Mol. Cells, Vol. 10, No. 5, pp. 525–532

Molecules and Cells

Characterization of Mutant HIV-1 Integrase Carrying Amino Acid Changes in the Catalytic Domain

Khampoune Sayasith[†,*], Gilles Sauvé, and Jocelyn Yelle

Armand-frappier Institute, Virology Research Center, Bldg 22, 531, des Prairies Bldv, Laval, Quebéc, Canada H7N 4Z3.

(Received on May 23, 2000)

To gain insight into the importance of conserved residues in the core domain of HIV-1 IN, we performed site-directed mutagenesis of the full-length enzyme, overexpressed the mutant proteins in *E. coli*, purified and analyzed their 3′-processing, integration and disintegration activities *in vitro*. Change of E152V in the DD(35)E motif abolished all detectable activities of IN. Alteration of two highly conserved residues, P145 and K156, by isoleucine, resulted in a substantial loss or completely abolished the three activities of the enzyme. Mutant P90D weakly reduced the 3′-processing but severely affected the two other IN activities. Results obtained from double and triple mutations, P90D/P145I and P145I/F185K/C280S, clearly suggest a crucial role of P145 in the catalytic function of IN, whereas the mutants V150E, L158F and L172M had no detectable effect on any of the IN activities. Taken together, these results allowed us to conclude that all the conserved amino acids in the core domain of IN are not equally important for catalytic functions: like D64, D116 and E152, our data suggest that P90, P145 and K156 are also essential for all three enzymatic activities of HIV-1 IN *in vitro*, whereas V150, L158 and L172 appear to be less critical.

Keywords: HIV-1 Integrase; Site-directed Mutagenesis; Integrase Catalytic Domain; Integrase Activity; Integrase Mutants.

Introduction

Integration of retroviral DNA into the host genome is an essential step for the replication cycle of human immunodeficiency virus type 1 (HIV-1) (Goff, 1992; Katz and Skalka, 1994). Following completion of reverse transcription *in vivo*, integrase (IN), a 32 kDa protein, performs a 3′-processing activity for removing GT nucleotides at each end of the viral LTR, leaving a 2-nucleotide 5′ end overhang and a 3′ end terminating with the phylogenetically conserved CA dinucleotide (Brown *et al*., 1989; Fujiwara and Mizuuchi, 1988; Roth *et al*., 1989). In the nucleus, HIV-1 IN integrates the 3′ end- processed viral DNA into the host genome during which IN makes a cut in the target DNA and covalently joins it with the viral DNA. This integration activity leaves a gap and two unpaired bases at the site of integration that could be repaired by cellular repair enzymes, resulting in a duplication of 5 bp which flanks the integrated proviral DNA (Chow *et al*., 1992; Farnet and Haseltine, 1990; Fujiwara and Mizuuchi, 1988; Kulkosky *et al*., 1992). *In vitro*, recombinant IN could reproduce these two activities using short oligonucleotides as a DNA substrate and a DNA target. In the reaction using a specific Y-shaped oligomer substrate, the purified IN could also catalyze the reverse of the integration or disintegration reaction to yield the recessed CA_{OH}-3′ end of viral DNA but its biological role *in vivo* is still unknown (Chow *et al*., 1992).

The full-length IN protein contains 288 amino acid residues and can be divided into three domains: the N-terminal zinc finger domain which seems to be required for the 3′-processing and integration activities

[†] Present address: CRRA, Faculty of Veterinary Medicine, University of Montreal, C.P. 5000, St-Hyacinthe, Québec, Canada J2S 7C6.

* To whom correspondence should be addressed.
Tel: 1-450-773-8521 Ext. 8428; Fax: 1-450-778-8103
E-mail: sayasik@yahoo.ca

Abbreviations: IN, integrase; HIV-1, human immunodeficiency virus type 1; RSV, Rous sarcoma virus; LTR, long terminal repeat; kDa, kilodalton; aa, amino acid; IPTG, isopropyl-1-thio-β-D-galactopyranoside; Hepes, *N*-2-hydroxyethylpiperazine-*N*-2-ethanesulfonic acid; MOPS, 3-[*N*-Morpholino]propanesulfonic acid; *E. coli*, *Escherichia coli*.

(Engelman and Craigie, 1992; Leavitt *et al*., 1993; van Gent *et al*., 1993); the C-terminal domain which is the least conserved region and has a major nonspecific DNA binding determinant of the enzyme (Bushman, *et al*., 1993; Engelman and Craigie, 1992; Leavitt *et al*., 1993; Lutzke, *et al*., 1995; van Gent *et al*., 1993; Woerner *et al*., 1992); and the core domain which contains three invariant residues (Asp64, Asp116, and Glu152), called the DD(35)E motif, which are critical for the enzyme activities *in vivo* and *in vitro* and are involved in the binding of IN to the divalent cations (Mn^{2+} or Mg^{2+}). Many studies have revealed details regarding the specificity of the DNA recognition and selection of integration sites in the core domain of IN. Individual expression of N- and C-termini suggests a limitation of the active site in the core domain (Engelman and Craigie, 1992; Engelman *et al*., 1995; Kulkosky *et al*., 1992; van Gent *et al*., 1993). IN is a highly hydrophobic protein, especially in the core domain containing D64, D116 and E152 residues, which seem to be close to one another in the active conformation, but the exact nature of their involvement in the enzymatic functions remains to be defined (Maignan *et al*., 1998). Adjacent residues to this region are also highly conserved and are believed to be important in the IN catalytic activities.

Presently, a lot of information has been gathered regarding the contribution of each of the IN regions and also some specific amino acids of the enzyme to its activities *in vitro*; however, little is known about the consequence of specific amino acid changes *in vivo*. Thus, it would be interesting to know which residues in the core domain are functionally involved in the enzymatic reaction. In order to gain more insight on the functional organization of HIV-1 IN, we first generated specific mutations in the catalytic domain of the full-length IN, neighboring the invariant residues, and analyzed their 3′-processing, integration and disintegration activities *in vitro*.

Materials and Methods

Cloning and site-directed mutagenesis The wild-type and soluble active mutant of recombinant plasmids, pINSD-His (Engelman and Craigie, 1992) and pINSD-His.sol (Jenkins *et al*., 1996), were kindly provided by Dr. Craigie (NIH). The full-length of the HIV-1 IN gene (strain IIIB) was PCR-amplified from the λHXB2 phage and cloned into the pET-16b expression vector (Novagen), generating the wild-type HIV-1 IN construction, called pET16-INHis. pET16-INHis was used for the production of various IN mutants. Site-directed mutations were introduced into the core domain of the IN coding region of the pET16-INHis construct by using the U.S.E. Mutagenesis Kit (Pharmacia Biotech), according to the manufacturer's instructions. All oligonucleotides were obtained from the Armand-Frappier Institute [Laval (Qc), Canada], purified on 20% denaturing polyacrylamide gels, and spectrophometrically quantified. The sequence of the oligonucleotide selection primer that converted the unique *PstI* site on the plasmid to *SacII* (bold) was PST: 5′ ACA CCA CGA TGC **CCG CGG** CAA TGG CAA CAA 3′. The sequences of the mutagenic oligonucleotides that, unless indicated, simultaneously introduced a specific site of the restriction enzyme and an amino acid change (bold) were as follow: E152V: 5′ AAG GAG TAG TA**G TCT CT**A TGA AT 3′ (*BsmaI*); V150E: 5′ AGT CAA GGA **GAA** GTA GAA TCT 3′ (no restriction site); K156I: 5′ TCT ATG **AAT ATT** GAA TTA AAG 3′ (*SspI*); P145I: 5′ TTC CCT AC**A ATA TT**C AAA GTC AAG G 3′ (*SspI*); L158F: 5′ AAT AAA GAA **TT**C AAG AAA ATT ATA 3′ (*Eco*RI); L172M: 5′ CAG GCT GAA **CAT ATG** AAG ACA GCA GTA 3′ (*NdeI*); P90D: 5′ AGA AGC AGA AGT **TAT CGA** TGC AGA AAC AGG GGC AGG 3′ (*ClaI*). The presence of mutation was confirmed by digestion with an appropriate restriction enzyme, and by dideoxy sequencing. The double mutants were also generated by the same strategy. P90D and P145I mutagenic oligonucleotides were used along with the PST selection primer to introduce the mutations P90D/P145I into the pET16-INHis vector during the synthesis process, while the P145I and PST primers and pINSD-His.sol vector were used for generating a triple mutation.

Protein expression and purification The 10xHis tagged HIV-1 IN protein was expressed in *Escherichia coli* strain BL21(DE3)pLysS and purified as described previously with minor modifications (Craigie *et al*., 1995). Briefly, bacteria were transformed with the construct and selected on agar plates containing 100 μg/ml ampicillin. The resulting colonies were screened for IN expression upon induction with 1.0 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at an $OD_{600\ nm}$ of 0.4 during 4 h at 37°C. All following steps were performed at 4°C. The cells were harvested and resuspended in a buffer A containing 25 mM Tris-HCl (pH 8.1), 1 M NaCl and 10% glycerol. Cells were lysed with lysozyme (1 mg/ml), sonicated, and cleared by a slow centrifugation. The lysates were collected by centrifugation at 20,000 × *g* for 1 h (Model JA-20). The supernatant (soluble IN fraction) was discarded. The pellet (insoluble IN fraction) was homogenized in the denaturing buffer B containing 25 mM HEPES (pH 8.1), 1 M NaCl and 6 M Gnd-HCl, stirred overnight and collected by centrifugation. A 2-ml chelating Sepharose Fast Flow column (Novagen) precharged with NiSO was used for purification of IN. The denaturing IN fraction was loaded twice and the column was washed with 30 ml of buffer B followed by three sequential washes with 30 ml of buffer B containing 5 mM, 20 mM and 60 mM imidazole. Proteins were eluted with 200 mM imidazole in buffer B and analyzed on a 12% SDS-PAGE. Fractions containing the IN were pooled, EDTA was added to a final concentration of 5 mM, and dialyzed for at least 12 h, with stirring, against buffer B (pH 7.5) containing 2 mM β-mercaptoethanol (β-me) and 2 mM EDTA. A second dialysis was then performed against the latter buffer containing 10 mM DTT instead of β-me. IN was sequentially refolded by dialysis against buffer C containing 25 mM HEPES (pH 7.5), 0.5 M NaCl, 1 mM DTT, 0.1 mM EDTA and 2 M urea, and buffer D containing 25 mM HEPES (pH 7.5), 1 M NaCl, 1 mM DTT, 1 mM EDTA, 10 mM CHAPS and 10% glycerol (wt/vol). For the soluble IN fraction, purification was performed as described above, except that no detergent was

added to the buffers. IN was identified by the Western blotting technique using rabbit anti-HIV-1 antiserum (NIH), and the protein concentration was estimated by SDS-PAGE and Coomassie Blue staining.

IN DNA substrate preparation Gel purified single-stranded oligonucleotides were labeled at their 5′ end with [γ-^{32}P]dATP using T4 polynucleotide kinase in a reaction volume of 20 µl. Following a 30 min incubation at 37°C, the reaction was stopped through addition of 25 mM EDTA and 100 mM NaCl in a volume of 50 µl, heated for 15 min, and purified through a G-25 Sephadex quick spin column. The labeled strand was then annealed to a 3-fold excess of the unlabeled complementary oligonucleotide by heating and slowly cooling to room temperature. The DNA substrate duplexes used in this study and their structure are shown in Table 1.

Assays of HIV-1 IN activities Unless otherwise indicated, the standard 20-µl reactions of 3′-processing and DNA integration activities were conducted in 25 mM MOPS (pH 7.2), 7.5 mM $MnCl_2$, 1 mM DTT, 50 mM NaCl, 10% glycerol (wt/vol), 10% DMSO, 0.1 mg/ml BSA, 1 pmol KD2020 substrate and 1 µl purified HIV-1 IN (400 nM). The reactions were initiated by addition of IN, incubated for 60 min at 37°C, stopped by adding an equal volume of loading buffer containing 95% formamide, 25 mM EDTA (pH 8.0), and 0.05% bromophenol blue and heated before resolving on a 20% denaturing polyacrylamide gel. Assays for the specific integration activity alone were carried out by using the KD1820 substrate containing a 5′ end-radiolabeled preprocessed strand, while disintegration reactions were performed with the Y-oligomer substrate (KD2012), which corresponds to the HIV-1 U5-LTR end-specific sequence in the viral portion of the integration intermediate product. Products of the reactions were visualized by autoradiography on wet gels at −70°C with Kodak X-Omat AR-2 films. The signal intensity was determined by densitometry (Bio-Rad Model GS-670 Imaging Densitometer).

Results

Generation and purification of HIV-1 IN proteins Several site-directed mutations in the core domain of the full-length IN were generated to assess the role of generally well-conserved residues among the retroviral IN and retrotransposase proteins (Engelman and Craigie, 1992; Kulkosky *et al.*, 1992). INs were expressed in *E. coli* as fusion 10xhistidine-tagged proteins by using a pET-16b vector. High levels of expression of the enzyme in the inclusion bodies were observed when induced with 1 mM IPTG (data not shown). All fusion proteins were purified using a nickel-chelated Sepharose column, analyzed on SDS-PAGE (Fig. 1) and confirmed by Western blots (data not shown). This method showed a homogeneous IN protein of more than 90% as judged by Coomassie Blue staining.

Effect of amino acid substitutions in the DD(35)E motif on HIV-1 IN endonucleolytic activity We analyzed the endonucleolytic activity of nine integrase mutants using a radiolabeled duplex DNA (20-mer) that mimics the U5 end of the viral DNA sequence. The specific 3′-processing removes a GT dinucleotide (Fig. 2), and the processed −2 products were analyzed on a 20% denaturing polyacrylamide gels and visualized by autoradiography. Table 2 presents the results of these

Table 1. Substrates for integrase.

Substrate designation	Oligonucleotide components	Structure[a]
KD2020	S50-20	5′ TGTGGAAAATCTCTAGCAGT -3′ (Δ)
		3′ - ACACCTTTTAGAGATCGTCA -5′ **KD2020**
	S05-20	
		5′ TGTGGAAAATCTCTAGCA -3′
		3′ - ACACCTTTTAGAGATCGTCA -5′ **KD1820**
KD1820	S50-18	
	S05-20	
KD2012	S05-D1	5′- ACTGCTAGAGAT*TTTCCACA-3′
	S05-12	3′- TGACGATCTCTAAAAGGTGT-5′
	S50-20	(Δ) **KD2012**

[a] Triangles and arrows indicate potential 3′-processing sites and disintegration sites, respectively. The asterisk in KD2012 represents a free 3′-OH group. Radiolabeling is indicated as a circle.

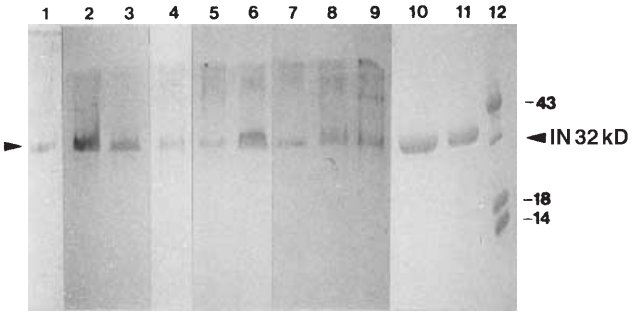

Fig. 1. SDS-PAGE analysis of wild-type and site-directed mutants of HIV-1 IN proteins. Transformed bacteria were induced for IN expression during 4 h at 37°C. IN was purified as described in the **Materials and Methods** section, analyzed by SDS-PAGE (12%) and revealed by Coomassie Blue staining: lane 1, K156I; lane 2, V150E; lane 3, E152V; lane 4, P145/F185K/280S; lane 5, P145I/P90D; lane 6, P90D; lane 7, L172M; lane 8, L158F; lane 9, P145I; lane 10, F185K/C280S; lane 11, wild-type IN; and lane 12, standard proteins. The arrow indicates the molecular weight of IN.

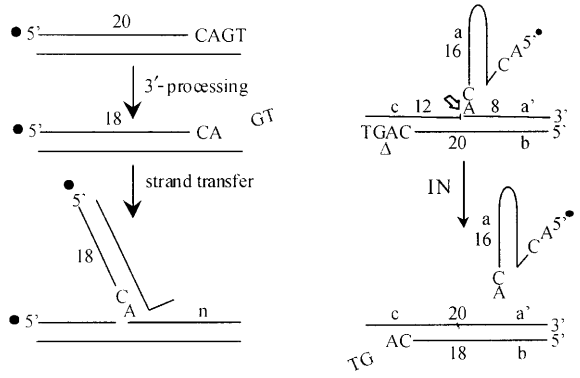

Fig. 2. *In vitro* activities of the retroviral integrase. (*Left*) Standard *in vitro* assays for 3′-processing and integration reactions. In the 3′-processing reaction, integrase cleaves the DNA substrate and removes the two terminal bases leaving the CA dinucleotide. During the integration process, integrase catalyzes the concerted and ligation reactions in which the recessed 3′ end of viral DNA is joined to the target DNA. Integrase can also catalyze an apparent reversal of the DNA integration, a process known as disintegration. The numbers refer to the oligonucleotides' length, and "*n*" varies in size depending on the site of integration. (*Right*) The disintegration reaction and products using the substrate KD2012 (a, a′, b, and c represent different segments of the substrate). The triangle and arrow point to the 3′-processing and disintegration sites, respectively.

assays. The 3′-processing reaction was efficiently observed with various concentrations of wild-type IN (Fig. 3A, lane 1; Fig. 3B and 3C, lanes 6 and 7; Fig. 3D, lane 1). Substitution of Glu152 with valine (E152V) resulted in a complete loss of the cleavage activity (Fig. 3A, lanes 2 and 3), while IN protein with a mutation at position Pro90, Val150, Leu158, or Leu172

Table 2. *In vitro* activities[a] of HIV-1 IN mutants.

Integrase mutants	*In vitro* activity		
	3′-processing	Integration	Disintegration
P90D	+ +	+/−	−
P145I	−	−	−
P145I/P90D	−	−	−
P145I/F185K/C280S	−	−	−
V150E	+ + +	+ +	+ + +
E152V	−	−	−
K156I	−	−	+/−
L158F	+ +	+ + +	+ +
L172M	+ + +	+ +	+ +

[a] + + +, 80–100% of WT IN activity; + +, 25–80% of WT IN activity; +, 5–25% of WT IN activity; +/−, < 5% of WT IN activity; − , no activity detected. Values are averages of at least two independent experiments.

(P90D, V150E, L158F, or L172M) was observed to be less detrimental for the same activity: approximately 10-fold and 20-fold reductions of this activity for L158F and P90D were noted, respectively, (Fig. 3B, compare lanes 1 and 3 with lane 6). Similar activity to the wild-type for L172M (Fig. 3D, lanes 6 and 7) and V150E mutants (Fig. 4) was also observed. A change of Pro145 or Lys156 (P145 or K156) with isoleucine (I) resulted in a protein with no processing activity above the background (Fig. 3B, lane 5, and Fig. 4A, lanes 11–13). P145 seems to be important for IN functions. Its effects were then investigated with the pINSD-His.sol vector containing a double mutation F185K/C280S, which possesses similar catalytic activities as the wild-type IN (Fig. 3D, lanes 2 and 3) and is useful for producing soluble IN (Jenkins *et al.*, 1996). The resulting IN mutant P145I/F185K/C280S showed a complete abolition of its cleavage activity (Fig. 3D, lanes 8 and 9). The double mutant P90D/P145I also caused a complete loss of the latter activity (Fig. 3D, lanes 4 and 5). Detection of the −1 product was somehow superior to the −2 products which probably reflects the presence of contaminating bacterial nucleases in this protein preparation. However, the pattern of cleaved products was also different from those exhibited by the wild-type IN (Fig. 3D, lane 1). The repeated experiment with the highly purified P90D/P145I mutant did not show these two products (data not shown). Taken together, these results indicate that only a few of the conserved amino acids are substantially involved in the IN endonucleolytic activity.

Effect of amino acid changes in the DD(35)E motif on the integration and disintegration activities We used the preprocessed substrate KD1820 for analyzing the DNA integration activity of IN mutants. This reaction is based

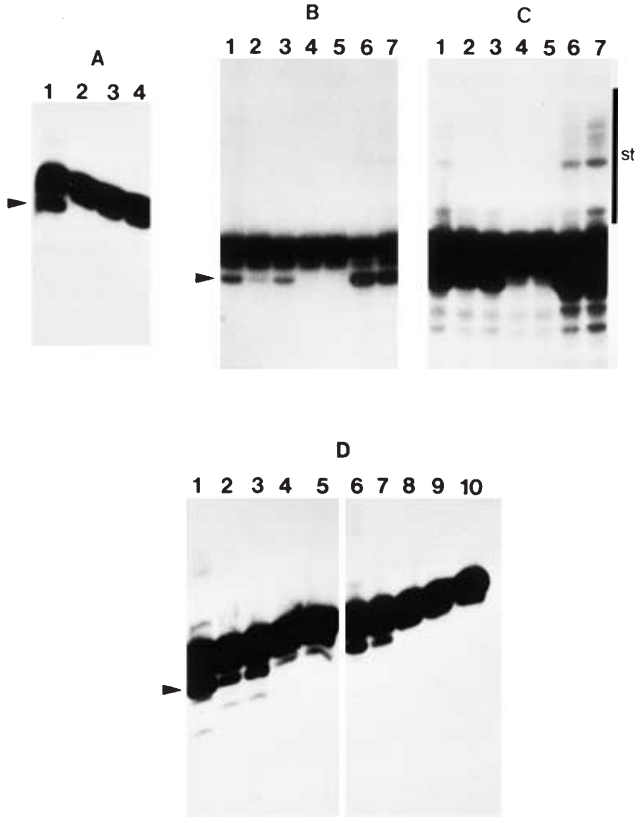

Fig. 3. Analysis of the 3′-processing activity of IN protein. The end-labeled KD2020 substrate was prepared as described in the **Materials and Methods** section. The standard reaction contained 1 pmole of DNA substrate, 25 μM MOPS (pH 7.2), 7.5 mM $MnCl_2$, 50 mM NaCl, 10% glycerol, 10% DMSO, 1 mg/ml BSA and various concentrations of IN. The reactions were incubated for 1 h at 37°C, stopped by addition of an equal volume of loading buffer and analyzed on a 20% denaturing polyacrylamide gel. The reaction products were visualized by autoradiography of the wet gel. **Panel A**: wild-type IN (400 nM), lane 1; E152V, lanes 2 (400 nM) and 3 (800 nM); and DNA substrate without IN, lane 4. **Panel B**: L158F, lanes 1 (400 nM) and 2 (800 nM); P90D lanes 3 (400 nM) and 4 (800 nM); P145I, lane 5 (400 nM); and wild-type IN, lanes 6 (400 nM) and 7 (800 nM). **Panel C**: same reactions as Panel **B** with a longer exposure time of the autodiographic film. **Panel D**: wild-type IN, lane 1 (400 nM); F185K/C280S, lanes 2 (400 nM) and 3 (800 nM); P90D/P145I, lanes 4 (400 nM) and 5 (800 nM); L172M, lanes 6 (800 nM) and 7 (400 nM); P145I/F185K/C280S, lanes 8 (800 nM) and 9 (400 nM); and DNA substrate without IN, lane 10. The films were exposed for 16 h without an intensifying screen, except for Panel **C**. The arrows and ST symbol indicate the 3′-processing and integration products, respectively.

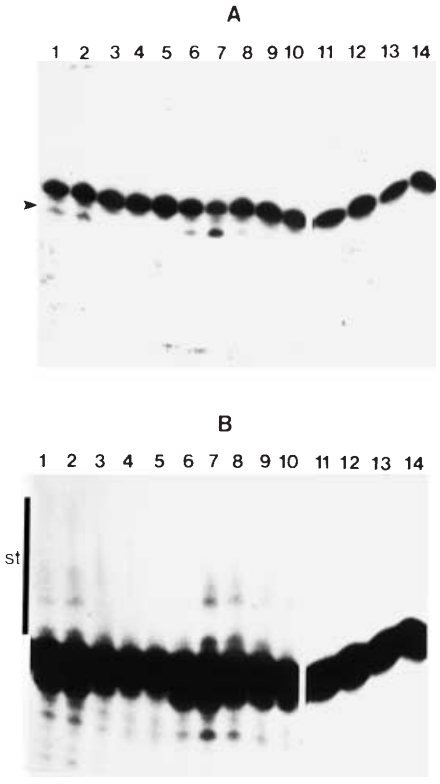

Fig. 4. Analysis of the 3′-processing activity of the mutants V150E and K156I. Reactions were performed in the same conditions as described in Fig. 3 with various concentrations of the mutants. **Panel A**: V150E, lanes 1 (1,2 μM), 2 (800 nM), 3 (400 nM), 4 (200 nM) and 5 (100 nM) wild-type IN, lanes 6 (1,2 μM), 7 (800 nM), 8 (400 nM), 9 (200 nM) and 10 (100 nM); K156I, lanes 11 (1,2 μM), 12 (800 nM) and 13 (400 nM); and DNA substrate without IN, lane 14. **Panel B**: same reactions as Panel **A** with a longer exposure time of the autoradiographic film. The arrows and ST symbol indicate the products of the 3′-processing integration reaction, respectively.

on the joining of the preprocessed DNA to other identical DNAs resulting in longer products than the starting DNA. Figure 5A shows the data obtained by these assays. No joined products were detected with the K156I, E152V and P145I mutant proteins (lanes 5, 6, and 9, respectively). The L172M and V150E mutants displayed approximately 25% of the joined products (lanes 3 and 4), while the L158F mutant conserved a comparable integration activity to that of the wild-type (lane 7). As detected in the endonucleolytic assays, P90D only showed 5% or less of the wild-type integration activity (lane 8). The double mutation F185K/C280S also resulted in the same product patterns as the wild-type. Therefore, the joining and 3′ end-nicking activities of P90D/P145I (lane 12) and P145I/F185K/C280S (lane 13) mutants were impaired. In our disintegration assay, the KD2012 substrate was constructed by hybridizing together the S05-D1, S20-12 and S05-20 oligonucleotides, generating a typical Y-shaped structure with an accessible disintegration site at the base of the hairpin and a 3′-endonucleolytic site at the end of the construct (Table 1). The disintegration and 3′-processing reactions catalyzed by HIV-1 IN on this substrate are represented in Fig. 1 (right panel). The 5′ end radiolabeling was introduced into the hairpin portion of the multiplex KD2012 substrate for monitoring only the IN disintegration

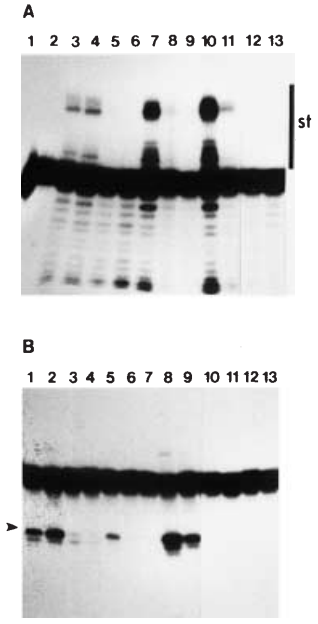

Fig. 5. Integration and disintegration activities of HIV-1 IN. **Panel A** (integration reaction): reactions were performed under similar conditions to those described in Fig. 3 using 1 pmole of preprocessed integration substrate (KD1820) and 400 nM IN. DNA substrate without IN, lanes 1 and 2; L172M, lane 3; V150E, lane 4; K156I, lane 5; E152V, lane 6; L158F, lane 7; P90D, lane 8; P145I, lane 9; wild-type IN, lane 10; F185K/C280S, lane 11; P90D/P145I, lane 12; and P145I/F185K/C280S, lane 13. **Panel B** (disintegration reaction): reactions were performed under similar conditions to those described in Fig. 3 using 1 pmole of disintegration substrate (KD2012) and 400 nM IN. L172M, lane 1; V150E, lane 2; K156I, lanes 3 (400 nM) and 4 (800 nM); L158F, lane 5; P90D, lane 6; P145I, lane 7; wild-type IN, lane 8; F185K/C280S, lane 9; P90D/P145I, lane 10; and P145I/F185K/C280S, lane 11; E152V, lane 12; and DNA substrate without IN, lane 13. The films were exposed for 16 h with an intensifying screen. The arrows and ST symbol indicate the disintegration and integration products, respectively.

activity, as shown by a band of the cleaved product (16-mer) (Fig. 5B). As indicated in this figure, the disintegration activity of wild-type IN was efficiently produced (lane 8), whereas this activity for the mutants K156I, P90D, P145I and E152V was severely affected (lanes 3, 4, 6, 7, and 12, respectively). Similar results were also observed with the multiple mutants P90D/P145I (lane 10) and P145I/F185K/C280S (lane 11). The mutants F185K/C280S and V150E also showed a comparable activity to that of the wild-type, while a significant decrease of this activity was noted for the mutants L172M and L158F: approximately a 5-fold and 8-fold reduction were recorded, respectively (lanes 1 and 5).

Discussion

We report the *in vitro* analysis of several HIV-1 IN mutants carrying amino acid changes at conserved residue positions close to the invariant residues in the core domain. In addition to their conservation and proximity to the invariant residues, our IN mutants were also chosen as a function of the literature data suggesting their potentially important role for the activities of IN (Engelman and Craigie, 1992). P90 and P145 were selected because of their location in the hydrophobic region and their essential role in the proteins refolding process. Val150 and Lys156 were targeted because of their position close to E152. Mutant E152V was used as a negative control of the wild-type IN. Leu158 and Leu172 were chosen because of their specific location along with V165 and V180 which could play an important role in the dimerization of IN. We generally produced the major amino acid changes for increasing the risk to induce important local modifications of the enzyme's structure, which could result in the loss of IN activities. All IN proteins were obtained without great difficulty and were analyzed for their activities with both fractions: soluble and inclusion bodies.

Consequence of conserved residue changes on IN activities In agreement with the literature data, our results displayed a critical role of E152 in the enzymatic functions of IN. Although, a change of this residue could alter the recognition and specificity of the protein, affecting all activities of the enzyme. According to the hypothesis advanced by others, the E152V mutant could destabilize the IN-substrate binding (Gerton *et al.*, 1998; Heuer and Brown, 1998). The P90D, P145I, or K156I mutant also severely affected the catalytic activities of IN. Significant diminution of these activities observed with the P90D mutant suggests that this proline may reside in an essential region of the active site. Tridimensional structure analysis showed the location of P90 at the beginning of the α-helix, which could possibly be deformed by such a mutation (van den Ent *et al.*, 1998), therefore affecting the recognition of the viral DNA, the catalytic mechanism of IN itself, or its multimerization process. In fact, it is likely that substitution of a neutral residue by a charged one affects, to a lesser extent, the protein structure, and allows the protein to retain partial catalytic activities. Although similar results were obtained with P109A and P109S (Drelich *et al.*, 1992), more work is required to clarify these observations. In addition, close to P90, only E92 substitution was reported and showed normal activity of IN for E92A and E92Q and a marked decrease of the 3′-processing and integration but the intact disintegration activities for E92K (Pommier *et al.*, 1997). Taken together, these data allowed us to suggest that P90 could play a more important role than E92 in the IN activities.

Pro145 is found in a well-conserved segment (residues 140 to 149) of retroviral IN (Heuer and Brown

1998). Close to P145, the mutants Y143N and Y143F were reported and showed a delay of the virion replication, whereas the S147I mutant appeared to be unable to replicate normally in cellular culture (Pommier *et al.*, 1997). These results suggest that the integrity of this segment in the core domain seems to be essential for the infectivity of virions. Thus, our change of P145 could alter this integrity, possibly affecting the hydrogen bonds between the adjacent residues of E152, and therefore decrease the IN activities. Analysis of the multiple mutants showing a total loss of the wild-type integrase activities also supported a critical role of P145 for the IN enzymatic functions. Taken together, our results clearly support the fundamental role of P145 in the catalytic functions of IN. Thus, it becomes extremely interesting to further study this amino acid for a better understanding of its exact role in the active site of the enzyme, particularly the recognition and specificity of IN to its DNA substrate.

In agreement with others (Jenkins *et al.*, 1997), the K156I mutant also showed a complete abolition of IN 3′-processing and integration but conserved its residual disintegration activity. In this particular case, the IN mutant K156A caused a significant diminution of the infectivity of the corresponding viral particles, whereas K159Q did not markedly affect the viral replication (Stevenson *et al.*, 1990; Wiskerchen and Muesing, 1995). However, for the full-length IN, the mutant K156E showed a 4-fold reduction of disintegration activity when compared to that of the wild-type (Jenkins *et al.*, 1997). Through the detection of the residual disintegration activity of mutant K156I, our results support this evidence and seem to experimentally confirm the hypothesis suggesting the essential role of K156 in the IN-DNA substrate interaction. Our data are also in agreement with the results obtained by other groups, who reported the possible location of the specificity and the recognition of viral DNA in the core domain of IN (Goulaouic and Chow, 1996; Kulkosky *et al.*, 1992; Pahl and Flüel, 1995).

Despite the mutant's proximity to K156 and K159, there was no effect of the substitution of L158 with phenylalanine on the IN functions. Similar results were observed with L172M. Contrary to our first thoughts for the implication of L158 and L172 in a dimerization process of the protein, our results indirectly suggest that these two leucines either are not part of the process or do not play an important role in the enzymatic functions. Mutation V150E caused a minimal loss of IN activities. Mutant V151A also showed a significant reduction of infectivity, whereas the virions containing the IN mutation S153A replicated normally (Pommier *et al.*, 1997). Based on these observations, we suggest that residues clustered around E152 are not easily replaceable without the loss of IN activity or virus infectivity. So, the exact nature of the residue at position 150 appears less critical in this context.

As shown by others, no binding or cross-linking experiment has been described in which the effect of IN mutations in the DNA substrate correlate well with their effects on the catalytic activity using the U5 viral DNA end as a model (Drelich *et al.*, 1993; Engelman *et al.*, 1994; Gerton *et al.*, 1998). Since Engelman and his coworkers had shown the cross-linking of D64N, D116N, and E152 Q mutants of IN to the model U5 viral DNA end, they suggested that these residues in the wild-type protein did not contribute to the non-specific DNA binding. Using a similar cross-linking assay, another group reported a seemingly paradoxical result: since D116N and D116A mutants of IN did not cross-link to a model U5 viral DNA end, but D64A and E152A mutants did, this suggests that D116 was involved in interacting with the viral DNA, but D64 and E152 were not (Drelich *et al.*, 1992). In our experiment, all mutants of IN, including E152V, were also analyzed for binding of the U5 viral DNA end and the disintegration substrate by a UV cross-linking assay. The results showed that all of them retained DNA binding activity (data not shown). In agreement with other reports (Gerton *et al.*, 1998), our results also demonstrated that the level of binding activity of our mutant proteins had no consistent correlation with the catalytic activity of the protein (data not shown).

In conclusion, our results are consistent with other reports suggesting that not all the amino acids conserved in the core domain of IN are equally important for catalytic reactions (Leavitt *et al.*, 1993). In accord with other studies, we showed that the residues P90, P145, and K156, like D64, D116 and E152 are essential for the activities of HIV-1 IN *in vitro*. Our results also support the hypothesis of a single active site localized in the core domain of the enzyme. The identification of residues having a critical role in the recognition of the DNA substrate and the specificity of IN to its target DNA remains of capital importance for the development of drugs directed against this enzyme. The loss of IN specificity through the action of antiviral agents could result in an inappropriate cleavage of the DNA intermediate during the life cycle of the virus, preventing the integration of the viral genome into the host DNA of infected cells and impeding the formation of new infectious virions.

Acknowledgments The authors wish to thank Dr. Robert Craigie for helpful discussions on HIV-1 IN activities, and his generosity in providing the plasmids pINSD-His and pINSD-His.sol, and to Dr. François Shareck for his supervision in the preparation of the oligonucleotides and his advice on the site-directed mutagenesis techniques. This work was supported by grants from the MRC of Canada.

References

Brown, P. O., Bowerman, B., Varmus, H. E., and Bishop, J. M. (1989) Retroviral integration: structure of initial covalent product and its precursor, and a role for the viral IN protein. *Proc. Natl. Acad. Sci. USA* **86**, 2525–2529.

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.

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.

Craigie, R., Hickman, A. B., and Engelman, A. (1995) Integrase; in *HIV*, *Volume II: A practical approach*, J. Karn (ed), pp. 53–71, Oxford University Press.

Drelich, M., Wilhelm, R., and Mous, J. (1992) Identification of amino acid residues critical forendonuclease and integration activities of HIV-1 IN protein *in vitro*. *Virology* **188**, 459–468.

Engelman, A. and Craigie, R. (1992) Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function *in vitro*. *J. Virol*. **66**, 6361–6169.

Engelman, A., Englund, G., Orenstein, J. M., Martin, M. A., and Craigie, R. (1995) Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication. *J. Virol*. **69**, 2729–2736.

Farnet, C. and Haseltine, W. A. (1990) Integration of human immunodeficiency virus type 1(HIV-1) DNA *in vitro*. *Proc. Natl. Acad. Sci. USA* **87**, 4164–4168.

Fujiwara, T. and Mizuuchi, K. (1988) Retroviral DNA integration: structure of an integration intermediate. *Cell* **54**, 497–504.

Gerton, J. L., Ohgi, S., Olsen, M., DeRisi, J., and Brown, P. O. (1998) Effect of mutations in residues near the active site of human immunodeficiency virus type 1 integrase on specifc enzyme-substrate interaction. *J. Virol*. **72**, 5046–5055.

Goff, S. P. (1992) Genetic of retroviral integration. *Annu. Rev. Genet*. **26**, 527–544.

Goulaouic, H. and Chow, S. A. (1996) Directed integration of viral DNA mediated by fusion protein consisting of human immunodeficiency virus type-1 integrase and *Escherichia coli* LexA protein. *J. Virol*. **73**, 37–46.

Heuer, T. S. and Brown, P. O. (1998) Photo-cross-linking studies suggest a model for the architecture of an active human immunodeficiency virus type 1 integrase-DNA-complex. *Biochemistry* **37**, 6667–6678.

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

Jenkins, T. M., Esposito, D., Engelman, A., and Craigie, R. (1997) Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking. *EMBO J*. **16**, 6849–6859.

Katz, R. A. and Skalka, A. M. (1994) The retroviral enzymes. *Annu. Rev. Biochem*. **63**, 133–173.

Kulkosky, J., Jones, K. S., Katz, R. A., Mack, J. P. G., 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 transposase. *Mol. Cell. Biol*. **12**, 2331–2338.

Leavitt, A. D., Shuie, L., and Varmus, H. E. (1993) Site-directed mutagenesis of HIV-1 integrase demonstrated differential effects on integrase function *in vitro*. *J. Biol. Chem*. **268**, 2113–2119.

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

Maignan, S., Guilloteau, J. P., Zhou-Liu, Q., and Mikol, V. (1998) Crystal structure of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: high level of similarity of the active site with other viral integrases. *J. Mol. Biol*. **282**, 359–368.

Pommier, Y., Pilan, A. A., Bajaj, K., Mazumder, A., and Neamati, N. (1997) HIV-1 integrase as a target for antiviral drugs. *Antiviral Chemistry & Chemotherapy* **8**, 463–483.

Pahl, A. and Flügel, R. M. (1995) Characterization of human spuma retrovirus integrase by site-directed mutagenesis, by complementation, and by swapping the zinc finger domain of HIV-1. *J. Biol. Chem*. **270**, 2857–2966.

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

Stevenson, M., Haggerty, S., Lamonica, C. A., Meiter, C. M., Welch, S. K., and Wasiak, A. J. (1990) Integration is not necessary for expression of human immunodeficiency virus type 1 protein products. *J. Virol*. **64**, 2121–2125.

van den Ent, F. M., Vos, A., and Plasterk, R. H. (1998) Mutational scan of the human immunodeficiency virus type 2 integrase protein. *J. Virol*. **72**, 3916–3924.

van Gent, D. C., Vink, C., Oude Groeneger, A. A. M., and Plasterk, R. H. A. (1993) Complementation between HIV integrase proteins mutated in different domains. *EMBO J*. **12**, 3261–3267.

Wiskerchen, M. and Muesing, M.A. (1995) Human immunodeficiency virus type 1 integrase: effects of mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA templates, and sustain viral propagation in primary cells. *J. Virol*. **69**, 376–386.

Woerner, A. M., Klutch, M., Levin, J. G., and Marcus-Sekura, C. J. (1992) Localization of HIV-1 integrase to the C-terminal half of the protein. *AIDS Res. Hum. Retroviruses* **8**, 297–304.