Multiple Integrase Functions Are Required to Form the Native Structure of the Human Immunodeficiency Virus Type I Intasome

THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 274, No. 24, Issue of June 11, pp. 17358–17364, 1999
Printed in U.S.A.

Multiple Integrase Functions Are Required to Form the Native Structure of the Human Immunodeficiency Virus Type I Intasome*

(Received for publication, December 17, 1998, and in revised form, March 21, 1999)

Hongmin Chen‡, Shui-Qing Wei§, and Alan Engelman‡¶

From the ‡Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute and the Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115 and the §Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

Mu-mediated polymerase chain reaction footprinting was used to investigate the protein-DNA structure of human immunodeficiency virus type I (HIV-I) preintegration complexes. Preintegration complexes were partially purified from cells after using an established coculture infection technique as well as a novel technique using cell-free supernatant from transfected cells as the source of virus. Footprinting revealed that bound proteins protected the terminal 200–250 base pairs of each viral end from nuclease attack. Bound proteins also caused strong transpositional enhancements near each end of HIV-I. In contrast, regions of viral DNA internal to the ends did not show evidence of strong protein binding. The end regions of preintegrative HIV-I apparently form a unique nucleoprotein structure, which we term the intasome to distinguish it from the greater preintegration complex. Our novel system also allowed us to analyze the structure and function of preintegration complexes isolated from cells infected with integrase mutant viruses. Complexes were derived from viruses defective for either integrase catalysis, integrase binding to the viral DNA substrate, or an unknown function in the carboxyl-terminal domain of the integrase protein. None of these mutant complexes supported detectable integration activity. Despite the presence of the mutant integrase proteins in purified samples, none of these nucleoprotein complexes displayed the native intasome structure detected in wild-type preintegration complexes. We conclude that multiple integrase functions are required to form the native structure of the HIV-I intasome in infected cells.

A pivotal step in the retroviral life cycle is forming the provirus, an integrated cDNA copy of the viral RNA genome. The key viral players in integration are the *trans*-acting integrase (IN)[1] protein and the *cis*-acting DNA attachment site. Integration proceeds through three steps, the first two of which are known to require IN function. The linear ends of the cDNA are initially processed adjacent to phylogenetically conserved CA dinucleotides, resulting in a pair of recessed 3′ ends. After nuclear localization, the exposed 3′-hydroxyls are joined to the 5′-phosphates of a double-stranded staggered cut in chromosomal DNA. The final step is DNA repair, wherein the single-stranded gaps at the sites of joining are sealed, resulting in the sequence duplication of the double-stranded cut flanking the integrated provirus (for a review, see Ref. 1).

In infected cells, integration is mediated by large nucleoprotein complexes known as preintegration complexes (PICs), which are derived from the cores of infecting virions (2). PICs isolated from infected cells can integrate their endogenous cDNA into an exogenously added target DNA *in vitro* (2–7). Additionally, recombinant IN proteins purified after expression in bacteria can cut and join oligonucleotide attachment site DNA substrates (reviewed in Ref. 1). These latter *in vitro* assays have been invaluable for deciphering the structure and function of retroviral IN proteins. IN can be divided into three distinct functional domains: the amino-terminal, catalytic core, and carboxyl-terminal domains (reviewed in Ref. 8). The central domain contains the highly conserved D,D(35)E amino acid motif that comprises the IN active site (9, 10).

Although simplified *in vitro* assays using purified IN proteins have provided essential knowledge toward understanding the overall integration process, these systems only partially mimic integration *in vivo*, because the predominant recombination products result from the insertion of only one viral DNA end into just one strand of target DNA. Virus replication requires integration of both DNA ends into both strands of target DNA; the single-ended activity typical of some *in vitro* systems would not yield a productive viral infection. Although altering the source of purified IN protein from bacterial to viral and/or modifying reaction conditions can increase the frequency of two-ended integration products (11–13), these systems still do not recapitulate the efficiency of two-ended integration activity displayed by PICs isolated from infected cells (14).

The discrepancy in reaction products catalyzed by PICs as compared with purified IN proteins suggests that efficient two-ended integration activity might require higher-order protein-protein and/or protein-DNA interactions specific to nucleoprotein complexes derived from infected cells. To begin to address this, we have used *in vitro* Mu-mediated polymerase chain reaction (MM-PCR) footprinting to analyze the protein-DNA structure of human immunodeficiency virus type I (HIV-I) PICs partially purified from infected cells. We have established an efficient HIV-I infection system initiated from transfected cell supernatant that yields active PICs. This has allowed us to analyze the structure and function of HIV-I PICs derived from a number of IN mutant viruses. Our results indicate that multiple IN functions are required to form the native protein-DNA structure of the HIV-I intasome.

* This work was funded in part by National Institutes of Health Grant AI39394, by funds from the G. Harold and Lelia Y. Mathers Foundation, by a gift from the Friends 10, and by the National Institutes of Health Intramural AIDS Targeted Antiviral Program. 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.

¶ To whom correspondence should be addressed: Dept. of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney St., Boston, MA 02115. Tel.: 617-632-4361; Fax: 617-632-3113; E-mail: alan_engelman@dfci.harvard.edu.

[1] The abbreviations used are: IN, integrase; PIC, preintegration complex; PCR, polymerase chain reaction; MM-PCR, Mu-mediated polymerase chain reaction; HIV-I, human immunodeficiency virus type I; LTR, long terminal repeat; MoMLV, Moloney murine leukemia virus.

MATERIALS AND METHODS

Cells and Viruses—MOLT IIIB (5), SupT1 (15), C8166 (16), MT-4 (17), Jurkat (18), and CEM-12D7 (19) T-cell lines were maintained in RPMI 1640 medium containing 10% fetal calf serum. 293T (20) and HeLa-CD4 (21) cells were grown in Dulbecco's modified Eagle's medium containing 10% fetal calf serum. Two different infection systems were used to isolate HIV-I PICs. In one system, chronically infected MOLT IIIB cells were cocultured with uninfected SupT1 cells, essentially as described previously (5). HIV-I production (HTLV-IIIB strain) was stimulated 2- to 3-fold by pretreating MOLT IIIB cells (4×10^5 cells/ml) for 24 h with phorbol 12-myristate 13-acetate (10 μg/ml). In the other infection system, cell-free HIV-I (NL4-3 strain) produced from transfected 293T cells was used to infect C8166 cells (see below).

Plasmids and Oligonucleotides—Wild-type pNL4-3 (22) and IN mutants D116N (23) and K156E/K159E (24) have been described previously. HIV-I (NL4-3) carrying the substitution of Lys for Gln-62 (Q62K) in IN displays a replication-defective phenotype similar to that of K156E/K159E.[2] The W235E IN change (25) was incorporated into pNL4-3 using overlapping PCR (23).

Oligonucleotides were gel-purified before MM-PCR. Mu50 (5′- GTTTTCGCATTTATCGTGAAACGCTTTCGCGTTTTTCGTGCGCCGCTTCA) and Mu54 (5′-CTGCTGAAGCGGCGCACGAAAAACGCGAAAGCGTTTCACGATAAATGCGAAAAC) were annealed to generate Mu right-end DNA; Mu25 (5′-GCATTTATCGTGAAACGCTTTCGCG) was the Mu-specific PCR primer. HIV-I primers were chosen to match both the NL4-3 and HXBc2 (26) strains. The primers and their nucleotide positions in HXBc2 are as follows. AE330 (referred to as LA9 in Ref. 27; 805 to 787; 5′-GACGCTCTCGCACCCATCTC) and AE529 (8690–8708; 5′-GCTGAGGGGACAGATAGGG) were used in the first round to amplify the 5′ and 3′ regions, respectively, of HIV-I. Minus-strand primers AE604 (454 to 432; 5′-CAGTACAGGCAAAAAGCAGCTGC) and AE584 (215 to 196; 5′-CACAGGGTGTAACAAGCTGG) were used in second-round PCRs to analyze the 5′-long terminal repeat (LTR). Plus-strand primers AE461 (9332–9356; 5′-AGTGTTAGAGTGGAGGTTTGACAGC) and AE459 (9522–9546; 5′-GCTTTTTGCCTGTACTGGGTCTCTC) were used in the second rounds to analyze the 3′-LTR. The plus-strand primer AE525 (3220–3240; 5′-ACCTCCATTCCTTTGGATGGG) was used in the first round to analyze a region of HIV-I internal to the LTR ends; AE524 (3264–3281; 5′-GGACAGTACAGCCTATAG) was used in this second-round PCR.

Infections, PIC Isolation, and in Vitro Integration Assay—SupT1 cells (4×10^7) were cocultivated with phorbol 12-myristate 13-acetate-treated MOLT IIIB cells in 20 ml of MOLT IIIB-conditioned medium (24-h supernatant of unstimulated cells seeded at 1×10^6 cells/ml) at a ratio of 10:1. After 5 h, cells were washed twice in buffer K (20 mM HEPES, pH 7.5, 5 mM $MgCl_2$, 150 mM KCl, 1 mM dithiothreitol, and 20 μg/ml aprotinin) and lysed in 1 ml of buffer K-0.025% (w/v) digitonin. Crude cytoplasmic extract cleared by brief centrifugation was treated with RNase A (0.1 mg/ml) for 30 min at room temperature. To partially purify PICs, 1.5 ml of lysate was passed through a bovine serum albumin-coated Sepharose CL-4B (Amersham Pharmacia Biotech) spin column (12 ml) equilibrated in buffer K-0.025% digitonin. The eluate was further purified on a 10-ml, 10–50% (w/v) Nycodenz gradient prepared in buffer K. Centrifugation was for 16 h at 274,000 $\times$ *g* in a Beckman SW 41 Ti rotor at 4 °C. Twelve fractions were collected from the top of the gradient.

293T cells were seeded at 5.8×10^4 cells/cm^2 in a 14-cm plate 24 h before transfection. Cells were transfected with plasmid DNA (50 μg) using calcium phosphate coprecipitation (28). Transfected cell supernatant was passed through 0.45 μm filters and treated with DNase I (Promega; 2 units/ml) in the presence of 10 mM $MgCl_2$ for 45 min at 37 °C. This step reduced the level of plasmid carry-over from transfection to the limit of detection by Southern blotting. For analyzing reverse transcription in cell lines, T cells (2×10^7 cells) and HeLa-CD4 cells (4×10^6 cells plated in a 10-cm plate 20 h before infection) were infected for 8 h with 10 ml of cell supernatant. Crude cytoplasmic extracts (0.5 ml) treated with RNase A were deproteinized by adding SDS, EDTA, and proteinase K to final concentrations of 0.5% (w/v), 6 mM, and 0.6 mg/ml, respectively. After incubating at 56 °C for 1 h, DNA was recovered by phenol/chloroform extraction and precipitation with ethanol. To isolate PICs, C8166 cells (3×10^7) were infected for 8 h with 20 ml of transfected cell supernatant. RNase A-treated cytoplasmic extract was loaded directly onto Nycodenz gradients.

PIC activity was assayed essentially as described previously (5). Briefly, 200 μl of crude cytoplasmic extract, spin column eluate, or gradient fraction was incubated with 600 ng of linearized ϕX174 DNA for 45 min at 37 °C. Reactions were deproteinized, and DNA was recovered by precipitation with ethanol. DNA was electrophoresed through 0.6% agarose gels in Tris acetate-EDTA (28), transferred to GeneScreen Plus membrane (NEN Life Science Products), and probed with a LTR-specific riboprobe (HXBc2 nucleotides 8896–9615). Integration activity was quantified as the percentage of cDNA substrate converted into product using either PhosphorImager analysis (Molecular Dynamics) or densitometry (IS-1000 Digital Imaging System).

Western Blotting—Gradient-purified C8166 cell extracts adjusted to 15% (w/v) glycerol were frozen in liquid N_2 and stored at −80 °C. SDS was added to a final concentration of 0.25% to thawed samples (200 μl), and proteins were recovered by methanol-chloroform-H_2O extraction essentially as described previously (29). Briefly, samples were mixed with 4 volumes of methanol, followed by 2 volumes of chloroform. Two phases were separated after mixing 3 volumes of H_2O and centrifuging at 9,000 $\times$ *g* for 1 min. Proteins at the interface were precipitated using 3 volumes of methanol and spinning at 9,000 $\times$ *g* for 2 min. Pelleted proteins were resuspended in 1× sample buffer (12 mM Tris-HCl, pH 6.8, 5% glycerol, 0.4% SDS, 2.88 mM 2-mercaptoethanol, and 0.02% bromphenol blue), boiled for 10 min, and electrophoresed through 10% SDS-polyacrylamide gels. Proteins were transferred to Hybond-C extra membrane (Amersham Pharmacia Biotech), and IN was detected using monoclonal antibody 8E5 (30) with an ECL Western blotting kit (Amersham Pharmacia Biotech). Recombinant HIV-I IN protein was purified after expression in *Escherichia coli* as described previously (31).

MM-PCR Footprinting—Mu A protein was kindly provided by Dr. Michiyo Mizuuchi (National Institute of Diabetes and Digestive and Kidney Diseases). Two different naked DNA footprinting controls were generally used. In one, plasmid DNA was added to buffer K to a final concentration of 1.25 ng/ml. The other control was deproteinized PICs resuspended in buffer K. Mu transpososomes were assembled by mixing the annealed Mu end (40 nM) with Mu A transposase (232 nM) in 40 mM Tris-HCl, pH 8.0, 0.005% (w/v) bovine serum albumin, 30% (w/v) glycerol, 30% (v/v) dimethyl sulfoxide, 0.2% (v/v) Triton X-100, and 150 mM KCl. After a 10-min incubation at 30 °C, the mixture was diluted 1:3 in the same buffer. Assembled Mu transpososomes (200 μl) were mixed with an equal volume of gradient-purified PICs or naked DNA controls, and transposition was initiated by adding $CaCl_2$ to a final concentration of 10 mM. After 30 min at 30 °C, reactions were stopped by adding SDS, EDTA, EGTA, and proteinase K to final concentrations of 0.5%, 6 mM, 10 mM, and 0.6 mg/ml, respectively. Deproteinized Mu transposition products were recovered by precipitation with ethanol.

DNA was dissolved in 50 μl of H_2O, of which 3 μl was used in nested PCR. The first round (50 μl) contained 20 mM Tris-HCl, pH 8.8, 10 mM KCl, 10 mM $(NH_4)SO_4$, 4 mM $MgSO_4$, 0.1% Triton X-100, 0.4 μM each of Mu25 and a viral-specific primer, 0.4 mM of each dNTP, 0.1 mg/ml bovine serum albumin, and 1 unit of Vent DNA polymerase (New England Biolabs). Twenty-five cycles were performed on a Perkin-Elmer GeneAmp PCR system with the following PCR profile: preheating at 96 °C for 4 min, denaturation at 96 °C for 15 s, annealing at 58 °C for 1 min, and elongation at 74 °C for 1 min. A final 10-min elongation at 74 °C was then performed. An aliquot (2 μl) was transferred for the second round (25 μl; 25 cycles), which contained 20 mM Tris-HCl, pH 8.8, 10 mM KCl, 10 mM $(NH_4)SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100, 0.2 μM each of Mu25 and a nested viral primer, 5×10^5 cpm of 5′ end ^{32}P-labeled viral primer, 0.2 mM of each dNTP, 0.1 mg/ml bovine serum albumin, and 1 unit of Vent (exo$^-$) DNA polymerase (New England Biolabs). PCR was carried out essentially as described above, except that the elongation time was adjusted from 10 to 45 s, based on the length of the fragment being amplified. Aliquots of second-round PCRs were analyzed on 5% denaturing sequencing gels.

RESULTS

Active HIV-I PICs from Molecularly Cloned DNA—In this study, HIV-I PICs were isolated from acutely infected cells using one of two different tissue culture infection systems. In one system, HIV-I (HTLV-IIIB strain) infection was initiated by mixing uninfected SupT1 cells with chronically infected MOLT IIIB cells, essentially as described previously (5). Cells were lysed 5 h after infection, and cytoplasmic extract containing HIV-I PICs was purified by spin column chromatography. This step removed 80–90% of the total protein present in the cell extract, yielding PICs that reproducibly supported a higher level of integration activity than those in the starting extract (Fig. 1*A*, *lanes 1–4*; data not shown). The column eluate was

[2] A. Engelman, manuscript in preparation.

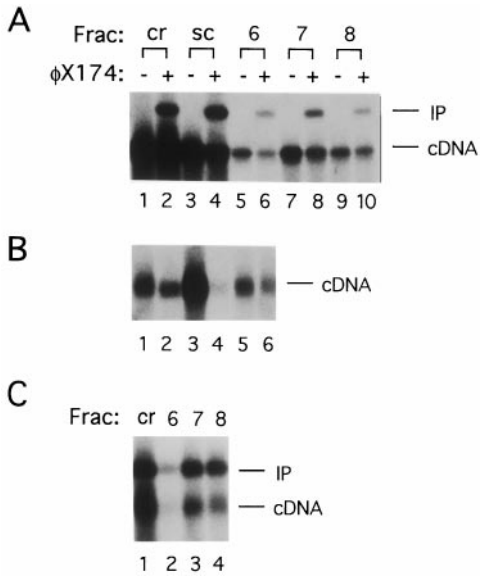

FIG. 1. **Integration activity of HIV-I PICs.** *A*, Southern blot of PICs isolated from infected SupT1 cells. ϕX174 target DNA was included in integration reactions as indicated. *Lanes 1* and *2*, crude cytoplasmic extract (about 24% of the substrate was converted to product in *lane 2*); *lanes 3* and *4*, spin column eluate (about 37% of the substrate was converted to product in *lane 4*.) *Lanes 5–10*, activity of gradient-purified PICs; about 37% of the substrate was converted to product in *lane 6*, about 32% of the substrate was converted to product in *lane 8*, and about 26% of the substrate was converted to product in *lane 10*. *B*, cDNA synthesis in various CD4-positive cell lines. *Lane 1*, SupT1 cells; *lane 2*, HeLa-CD4 cells; *lane 3*, C8166 cells; *lane 4*, MT-4 cells; *lane 5*, Jurkat cells; *lane 6*, CEM-12D7 cells. *C*, activity of PICs isolated from C8166 cells. *Lane 1*, activity of crude cytoplasmic extract. *Lanes 2–4*, activities of gradient fractions 6–8, respectively. About 60% of HIV-I cDNA was converted to product in each of these reactions. *cDNA*, 9.7-kilobase pair HIV-I substrate; *IP*, 15.1-kilobase pair integration product.

further purified on a Nycodenz gradient, which was fractionated and analyzed for HIV-I DNA content and PIC activity. DNA and integration activity co-sedimented to fractions 6, 7, and 8 (Fig. 1*A*, *lanes 5–10*). The integration activity of the crude cell extract varied from 10% to 50%, depending on the individual preparation.

One goal of this study was to analyze the structure and function of HIV-I PICs derived from IN mutant viruses. Thus, we set out to establish an infection technique using transfected cell supernatant as the source of cell-free virus. Preliminary experiments tested the ability of six different CD4-positive cell lines to support reverse transcription after infection with molecularly cloned HIV-I (NL4-3 strain). Of these six cell lines, C8166 T cells (16) supported the highest level of cDNA synthesis (Fig. 1*B*). A kinetic analysis revealed maximum levels of reverse transcription approximately 8 h after infecting C8166 cells (data not shown). Because HIV-I cDNA in the crude cytoplasmic extract of C8166 cells was lost during spin column chromatography (data not shown), these extracts were directly purified by Nycodenz gradient centrifugation. The gradients were fractionated and analyzed for HIV-I DNA content and PIC activity. Both cDNA and integration activity colocalized to fractions 6–8 (Fig. 1*C*, *lanes 2–4*). PICs from C8166 cells routinely integrated more of their cDNA substrate into the target (30–80% of substrate converted to product) as compared with PICs isolated from the coculture infection system.

MM-PCR Analysis of Wild-type HIV-I PICs—MM-PCR footprinting was developed to analyze the native protein-DNA structure of Moloney murine leukemia virus (MoMLV) PICs (32). In this footprinting technique, preassembled Mu transpososomes are the DNA cleavage reagent. Mu A transposase inserts Mu-end DNA at the site of cleavage, and, as in any protein-DNA footprinting technique, bound protein interferes with the ability of the cleavage reagent to cut the nucleic acid. The advantage of this coupled cutting and DNA joining technique over other footprinting methods is that the DNA cleavage reagent itself becomes a substrate for subsequent PCR amplification (Fig. 2). We thus applied this technique to analyze the native protein-DNA structure of HIV-I PICs isolated from cells using two different tissue culture infection systems.

Three different substrates, naked plasmid, deproteinized PICs, and native PICs, were generally analyzed. PICs partially purified after the coculture infection were initially tested. As predicted, deproteinized PICs supported a pattern of Mu transposition similar to the pattern detected using naked plasmid DNA (Fig. 3, *A* and *C*, compare *lanes 2* to *lanes 1*). Native PICs, however, revealed dramatically different transpositional patterns. Bound proteins protected more than 100 base pairs from each LTR end (Fig. 3, *A* and *C*, *lanes 3*). In addition to these large footprinted regions, transpositional enhancements were detected near the very ends of HIV-I (Fig. 3, *A* and *C*, *lanes 3*). Both the footprinted and enhancement regions were characteristic of the viral DNA ends; internal regions of deproteinized and native PICs supported similar distributions of Mu transposition (Fig. 3*B*). However, subtle differences were detected (Fig. 3*B*), suggesting that proteins may loosely associate with internal regions of HIV-I. PICs partially purified from infected C8166 cells yielded the same overall results. These PICs generally displayed clearer footprinted regions (Fig. 3, *D* and *E*), perhaps due to their higher integration activity.

The internal boundaries of the footprints were next identified. Whereas the U3 footprint extended approximately 250 base pairs from the end of HIV-I, the U5 footprint extended about 200 base pairs (Fig. 4). PICs isolated from infected C8166 cells displayed sharper footprint boundaries than those isolated from the coculture infection (Fig. 4; data not shown).

Multiple IN Functions Are Required to Form the HIV-I Intasome—Because IN is essential for retroviral DNA integration, we tested the effects of different IN mutations on the structure and function of HIV-I PICs. HIV-I IN mutants can be divided into two classes, class I and II, based on their effects on the virus replication cycle (33). Whereas class I mutants are blocked specifically at the integration step, class II mutants display reverse transcription and/or virus assembly defects. Thus, measurable quantities of PICs can only be derived from class I IN mutant viral infections. Most HIV-I deletions, as well as a number of point changes, unfortunately fall into class II (33).

HIV-I IN contains three functional domains, the amino-terminal, catalytic core, and carboxyl-terminal domains (8), and each domain contains at least one amino acid residue that is conserved among all retroviruses (33). The amino-terminal domain contains two conserved His residues and two conserved Cys residues. Viral mutants carrying substitutions of any one of these residues display the class II replication phenotype (33), therefore measurable quantities of these PICs were not recoverable (data not shown). The catalytic domain contains the conserved active site Asp and Glu residues of the D,D(35)E motif as well as residues such as Lys-159 (24, 34) and Gln-62 (31, 34, 35) implicated in viral DNA end binding. Substituting any of the three active site residues yields the class I mutant viral phenotype, as do certain substitutions of Lys-159 (the double mutant K156E/K159E was used here) and Gln-62 (Q62K). The substitution of Glu for the conserved carboxyl-terminal domain residue Trp-235 (W235E) also yields the class I mutant viral phenotype (25). At present, the exact defect of the W235E mutant virus is unknown. We thus analyzed four

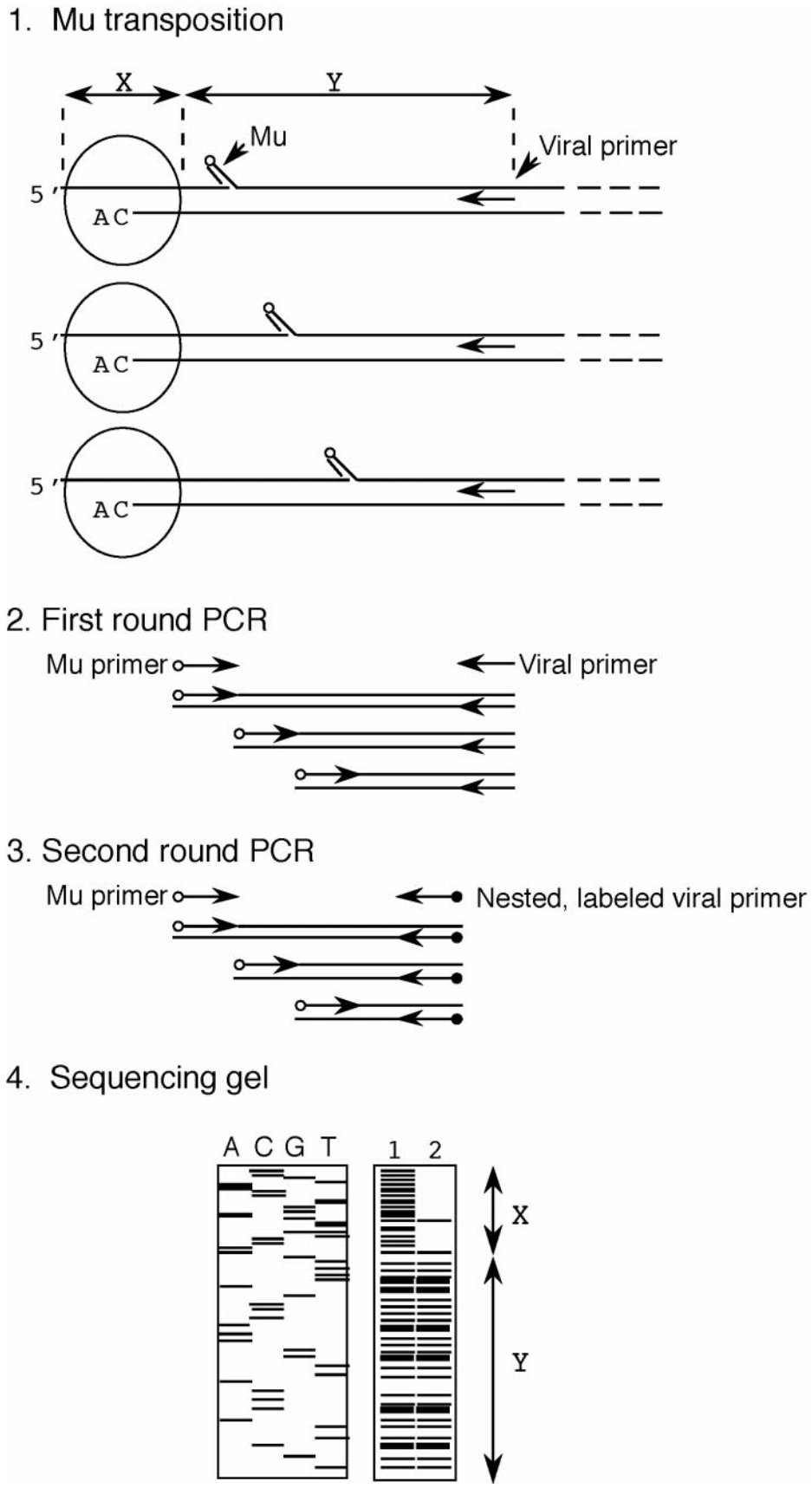

FIG. 2. **MM-PCR footprinting.** Transposition results in the insertion of Mu end DNA at the site of nuclease attack (step 1). The frequencies and distribution of Mu cutting are detected using two rounds of PCR (steps 2 and 3). Regions of protein binding that prevent Mu from cutting the DNA (indicated by *X*) are revealed after denaturing polyacrylamide gel electrophoresis (step 4). *Lane 1*, MM-PCR pattern of deproteinized control; *lane 2*, pattern of native nucleoprotein complex.

different class I mutant viruses that affected three different IN functions: D116N, which was defective for catalysis; Q62K and K156E/K159E, each of which was defective for viral DNA end binding; and W235E, which was defective for an unknown carboxyl-terminal domain function.

Cytoplasmic extracts of C8166 cells were prepared after in-

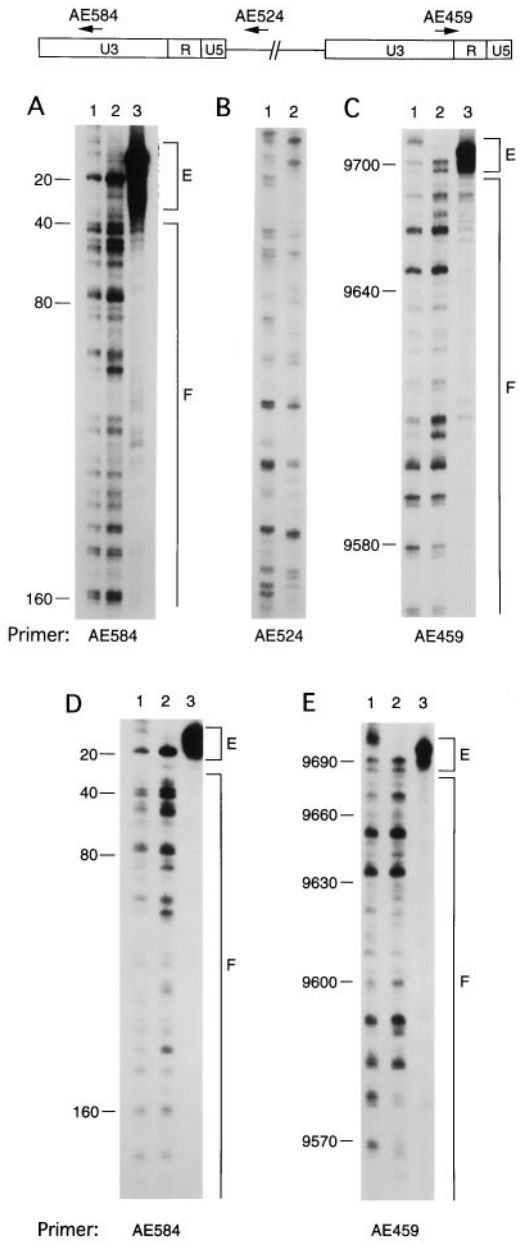

FIG. 3. **MM-PCR footprinting of wild-type HIV-I PICs.** Mu transposition reactions were deproteinized and analyzed by sequencing gels after two rounds of PCR. The *top panel* shows the relative positions of HIV-I primers used in second-round PCRs. *A*, native structure of the U3 end of HTLV-IIIB. *Lane 1*, naked plasmid DNA; *lane 2*, deproteinized PICs; *lane 3*, native PICs. *B*, internal HIV-I region. *Lanes 1* and *2*, deproteinized and native samples, respectively. *C*, native structure of the U5 end of HTLV-IIIB. *Lanes 1–3* were the same as in *A*. *D* and *E*, native structure of the U3 and U5 ends, respectively, of NL4-3. *Lanes 1–3* in each panel were the same as in *A*. Nycodenz gradient fraction 7 was analyzed in *A–E*; MM-PCR of fraction 6 from infected SupT1 cells revealed footprinting and enhancement patterns indistinguishable from those of fraction 7. *Numbers* to the *left* of the panels refer to the nucleotide position in HIV-I. The HXBc2 molecular clone of HTLV-IIIB terminates at nucleotide 9718; NL4-3 U5 ends at nucleotide 9709. *E*, regions of transpositional enhancement; *F*, footprinted regions.

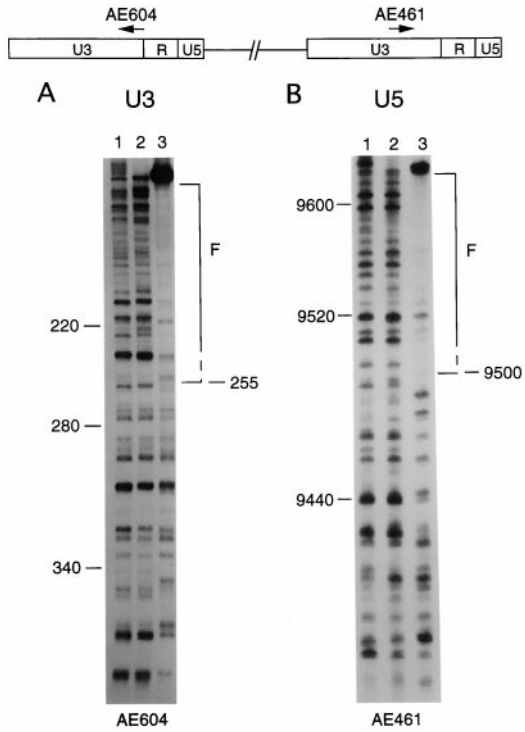

FIG. 4. **Extent of protein-DNA footprint.** The *top panel* shows the relative positions of viral primers used in the second-round PCRs. *A*, structure of the U3 end of NL4-3. *B*, the U5 end of NL4-3. Other labeling is as described in the legend to Fig. 3.

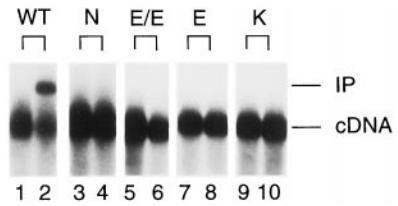

FIG. 5. **Integration activity of HIV-I IN mutant PICs.** PICs isolated after infection with the indicated IN mutant viruses were assayed for *in vitro* integration activity. *WT*, wild-type NL4-3; *N*, D116N IN mutant; *E/E*, K156E/K159E; *E*, W235E; *K*, Q62K. The samples in the *even-numbered lanes* were reacted with ϕX174 target DNA. Other labeling is the same as that described in the legend to Fig. 1.

fection with either wild-type or class I IN mutant viruses. Whereas wild-type PICs converted about 50% of the cDNA substrate to the integration product, none of the mutant viral PICs supported detectable integration activity (Fig. 5). The preparations were then centrifuged into Nycodenz gradients. Each mutant sedimented in Nycodenz to the same position as did wild-type, indicating that the nucleoprotein structure of each PIC was largely intact (data not shown). The gradient-purified samples were then subjected to MM-PCR footprinting. The results of this analysis showed that none of the mutant PICs displayed the pattern of protein-DNA footprinting and transpositional enhancements indicative of the wild-type HIV-I intasome (Fig. 6). To ensure that the lack of intasome structure for each of the mutants was not simply due to the absence of the IN protein, gradient-purified samples were analyzed by Western blotting using an anti-IN monoclonal antibody. The results of this experiment showed that each of the mutant PICs contained IN protein at a level comparable to that of wild-type (Fig. 7).

DISCUSSION

In this report, we describe the native protein-DNA structure of preintegrative HIV-I as detected by MM-PCR footprinting. The structure has two distinguishable characteristics. Firstly, large regions (200–250 base pairs) of protein binding were detected at the ends of HIV-I. Secondly, enhanced regions of nuclease attack were detected near the very ends of the virus. The overall structure is quite similar to the one previously reported for MoMLV (32). Thus, similar intasome structures may be common to the preintegrative DNAs of a variety of retroviruses.

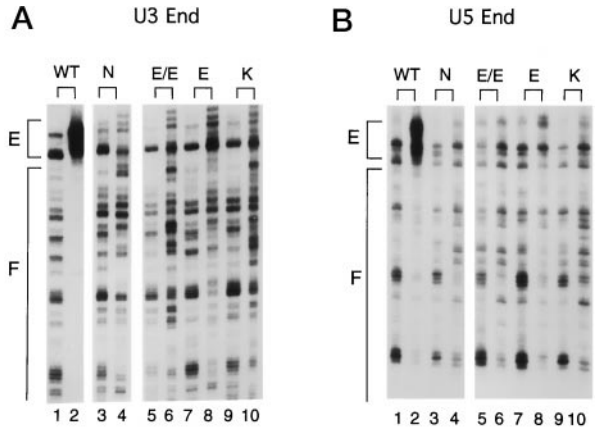

FIG. 6. **MM-PCR analysis of HIV-I IN mutant PICs.** IN mutant PICs partially purified by Nycodenz gradient centrifugation were analyzed by MM-PCR footprinting. *A*, structure of the U3 end of HIV-I. *B*, structure of the U5 end. The samples in the *odd-numbered lanes* were deproteinized before Mu transposition; *even-numbered lanes* contained the matched native samples. Other labeling is the same as that described in the legends to Figs. 3 and 5.

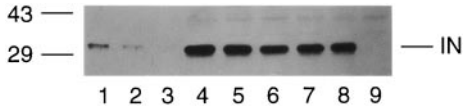

FIG. 7. **IN protein content of HIV-I mutant PICs.** Gradient-purified samples were analyzed by Western blotting. *Lanes 1–3* contained 50, 25, and 12.5 ng, respectively, of recombinant HIV-I IN protein. The samples in *lanes 4–8* were isolated after infection with wild-type, D116N, K156E/K159E, W235E, and Q62K HIV-I, respectively. The sample in *lane 9* contained a gradient-purified extract from mock-infected cells. The positions of molecular mass standards in kilodaltons are indicated on the *left*.

Numerous results point to the functional relevance of the retroviral intasome. Firstly, functionally intact IN was required to form both the MoMLV (32) and HIV-I (Fig. 6) structures. Also, both the MoMLV (32) and HIV-I (36) structures, as well as PIC function, were undetectable after stripping bound proteins with high concentrations of salt, and both structure and function were restored in parallel by adding back extracts from uninfected cells. Finally, Mu insertion into the enhanced regions destroyed the ability of MoMLV PICs to undergo subsequent *in vitro* integration (32). This interference highlights the functional relevance of the retroviral intasome as detected by MM-PCR footprinting.

The basis for the transpositional enhancements at the ends of MoMLV and HIV-I DNA is unknown. Retroviral integration and Mu transposition share numerous features, notably their mechanisms of polynucleotidyl transfer (reviewed in Ref. 37). Recombinant HIV-I IN protein preferentially integrates oligonucleotide attachment DNA substrates into regions of target DNA distortion (38), suggesting that protein binding in native PICs may distort viral end regions and create hot spots for Mu insertion. Alternatively, it is possible that proteins bound near the ends of retroviral DNA physically interact with the Mu nuclease. Bacterially expressed MoMLV IN protein bound to recombinant viral DNA displayed the viral end footprinting and enhancement patterns associated with the endogenous intasome isolated from infected cells (39), suggesting that IN is solely responsible for the MoMLV structure as detected by MM-PCR footprinting. However, we have been unable to detect evidence for the HIV-I intasome structure using purified recombinant IN with either recombinant viral DNA or deproteinized cDNA substrates purified from infected cells (data not shown).

To study the role of IN in the formation of the HIV-I intasome, we established a highly efficient infection system initiated with molecularly cloned virus. Numerous CD4-positive cell lines were screened for their ability to support reverse transcription after infection (Fig. 1*B*), and one T-cell line, C8166 (16), was not only found to support efficient DNA synthesis, but the resulting HIV-I PICs displayed efficient *in vitro* integration activity (Fig. 1*C*). As predicted, PICs isolated after infection with class I viral mutants defective for either IN catalysis (D116N), viral DNA end binding (Q62K and K156E/K159E), or an unknown carboxyl-terminal function (W235E) did not display detectable levels of *in vitro* integration activity (Fig. 5). Despite containing their mutant IN proteins (Fig. 7), none of the mutant PICs displayed the protein-DNA footprints and viral end enhancements indicative of the wild-type HIV-I intasome (Fig. 6). Although this was somewhat expected for IN core domain mutants K156E/K159E and Q62K, whose defects are predicted to disrupt HIV-I attachment site DNA binding (24, 34, 35), this was unanticipated for both the carboxyl-terminal domain mutant (see below) and the D116N active site mutant. One IN active site residue, Glu-152, has been implicated in interacting with the HIV-I attachment site (34). In contrast, recombinant D116N mutant proteins have not revealed evidence of DNA binding defects using either functional (34) or physical (40) *in vitro* assay systems. Thus, we speculate that IN catalysis may be closely linked to HIV-I intasome formation in infected cells. There is a precedent in other DNA recombination systems, notably Mu transposition, that the recombinationally active Mu A protein-Mu DNA transpososome becomes more stable with each successive step along the recombination pathway (37). Thus, 3′ processing of the viral ends by IN may stabilize the HIV-I nucleoprotein complex. A kinetic analysis revealed that the MoMLV intasome formed relatively slowly (32), consistent with the interpretation that IN catalysis may be required for intasome formation in infected cells.

The W235E defect in infected cells is unknown, in part because recombinant W235E mutant IN protein displays wild-type levels of 3′ processing and DNA strand transfer activities in *in vitro* integration assays (41). A region of the carboxyl terminus between residues 247 and 270 has been implicated in binding to the HIV-I attachment site (35, 42). In contrast, the region between residues 213 and 246, where Trp-235 resides, cross-linked to the target DNA portion of an *in vitro* disintegration DNA substrate (42). It is therefore possible that the W235E mutant virus is defective for interacting with chromosomal DNA in infected cells. However, because the structure of W235E PICs purified from cytoplasmic extracts was indistinguishable from the structure of either the D116N or attachment site DNA binding mutants, we propose that the W235E virus is blocked at a step that precedes target DNA interaction in the nuclei of infected cells. Determining the precise defect of the W235E mutant may reveal higher-order protein-protein and/or protein-DNA interactions important for PIC activity that are dispensable for recombinant IN function in more simplified *in vitro* integration assays. We also plan to continue to analyze the structure and function of HIV-I IN and attachment site mutant PICs isolated from infected cells.

Acknowledgments—We thank M. Mizuuchi for purified Mu A protein, X. Wu and J. Kappes for HeLa-CD4 cells, D. Helland for 8E5-producing hybridoma cells, and D. Harris for critical review of the manuscript.

REFERENCES

1. Brown, P. O. (1997) *Retroviruses* (Coffin, J. M., Hughes, S. H., and Varmus, H. E., eds), pp. 161–203, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
2. Bowerman, B., Brown, P. O., Bishop, J. M., and Varmus, H. E. (1989) *Genes Dev.* **3,** 469–478
3. Brown, P. O., Bowerman, B., Varmus, H. E., and Bishop, J. M. (1987) *Cell* **49,** 347–356

4. Fujiwara, T., and Mizuuchi, K. (1988) *Cell* **54,** 497–504
5. Farnet, C. M., and Haseltine, W. A. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 4164–4168
6. Ellison, V., Abrams, H., Roe, T., Lifson, J., and Brown, P. (1990) *J. Virol.* **64,** 2711–2715
7. Lee, Y. M. H., and Coffin, J. M. (1990) *J. Virol.* **64,** 5958–5965
8. Andrake, M. D., and Skalka, A. M. (1996) *J. Biol. Chem.* **271,** 19633–19636
9. Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R., and Davies, D. R. (1994) *Science* **266,** 1981–1986
10. Bujacz, G., Jaskolski, M., Alexandratos, J., Wlodower, A., Merkel, G., Katz, R. A., and Skalka, A. M. (1995) *J. Mol. Biol.* **253,** 336–346
11. Vora, A. C., McCord, M., Fitzgerald, M. L., Inman, R. B., and Grandgenett, D. P. (1994) *Nucleic Acids Res.* **22,** 4454–4461
12. Goodarzi, G., Im, G.-J., Brackmann, K., and Grandgenett, D. P. (1995) *J. Virol.* **69,** 6090–6097
13. Aiyar, A., Hindmarsh, P., Skalka, A. M., and Leis, J. (1996) *J. Virol.* **70,** 3571–3580
14. Miller, M. D., Farnet, C. M., and Bushman, F. D. (1997) *J. Virol.* **71,** 5382–5390
15. Smith, S. D., Shatsky, M., Cohen, P. S., Warnke, R., Link, M. P., and Glader, B. E. (1984) *Cancer Res.* **44,** 5657–5660
16. Salahuddin, S. Z., Markham, P. D., Wong-Staal, F., Franchini, G., Kalyanaraman, V. S., and Gallo, R. C. (1983) *Virology* **129,** 51–64
17. Miyoshi, I., Taguchi, H., Kubonishi, I., Yoshimoto, S., Ohrauki, Y., Shiraishi, Y., and Akagi, T. (1982) *Jpn. J. Cancer Res.* **28,** 219–228
18. Weiss, A., Wiskocil, R. L., and Stobo, J. D. (1984) *J. Immunol.* **133,** 123–128
19. Ross, E. K., Buckler-White, A. J., Rabson, A., Englund, G., and Martin, M. A. (1991) *J. Virol.* **65,** 4350–4358
20. Pear, W. S., Nolan, G. P., Scott, M. L., and Baltimore, D. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 8392–8396
21. Wu, X., Liu, H., Xiao, H., Conway, J. A., Hehl, E., Kalpana, G. V., Prasad, V., and Kappes, J. C. (1999) *J. Virol.* **73,** 2126–2135
22. Adachi, A., Gendelman, H. E., Koenig, S., Folks, T., Willey, R., Rabson, A., and Martin, M. A. (1986) *J. Virol.* **59,** 284–291
23. Engelman, A., Englund, G., Orenstein, J. M., Martin, M. A., and Craigie, R. (1995) *J. Virol.* **69,** 2729–2736
24. Jenkins, T. M., Esposito, D., Engelman, A., and Craigie, R. (1997) *EMBO J.* **16,** 6849–6859
25. Leavitt, A. D., Robles, G., Alesandro, N., and Varmus, H. E. (1996) *J. Virol.* **70,** 721–728
26. Ratner, L., Haseltine, W., Patarca, R., Livac, K. J., Starcich, B., Josephs, S. J., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, R. R., Jr., Pearson, M. L., Lauteenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C., and Wong-Staal, F. (1985) *Nature* **313,** 227–283
27. Zack, J. A., Arrigo, S. J., Weitsman, S. R., Go, A. S., Haislip, A., and Chen, I. S. Y. (1990) *Cell* **61,** 213–222
28. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
29. Wessel, D., and Flügge, U. I. (1984) *Anal. Biochem.* **138,** 141–143
30. Nilsen, B. M., Haugan, I. R., Berg, K., Olsen, L., Brown, P. O., and Helland, D. E. (1996) *J. Virol.* **70,** 1580–1587
31. Engelman, A., Liu, Y., Chen, H., Farzan, M., and Dyda, F. (1997) *J. Virol.* **71,** 3507–3514
32. Wei, S.-Q., Mizuuchi, K., and Craigie, R. (1997) *EMBO J.* **16,** 7511–7520
33. Engelman, A. (1999) *Adv. Virus Res.* **52,** 411–426
34. Gerton, J. L., Ohgi, S., Olsen, M., Derisi, J., and Brown, P. O. (1998) *J. Virol.* **72,** 5046–5055
35. Esposito, D., and Craigie, R. (1998) *EMBO J.* **17,** 5832–5843
36. Chen, H., and Engelman, A. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 15270–15274
37. Lavoie, B. D., and Chaconas, G. (1996) *Curr. Topics Microbiol. Immunol.* **204,** 83–102
38. Pruss, D., Bushman, F. D., and Wolffe, A. P. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 5913–5917
39. Wei, S.-Q., Mizuuchi, K., and Craigie, R. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 10535–10540
40. Engelman, A., Hickman, A. B., and Craigie, R. (1994) *J. Virol.* **68,** 5911–5917
41. Leavitt, A. D., Shuie, L., and Varmus, H. E. (1993) *J. Biol. Chem.* **268,** 2113–2119
42. Heuer, T. S., and Brown, P. O. (1997) *Biochemistry* **36,** 10655–10665

Multiple Integrase Functions Are Required to Form the Native Structure of the Human Immunodeficiency Virus Type I Intasome

Hongmin Chen, Shui-Qing Wei and Alan Engelman

J. Biol. Chem. 1999, 274:17358-17364.
doi: 10.1074/jbc.274.24.17358