69, 5908–5911
45. LaFemina, R. L., Schneider, C. L., Robbins, H. L., Callahan, P. L., LeGrow, K., Roth, E., Schleif, W. A., and Emin, E. A. (1992) *J. Virol.* 66, 7414–7419
46. Sakai, H., Kawamura, M., Sakuragi, J.-I., Sakuragi, S., Shibata, R., Ishimoto, A., Ooi, N., Ueda, S., and Adachi, A. (1993) *J. Virol.* 67, 1169–1174
47. Englund, G., Theodore, T. S., Freed, E. O., Engelman, A., and Martin, M. A. (1995) *J. Virol.* 69, 3216–3219