

# Identification of a negative regulatory mechanism for the repair of U5 long terminal repeat DNA by the human immunodeficiency virus type 1 integrase DNA polymerase

Research Article

Brian E. Udashkin, Andrea Acel, Avi Shtvi, Benjamin Alt, Henry Triller, Mark A. Wainberg<sup>2</sup> and Emmanuel A. Faust<sup>1\*</sup>

Lady Davis Institute for Medical Research, SMBD-Jewish General Hospital and Departments of Medicine<sup>1</sup> and Microbiology and Immunology<sup>2</sup>, and McGill AIDS Center, McGill University, Montreal, Quebec, Canada H3T 1E2

---

\*Corresponding author: Tel: (514) 340-8260; Fax: (514) 340-7502; E-mail: mdef@musica.mcgill.ca

Received: 2 October 1998; accepted: 13 November 1998 (Communicated by Allan Wolffe)

## Summary

The quasi-random integration of retroviral DNA elements into the chromosomes of infected cells is believed to proceed by a four-step mechanism. The 3'- ends of the long terminal repeats (LTRs) are processed by the endonucleolytic cleavage and removal, usually, of a GT dinucleotide (step 1); the 3'-processed DNA ends are inserted at staggered nicks in the host DNA via a DNA strand transfer reaction, simultaneously generating short gaps at the sites of insertion (step 2); the gaps are repaired by a DNA polymerase (step 3); and the 5'-ends of the viral DNA are joined to the host DNA (step 4). Human immunodeficiency virus type 1 (HIV-1) integrase was previously reported to possess enzymatic activities capable of performing at least the first 3 steps of the integration process including an intrinsic DNA - dependent polymerase activity capable of short gap repair (Acel et al., 1998 J. Virol. 72: 2062-2071). In the present study, the behavior of the integrase DNA polymerase was examined in a DNA end-repair assay in which the frequency of polymerization on 3'-processed HIV-1 U5 LTRs was examined. The frequency of polymerization was negatively regulated by the 5'-AC sequence comprising the 2-nucleotide template and by the sequence of the adjacent conserved 5'TG/CA dinucleotide. Mutations within these DNA elements of the LTR enhanced the polymerization frequency on 2-nucleotide templates between 3- and 100-fold. In most cases, the integrase DNA polymerase added only one nucleotide to 3'-processed LTRs even though the DNA was comprised of 2-nucleotide template-primers. This level of regulation was controlled by a DNA binding and/or zinc finger domain in the integrase protein. By contrast, the integrase DNA polymerase behaved in a processive manner with homopolymeric pyrimidine templates, extending nascent DNA chains up to at least 20 nucleotides, whereas DNA polymerization with an oligo dA template exhibited a lower processivity of 1-7 nucleotides. The results suggest a model whereby an interaction between integrase and specific DNA elements in the HIV-1 LTRs prevents the repair of 3'-processed LTRs by the integrase DNA polymerase. Drugs with the ability to alter this regulatory aspect of integrase DNA polymerase function and thus induce repair of processed LTRs are predicted to block integration of HIV-1 DNA, and thus have potentially lethal consequences for HIV-1 replication.

## I. Introduction

The success of gene therapy in treating disease will ultimately depend on the ability of the target tissue to integrate foreign DNA into its chromosomes and therefore

model systems for the study of the mechanisms of repair and recombination in mammalian cells are needed to provide a means to study DNA integration processes. Retroviruses provide accessible model systems to study the integration of DNA elements into mammalian

chromosomes. The process of integration, whereby a double stranded DNA copy of a retroviral genome is joined to host cell DNA, is typified in the human immunodeficiency virus type1 (HIV-1) replicative cycle (Ansari-Lari et al., 1995; Cannon et al., 1994; Cara et al., 1995; Engelman et al., 1995; Englund et al., 1995; Lafemina et al., 1992; Sakai et al., 1993; Shin et al., 1994; Taddeo et al., 1994; Wiskerchen and Muesing, 1995a; Wiskerchen and Muesing, 1995b). In this process, two nucleotides are cleaved from the ends of the U3 and U5 portions of the HIV-1 long terminal repeats (LTRs) by the HIV-1 encoded integrase protein (Engelman et al., 1991; Tramontano et al., 1998; Vink, 1991; Vink et al., 1991). Cleavage occurs next to an invariant CA dinucleotide within each copy of the LTR, thus creating recessed 3' ends and 5' overhangs (Bushman and Craigie, 1991; LaFemina et al., 1991; Leavitt et al., 1992; Sherman and Fyfe, 1990; Vink et al., 1991). The 3' ends of the cleaved preintegration intermediate are then joined pairwise in a concerted manner to the host cell DNA (Brown, 1990; Craigie, 1992; Grandgenett and Mumm, 1990). The mechanism of DNA strand transfer, sometimes referred to as 3' end joining, involves a nucleophilic attack of adenylate hydroxyls on phosphodiester bonds located 5 base pairs apart on either strand of the host DNA. This concerted transesterification reaction produces an intermediate in which the viral genome is flanked by 5-nucleotide gaps and is linked to the host DNA by its 3' ends (**Fig. 1**). Gap repair and removal of adjoining mispaired nucleotides (the so-called 5'-joining reaction) may be carried out either by the integrase itself (Acel et al., 1998) or by host cell DNA repair enzymes, although experimental evidence to support the latter notion is lacking (Roe et al., 1997). Characteristic 5-bp repeats flank the inserted provirus (Ellison et al., 1990; Vink et al., 1990).

Cleavage and integration reactions are catalyzed *in vitro* by purified recombinant HIV-1 integrase produced in bacteria (Asante-Appiah and Skalaka, 1997). The *in vitro* reactions utilize synthetic oligonucleotide substrates homologous to terminal portions of the U3 and U5 regions of the HIV-1 LTR (Bushman and Craigie, 1991; Bushman et al., 1990; Carteau et al., 1993; LaFemina et al., 1991; Sherman and Fyfe, 1990; Vincent et al., 1993; Vink and Plasterk, 1993; Vink et al., 1991). Integrase also catalyzes bimolecular disintegration reactions using cross bone substrates indicating that integrase forms dimeric structures that are able to coordinate two substrate molecules perhaps by utilizing two appropriately positioned DNA binding sites (Chow and Brown, 1994a; Chow and Brown, 1994b; Mazumder et al., 1994). The crystal structure of a central catalytic core domain (amino acid residues 50-212) confirms the existence of dimeric contacts (**Fig. 2**) (Dyda et al., 1994; Goldgur et al., 1998; Maignan et al., 1998).

There is a sequence requirement for approximately 5 nucleotides in line to the CA dinucleotide at which cleavage occurs (Balakrishnan and Jonsson, 1997; Katzman et al., 1989; LaFemina et al., 1991; Vink et al., 1991). Cleavage of LTR sequences is specific for the homologous IN protein so that HIV-1 IN does not cleave the LTR sequences of either the avian sarcoma virus or the murine leukemia virus (Katzman et al., 1989; LaFemina et al., 1991).

Purified integrase exists as a homodimer in solution (Vincent et al., 1993) but likely functions as a multimer in which individual subunits provide separate functions for DNA binding, cleavage and integration (Engelman et al., 1993; Kalpana and Goff, 1993; van Gent et al., 1993).

**Fig. 1. Schematic illustration of a gapped retroviral DNA integration intermediate.** The diagram depicts the structure at the ends of the retroviral U3 and U5 LTRs after the 3'-ends of the viral DNA have been joined to host cell DNA. Also shown are 5 -

nucleotide gaps and unpaired 5' -AC tails, representative of HIV-1. In HIV-1 LTRs the base pair adjacent to the unpaired 5'-AC tails is 5' - TG/CA.

**Fig. 2. Ribbon diagram of the 3-dimensional structure of the HIV-1 integrase catalytic core domain (Dyda et al., 1994).** The illustration depicts a dimeric structure in which aspartate and glutamate residues (D64, D116 and E152) form a pocket believed to be required for catalysis. The 3-dimensional structure has features in common with RNase H and other polynucleotidyl transferases.

Dimeric, tetrameric and oligomeric forms of avian sarcoma virus integrase have been detected by crosslinking experiments and by gel exclusion chromatography (Andrake and Skalka, 1995). Only multimeric forms of HIV-1 integrase display DNA polymerase activity (Acel et al., 1998). Specific regions of integrase required for multimerization have been localized to the catalytic and C-terminal domains (Donzella et al., 1998). *In vivo* integrase may be part of a preintegration complex in which it interacts with other viral and cellular proteins (Carteau et al., 1997; Miller et al., 1997) .

Detailed analysis of the effects of a variety of mutations in integrase on the *in vitro* cleavage and integration reactions has allowed the definition of distinct protein domains (Vink and Plasterk, 1993). Separate N- and C-terminal domains in the integrase molecule are responsible for DNA binding. The N-terminal domain contains an HHCC region resembling a zinc finger motif

(Bushman et al., 1993; Cai et al., 1997; Woerner and Marcus-Sekura, 1993) which may play a role in positioning integrase on the viral DNA together with the C-terminal nonspecific DNA binding domain (Engelman et al., 1993; Vincent et al., 1993; Vink et al., 1993; Woerner et al., 1992; Woerner and Marcus-Sekura, 1993). A role for the zinc finger motif in the cleavage reaction has also been recognized (Asante-Appiah and Skalka, 1997; Engelman and Craigie, 1992; Khan et al., 1991; Lafemina et al., 1992; Van Gent et al., 1992). Binding of zinc ions to the N-terminal domain promotes tetramerization of full length integrase (Lee et al., 1997). A highly conserved central domain (amino acids 50-212) contains the catalytic center upon which all cleavage and strand transfer reactions are dependent (Skalka, 1993). Indeed, there is an absolute requirement in catalysis for conserved Asp64, Asp116 and Glu152 residues which make up the so called D,D,(35)E region (Bushman et al.,

1993; Drelich et al., 1993; Drelich et al., 1992; Engelman and Craigie, 1992; Kulkosky et al., 1992; Lafemina et al., 1992; Leavitt et al., 1993). These residues form a pocket which likely binds divalent cations.

In the present study we sought to define in greater detail the role of the integrase DNA polymerase in the integration process as exemplified by HIV-1. In terms of the current models for integration, it is assumed that multimeric forms of integrase are assembled at the ends of the viral DNA where they carry out the enzymatic reactions that are necessary to effect the integration process. In considering this model, it became clear that a mechanism would have to exist to regulate the repair of the viral DNA ends once they had been processed by the endonuclease activity of integrase. Model DNA substrates representing processed viral DNA ends were therefore used to examine the behavior of the integrase DNA polymerase to determine if there exists a regulatory mechanism governed strictly by the structure of HIV-1 integrase and/or the DNA sequence of the HIV-1 U5 LTR.

## II. Results

### A. Repair of a 3'-processed HIV-1 U5 LTR terminus by the integrase DNA polymerase is dependent on template sequence

Template primers consisted of a U5 duplex DNA segment attached to a 2-nucleotide 5'-tail such as would exist in a DNA end that had undergone 3'-processing by integrase (**Fig. 3A**). All 16 2-nucleotide combinations were tested as substrates for the integrase DNA polymerase. Repair reactions were conducted in the presence of all 4 dNTPs and one radiolabeled dNTP. The latter was complementary to the first base next to the primer terminus and varied according to the sequence of the template - primer. Thus, we could ensure that if only the first base was utilized as a template by the integrase DNA polymerase, which was the case in many instances (see below), a radiolabeled product would still be produced. The results indicated that repair is controlled by the sequence of the DNA template. The frequency of polymerization observed for 13 out of the 16 template sequence combinations examined was increased up to 4-fold relative to the wild type sequence 5'-AC. (**Fig. 4, Table 1**).

Analysis of the various reaction products by single-nucleotide resolution PAGE revealed that in many cases the integrase DNA polymerase appeared to add only a single nucleotide to the 2-nucleotide template (**Fig. 4**). In some cases the second nucleotide was added as well, but in these cases the enzyme paused frequently after the first nucleotide, yielding a doublet in the PAGE analysis (**Fig. 4, lanes 8, 10, 12, 14, 15, 16**). By contrast to this behavior of the integrase DNA polymerase, the Klenow fragment of *E. coli* DNA polymerase I and the HIV-1 reverse

transcriptase copied both nucleotides almost quantitatively regardless of sequence (see below). It seemed unlikely that these results were due in part to the action of the integrase mediated 3'-processing activity since we have also observed this behavior of the integrase DNA polymerase with DNA substrates lacking the conserved TG/CA dinucleotide known to be necessary for the processing reaction (see below).

**Fig. 3. Diagram of hairpin 2-nucleotide template-primers used as substrates for DNA repair.** The hairpin stem region designated as U5 in the diagram is 30 nucleotides in length including the two 5' -TG/CA terminal base pairs and its sequence is homologous to the terminal 30 nucleotides of the HIV-1 U5 LTR (see Materials and Methods). The diagram also indicates how the template sequences (A) or the internal dinucleotide sequences (B) were varied to produce different substrates for end repair reactions.

### B. Effect of mutations in the conserved 5'-TG/CA dinucleotide on the repair of 3'-processed U5 LTRs

All template primers used in this series of experiments had a 5'-AC unpaired tail and differed in the sequence of the adjacent conserved 5' -TG/CA dinucleotide (**Fig. 3B**). Repair by integrase DNA polymerase was conducted in the presence of all four dNTPs and radiolabeled dGTP. Thus, the first incorporated nucleotide would be radiolabeled in all DNA products. Mutations in the wild type TG/CA conserved dinucleotide had profound effects on the polymerization frequency occurring on the same 5'-AC unpaired tail (**Fig. 5, Table 2**). The wild type

sequence 5'-ACTG exhibited a polymerization frequency that was between 3 and 100-fold lower than that obtained with 14 other sequences. The greatest polymerization frequency was obtained with the template-primer bearing the sequence, 5'-ACAA (Fig. 5, lane 9). Only one

mutation, 5'-ACGC, reduced the polymerization frequency relative to the wild type sequence ACTG (Fig. 5, lanes 1 and 7). Thus, the overwhelming overall effect of mutations in the conserved

**Fig. 4. Effect of template sequence on end repair by the wild type integrase DNA polymerase.**

Reaction mixtures (described in Materials and Methods) contained 4dNTPs and either [ $\alpha$ - $^{32}$ P] dCTP (lanes 1-4), [ $\alpha$ - $^{32}$ P] dGTP (lanes 5-8), [ $\alpha$ - $^{32}$ P]TTP (lanes 9-12) or [ $\alpha$ - $^{32}$ P] dATP (lanes 13-16). The template sequence of each 2-nucleotide template-primer is given at the top of each lane.

The numbers to the right of the Fig. refer respectively to the incorporation of either 1 (+1) or 2 (+2) nucleotides into the DNA substrate.

**Fig. 5 Effect of changes in the conserved dinucleotide on end repair by the wild type integrase DNA polymerase.** The reaction mixtures as described in Materials and Methods contained all 4dNTPs and [ $\alpha$ - $^{32}$ P] dGTP as the only radiolabeled nucleotide. The boxed 5' -AC at the top left of each section of the Fig. represents the 2-nucleotide template sequence in each of the 16 different template-primers. The sequences of the dinucleotides adjacent to the 5'-AC tail are given at the top of each lane of the gel. The complementary nucleotides to each of these sequences (not shown) were also included in the oligonucleotide template-primers. The number (+1)

to the right of the Fig. refers to the incorporation of 1 nucleotide into the 32 -nucleotide template-primers.

**Table 1.** Relative frequency of polymerization by integrase DNA polymerase using 2-nucleotide NNTG templates<sup>a,b</sup>

Template sequence 5'→3'	AG	TG	GG	CG	AC	TC	GC	CC	AA	TA	GA	CA	AT	TT	GT	CT
Polymerization frequency <sup>c</sup>	0.5	2.7	3.4	1.0	1.0	2.8	1.9	1.2	1.6	4.2	0.1	2.2	3.8	2.3	3.0	1.0
Polymerization frequency <sup>d</sup>	0.2	1.0	2.2	0.3	1.0	3.1	1.4	2.8	0.9	4.0	0.7	1.7	2.4	1.6	0.9	0.6

<sup>a</sup>Phosphorimager adjusted volumes normalized to 1.0; <sup>b</sup>DNA concentration - 0.02 µg/ml; <sup>c</sup>Wild type; <sup>d</sup>Core domain.

**Table 2.** Relative frequency of polymerization by integrase DNA polymerase using 2-nucleotide ACN'N' templates<sup>a,b</sup>

Template sequence 5'→3'	AC AG	AC TG	AC GG	AC CG	AC AC	AC TC	AC GC	AC CC	AC AA	AC TA	AC GA	AC CA	AC AT	AC TT	AC GT	AC CT
Polymerization frequency <sup>c</sup>	5.5	1.0	6.7	9.0	25	9.7	0.2	10	100	12	8.7	9.5	18	12	3.3	17
Polymerization frequency <sup>d</sup>	1.2	1.0	1.4	1.4	2.6	1.6	.01	1.8	5.9	1.0	1.0	0.1	2.0	1.0	1.1	1.9

<sup>a</sup>Phosphorimager adjusted volumes normalized to wild type = 1.0; <sup>b</sup>DNA concentration - 0.02 µg/ml; <sup>c</sup>Wild type; <sup>d</sup>Core Domain

TG/CA dinucleotide was to enhance the polymerization frequency at the 5'-AC unpaired tail by as much as 100-fold. Despite this significant increase in the polymerization frequency, only the first template nucleotide was consistently copied in every case.

### C. The integrase DNA polymerase is processive with homopolymeric pyrimidine templates

Next, the integrase DNA polymerase was incubated with a homopolymeric template-primer in which the primer stem consisted of a nonviral sequence and the template was oligo dC<sub>21</sub> (Fig. 6, lanes 3 and 4). The reaction products constituted a heterogeneous mixture of radiolabeled DNA molecules that formed a ladder after gel electrophoresis whose pattern was consistent with the addition to the primer terminus of between 1 and 21 deoxyguanylate residues. Although there was clear evidence of pausing at many positions, DNA products representing the addition of between 17 and 20 nucleotides to the DNA primer terminus predominated. When the template contained adenylate residues interspersed among

oligo dC at 7 nucleotide intervals the integrase DNA polymerase appeared to pause less frequently before reaching the end of the template (Fig. 6, lane 2). Integrase DNA polymerase activity with an oligo dT<sub>20</sub> template, primed from an HIV-1 U5 LTR hairpin primer, likewise supported the synthesis of relatively long 19-20 oligonucleotide DNA chains and the size distribution of the DNA products remained the same over an 8-fold range of enzyme concentration indicating that the reaction mechanism was processive (data not shown). Template pyrimidines therefore can be said to support DNA chain elongation largely by a processive mechanism. The final nucleotide of these homopolymeric templates was not utilized efficiently in either case, as reported previously (Acel et al., 1998), so that polymerization with relatively short DNA templates may be governed by different mechanisms as mentioned earlier. The results show clearly, that in contrast to the general inability to polymerize both nucleotides of a 2-nucleotide template, the integrase DNA polymerase is capable of polymerizing relatively longer DNA chains when presented with the appropriate template-primer.

#### **D. Evidence for quasi-processive (distributive) chain elongation with a homopolymeric purine template**

When incubated with a homopolymeric oligo dA<sub>21</sub> DNA template (**Fig. 7**, lane 2), the integrase DNA

polymerase generally added between 1-7 nucleotides to this DNA as template - primer. Products consistent with the polymerization of up to 19 nucleotides, were not produced in significant amounts. There was a tendency to make longer DNA chains at intermediate DNA concentrations,

**Fig. 6 Behaviour of the integrase DNA polymerase with a homopolymeric oligo dC template.** Reaction mixtures contained either dGTP alone or all 4 dNTPs together with [ $\alpha$ -<sup>32</sup>P] dGTP. The letters to the right and left of the Fig. represent the number of guanylate residues incorporated into DNA. G357CA and G357C21 oligonucleotides consist of the same nonviral hairpin stem of 30 nucleotides and 21 nucleotide templates as defined in Table 3.

**Fig. 7 Effect of template pyrimidines on DNA chain elongation by integrase DNA polymerase.** The different template-primers used all were comprise of a common hairpin stem as described in Materials and Methods and a 21 nucleotide template sequence as designated at the top of each lane and defined in Table 3. A template consisting of oligodA 21 is listed in lane 2 and designated AAA. The numbers at the side of the Fig. refer to the number of nucleotides incorporated into DNA whereas the letters to the left and right of the Fig. indicate the identity of nucleotides incorporated into DNA in lanes 1 and 6, respectively.

albeit inefficiently, but this was abrogated at high DNA concentrations (not shown). The mechanism of chain elongation with an oligo dA template therefore appears to be either quasi-processive or distributive in nature. The longer DNA products obtained with an oligo dA<sub>21</sub> template were comparatively in low abundance versus the products of polymerization on a pyrimidine template (compare Fig. 6, lanes 3 and 4 with Fig. 7, lane 2). Hence, the processivity of the integrase DNA polymerase may be enhanced by pyrimidine tracts in the DNA template, a suggestion we explore further below.

### E. Utilization of mixed nucleotide sequence templates

We explored the possibility that the processivity of the integrase DNA polymerase may be enhanced by pyrimidine tracts in the DNA template. When 3 dC residues were placed at 7-nucleotide intervals within an oligo dA<sub>18</sub> template, processivity increased dramatically to a value of 20 (Fig. 7, lane 1). Also, there appeared to be many pause sites with the strongest ones corresponding to the location of the dC residues in the template (Fig. 7, lane 1). A parallel experiment in which Ts were used in place of dC yielded an almost identical effect on processivity while increasing the polymerization frequency several fold (Fig. 7, lane 6). Once again many pause sites were observed with the strongest ones located at A5 and A13, the position of these latter sites being apparently unrelated to the template T residues at positions 7 and 14 (Fig. 7, lane 6).

When the dinucleotides CT and GT were included in an oligodA<sub>15</sub> template (Table 3) much the same results were obtained; the inclusion of these extra nucleotides enhanced the overall extent of DNA chain elongation (Fig. 7, lanes 3 and 4). Finally, when 3 dG residues were included in the oligo dA<sub>18</sub> template (see Table 3), the extent of DNA chain elongation was relatively low and resembled the result obtained with an oligo dA<sub>21</sub> template (Fig. 7, lane 5). This result confirms the suspicion that template purines tend to contribute to low processivity by the integrase DNA polymerase. Nevertheless, the integrase DNA polymerase is clearly capable of chain elongation beyond one or two nucleotides, regardless of template sequence, as seen with 2-nucleotide template-primers.

**Table 3.** List of oligonucleotide template sequences used in this study to examine DNA chain elongation by the integrase DNA polymerase<sup>a</sup>

G357C21.....cccccccccccccccccc - 5'

G357CA.....ccccccccccccccac - 5'  
 CAA.....aaaaaacaaaaaac - 5'  
 AAA.....aaaaaaaaaaaaaaaaaa - 5'  
 CGA.....aaaagcaaaagc - 5'  
 CTA.....aaaatcaaaatc - 5'  
 GAA.....aaaagaaaaag - 5'  
 TAA.....aaaaataaaaaat - 5'

<sup>a</sup>the sequences listed represent extensions at the 5' - end of a common hairpin stem whose sequence is given in Materials and Methods.

### F. Sequence specificity exhibited by a core domain integrase DNA polymerase

Next we determined whether the amino or carboxyl terminus of integrase might play a role in the mechanism of sequence recognition at 2-nucleotide DNA template-primers. We isolated a core domain segment of integrase that included the first 3 N-terminal amino acids, linked to amino acids 51-190 and termed IN<sup>1-3X451-190</sup>. This truncated form of integrase displayed virtually the same specific activity for DNA polymerase as the wild type enzyme (data not shown).

The core domain deletions had a profound effect on the relative polymerization frequencies observed for template-primers with mutations in the conserved dinucleotide 5'-TG/CA (Fig. 8, Table 2). The main effect was an overall tendency to reduce the variation in polymerization frequency among the various mutant sequences. Thus, while the wild type enzyme exhibited a relatively large variation in polymerization frequency, depending on the sequence of the dinucleotide adjacent to the 5'AC template, the core domain exhibited only a 2-5-fold variation in this regard. There were two exceptions to this, namely the 5'-ACGC sequence which was not utilized at all by both enzymes and the 5' - ACCA sequence which exhibited a 50-fold reduction in polymerization frequency in the case of the core domain (Fig. 8, lanes 10 and 16, Table 2).

An amino terminal deletion construct IN<sup>1-3X4 51-288</sup> missing specifically the zinc finger domain (amino acids 4-50) gave identical results to that obtained with the core domain indicating that the amino terminus is responsible for maintaining the large differences in polymerization frequency observed with the wild type enzyme.

### G. Effect of core domain deletions on DNA chain elongation

The polymerization frequency of the core domain was not affected differently from wild type integrase in respect

to changes in the 5'-AC template sequence (**Fig. 9, Table 1**). The core domain deletions seemed to enhance the ability of the integrase DNA polymerase to copy both nucleotides of various 2-nucleotide template-primers (**Fig. 9, lanes 3, 6 and 7**). This capability of the core domain was further enhanced at higher enzyme concentrations but even under these conditions wherein the wild type enzyme also was more efficient in copying both nucleotides (data not shown) the core domain (and the wild type enzyme) paused frequently after inserting one nucleotide in the

various 2-nucleotide template-primers. It is noteworthy that even at relatively high enzyme concentrations, the wild type template sequence 5'-AC was relatively inefficiently repaired either by the core domain or the wild type DNA polymerase (**Fig. 9, lane 5**). By contrast, the HIV-1 reverse transcriptase readily repaired both nucleotides of all the 2-nucleotide template-primers examined even when the enzyme was present at a relatively low concentration (compare **Figures 4 and 10**). The results add to the evidence for an inherent regulatory

**Fig. 8. Effect of changes in the conserved 5'-TG/CA dinucleotide on end repair by the integrase core domain.** All reaction mixtures contained all 4 dNTPs and [ $\alpha$ -<sup>32</sup>P] dGTP as the only radiolabeled nucleotide. The boxed 5' -AC at the top left of each section of the Fig. represents the 5'-AC 2-nucleotide template sequence in each of the 16 different template-primers. The sequences of the dinucleotides adjacent to the 5'-AC tail are given at the top of each lane of the gel. The complementary nucleotides to each of these sequences (not shown) were also included in the oligonucleotide template-primers. The number (+1) to the right of the Fig. refers to the incorporation of 1 nucleotide into the 32 - nucleotide template-primer.

**Fig. 9 Effect of template sequence on end repair by the integrase core domain.** All reaction mixtures contained all 4 dNTPs and one [ $\alpha$ -<sup>32</sup>P] dNTP as described in Materials and Methods and as indicated in Fig. 4. The template sequence of each 2-nucleotide template-primer is given at the top of each lane. The

numbers to the right of the Fig. refer respectively to the incorporation of either 1 (+1) or 2 (+2) nucleotides into the DNA substrate.

**Fig. 10 Effect of template sequence on end repair by HIV-1 reverse transcriptase.** All reaction mixtures contained all 4 dNTPs and one [ $\alpha$ - $^{32}$ P] dNTP as described in Materials and Methods and as indicated in Fig. 4. The template sequence of each 2-nucleotide template-primer is given at the top of each lane. The numbers to the right of the Fig. refer respectively to the incorporation of either 1 (+1) or 2 (+2) nucleotides into the DNA substrate.

mechanism intrinsic to the structure of the HIV-1 integrase.

### III. Discussion

The relatively poor ability of the HIV-1 integrase DNA polymerase to repair the ends of DNA molecules is one of the most striking observations presented in this study. This is in stark contrast to the behavior displayed by this

enzyme with gapped DNA substrates, reported previously, where gaps of 1,2,5 or 7 nucleotides were efficiently and completely repaired, regardless of sequence (Acel et al., 1998). As far as we know, this differentiation of function is unique among DNA polymerases and would be consistent with a role for the integrase DNA polymerase in the repair of 5-nucleotide gaps that arise during the integration of HIV-1 DNA into cellular chromosomes. The integrase DNA polymerase therefore is likely designed to function exclusively as a repair enzyme. The integrase

DNA polymerase clearly would be incapable of acting as a replicative enzyme based on the results presented here because of a strong tendency to pause frequently during DNA chain elongation and especially because of its selectively poor ability to copy stretches of template purines (as compared to pyrimidines). The significance of the preference of the integrase DNA polymerase for template pyrimidines and their ability to increase the processivity of the enzyme is not clear at the present time.

The influence of the conserved dinucleotide 5'-TG/CA at the end of the HIV-1 LTR on the frequency of polymerization of the unpaired 5'AC tail is most significant. The dinucleotide sequence has a profound negative regulatory effect on the integrase DNA polymerase since mutations in either of these two nucleotides of the HIV-1 U5 LTR yields a much greater polymerization frequency, albeit only with respect to the insertion of the first nucleotide. This negative regulatory mechanism appears to be ideal as a way of further preventing the repair of 3'-processed ends of the HIV-1 LTRs while permitting the integrase DNA polymerase to repair gaps fully at a later stage in the integration process. The conserved 5' - TG/CA dinucleotide plays an important role in the 3' - processing reaction as well; mutations of this sequence have the effect of reducing the efficiency of endonucleolytic cleavage substantially. Taken together with our results this effect would be consistent with a model in which mutations in the LTR decrease 3' - processing by the integrase endonuclease while at the same time, increase the frequency of repair by the integrase DNA polymerase. The combination of these effects would tend to maintain the integrity of the ends of the viral DNA and perhaps decrease the probability of integrating a defective viral genome.

We have also reported here for the first time the effect of core domain deletions on the integrase DNA polymerase activity. The core domain retains all of the determinants required for the DNA polymerase activity of integrase including some of the regulatory determinants for DNA chain elongation. Some functions may however be lost such as the apparent negative regulatory effects of the conserved 5' TG/CA dinucleotide on polymerization frequency. Either the amino terminal zinc finger domain, or the C-terminal DNA binding domain may therefore play a regulatory role in this aspect of DNA polymerase function.

In terms of mechanism then, what can explain the inability of the integrase DNA polymerase to repair short templates at the ends of DNA molecules? How is chain elongation abrogated and what determines the sequence specificity with regard to polymerization frequency? The answers to these questions have not been obtained in the present study. Possibilities include a high  $K_m$  for certain template-primers leading to a low initial polymerization rate or in the absence of differences in the affinity of the enzyme for the DNA template, the results could be

explained in terms of a very slow off rate after incorporation of the penultimate nucleotide. Further detailed kinetic experiments are necessary to distinguish among these possibilities.

The regulation of end-repair by the integrase DNA polymerase emphasizes the importance of 3' - processing for the integration process and may have significance for the use of retroviruses as vectors or for the virus-free delivery of foreign genes into mammalian cells and into humans. The results presented here serve to emphasize the importance of preventing mutations in the LTRs in order to maintain a high efficiency of integration. It is possible to envision the design of artificial gene delivery systems which incorporate some features of the natural process of retroviral integration without the use of live viruses. Such systems might include the retroviral LTRs linked to a foreign therapeutic gene of interest in combination with an integrase expression system. Rudimentary designs of systems in which integrase, fused to another retroviral gene product, acts in *trans* to influence the properties of cotransfected LTR plasmids or mutant viral genomes missing the integrase gene, have been described (Faust et al., 1995; Liu et al., 1997; Wu et al., 1997).

Integrase is the newest and perhaps most promising target for antiviral chemotherapy (Pommier et al., 1997). Although many compounds have been shown to inhibit integrase function *in vitro* it is not clear that a good *in vivo* antiviral agent targeting integrase has been developed as yet (Burke et al., 1995; Cherepanov et al., 1997; King and Robinson, 1998; Mazumder et al., 1995; Neamati et al., 1997; Neamati et al., 1998; Robinson et al., 1996). Our study of end repair suggests a novel strategy for interfering with the integration process. Clearly, a mechanism exists preventing the integrase DNA polymerase from filling in the 3'-processed ends of the HIV-1 LTR. If the determinants on integrase that regulate this property of the enzyme could be clearly defined, then it might be possible to design drugs capable of 'flipping' the enzyme into a mode in which the regulation is lost, while maintaining the DNA polymerase activity intact. Under these circumstances the efficient repair of the 3'-processed ends could be effected with lethal consequences to the virus due to a failure to complete the integration process.

## IV. Materials and Methods

### A. Purification of the integrase DNA polymerase

The expression plasmid pQE30 $\Delta$ IN (Faust et al., 1995; Faust et al., 1996) was propagated in the Kan<sup>r</sup> *E. coli* strain M15 pREP(Qiagen) at 37°C and integrase expression was induced by adding IPTG as described (Faust et al., 1995; Faust et al., 1996). All further steps of purification were carried out at room temperature unless otherwise specified. Frozen bacterial pellets

derived from 1.2 L of bacterial culture were thawed and resuspended in 100 ml of B-Per solution (Pierce Chemical Co.) with the aid of a glass Dounce homogenizer. The suspension was centrifuged in a Beckman J2-21 centrifuge at 27,000g for 15 minutes using a Beckman JA-14 rotor. Pellets were resuspended in 100 ml B-Per solution by Dounce homogenization, lysozyme (Sigma) was added (200 micrograms/ml) and inclusion bodies were collected by centrifugation. The pellet was washed once more using a 1:10 dilution of the B-Per solution, left overnight at room temperature, and resuspended finally in 60 ml of buffer A (6M guanidine HCl, 0.1M Na-phosphate, 0.01M Tris-HCl pH 8.0). Qiagen nickel nitrilotriacetate (Ni<sup>++</sup>NTA) resin (6 ml of a 50% slurry) equilibrated in buffer A, was added to the protein solution and stirred for 1 hour at room temperature. The mixture was poured into a polypropylene column and the Ni<sup>++</sup>NTA beads allowed to pack under gravity. The resin bed was then washed with 120 ml of buffer A followed by 60 ml of 8M urea, 0.1M sodium phosphate, 0.01M Tris-HCl pH 8.0 and 18 ml each of the latter solution adjusted to pH 6.3 and pH 5.9. Integrase was eluted from the Ni<sup>++</sup>NTA resin in 21 ml of 8M urea, 0.1M sodium phosphate, 0.01M Tris-HCl pH 4.5 and renatured at 4°C by step-wise dialysis over a period of 3 days. Dialysis was done sequentially against 4M and 2M urea in 50mM Hepes-HCl pH 7.5, 1.0M NaCl and 1mM DTT. The sample was then dialyzed against final dialysis buffer (FDB) (50mM HEPES-HCl pH 7.5, 1.0M NaCl, 1mM DTT, 10% glycerol, 1mM CHAPS and 0.1mM EDTA) and then against 50mM imidazole, 1M NaCl, 50 mM HEPES-HCl pH 7.5, 10 mM β-mercaptoethanol, 1mM CHAPS and 10% glycerol. Dialyzed samples were applied to a 6 ml column of Ni<sup>++</sup>NTA beads equilibrated in the same buffer. The flow through was collected and the column washed with an additional 3.5 ml of the equilibrating buffer. The wash was combined with the flow through fraction and the combined sample was concentrated to a final volume of 0.5-1.0 ml using an Amicon Centricon Plus 80 filtration unit (80 ml capacity). The concentrated sample was applied immediately to a 92ml S-300 column (Pharmacia) 0.5x90cm that had been equilibrated in FDB. The column was developed at a rate of 5.5 ml/h in FDB, fractions of 1 ml were collected and the DNA polymerase activity was located using the trichloroacetic acid precipitation method as described previously (Acel et al., 1998). Fractions containing integrase DNA polymerase activity were pooled and stored at 4°C.

### B. Zinc finger N-terminal deletion

The amino terminal deletion mutant pQE 30ΔINΔZ consisted of the following format in respect to integrase IN<sup>1-3 X4 51-288</sup> where X refers to ValValArgLeu amino acids derived from the insertion of a linker(see below) and the numbers refer to the position of wild type amino acids in the integrase protein. Thus, this construct has a deletion of integrase amino acids 4-50 inclusive. It was derived by isolating the large fragment from a partial Xba I digest of pQE30Δ IN, cutting the isolated fragment with NsiI and isolating the large fragment once again. The latter consists of pQE30Δ IN missing the XbaI/NsiI region. The deleted plasmid was recircularized in the presence of a synthetic oligonucleotide linker 5'- CTA GAC GTA GTC CGT CTG CA - 3' hybridized with 3'- TG CAT CAG GCA G - 5' to produce the deleted plasmid pQE 30Δ INΔZ.

### C. Carboxy-terminal deletion

pCMV IN was cleaved with PstI and EcoRV. The linearized plasmid DNA was digested with exonuclease III and S1 nuclease using the Erase-a base kit (Promega) and circularized by ligation. The extent of deletion in the integrase gene was determined by DNA sequence analysis. A clone with a deletion end point at nucleotide 570 of the integrase coding region (includes amino acids 1-190) was cleaved with SmaI and recircularized with DNA ligase in the presence of HindIII linkers. The resulting plasmid was digested with HindIII and the small fragment was subcloned at the HindIII site of pQE30 Δ to produce pQE30 Δ IN 8.4.

### D. Core domain construct

The pQE 30Δ IN ΔZ plasmid was cleaved with BfrI and XhoI and the small fragment ligated to the large fragment derived from a BfrI/XhoI digest of pQE30 Δ IN 8.4. The resulting plasmid was comprised of the deleted N-terminus of the ΔZ plasmid and the deleted C-terminus of the 8.4 plasmid. Integrase was expressed from this construct as a core domain fusion protein that included the integrase amino acids IN<sup>1-3X451-190</sup>.

### E. DNA polymerase reactions

Unless stated otherwise standard DNA polymerase reactions were conducted using synthetic oligonucleotide template-primers with an oligodT<sub>20</sub> template as described previously (Acel et al., 1998) and quantified by precipitation of nascent radiolabeled DNA in trichloroacetic acid followed by liquid scintillation counting. Reaction mixtures contained 10 mM Tris-HCl pH 7.5, 5mM MgCl<sub>2</sub>, 5mM DTT, 200μg/ml bovine serum albumin, DISPOL 17 DNA (0.25mg/ml) 5'<sup>T</sup><sub>20</sub>ACTGCTAGAGATTTTAAAATCTCTAGCAGT 3' and 1μM dATP with 4μCi [α-<sup>32</sup>P] dATP (Mandel Scientific Co) in a total volume of 25 μl. Integrase (1 unit =1 pmol dAMP incorporated into DNA) was added in 1μl FDB. Reaction mixtures were incubated at 37°C for 1h.

For 2-nucleotide repair reactions, synthetic oligonucleotides with a 5'-AC template sequence and a 15-base pair hairpin primer stem matching the U5 LTR were used and reactions were conducted in the presence of all four dNTPs under standard conditions. The oligonucleotide with a wild type U5 sequence was:

5'-ACTGCTAGAGATTTTCCGGAAAATCTCTAGCA-3'.

Oligonucleotides used for end-repair reactions varied either in the sequence of first two nucleotides (5'-AC) or in the 4 nucleotides comprising the two adjacent base pairs 5'-TG/CA; in the former series of reactions the [α-<sup>32</sup>P] dNTP added (1μCi) was complementary to the penultimate 5' nucleotide (the first nucleotide adjacent to the primer stem) and in the latter series [α-<sup>32</sup>P] dGTP was added (1μCi) in all cases. Other oligonucleotides used in chain elongation studies consisted of a nonviral hairpin stem with the sequence

5'-  
GTAGCTCCGATCCGGTATATACCGATCCGAGCTAC- 3'  
extended at the 5'-end by the template sequences listed in **Table 3**. DNA polymerase reactions conducted with these

oligonucleotides contained all 4 dNTPs and 1 $\mu$ Ci of either [ $\alpha$ -<sup>32</sup>P] TTP or [ $\alpha$ -<sup>32</sup>P] dGTP depending on the sequence of the DNA template. Reaction products in this latter series of experiments were cleaved with MboI prior to PAGE analysis which was done at single-nucleotide resolution as described previously (Acel et al., 1998).

## Acknowledgements

We gratefully acknowledge Michael Parniak for his generous gift of HIV-1 reverse transcriptase and for many helpful discussions. This work was supported by grants from the National Health and Research Development Project of Health and Welfare Canada, the Medical Research Council of Canada and the Canadian Foundation for AIDS Research.

## References

- Acel, A., Udashkin, B.E., Wainberg, M. and Faust, E.A. **1998** Efficient gap repair catalyzed in vitro by an intrinsic DNA polymerase activity of human immunodeficiency virus type 1 integrase. **J. Virol.** 72, 2062-2071.
- Andrake, M.D. and Skalka, A.M. **1995** Multimerization determinants reside in both the catalytic core and C terminus of avian sarcoma virus integrase. **J Biol Chem** 270(49), 29299-29306.
- Ansari-Lari, M.A., Donehower, L.A. and Gibbs, R.A. **1995** Analysis of human immunodeficiency virus type 1 integrase mutants. **Virology** 211(1), 332-335.
- Asante-Appiah, E. and Skalka, A.M. **1997** Molecular mechanisms in retrovirus DNA integration. **Antiviral Res.** 36, 139-156.
- Balakrishnan, M. and Jonsson, C.B. **1997** Functional identification of nucleotides conferring substrate specificity to retroviral integrase reactions. **J. Virol.** 71, 1025-1035.
- Brown, P.O. (1990) Integration of retroviral DNA, pp. 19-48. Vol. 157. C. Springer-Verlag, Berlin-Heidelberg.
- Burke, T.R., Jr., Fesen, M.R., Mazumder, A., Wang, J., Carothers, A.M., Grunberger, D., Driscoll, J., Kohn, K. and Pommier, Y. **1995** Hydroxylated aromatic inhibitors of HIV-1 integrase. **J Med Chem** 38(21), 4171-4178.
- Bushman, F., 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.
- Bushman, F.D. and Craigie, R. **1991** Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. **Proc. Natl. Acad. Sci. USA** 88, 1339-1343.
- Bushman, F.D., Fujiwara, T. and Craigie, R. **1990** Retroviral DNA integration directed by HIV integration protein in vitro. **Science** 249, 1555-1558.
- Cai, M., Zheng, R., Caffrey, M., craigie, R., Clore, G.M. and Gronenborn, A.M. **1997** Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. **Nature structural biology** 4, 567-577.
- Cannon, P.M., Wilson, W., Byles, E., Kingsman, S.M. and Kingsman, A.J. **1994** Human immunodeficiency virus type 1 integrase: Effect on viral replication of mutations at highly conserved residues. **J. Virol.** 68, 4768-4775.
- Cara, A., Guarnaccia, F., Reitz, M.S., Jr., Gallo, R.C. and Lori, F. **1995** Self-limiting, cell type-dependent replication of an integrase-defective human immunodeficiency virus type 1 in human primary macrophages but not T lymphocytes. **Virology** 208, 242-248.
- Carteau, S., Batson, S.C., Poljak, L., Mouscadet, J.-F., de Rocquigny, H., Darlix, J.-L., Roques, B.P., Kas, E. and Auclair, C. **1997** Human Immunodeficiency Virus Type 1 nucleocapsid protein specifically stimulates Mg<sup>2+</sup>-dependent DNA integration in vitro. **J. Virol.** 71, 6225-6229.
- Carteau, S., Mouscadet, J.F., Goulaouiac, H., Subra, F. and Auclair, C. **1993** Quantitative assay for human immunodeficiency virus deoxyribonucleic acid integration. **Arch. Biochem. Biophys.** 300, 756-760.
- Cherepanov, P., Este, J.A., Rando, R.F., Ojwang, J.O., Reekmans, G., Steinfield, R., David, G., de Clercq, E. and Debyser, Z. **1997** Mode of interaction of G-quartets with the integrase of human immunodeficiency virus type 1. **Mol. Pharmacology** 52, 771-780.
- Chow, S.A. and Brown, P.O. **1994a** Juxtaposition of two viral DNA ends in a bimolecular disintegration reaction mediated by multimers of human immunodeficiency virus type 1 or murine leukemia virus integrase. **J Virol** 68, 7869-7878.
- Chow, S.A. and Brown, P.O. **1994b** Substrate features important for recognition and catalysis by human immunodeficiency virus type 1 integrase identified by using novel DNA substrates. **J Virol** 68, 3896-3907.
- Craigie, R. **1992** **Trends Genet.** 8, 187-190.
- Donzella, G.A., Leon, O. and Roth, M.J. **1998** Implication of a central cysteine residue and the HHCC domain of moloney murine leukemia virus integrase protein in functional multimerization. **J. Virol.** 72, 1691-1698.
- Drelich, M., Haenggi, M. and Mous, J. **1993** Conserved residues Pro-109 and Asp-116 are required for interaction of the human immunodeficiency virus type 1 integrase protein with its viral DNA substrate. **J. Virol.** 67, 5041-5044.
- Drelich, M., Wilhelm, R. and Mous, J. **1992** Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein in vitro. **Virology** 188, 459-468.
- Dyda, F., Hickman, A.B., Jenkins, T.M., Engelman, A., Craigie, R. and Davies, D.R. **1994** Crystal structure of the catalytic domain of HIV-1 integrase: Similarity to other polynucleotidyl transferases. **Science** 266, 1981-1986.
- Ellison, V., Abrams, H., Roe, T., Lifson, J. and Brown, P. **1990** Human immunodeficiency virus integration in a cell-free system. **J. Virol.** 64, 2711-2715.
- Engelman, A., Bushman, F.D. and Craigie, R. **1993** Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. **EMBO J.** 12, 3269-3275.

- 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, 6361-6369.
- 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.
- Engelman, A., Mizuuchi, K. and Craigie, R. **1991** HIV-1 DNA integration: Mechanism of viral DNA cleavage and DNA strand transfer. **Cell** 61, 1211-1221.
- Englund, G., Theodore, T.S., Freed, E.O., Engelman, A. and Martin, M.A. **1995** Integration is required for productive infection of monocyte-derived macrophages by human immunodeficiency virus type 1. **J. Virol.** 69, 3216-3219.
- Faust, E.A., Acel, A., Udashkin, B. and Wainberg, M.A. **1995** Human immunodeficiency virus type 1 integrase stabilizes a linearized HIV-1 LTR plasmid in vivo. **Biochem Mol Biol Int** 36, 745-758.
- Faust, E.A., Garg, A., Small, L., Acel, A., Wald, R. and Udashkin, B. **1996** Enzymatic capability of HIS-tagged HIV-1 integrase using oligonucleotide disintegration substrates. **J. Biomed. Sci.** 3, 254-265.
- Goldgur, Y., Dyda, F., Hickman, A.B., Jenkins, T.M., Craigie, R. and Davies, D.R. **1998** Three new structures of the core domain of HIV-1 integrase: An active site that binds magnesium. **Proc Natl Acad Sci USA** 95, 9150-9154.
- Grandgenett, D.P. and Mumm, S.R. **1990** Unravelling retrovirus integration. **Cell** 60, 3-4.
- Kalpana, G.V. and Goff, S.P. **1993** Genetic analysis of homomeric interactions of human immunodeficiency virus type 1 integrase using the yeast two-hybrid system. **Proc Natl Acad Sci USA** 90, 10593-10597.
- Katzman, M., Katz, R.A., Skalka, A.M. and Leis, J. **1989** The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the in vivo sites of integration. **J. Virol.** 63, 5319-5327.
- Khan, E., Mack, J.P., Katz, R.A., Kulkosky, J. and Skalka, A.M. **1991** Retroviral integrase domains: DNA binding and the recognition of LTR sequences. **Nucleic Acids Res.** 19, 851-860.
- King, P.J. and Robinson, W.E., Jr. **1998** Resistance to the anti-human immunodeficiency virus type 1 compound L- chicoric acid results from a single mutation at amino acid 140 of integrase [In Process Citation]. **J Virol** 72, 8420-8424.
- Kulkosky, J., Jones, K.S., Katz, R., Mack, J. and Skalka, A.M. **1992** Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. **Mol. Cell Biol.** 12, 2331-2338.
- LaFemina, R.L., Callahan, P.L. and Cordingley, M.G. **1991** Substrate specificity of recombinant human immunodeficiency virus integrase protein. **J. Virol.** 65, 5624-5630.
- LaFemina, R.L., Schneider, C.L., Robbins, H.L., Callahan, P.L., LeGrow, K., Roth, E., Schleif, W.A. and Emimi, E.A. **1992** Requirement of active human immunodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. **J. Virol.** 66, 7414-7419.
- Leavitt, A.D., Rose, R.B. and Varmus, H.E. **1992** Both substrate and target oligonucleotide sequences affect in vitro integration mediated by human immunodeficiency virus type 1 integrase protein produced in *Saccharomyces cerevisiae*. **J. Virol.** 66, 2359-2368.
- Leavitt, A.D., Shuie, L. and Varmus, H.E. **1993** Site-directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions in vitro. **J. Biol. Chem.** 268, 2113-2119.
- Lee, S.P., Xiao, J., Knutson, J.R., Lewis, M.S. and Han, M.K. **1997** Zn<sup>++</sup> promotes the self-association of human immunodeficiency virus type 1 integrase in vitro. **Biochemistry** 36, 173-180.
- Liu, H., Wu, X., Xiao, H., Conway, J.A. and Kappes, J.C. **1997** Incorporation of functional Human Immunodeficiency Virus Type 1 integrase into virions independent of the gag-pol precursor protein. **J. Virol.** 71, 7704-7710.
- Maignan, S., Guilloteau, J.P., Zhou-Liu, Q., Clement-Mella, C. and Mikol, V. Crystal structures 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.
- Mazumder, A., Engelman, A., Craigie, R., Fesen, M. and Pommier, Y. **1994** Intermolecular disintegration and intramolecular strand transfer activities of wild-type and mutant HIV-1 integrase. **Nucleic Acids Res** 22, 1037-1043.
- Mazumder, A., Gazit, A., Levitzki, A., Nicklaus, M., Yung, J., Kohlhagen, G. and Pommier, Y. **1995** Effects of typhostins, protein kinase inhibitors, on human immunodeficiency virus type 1 integrase. **Biochemistry** 34, 15111-15122.
- Miller, M.D., Farnet, C.M. and Bushman, F.D. **1997** Human Immunodeficiency Virus Type 1 preintegration complexes: Studies of organization and composition. **J. Virol.** 71, 5382-5390.
- Neamati, N., Hong, H., Sunder, S., Milne, G.W.A. and Pommier, Y. **1997** Potent inhibitors of human immunodeficiency virus type 1 integrase: Identification of a novel four-point pharmacophore and tetracyclines as novel inhibitors. **Mol Pharmacol** 52, 1041-1055.
- Neamati, N., Mazumder, A., Sunder, S., Owen, J.M., Tandon, M., Lown, J.W. and Pommier, Y. **1998** Highly potent synthetic polyamides, bisdistamycins, and lexitropsins as inhibitors of human immunodeficiency virus type 1 integrase. **Mol Pharmacol** 54, 280-90.
- Pommier, Y., Pilon, A.A., Bajaj, K., Mazumder, A. and Neamati, N. **1997** HIV-1 integrase as a target for antiviral drugs. **Antiviral Chem Chemother** 8, 463-483.
- Robinson, J.W.E., Cordeiro, M., Abdel-Malek, S., Jia, Q., Chow, S.A., Reinecke, M.G. and Mitchell, W.M. **1996** Dicafeoylquinic acid inhibitors of human immunodeficiency virus integrase: Inhibition of the core catalytic domain of human immunodeficiency virus integrase. **Mol Pharmacol** 50, 846-855.
- Roe, T., Chow, S.A. and Brown, P.O. **1997** 3'- End processing and kinetics of 5'-end joining during retroviral integration in vivo. **J. Virol.** 71, 1334-1340.

- Sakai, H., Kawamura, M., Sakuragi, J.-I., Sakuragi, R., Shibata, R., Ishimoto, A., Ono, N., Ueda, S. and Adachi, A. **1993** Integration is essential for efficient gene expression of human immunodeficiency virus type I. **J. Virol.** 67, 1169-1174.
- Sherman, P.A. and Fyfe, J.A. **1990** Human immunodeficiency virus integration protein expressed in E. coli possesses selective DNA cleaving activity. **Proc. Natl. Acad. Sci. USA** 87, 5119-5123.
- Shin, C.G., Taddeo, B., Haseltine, W.A. and Farnet, C.M. **1994** Genetic analysis of the human immunodeficiency virus type 1 integrase protein. **J Virol** 68, 1633-1642.
- Skalka, A.M. **1993** Retroviral DNA integration: lessons for transposon shuffling. **Gene** 135, 175-182.
- Taddeo, B., Haseltine, W.A. and Farnet, C.M. **1994** Integrase mutants of human immunodeficiency virus type 1 with a specific defect in integration. **J Virol** 68, 8401-8405.
- Tramontano, E., LaColla, P. and Cheng, Y.-C. **1998** Biochemical characterization of the HIV-1 integrase 3'-processing activity and its inhibition by phosphorothioate oligonucleotides. **Biochemistry** 37, 7237-7243.
- Van Gent, D.C., Groeneger, A.A.M.O. and Plasterk, R.H.A. **1992** Mutational analysis of the integrase protein of human immunodeficiency virus type 2. **Proc. Natl. Acad. Sci. USA** 89, 9598-9602.
- van Gent, D.C., Vink, C., Groeneger, A.M.O. and Plasterk, R.H.A. **1993** Complementation between HIV integrase proteins mutated in different domains. **EMBO J.** 12, 3261-3267.
- Vincent, K.A., Ellison, V., Chow, S.A. and Brown, P.O. **1993** Characterization of human immunodeficiency virus type 1 integrase expressed in Escherichia coli and analysis of variants with amino terminal mutations. **J. Virol.** 67, 425-437.
- Vink, C. **1991** Site-specific hydrolysis and alcoholysis of human immunodeficiency viral DNA termini mediated by the viral integrase protein. **Nucleic Acids Res.** 19, 6691-6698.
- Vink, C., Groeneger, A.A.M.O. and Plasterk, R.H.A. **1993** Identification of the catalytic and DNA binding region of the human immunodeficiency virus type I integrase protein. **Nucleic Acids Res.** 21, 1419-1425.
- Vink, C., Groenink, M., Elgersma, Y., Fouchier, R.A.M., Tersmette, M. and Plasterk, R.H.A. **1990** Analysis of the junctions between human immunodeficiency virus type 1 proviral DNA and human DNA. **J. Virol.** 64, 5626-5627.
- Vink, C. and Plasterk, R.H.A. **1993** The human immunodeficiency virus integrase protein. **Trends Genet** 9, 433-437.
- Vink, C., van Gent, D.C., Elgersma, Y. and Plasterk, R.H. **1991** Human immunodeficiency virus integrase protein requires a subterminal position of its viral DNA recognition sequence for efficient cleavage. **J. Virol.** 65, 4636-4644.
- Wiskerchen, M. and Muesing, M.A. **1995a** 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.
- Wiskerchen, M. and Muesing, M.A. **1995b** Identification and characterization of a temperature-sensitive mutant of human immunodeficiency virus type 1 by alanine scanning mutagenesis of the integrase gene. **J Virol** 69, 597-601.
- Woerner, A.M., Klutch, M., Levin, J.G. and Marcus-Sekura, C.J. **1992** Localization of DNA binding activity of HIV-1 integrase to the C-terminal half of the protein. **AIDS-Res-Hum-Retroviruses** 8, 2433-2437.
- Woerner, A.M. and Marcus-Sekura, C.J. **1993** Characterization of a DNA binding domain in the C-terminus of HIV-1 integrase by deletion mutagenesis. **Nucleic Acids Res.** 21, 3507-3511.
- Wu, X., Liu, H., Xiao, H., Conway, J.A., Hunter, E. and Kappes, J.C. **1997** Functional RT and IN incorporated into HIV-1 particles independently of the gag/pol precursor protein. **EMBO J** 16, 5113-5122.

