Articles

Loop$^{202-208}$ in Avian Sarcoma Virus Integrase Mediates Tetramer Assembly and Processing Activity†

Mary A. Bosserman,‡ Daniel F. O'Quinn,‡ and Isaac Wong*,‡,§

Department of Molecular and Cellular Biochemistry and Markey Cancer Center, University of Kentucky, Lexington, Kentucky 40536

Received January 30, 2007; Revised Manuscript Received July 24, 2007

ABSTRACT: Integrase (IN) catalyzes insertion of the retroviral genome into the host via two sequential reactions. The processing activity cleaves the 3′-dinucleotides from the two ends of the viral DNA which are then inserted into the host DNA. Tetramers are required for the joining step. While dimers have been shown to catalyze processing, they do so inefficiently, and the oligomeric requirement for processing is unknown. We have replaced loop$^{202-208}$ at the putative dimer−dimer interface of the avian sarcoma virus IN with its analogue, loop$^{188-194}$, from human immunodeficiency virus IN. The mutation abolished disintegration activity and a 2×10^{-2} s^{-1} fast phase during single-turnover processing. A 3×10^{-4} s^{-1} slow processing phase was unaffected. Preincubation with a DNA substrate known to promote tetramerization increased products formed during the fast phase by 2.5-fold only for wild-type IN, correlating the fast and slow phases with processing by tetramers and dimers, respectively. We propose a novel tetramer model for coupling processing and integration based on efficient processing by the tetramer. We provide for the first time direct evidence of the functional relevance of a structural element, loop$^{202-208}$, which appears to be required for mediating the interaction between dimer halves of the active tetramer.

Integration of a DNA copy of the RNA viral genome into the host DNA is a required step in the replication cycle of retroviruses such as the human immunodeficiency virus (HIV). Integrase (IN) catalyzes this in two sequential reactions. First, in the processing reaction, IN catalyzes a hydrolysis reaction to remove 3′-terminal dinucleotides from both U5 and U3 long terminal repeats (LTR's) of the duplex viral DNA generated by reverse transcriptase (RT). IN then catalyzes the nucleophilic attack by the processing-generated recessed 3′-ends on the phosphate backbone of opposite strands of the host DNA, fusing the viral sequence into the host genome. Selection of the insertion sites on the host is not sequence-specific. However, the two viral LTR's are inserted on opposite strands of the host target separated by a virus-specific signature spacing, 6 bp for the avian sarcoma virus (ASV) (*1*−*3*). This spatial constraint suggests that full-site integration likely requires the simultaneous binding of both processed viral DNA ends with the host DNA in a multimeric IN−DNA complex.

The integration reaction is mechanistically better understood than the processing reaction with respect to its quaternary structural requirements. In Bao et al. (*4*), we showed by active site titrations that ASV IN catalyzed disintegration, an in vitro mimic of the microscopic reversal of the integration reaction, with a 4:1 [IN]:[DNA] stoichiometry. We further demonstrated by atomic force microscopy that the responsible nucleoprotein complex is tetrameric and that assembly is likely induced by the binding of host DNA. Subsequently, cross-linked HIV-1 IN tetramers (*5*) as well as non-cross-linked, stable tetrameric synaptic complexes of HIV-1 IN bound to pairs of viral ends (*6*) have been isolated and shown to be active in full-site integration of both viral LTR's, whereas dimers catalyze only half-site reactions. These functional studies directly support an active tetramer model during the integration reaction.

The structural requirement for processing the LTR's prior to integration, however, remains less clear. Biophysical studies correlating activity with oligomeric states (*7*, *8*) show that dimers are sufficient to support processing activity, leading to the proposal that processing at the two ends of the viral DNA may be independently catalyzed by two different dimers which then assemble into a tetramer on target DNA for the subsequent integration reaction. However, IN dimers catalyze processing extremely inefficiently, measured at 0.004 min^{-1} per end for HIV-1 IN, and requiring two slow steps in sequence prior to full-site integration would severely limit the overall efficiency of the reaction. Furthermore, mutational analysis of amino acids thought to be in contact with LTR's based on a tetramer molecular model (*9*) and

† This work was supported by the Markey Cancer Center, University of Kentucky Medical Center, Lexington, KY.

* To whom correspondence should be addressed: Markey Cancer Center, University of Kentucky, 800 Rose St., Combs Research Bldg. Rm. 124-C, Lexington, KY 40536. Telephone: (859) 323-3036. Fax: (859) 257-8940. E-mail: iwong3@email.uky.edu.

‡ Department of Molecular and Cellular Biochemistry.

§ Markey Cancer Center.

10.1021/bi700197a CCC: $37.00
Published on Web 09/11/2007

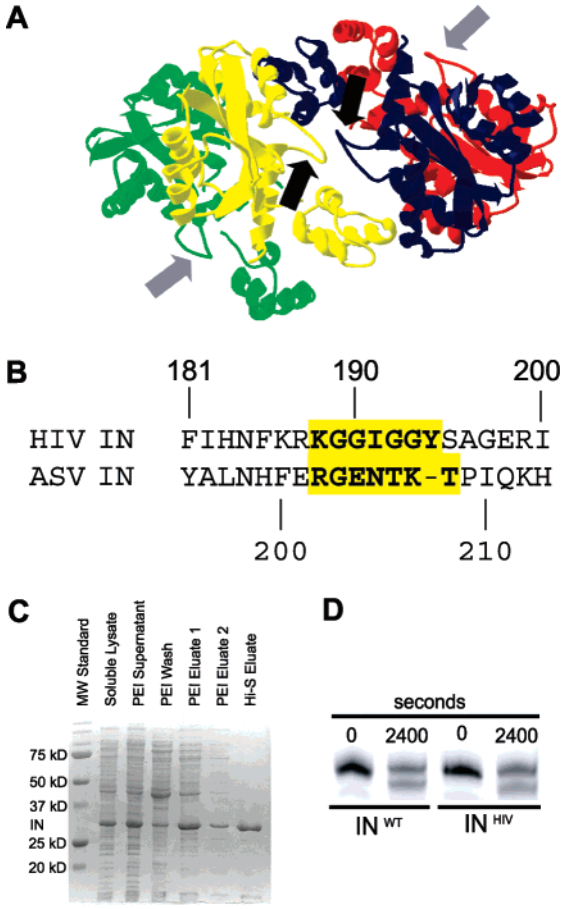

FIGURE 1: IN^{HIV} construction and activity. (A) A dimer-of-dimer tetramer model of HIV-1 IN has two $loop^{188-194}$ forms at the putative dimer−dimer interface (black arrows) and two externally facing loops (gray arrows). (B) Structural alignment of HIV and ASV IN sequences shows the analogous $loop^{202-208}$ of ASV IN (*23*). (C) Coomassie-stained SDS−PAGE gel showing purification of mutant IN^{HIV} from the cleared lysate by PEI pull-down, extraction, and Hi-S column chromatography. (D) Initial end-point analysis (2400 s) of processing shows comparable activities with wild type and mutant enzymes.

earlier cross-linking studies (*10*, *11*) implicate amino acids from both dimer halves of a tetramer in the formation of each LTR binding site. Finally, data from a number of other studies also suggest coordination of LTR processing (*6*, *12*, *13*) and full-site integration (*14*) mediated by higher-order oligomers, likely tetramers. However, correlation of processing activity, enhanced or otherwise, with tetramers has not been demonstrated.

Crystallographic and NMR structures for three subdomains of IN (N-terminus, catalytic core, and C-terminus) and several two-domain constructs (*15*−*22*) are dimeric. Thus, it remains to be seen if the full-length enzyme and/or the cocrystal with an appropriate substrate DNA would reveal a tetrameric model. Several dimer-of-dimer models have been suggested for the tetramer on the basis of contacts between crystallographic dimers in the two-domain structures (*9*, *21*, *22*). In the HIV-1-based model of Wang et al., $loop^{188-194}$ (KGGIGGY) from one dimer appears to interact reciprocally with its counterpart from a neighboring dimer (Figure 1A, black arrows), defining an apparent dimer−dimer assembly interface. Structural alignment (*23*) further identifies an analogous interfacial $loop^{202-208}$ (RGENTKT) in ASV IN (Figure 1B). A molecular dynamics modeling study suggested a different role for this loop in recognizing and binding the flanking DNA sequence one helical turn away from the cognate CAGT termini of the pre-insertion viral DNA (*24*). Interestingly, a different structural model hypothesized by Chen et al. (*9*) does not include this loop segment within the LTR binding trench. We have chosen to study the function of ASV $loop^{202-208}$ by replacing it with $loop^{188-194}$ from the HIV-1 IN. The effects of the loop swap on the pre-steady-state kinetic properties of IN-catalyzed reactions demonstrate a significant functional role of the loop during integration and further illustrate how structural features of both protein and DNA are required in mediating the catalytic function of the enzyme.

EXPERIMENTAL PROCEDURES

Oligodeoxyribonucleotides and Enzymes. Oligodeoxyribonucleotides (ODNs) were synthesized by Integrated DNA Technologies, Inc. (Coralville, IA), with the sequences listed in Table 1. Fluorescent ODNs were 5′-labeled with hexachlorofluorescein (HEX). Purification by denaturing PAGE and spectrophotometric determination of concentrations were as described previously (*12*, *13*) using the extinction coefficients listed in Table 1. Substrates were annealed at equimolar concentrations of component ODNs in a MasterCycler (Eppendorf) by being heated to 95 °C for 2 min and then cooled by 5 °C every 2 min until the temperature reached 20 °C in 20 mM Tris (pH 8.0), 5 mM Na-HEPES (pH 7.5), and 500 mM NaCl. Stoichiometric annealing were tested by native PAGE. The use of the thermocycler ensured efficient (40 min) and reproducible annealing. Wild-type integrase (IN^{WT}) and mutant integrase (IN^{HIV}) were overexpressed in *Escherichia coli* BL21(DE3), purified, and stored at −80 °C in 50 mM Na-HEPES (pH 7.5), 500 mM NaCl, and 40% glycerol as described previously (*12*, *13*).

Mutagenesis. The IN^{HIV} mutant enzyme was constructed through a two-step mutagenesis process using the Stratagene (La Jolla, CA) QuikChange II kit according to the manufacturer's specifications. A half-loop mutation was first created using the primer 5′-CGTGGTGAAATTGGTGGCTATCCGATACAGAAACACTGGAGACCTAC. The final mutation was then generated with a second primer, 5′-CACTTTGAGAAAGGTGGCATTGGTGGCTATCCGATACAGAAAC. Complementary primers were also synthesized. Targeted mutations and the absence of other mutations were confirmed by direct DNA sequencing (Functional Biosciences, Inc., Madison, WI).

Single-Turnover Disintegration and Processing Assays. Standard reactions (100 μL) were carried out at 37 °C in 20 mM Tris (pH 8.0), 10 mM Na-HEPES (pH 7.5), 4% glycerol, 10 or 5 mM $MnCl_2$, and 100 or 450 mM NaCl for processing or disintegration, respectively. Integrase was initially incubated with DNA substrate in the absence of $MnCl_2$ for 30 min at 37 °C. In processing reactions in the presence of the disintegration Y-substrates, initial incubation with Y (30 min) was followed by an additional incubation for 30 min at 37 °C in the presence of HEX-labeled processing substrate. Reactions were initiated by the addition of $MnCl_2$. Final concentrations of IN and DNA substrates were 1 μM and 250 nM for disintegration reactions and 3.65 μM and 150 nM for processing reactions, respectively. Aliquots were

Table 1: Oligodeoxyribonucleotide Sequences Used in This Study Listed with Their Extinction Coefficients

Name[a]	Nucleotide Sequence (5'→3')	ε_{260} ($mM^{-1}cm^{-1}$)
asvU5T[b]	GCTGAAGCAGAAGGCTTCATT[b]	205
asvU5T-2	GCTGAAGCAGAAGGCTTCA	189
asvU3T[b]	TGCTATTGCATAAGACTACATT[b]	246
asvU3T-2	TGCTATTGCATAAGACTACA	191
asvY3B	GCTATTGCATAAGACTACAACACGATTAGATGGTATTCAACAGC	441
hivU5B	ACTGCTAGACATTTTCCACACTG	218
hivY5B	AGTCAGTGTGGAAAATCTCTAGCAACACGATTAGATGGTATTCA	479
targetY44	GCTGTTGAATACCATCTAATCGTGTCGGGTCTCGTACTGCGGAA	419
targetY19	TTCCGCAGTACGAGACCCG	211

[a] ODNs derived from the ASV or HIV-1 LTR's are designated by asv or hiv, respectively. Nonvirally derived ODNs are denoted by target. [b] The scissile strands shown are denoted by "T". Sequences of asvU5B and asvU3B are not shown but are completely complimentary to asvU5T and asvU3T.

withdrawn at varying time points and quenched 1:1 (v/v) to achieve final concentrations of 4 M urea, 10 mM EDTA, and 0.1% SDS. Reaction products were separated by PAGE on 12 or 20% denaturing gels for disintegration or processing assays, respectively. Fluorescent bands were visualized on a Typhoon 9410 instrument (Amersham Biosciences). Image quantification and nonlinear least-squares fittings were performed with ImageQuant version 5.0 (Molecular Dynamics) and Kaleidagraph (Synergy, Reading, PA), respectively, as described previously (*12*, *13*).

Chemical Cross-Linking. The chemical cross-linking reagent, *cis*-aquahydroxydiaminoplatinum (AHDAP), was prepared as described previously (*5*, *25*). IN^{WT} or IN^{HIV} (20 μM) was preincubated in the presence of the Y^{ASV} substrate (4 μM) in a 60 μL reaction volume for 30 min at 37 °C before the addition of freshly prepared AHDAP to a final concentration of 100 μM. Cross-linking was allowed to progress for 10 min at 37 °C in the dark followed by the addition of 4× SDS gel loading buffer. Cross-linked oligomers were separated by PAGE on 12% SDS gels.

Fitting of Kinetic Time Courses. Time courses were generally fitted to a sum of exponential functions according to eq 1:

$$y = \sum_{i=1}^{n} A_i(1 - e^{-k_i t}),\ n = 1 \text{ or } 2 \qquad (1)$$

using the minimum number of exponential phases. When the second phase was slow on the time scale of the experiment and appeared linear, the time courses were fitted to the sum of a single-exponential function and a linear phase as in eq 2:

$$y = A(1 - e^{-k_1 t}) + k_2 t \qquad (2)$$

Time courses with a lag phase were fitted to eq 3:

$$y = A\left(1 + \frac{k_2}{k_1 - k_2}e^{-k_1 t} + \frac{k_1}{k_2 - k_1}e^{-k_2 t}\right) + k_3 t \qquad (3)$$

which describes the evolution of the product through two sequential intermediate steps with comparable rate constants,k_1 and k_2, followed by a much slower third step as shown in eq 4:

$$E \xrightarrow[(+S)]{k_1} EI \xrightarrow{k_2} EP \xrightarrow[(-P)]{k_3} E \qquad (4)$$

Without concentration dependent data, values for k_1 and k_2 are interchangeable due to the symmetry of eq 3 and cannot be assigned to specific steps.

RESULTS

Mutagenesis and Initial Characterization of IN^{HIV}. The mutant enzyme, IN^{HIV}, was constructed by swapping $loop^{202-208}$ in ASV IN with $loop^{188-194}$ from HIV-1 IN. The mutation was verified by direct sequencing of the plasmid clone, and IN^{HIV} was overexpressed and purified. Figure 1C shows that overexpressed IN^{HIV} was soluble and was purified in a manner identical to that for wild-type ASV (*13*). IN^{HIV} was assayed for processing activity using a short 21 bp duplex containing the ASV U3 end sequence. Processing would remove the 3′-terminal TT pair from the cognate CATT sequence, leaving a 19-nucleotide product which was separated from the substrate 21-mer by denaturing PAGE. Comparable processing products were observed after a reaction time of 2400 s for both IN^{WT} and IN^{HIV} (Figure 1D).

IN^{HIV} Is Disintegration and Tetramerization Defective. Because the disintegration reaction has been demonstrated to require tetramer assembly and can further be stoichiometrically correlated to IN activity in active site titrations (*4*), it provides a robust functional measure of the active tetramer population. Figure 2A schematically illustrates the

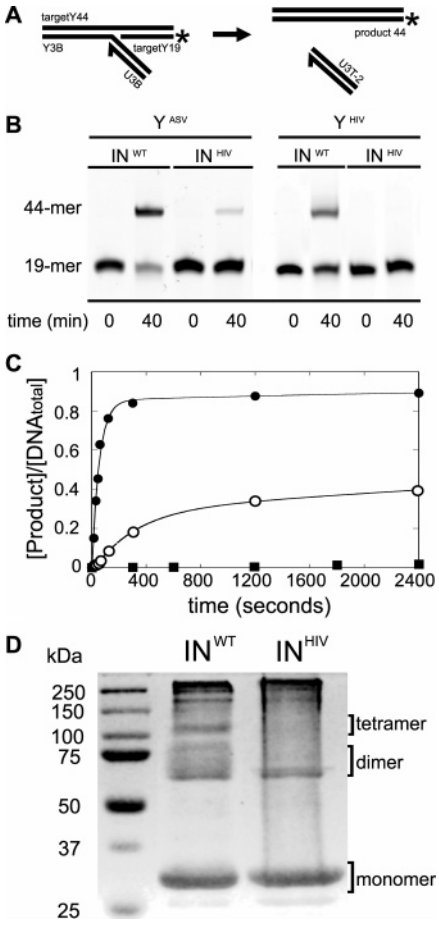

FIGURE 2: Reduced disintegration activity with IN^{HIV}. (A) Schematic of the disintegration reaction shows the transformation of a 5′-HEX-labeled 19-mer (targetY19) at the start of the reaction to a 44-mer product. The HEX label is denoted with an asterisk. Component ODNs are as labeled (see Table 1 for sequence information). (B) End-point gel analysis of the disintegration reaction by IN^{WT} and IN^{HIV} using Y^{ASV} showed ~90 and ~2% conversion, respectively, at 2400 s. Reactions using Y^{HIV} with IN^{WT} and IN^{HIV} showed reduced and complete loss of activity with IN^{WT} and IN^{HIV}, respectively. (C) Kinetic time courses of disintegration reactions for Y^{ASV} by IN^{WT} (●) and IN^{HIV} (■) are shown. The full time course of IN^{WT}-catalyzed disintegration on Y^{ASV} showed a characteristic rapid burst phase accounting for >84% conversion to products within 200 s. The solid line represents the best fit to a double-exponential function with the following rate constants for the fast and slow phases: $k_f = 1.8 \times 10^{-2}$ s^{-1} and $k_s = 2.3 \times 10^{-4}$ s^{-1}, respectively. The time course for IN^{HIV}-catalyzed disintegration on Y^{ASV} was linear with a rate constant of $<8 \times 10^{-6}$ s^{-1} with <2% product accumulation after 40 min. The time course of IN^{WT}-catalyzed disintegration on Y^{HIV} (○) showed a lag phase that was best fit to eq 3 with the following rate constants: $k_1 = 3.2 \times 10^{-3}$ s^{-1}, $k_2 = 3.6 \times 10^{-2}$ s^{-1}, and $k_3 = 4.4 \times 10^{-5}$ s^{-1}. (D) IN^{WT} and IN^{HIV} preincubated (20 μM) with Y^{ASV} substrate (4 μM) were chemically cross-linked and resolved by SDS−PAGE to identify the multimeric states of the wild-type and mutant enzymes during disintegration. Oligomers with molecular weights appropriate for the tetramer were present only in the wild-type lane. Reduced dimers for the mutant enzyme was also observed. Significant accumulation of aggregates larger than the tetramers was observed for both enzymes, and the levels of these always increased with cross-linker concentration or cross-linking time and likely reflect nonspecific aggregation due to the high concentrations of IN and/or cross-linker-induced aggregation.

disintegration reaction with a Y-shaped substrate, Y^{ASV}. This substrate mimics the intermediate product of integrating a single viral DNA end into a target sequence. The disintegration reaction represents the microscopic reversal of this reaction by expelling the viral DNA arm to regenerate a duplex 44-mer target and a 19/21-mer viral end sequence.

Initial analysis of total disintegration products (Figure 2B) formed after 40 min showed ~90 and ~2% conversion to products by IN^{WT} and IN^{HIV}, respectively. Similar end-point assays were also performed using a Y-substrate, Y^{HIV}, containing HIV LTR sequences. Figure 2B shows that the wild-type IN could perform limited disintegration on this "noncognate" substrate, but the HIV loop mutant, IN^{HIV}, had no activity on the HIV substrate.

Analysis of the actual reaction time courses revealed even more dramatic differences in the reaction rates (Figure 2C). With the ASV Y-substrate, Y^{ASV}, the IN^{WT}-catalyzed reaction was best described by a double-exponential function (eq 1) with >84% substrate converted within a 1.8×10^{-2} s^{-1} burst phase lasting ~200 s. A second slow phase followed at 2.3×10^{-4} s^{-1}. In contrast, IN^{HIV} exhibited a $>2.3 \times 10^3$-fold slower linear rate of $<8 \times 10^{-6}$ s^{-1}. With the HIV Y-substrate, Y^{HIV}, and wild-type IN, reductions in both amplitude and rate were apparent relative to those of the reaction with IN^{WT} and Y^{ASV}. The time course additionally exhibited an initial lag instead of the characteristic burst phase. The lag was fitted by eq 3 for two sequential steps with comparable rate constants ($k_1 = 3.2 \times 10^{-3}$ s^{-1}; $k_2 = 3.6 \times 10^{-2}$ s^{-1}) followed by a linear phase ($k_3 = 4.4 \times 10^{-5}$ s^{-1}). The combination of IN^{HIV} and Y^{HIV} showed no detectable activity, and its time course was therefore omitted.

To verify the structural correlation between the loss of activity and the failure to form tetramers, we attempted to cross-link protein−DNA complexes formed with Y^{ASV} and wild-type or mutant IN. Cross-linking reactions were carried out by addition of 100 μM AHDAP to preincubated mixtures of enzyme and Y^{ASV} followed by SDS−PAGE analysis to resolve cross-linked oligomers (Figure 2D). With IN^{WT}, bands representing proteins in the appropriate size range for monomer, dimers, tetramers, and higher-order aggregates were visible. In experiments with IN^{HIV}, however, only monomer, dimer, and higher-order aggregate bands were observed; bands corresponding to tetramers were completely absent. Attempts to increase cross-linking efficiency by increasing either reagent concentration or reaction time resulted only in conversion of monomers directly to highly molecular aggregated forms (darker bands near the top of the gel). However, direct comparison of the tetramer region between wild-type and mutant IN clearly shows a loss of tetramerization, although, interestingly, dimer formation appeared to be less affected.

Biphasic Processing by IN^{WT}. To better detect the processing contribution by tetramers, we re-examined the single-turnover time course of the wild-type processing reaction. The reaction profiles for processing of ASV U3- and ASV U5-derived substrates with IN^{WT} (Figure 3A) were biphasic with 10−13% fast phases at 2.7×10^{-2} and 2.0×10^{-2} s^{-1}, respectively, followed by slower phases in the range of 3.7−3.8×10^{-4} s^{-1}. The small differences between the ASV U3 reaction and the ASV U5 reaction are consistent with the known preferences of ASV IN for the U3 sequence (*12*, *13*).

To test if tetramers contributed the burst of efficient processing, we preincubated IN^{WT} with Y^{ASV} prior to adding the labeled ASV U3 substrate as Y^{ASV} should stimulate tetramer assembly. As shown in Figure 3A, a 2.5-fold increase in the amplitude of the fast phase was observed

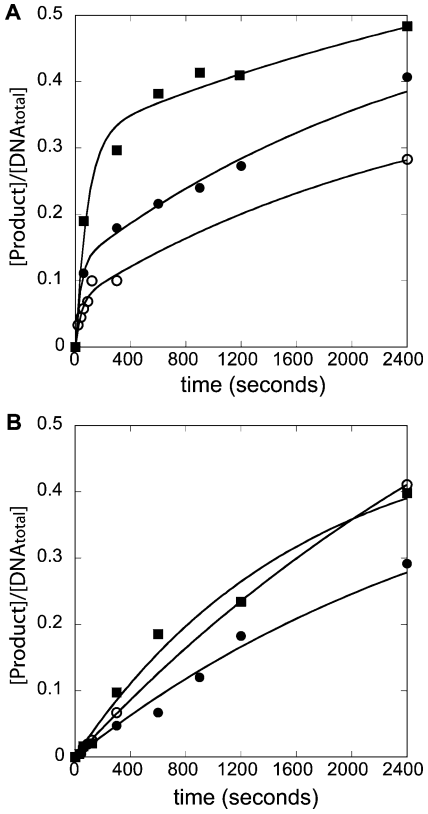

FIGURE 3: Processing activity of IN^{WT} and IN^{HIV}. (A) IN^{WT}-catalyzed processing of both ASV U3 (●) and ASV U5 (○) substrates is biphasic with the following best-fit rate constants: $k_f = 2.7 \times 10^{-2}$ s^{-1} and $k_s = 3.7 \times 10^{-4}$ s^{-1} for ASV U3 and $k_f = 2.0 \times 10^{-2}$ s^{-1} and $k_s = 3.8 \times 10^{-4}$ s^{-1} for ASV U5. Addition of the Y^{ASV} substrate to the ASV U3 processing reaction mixture increased the burst amplitude of the reaction from 12% in the absence of Y^{ASV} to 31% in the presence of Y^{ASV} (■) with the following rate constants: $k_f = 1.2 \times 10^{-2}$ s^{-1} and $k_s = 3.8 \times 10^{-4}$ s^{-1}. (B) IN^{HIV}-catalyzed processing of both the ASV U3 (●) and ASV U5 (○) substrates is monophasic with the following best-fit rate constants: $k_s = 3.3 \times 10^{-4}$ s^{-1} for U3 and $k_s = 2.4 \times 10^{-4}$ s^{-1} for U5. Addition of the Y^{ASV} substrate to the U3 processing reaction with IN^{HIV} did not significantly alter the time course (■) and yielded a single-exponential rate constant (k_s) of 3.3×10^{-4} s^{-1}.

together with a marginal decrease of its rate constant to 1.2×10^{-2} s^{-1}. The incomplete stimulation of the fast phase amplitude as well as the 2-fold reduction in the apparent rate constant likely reflects the occupation of one of the two functional active sites on the tetramer by the Y-substrate, halving the number of available active sites for processing. The correlation of the rapid processing phase with tetramer assembly demonstrates that tetramers catalyze processing 100-fold more efficiently than dimers.

IN^{HIV} Exhibits Only Slow-Phase Processing Kinetics. The time courses of processing by IN^{HIV} showed a distinct absence of fast phases (Figure 3B). Instead, we observed only single-exponential behavior with rate constants of 3.3×10^{-4} and 2.4×10^{-4} s^{-1} for U3 and U5 substrates, respectively, comparable to those observed in the slow phases of the wild-type reactions attributable to dimer activity. Finally, addition of Y^{ASV} failed to stimulate a burst phase in mutant-catalyzed processing, confirming the loss of kinetic behaviors characteristic of the tetramer in the loop mutant. We note further that comparable amounts of processing products were observed for mutant and wild-type enzymes at long time points (Figure 1D) and that the tetramer activity was distinguishable only within the initial 100 s of the reaction, likely the cause for prior failures to detect the enhanced tetramer activity.

Table 2: Reaction Rate Constants[a]

	DNA		rate constant (s^{-1}) IN^{WT}	IN^{HIV}
disintegration	Y^{ASV}	k_f	$(1.8 \pm 0.2) \times 10^{-2}$	–
		k_s	$\sim 2.3 \times 10^{-4}$	$< 8 \times 10^{-6}$
	Y^{HIV}	k_1	$(3.2 \pm 0.4) \times 10^{-3}$	undetectable
		k_2	$(3.6 \pm 1.1) \times 10^{-2}$	–
		k_3	$(4.4 \pm 0.8) \times 10^{-5}$	–
processing	ASV U3	k_f	$(2.7 \pm 1.4) \times 10^{-2}$	–
		k_s	$(3.7 \pm 1.5) \times 10^{-4}$	$(3.3 \pm 0.7) \times 10^{-4}$
	U3 + Y^{ASV}	k_f	$(1.2 \pm 0.4) \times 10^{-2}$	–
		k_s	$(2.4 \pm 0.4) \times 10^{-4}$	$(6.2 \pm 1.5) \times 10^{-4}$
	ASV U5	k_f	$(2.3 \pm 0.8) \times 10^{-2}$	–
		k_s	$(3.8 \pm 0.5) \times 10^{-4}$	$(2.4 \pm 0.4) \times 10^{-4}$

[a] Apparent rate constants derived from exponential fits to time courses are listed for IN^{WT} and IN^{HIV}. Subscripts f and s denote rate constants for fast and slow phases, respectively.

DISCUSSION

Table 2 summarizes our pre-steady-state comparison of the kinetics of disintegration and processing by the wild-type, IN^{WT}, and HIV-loop-swapped chimeric enzyme, IN^{HIV}. Comparison of the observed rate constants shows clearly that the mutation abolished disintegration activity and rapid processing but preserved slow processing activity. Cross-linking further confirms that the loss of disintegration can be correlated with the loss of tetramer assembly. These results address two critical issues in the IN reaction mechanism. First, the loss of the signature tetramer kinetic characteristics in IN^{HIV} functionally confirms the putative position of $loop^{202-208}$ at the assembly interface, providing direct support to the dimer-of-dimer model proposed by Wang et al. (*9*, *22*). Second, the loss of rapid processing activity in the tetramer-defective mutant suggests that the tetramer catalyzes processing 100-fold faster than the dimers. Mechanistic implications of these findings are discussed below, leading to the proposal of a mechanism in which processing and integration are coordinately coupled by tetramer assembly to ensure full-site integration of both processed viral ends.

$Loop^{202-208}$ of ASV IN Mediates Tetramer Formation. Assembly of an IN tetramer for full-site integration, i.e., concerted integration of both viral LTR's, is now well supported by biochemical evidence (*4*−*6*, *8*, *9*). Quaternary structural models have been proposed on the basis of interactions observed in neighboring crystallographic dimers of two-domain constructs (*9*, *21*, *26*). Additionally, modeling based on cross-linking data has led to a proposed tetramer model for the full-length polypeptide. Mutational analysis of amino acids in putative LTR binding sites of this model further established a functional correlation with processing activity. However, knowledge of the tetramer structure is lacking, due largely to the absence of a tetrameric crystal structure, and structural elements directly responsible for tetramer assembly have not been identified.

Ren et al. (*27*) recently obtained electron microscopy images of tetramer particles with a substantial structural asymmetry which cannot be explained by any of the current

structural models based on neighboring crystallographic dimers. This asymmetry further resonates with previous biochemical studies that have demonstrated functional asymmetry in the catalytic mechanism (*12*, *13*). On the other hand, these electron microscopy images also cannot distinguish between any of the putative dimer−dimer interfaces proposed in these models, provided that the two dimer halves assemble asymmetrically about that interface.

To directly test one such interface, we have chosen to mutate a unique loop at the center of one such interfacial model (*22*) by swapping the native ASV sequence with the structurally analogous HIV-1 IN sequence. Results of pre-steady-state kinetic analysis of the loop mutant, IN^{HIV}, compared with wild-type ASV IN^{WT} clearly demonstrate that replacement of $loop^{202-208}$ does not affect dimer-catalyzed processing activity but completely disrupts disintegration, an activity characteristic of the tetramer. Additionally, while chemical cross-linking led to tetramer-sized oligomers for wild-type IN , these were completely absent in parallel experiments performed using IN^{HIV}.

Because of the large number of amino acid changes in the chimeric construct, we had concerns that the mutation may disrupt the overall fold of the dimer. However, the outward-facing position of the loop in the dimer structure (Figure 1A, gray arrows) would suggest that the loop itself contributed minimally to the fold and stability of the dimer half to which it belonged. This was further validated experimentally when we found that the mutant enzyme was able to catalyze the processing reaction despite being completely defective in disintegration. Our results, therefore, directly validate the dimer-of-dimer models suggested by Wang et al. (*22*) and Chen et al. (*9*) and provide the first direct identification of a central structural element at the tetramer assembly interface.

Interestingly, we had designed the loop swap originally with the naïve presupposition that the reciprocal nature of the interaction by the loops and the symmetry at the assembly interface would lead to retention of activity; i.e., HIV loop can interact with HIV loop just as readily as ASV loop with ASV loop. We had thus also constructed half-loop mutants for the explicit purpose of abolishing the reciprocity of the interaction. However, all loop mutations severely compromised activity (M. A. Bosserman, personal communication), suggesting other functionally important interactions and structural constraints at play. We note, however, that the amino acid side chains of the ASV and HIV loops differ dramatically from each other: while the HIV sequence is characterized by a uniquely small "footprint" consisting almost entirely of four glycine residues, the ASV loop is much bulkier and highly charged. Thus, the loss of activity may simply reflect some spatial constraints imposed by the rest of the structure at the interface which may, for example, prevent two "downsized" HIV loops from reciprocally contacting each other.

Flanking Sequence Recognition and Binding. Wielens et al. (*24*) have suggested that this loop may be involved in recognizing and binding the flanking DNA sequence one helical turn away from the cognate CAGT termini of the viral DNA. In contrast, Chen et al. (*9*) propose LTR-binding trenches which do not appear to include the loop sequence; however, they do not directly address any functions for the loop in question in either assembly or DNA binding. Our results are consistent with the latter model and do not support a loop function for LTR binding.

If the primary function of the loop were to recognize viral flanking sequences, then changing the Y-substrate to the matching HIV DNA sequence in Y^{HIV} should have restored, at least partially, activity to the IN^{HIV} mutant. Conversely, the wild-type ASV enzyme, IN^{WT}, exhibited altered but significant reactivity on the HIV Y-substrate, Y^{HIV}, despite the clearly incorrect cognate terminal sequence (*28*). Finally, mutant IN^{HIV} was able to process ASV U3 and U5 substrates, albeit at the slower "dimer" rate. These results demonstrates a distinct lack of correlation between the loop and LTR binding and, therefore, support the structural model for LTR binding proposed by Chen et al. (*9*).

Our results do not necessarily rule out the possibility that the loop may play some other secondary role because a total of four loops exist in the tetramer model. Only two of these would be interfacial (Figure 1A, black arrows) per tetramer, leaving two externally facing loops (Figure 1A, gray arrows) available, at least in theory, to serve some other unknown function. However, any such roles must necessarily be minor relative to that of tetramerization, as their effects, if any, are completely masked by the overriding effects of disrupting the tetramer-promoting function of the interfacial loops.

Functional Role of Tetramers in Processing. While the tetramer is clearly implicated in full-site integration, processing of the two LTR ends is thought to be independently catalyzed by different IN dimers (*5*, *6*, *8*, *14*). These dimers at the two ends would then assemble on the target DNA into the required tetramer for concerted full-site joining (*5*, *6*, *8*, *14*; Figure 4A). This model, however, lacks the experimental support of a cooperative mechanism that ensures complete processing at both ends prior to tetramer assembly. Thus, full-site integration, which would require the processing of both viral LTR's, would be rate-limited by two consecutive, slow processing steps (2×10^{-4} s^{-1} for ASV IN or 0.004 min^{-1} for HIV IN) (*7*). On the other hand, tetramer assembly occurs rapidly on this time scale, being complete within the 20−30 min preincubation period at 37 °C (*5*, *6*, *14*). Preliminary stopped-flow data suggest that assembly may be 90% complete within 5 min (M. A. Bosserman, personal communication). The catalytic flux through such a mechanism would, therefore, favor the formation of a long-lived tetramer intermediate with one processed and one unprocessed LTR. Rapid tetramer-catalyzed insertion would then disrupt concerted full-site integration, yielding abortive half-site integration products instead.

We have previously shown using synapsed processing substrates containing two linked cognate end sequences that processing was facilitated by the cobinding of both U3 and U5 ends. Although this suggested that higher-order assembly stimulated processing (*12*), due to the poor extent of reactions in those experiments (<2%), the actual oligomeric state could not be defined. Recently, we have found that tetramer assembly was enhanced by a 20 min preincubation at 25−37 °C (*6*, *7*; M. A. Bosserman, personal communication), and we now routinely observe significant processing products (12−15%) formed at 2×10^{-2} s^{-1} in early time points. Correlation of this enhanced processing with tetramers is based on two separate and independent lines of evidence. First, a 2.5-fold enhancement of the amount of rapid processing was observed in the presence of the Y^{ASV} substrate

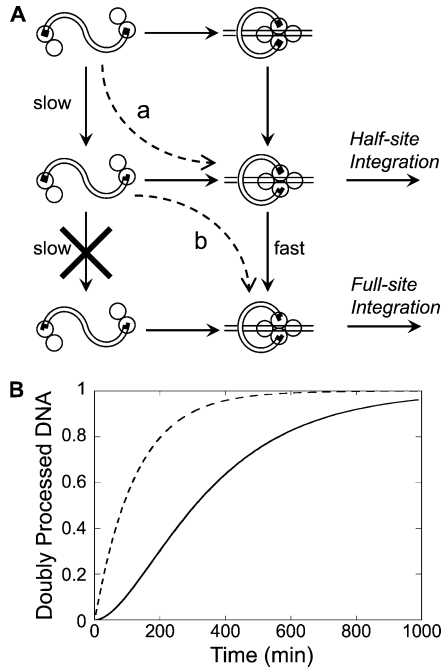

FIGURE 4: Dimer vs tetramer models for coupling processing and full-site integration. (A) Viral DNA is shown bound with IN dimers at the ends. If processing of the two ends of the viral DNA were to occur independently, then the slow rate of processing would favor tetramer assembly after a single processing event (path a). Without enhanced processing of the other end by the tetramer, insertion of the single processed end would result in half-site integration and an uncoupling of concerted integration. Path b shows how coupling of tetramer assembly with enhanced processing would permit the rapid processing of the unprocessed end needed for full-site integration. (B) Simulation of the time courses for generating doubly processed viral DNA showing the presence and absence of a lag in the dimer model (—) and tetramer model (– – –), respectively. The lag results from the need to undergo two consecutive rate-limiting processing steps to generate a substrate intermediate competent for concerted full-site integration. Simulations were generated assuming 0.004 min^{-1} processing of all ends for the dimer model vs 1.2 min^{-1} processing of the second end for the coupled tetramer model.

under conditions that would stimulate tetramer formation. Second, the enhanced processing was absent with the IN^{HIV} loop mutant, which independent cross-linking and disintegration results have shown to be defective in tetramer formation.

Inefficient processing by HIV-1 IN dimers has been well characterized by Smolov et al. (*7*) and Guiot et al. (*8*). Neither study detected discernible activity attributable to the tetramer. The single-turnover rate of 0.004 min^{-1} reported for HIV-1 IN dimer-catalyzed 3′-end processing (*7*) is consistent with the slow 3×10^{-4} s^{-1} phase in our experiments with ASV IN which we also attribute to the dimer. While Smolov et al. use a 2:1 [IN]:[DNA] processing stoichiometry to argue for a dimer reactive species, the same 2:1 reaction stoichiometry is also consistent with a tetramer binding two DNA ends, as suggested elsewhere (*4*−*6*, *21*, *22*).

Guiot et al. (*8*) correlated processing activity of HIV-1 IN with particle size using time-resolved anisotropy. However, the processing activity was measured using a single end-point assay at 400 min (Δr_{SDS} at $t = 400$ min) which is well past the stated linear range of the assay of <180 min. Because of the single-turnover nature of processing (*7*, *12*), differences in activities are not propagated and amplified through multiple rounds of turnovers as during the steady state but are instead manifested as distinct and characteristic exponential phases. Thus, an end-point assay would fail to reveal a mixture of activities contributed by distinct subpopulations, as we have demonstrated (compare Figure 1D with Figure 3A). Additionally, in their experiments, reactions were initiated by increasing the temperature of preincubated mixtures from 25 to 37 °C, further compromising the detection of rapid "bursts" because of the ill-defined time zero.

Despite the vast structural and functional similarities between HIV and ASV IN, there are differences as well, especially with respect to metal requirements for conformational changes and assembly (*17*, *29*). Thus, it is possible that the discrepancy between our results and those published for HIV-1 IN may simply reflect a more fundamental difference between the two enzymes.

Coupling Processing and Assembly Facilitates Full-Site Integration. The ability of tetramers to catalyze processing 2 orders of magnitude faster than dimers provides a kinetically attractive mechanism for coupling processing, assembly, and integration to ensure full-site integration of both viral ends. As illustrated in Figure 4A, premature assembly of a tetrameric integration-enabled complex on the host DNA with a singly processed end (path a) would lead to half-site integration in the absence of a coordinately increased speed of processing at the other end. However, by enhancing the tetramer-catalyzed rate of processing, a tetramer formed with a singly processed viral DNA would be rapidly converted to a doubly processed intermediate (path b) ready for full-site integration. This coupling of processing activity with tetramer assembly would therefore direct flux in the reaction mechanism toward the desired full-site integration reaction while bypassing the need to wait for the slow processing of the second LTR.

To illustrate the effects of this kinetic coupling, we simulated (*30*, *31*) time courses (Figure 4B) for forming the doubly processed intermediate required for full-site integration with (dashed line) and without (solid line) enhanced processing by the tetramer. In this example, *singly* processed DNA appears with an apparent rate constant of 0.008 min^{-1} (the apparent doubling of the input value of 0.004 min^{-1} is a statistical consequence of having two LTR ends in each viral DNA). The generation of the doubly processed DNA, however, shows a lag resulting from having to wait for two consecutive 0.004 min^{-1} processing events to produce the doubly processed viral DNA. The lag, a kinetic consequence of two consecutive rate-limiting steps in the reaction pathway, demonstrates the detriment of uncoupled processing of the two ends. Assembly into a tetramer during this delay (path a) without a coordinated increase in processing rate by the tetramer would highjack the reaction in favor of abortive half-site integration of the singly processed LTR. Our simulation of this model does not take into account the fact that dimers are likely capable of catalyzing half-site integrations (*6*), which would serve only to exacerbate the situation.

On the other hand, in the tetramer processing model, the increased efficiency of tetramer-catalyzed processing would allow assembly to be kinetically coupled to full-site integration by ensuring completion of processing upon tetramer assembly. As shown in the simulation, because processing

of the second end by the tetramer occurs much faster (modeled at 1.2 min^{-1} on the basis of the ASV rate constant measured here), a full-site integration complex can be generated from a singly processed DNA concurrently with assembly, thereby eliminating the lag observed with the dimer model. The resulting time course would be, therefore, identical to that for the processing of the first end. Coupling would therefore result in a significant reduction in half-life for complete processing of both ends in the given example. The absence of the lag indicates the restoration of flux toward the desired full-site integration product.

We have arbitrarily chosen a singly processed end as a starting point for our simulations. Because tetramer assembly occurs faster, being complete within 20 min of the preincubation period, the slower dimer-catalyzed processing rate may not be relevant at all if assembly of a tetramer were to precede any processing. However, assembly likely requires the binding of target DNA (*4*), which would in turn require transport into the nucleus. The slow processing of at least one end by dimers during this latent period is likely not detrimental and may even prove to be useful in suppressing aberrant premature integration of ends, e.g., into the middle of the viral DNA itself as dimers are incapable of joining.

Interestingly, Li et al. (*14*) observed that unprocessed DNA substrates led to an increased frequency of full-site integration whereas "preprocessed" DNA substrates tended to generate half-site integration products. These results further implicate a mechanistic coupling among processing, assembly, and integration. However, direct evidence for how this could be achieved had been puzzling because of the slow rate of dimer processing. The demonstration that assembled tetramers can process efficiently, therefore, provides a key mechanistic link for establishing a coordinated model in ASV IN. Whether the HIV IN reaction will prove to be similarly coordinated remains to be tested.

Structure and Sequence Specificity. In the proposed tetramer model (Figure 1A), the dimer−dimer interface runs almost parallel to a basic groove proposed to be the target DNA binding site (*4*). Because of the abundance of positively charged residues at the assembly interface, we had further hypothesized that the target DNA would facilitate assembly by neutralizing some of the like-charge repulsion between the dimer halves (*4*). Subsequently, HIV IN tetramers have been identified in vitro using only LTR ends (*5*, *6*). Thus, it is clear that LTR ends by themselves are sufficient at inducing assembly (*6*). However, these studies either required very specific incubation conditions or have relied upon chemical cross-linking to trap the tetramers formed. In the work of Ren et al. (*27*), a three-way junction DNA substrate similar to the Y-substrate was used to form a homogeneous complex of tetramers for single-particle imaging.

Our results support recognition and specificity for structural features of the Y-substrate in addition to the LTR sequence information in facilitating IN tetramer formation. Like the three-way junction DNA of Ren et al. (*27*), the Y-substrate mimics both structural and sequence elements of the integration intermediate recognized by the active site. The superior ability of the Y-substrate to induce assembly is clearly demonstrated in the experiment shown in Figure 3A, where the preincubation with the Y-substrate increased the extent of tetramer-catalyzed processing. More importantly, this result also demonstrated that the Y-substrate-induced tetramer supports processing at the enhanced rate of 1.2×10^{-2} s^{-1} and thus represents a physiologically relevant intermediate. Conversely, a Y-construct with the incorrect sequence information was shown to support a hobbled disintegration reaction, just as the viral LTR's alone could partially support tetramerization. Thus, both sequence information from the viral LTR and structural features of its junction to the target DNA are recognized by IN during tetramer assembly.

ACKNOWLEDGMENT

We thank Dr. Kogan Bao for his helpful discussion.

REFERENCES

1. Hindmarsh, P., and Leis, J. (1999) Retroviral DNA integration, *Microbiol. Mol. Biol. Rev. 63*, 836−843.
2. Craigie, R. (2001) HIV integrase, a brief overview from chemistry to therapeutics, *J. Biol. Chem. 276*, 23213−23216.
3. Asante-Appiah, E., and Skalka, A. M. (1997) Molecular mechanisms in retrovirus DNA integration, *Antiviral Res. 36*, 139−156.
4. Bao, K. K., Wang, H., Miller, J. K., Erie, D. A., Skalka, A. M., and Wong, I. (2003) Functional oligomeric state of avian sarcoma virus integrase, *J. Biol. Chem. 278*, 1323−1327.
5. Faure, A., Calmels, C., Desjobert, C., Castroviejo, M., Caumont-Sarcos, A., Tarrago-Litvak, L., Litvak, S., and Parissi, V. (2005) HIV-1 integrase crosslinked oligomers are active in vitro, *Nucleic Acids Res. 33*, 977−986.
6. Li, M., Mizuuchi, M., Burke, T. R., Jr., and Craigie, R. (2006) Retroviral DNA integration: Reaction pathway and critical intermediates, *EMBO J. 25*, 1295−1304.
7. Smolov, M., Gottikh, M., Tashlitskii, V., Korolev, S., Demidyuk, I., Brochon, J. C., Mouscadet, J. F., and Deprez, E. (2006) Kinetic study of the HIV-1 DNA 3′-end processing, *FEBS J. 273*, 1137−1151.
8. Guiot, E., Carayon, K., Delelis, O., Simon, F., Tauc, P., Zubin, E., Gottikh, M., Mouscadet, J. F., Brochon, J. C., and Deprez, E. (2006) Relationship between the oligomeric status of HIV-1 integrase on DNA and enzymatic activity, *J. Biol. Chem. 281*, 22707−22719.
9. Chen, A., Weber, I. T., Harrison, R. W., and Leis, J. (2006) Identification of amino acids in HIV-1 and avian sarcoma virus integrase subsites required for specific recognition of the long terminal repeat ends, *J. Biol. Chem. 281*, 4173−4182.
10. 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.
11. Heuer, T. S., and Brown, P. O. (1997) Mapping features of HIV-1 integrase near selected sites on viral and target DNA molecules in an active enzyme-DNA complex by photo-cross-linking, *Biochemistry 36*, 10655−10665.
12. Bao, K. K., Skalka, A. M., and Wong, I. (2002) Presteady-state Analysis of Avian Sarcoma Virus Integrase. II. Reverse-Polarity Substrates Identify Preferential Processing of the U3-U5 Pair, *J. Biol. Chem. 277*, 12099−12108.
13. Bao, K. K., Skalka, A. M., and Wong, I. (2002) Presteady-state Analysis of Avian Sarcoma Virus Integrase. I. A Splicing Activity and Structure-Function Implications for Cognate Site Recognition, *J. Biol. Chem. 277*, 12089−12098.
14. Li, M., and Craigie, R. (2005) Processing of viral DNA ends channels the HIV-1 integration reaction to concerted integration, *J. Biol. Chem. 280*, 29334−29339.
15. 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.
16. Bujacz, G., Jaskolski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R. A., and Skalka, A. M. (1995) High-resolution structure of the catalytic domain of avian sarcoma virus integrase, *J. Mol. Biol. 253*, 333−346.
17. Bujacz, G., Jaskolski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R. A., and Skalka, A. M. (1996) The catalytic domain

of avian sarcoma virus integrase: Conformation of the active-site residues in the presence of divalent cations, *Structure 4*, 89−96.

18. 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. U.S.A. 95*, 9150−9154.
19. Chen, J. C., Krucinski, J., Miercke, L. J., Finer-Moore, J. S., Tang, A. H., Leavitt, A. D., and Stroud, R. M. (2000) Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: A model for viral DNA binding, *Proc. Natl. Acad. Sci. U.S.A. 97*, 8233−8238.
20. Chen, Z., Yan, Y., Munshi, S., Li, Y., Zugay-Murphy, J., Xu, B., Witmer, M., Felock, P., Wolfe, A., Sardana, V., Emini, E. A., Hazuda, D., and Kuo, L. C. (2000) X-ray structure of simian immunodeficiency virus integrase containing the core and C-terminal domain (residues 50−293): An initial glance of the viral DNA binding platform, *J. Mol. Biol. 296*, 521−533.
21. Yang, Z. N., Mueser, T. C., Bushman, F. D., and Hyde, C. C. (2000) Crystal structure of an active two-domain derivative of Rous sarcoma virus integrase, *J. Mol. Biol. 296*, 535−548.
22. Wang, J. Y., Ling, H., Yang, W., and Craigie, R. (2001) Structure of a two-domain fragment of HIV-1 integrase: Implications for domain organization in the intact protein, *EMBO J. 20*, 7333−7343.
23. Snasel, J., Krejcik, Z., Jencova, V., Rosenberg, I., Ruml, T., Alexandratos, J., Gustchina, A., and Pichova, I. (2005) Integrase of Mason-Pfizer monkey virus, *FEBS J. 272*, 203−216.
24. Wielens, J., Crosby, I. T., and Chalmers, D. K. (2005) A three-dimensional model of the human immunodeficiency virus type 1 integration complex, *J. Comput.-Aided Mol. Des. 19*, 301−317.
25. Dufour, E., El Dirani-Diab, R., Boulme, F., Fournier, M., Nevinsky, G., Tarrago-Litvak, L., Litvak, S., and Andreola, M. L. (1998) p66/p51 and p51/p51 recombinant forms of reverse transcriptase from human immunodeficiency virus type 1: Interactions with primer tRNA(Lys3), initiation of cDNA synthesis, and effect of inhibitors, *Eur. J. Biochem. 251*, 487−495.
26. Long, Y. Q., Jiang, X. H., Dayam, R., Sanchez, T., Shoemaker, R., Sei, S., and Neamati, N. (2004) Rational design and synthesis of novel dimeric diketoacid-containing inhibitors of HIV-1 integrase: Implication for binding to two metal ions on the active site of integrase, *J. Med. Chem. 47*, 2561−2573.
27. Ren, G., Gao, K., Bushman, F. D., and Yeager, M. (2007) Single-particle image reconstruction of a tetramer of HIV integrase bound to DNA, *J. Mol. Biol. 366*, 286−294.
28. Johnson, A. A., Santos, W., Pais, G. C., Marchand, C., Amin, R., Burke, T. R., Jr., Verdine, G., and Pommier, Y. (2006) Integration requires a specific interaction of the donor DNA terminal 5′-cytosine with glutamine 148 of the HIV-1 integrase flexible loop, *J. Biol. Chem. 281*, 461−467.
29. Bujacz, G., Alexandratos, J., Wlodawer, A., Merkel, G., Andrake, M., Katz, R. A., and Skalka, A. M. (1997) Binding of different divalent cations to the active site of avian sarcoma virus integrase and their effects on enzymatic activity, *J. Biol. Chem. 272*, 18161−18168.
30. Waga, S., Bauer, G., and Stillman, B. (1994) Reconstitution of complete SV40 DNA replication with purified replication factors, *J. Biol. Chem. 269*, 10923−10934.
31. Wong, I., Lundquist, A. J., Bernards, A. S., and Mosbaugh, D. W. (2002) Presteady-state analysis of a single catalytic turnover by *Escherichia coli* Uracil-DNA glycosylase reveals a "Pinch-Pull-Push" mechanism, *J. Biol. Chem. 277*, 19424−19432.

BI700197A