Coordinated Disintegration Reactions Mediated by Moloney Murine Leukemia Virus Integrase

GEORGE A. DONZELLA, COLLEEN B. JONSSON,[†] AND MONICA J. ROTH*

Department of Biochemistry, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, Piscataway, New Jersey 08854

Received 21 September 1995/Accepted 19 March 1996

The protein-DNA and protein-protein interactions important for function of the integrase (IN) protein of Moloney murine leukemia virus (M-MuLV) were investigated by using a coordinated-disintegration assay. A panel of M-MuLV IN mutants and substrate alterations highlighted distinctions between the intermolecular and intramolecular reactions of coordinated disintegration. Mismatching of the crossbone single-strand region and altered long terminal repeat (LTR) positioning affected the intermolecular, but not the intramolecular, reactions of coordinated disintegration. Partial components of the crossbone substrate were coordinated by M-MuLV IN, indicating a reliance on both LTR and target DNA determinants for substrate assembly. The intramolecular reaction was dependent on the presence of either the HHCC domain or a crossbone LTR 5' single-stranded tail. An M-MuLV IN mutant without the HHCC domain (NΔ105) catalyzed reduced levels of double disintegration but not single disintegration. A separately purified HHCC domain protein (CΔ232) stimulated double disintegration mediated by NΔ105, suggesting a role of the N-terminal HHCC domain in stable IN-IN and IN-DNA interactions. Significantly, crossbone substrates lacking the LTR 5' tails were not recognized by the fingerless NΔ105 protein. Collectively, these data suggest similar roles of the HHCC domain and 5' LTR tail in substrate recognition and modulation of IN activity.

Integration of the reverse-transcribed retroviral genome into host chromosomal DNA is an obligatory step in the establishment of productive infections by retroviruses. This step in the retroviral life cycle is dependent on the virus-encoded integrase (IN) protein and DNA sequences present in the long terminal repeat (LTR) termini of the viral DNA (1, 2, 29, 30, 34, 48, 50). Following reverse transcription *in vivo*, IN recesses the 3' ends of both LTR termini to expose a subterminal 5'-CA-3' dinucleotide that is conserved among all retroviruses and retrotransposons (1, 2, 22, 29, 30, 39, 49). The 3'-processing reaction leaves a 5' overhang (5' single-stranded [ss] tail) on the LTR, which has been implicated in the stability of IN-LTR DNA complexes *in vitro* (17, 60). IN integrates the 3'-processed viral DNA into a chromosomal site by isoenzymatic, staggered transesterification of the LTR termini and target DNA backbone (21). *In vivo* studies of Moloney murine leukemia virus (M-MuLV) integration have demonstrated that mutation of a single LTR terminus prevents processing of both viral DNA ends, suggesting that 3' processing and integration are strongly dependent on bimolecular coordination of the viral termini by IN (40). A kinetic analysis of Rous sarcoma virus IN activity suggests that multimerization is functionally critical for 3'-processing and strand transfer reactions of integration (24). However, the specific interactions that promote IN multimerization and the mechanism of LTR and target DNA coordination are not well understood. The strand transfer reaction results in ss target sequences juxtaposed to the integrated DNA. Repair of these ss gaps generates the characteristic duplication of target DNA sequences that flank the

integrated provirus (1, 2, 30). The size of this target DNA duplication is specific to each retrovirus. M-MuLV generates a 4-bp target duplication upon chromosomal integration (1, 2, 30).

Mutational analyses of the IN proteins of human immunodeficiency virus type 1 (HIV-1) (5, 15, 19, 35, 51, 54, 59, 64, 65), HIV-2 (56, 57), avian retroviruses (6, 29, 32, 33, 62), and M-MuLV (26, 27, 48, 49) have defined three distinct functional domains required for *in vitro* and *in vivo* activities. The N termini of all retroviral IN contain an HHCC zinc finger-like domain. The HHCC domain of HIV-1 IN was shown by *in vitro* blotting experiments to coordinate a Zn²⁺ cation (3, 5, 38). Coordination of Zn²⁺ by a peptide comprising the HHCC domain of HIV-1 IN induced a helical structure transition as discerned by circular dichroism spectroscopy (3). Mutational analysis of the HHCC domain of IN has not clearly defined its functional role. While some HIV-1, HIV-2, Rous sarcoma virus, and M-MuLV IN HHCC finger mutants are inactive, others retain variable levels of *in vitro* activity (5, 6, 8, 19, 26, 27, 35, 50, 51, 54, 56–58). The central core of IN contains the conserved active-site triad of acidic amino acids, DD(X₃₅)E, each of which is required for activity *in vitro* and *in vivo* (15, 19, 20, 32, 33, 52, 54, 63). Recently, the crystal structure of a mutant analog of the HIV-1 IN catalytic core (amino acids 50 to 212) was solved and found to be dimeric in the absence of a divalent cation and DNA and displays homology to the RNase H domain of HIV-1 reverse transcriptase (16). Lastly, the C-terminal half of IN is least conserved among the IN proteins and has nonspecific DNA binding activity (47, 59, 64, 65). The C termini of HIV-1 and HIV-2 are dispensable for disintegration but required for *in vitro* integration activity (5, 56, 57). In contrast, mutant M-MuLV INs that lack greater than 28 amino acids from the C terminus are nonviable *in vivo* (48) and are inactive for *in vitro* 3' processing, strand transfer, and disintegration assays (26).

In vitro assays using purified IN and oligonucleotide substrates have been developed to study the 3' processing and

* Corresponding author. Mailing address: Department of Biochemistry, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, 675 Hoes Ln., Piscataway, NJ 08854. Phone: (908) 235-5048. Fax: (908) 235-4783. Electronic mail address: Roth@mbcl.rutgers.edu.

[†] Present address: Department of Chemistry-Biochemistry, New Mexico State University, Las Cruces, NM 88003.

strand transfer reactions of integration (6, 14, 15, 19, 21, 22, 24, 25, 27, 29, 31, 32, 35, 42, 44–46) and disintegration (6, 8, 10, 11, 14, 19, 25–27, 35, 57, 58). Several distinctions regarding integration and disintegration are notable. First, the integration reaction is bimolecular in that it requires coordination of both viral DNA termini and target DNA, whereas disintegration is a unimolecular reaction. Second, the disintegration substrate only partially mimics the integration intermediate (1, 2, 22) in that the three-way DNA junction contains only one of the intermediate's two viral LTR determinants. Third, many IN mutant proteins with impaired or inactive integration are capable of mediating unimolecular disintegration (6, 19, 26, 30, 32, 35, 51, 56–58). Thus, the *in vitro* integration and disintegration assays alone do not readily discern the complexity of protein-DNA and protein-protein interactions important for IN function.

A bimolecular disintegration assay has been developed to address some of the limitations of the standard integration and disintegration assays (9, 37, 55). In this assay, the integration intermediate (crossbone) substrate is coordinated by IN from independent component oligonucleotides. In the experiments described here, we used a large series of M-MuLV IN mutants and substrate modifications to study these coordinated-disintegration reactions. The results presented highlight the functional domains of the M-MuLV IN protein and substrate features that influence IN-DNA and IN-IN interactions.

MATERIALS AND METHODS

Materials. Exonuclease-free Klenow fragment of DNA polymerase I was obtained from United States Biochemical, and terminal deoxynucleotidyl transferase was purchased from Promega. [α - 32 P]dATP and [α - 32 P]TTP were purchased from Amersham. Ni^{2+} -nitrilotriacetic acid agarose was obtained from Qiagen. Heparin (sodium salt) was obtained from United States Biochemical. *Escherichia coli* exonuclease III was obtained from New England BioLabs.

Purification of MuLV IN proteins. Hexahistidine-tagged M-MuLV IN proteins were expressed in *E. coli* BL21(DE3), purified by denaturing Ni^{2+} -nitrilotriacetic acid chromatography, and renatured as previously described (25), with the following modifications. IN was extracted and purified at either 20 or 4°C under denaturing conditions. In some cases, IN proteins were batch purified on Ni^{2+} -nitrilotriacetic acid resin prior to column chromatography. No difference in IN purity or activity was observed, regardless of the extraction and purification temperature. A minimum of two preparations of the M-MuLV IN mutants described in this report were assayed in integration (26), disintegration (26), and coordinated-disintegration reactions. The purified N-terminal deletions, C209A, and K227A proteins contain the hexahistidine tag at the C terminus, and the C-terminal deletion mutants contain the tag at the N terminus. The position of the affinity tag on wild-type (WT) M-MuLV IN did not affect any activity (26). The influence of the tag on the M-MuLV IN mutants was not determined.

Oligonucleotides. DNA oligonucleotides were prepared by the University of Medicine and Dentistry of New Jersey Department of Biochemistry DNA synthesis facility and purified by electrophoresis through 20% denaturing polyacrylamide gels. Oligonucleotides used in this study are referred to by synthesis number. The ss region of each oligonucleotide is represented in lowercase letters, and the target DNA sequence is underlined; 4166, 5'-AATGAAAGT TCTTTCAcgetAGTCCTTGGAC; 4167, 5'-AATGAAAGTTCCTTTCAagcgAGTCCTTGGAC; 5354, 5'-TGAAAGTTCCTTTCAagcgAGTCCTTGGAC; 5355, 5'-TGAAAGTTCCTTTCAagcgAGTCCTTGGAC; 4408, 5'-AATGAAAGTTCCTTTCAaattTACCTTTAGGT; 4409, 5'-AATGAAAGTTCCTTTCA(g/c)TACCTTTAGGT; 4892, 5'-AATGAAAGTTCCTTTCAttttTACCTTTAGGT; 5699, 5'-AATGAAAGTTCCTTTCAaaaaTACCTTTAGGT; 4896, 5'-AATGAAAGTTCCTTTCAttttttTACCTTTAGGT; 4897, 5'-AATGAAAGTTCCTTTCAaagaagaaTACCTTTAGGT; 4898, 5'-AATGAAAGTTCCTTTCAACCTTTAGGTATccca; 4893, 5'-AATGAAAGTTCCTTTCAACCTTTAGGTATgga; 5467, 5'-AATGAAAGTTCCTTTCAcget; 5527, 5'-TGAAAGTTCCTTTCAcget; 4985, 5'-cgctTACCTGTTTACAGGTA.

Crossbone substrate preparation. Oligonucleotides 4167 and 5355 were 3' end labeled at 20°C for 45 min in a 10- μ l volume containing 10 mM Tris (pH 7.4), 5 mM MgCl_2 , 2 mM dithiothreitol, 5 to 8 U of the exonuclease-free Klenow fragment of DNA polymerase I, and 3.7 μ M [α - 32 P]dATP. Oligonucleotides 4166 and 5354 were similarly labeled with 3.7 μ M [α - 32 P]TTP. For the experiments in which the effects of changing the ss region length and sequence were investigated, substrate oligonucleotides were 3' end labeled with terminal deoxynucleotidyl transferase. The conditions established allowed limited 3' end addition of the labeled nucleotide. Briefly, oligonucleotides 4167, 4892, 4893, 4898, 4408, and 4409 were labeled at 37°C for 30 to 45 min in a 10- μ l volume

containing 100 mM sodium cacodylate (pH 6.8), 1 mM CoCl_2 , 0.1 mM dithiothreitol, 2 to 5 U of terminal deoxynucleotidyl transferase, and 3.7 μ M [α - 32 P]dATP; oligonucleotide 4166 was similarly labeled with 3.7 μ M [α - 32 P]TTP. Labeling reactions were terminated by addition of an equal volume of stop buffer (95% formamide, 25 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol) and subsequently purified by electrophoresis through 20% denaturing polyacrylamide gels. Under these conditions, single nucleotide additions were the predominant product (estimated at >95%). The labeled oligonucleotides were gel purified, eluted, and precipitated as previously described (11). After resuspension in 0.1 M NaCl, the labeled oligonucleotides were self-annealed by heating at 80°C for 10 min and slow cooling to room temperature. Substrate specific activities typically ranged from 40 to 200 Ci/mmol (1×10^5 to 4×10^5 cpm/pmol).

Coordinated-disintegration assays. The conditions for optimal activity in the coordinated-disintegration assay were established by titration of each reaction component. The concentrations of IN used were within the linear range for disintegration activity, and the percentage of active molecules is not known. Reactions were quantitated by scanning dried gels with a Molecular Dynamics PhosphorImager and ImageQuant 3.1 software. Reaction mixtures contained 67 nM (1 pmol) substrate and 0.6 to 1 μ M (10 to 15 pmol) M-MuLV IN protein(s) in a 15- μ l final volume containing 20 mM piperazine- N,N' -bis(2-ethanesulfonic acid) (PIPES) (pH 6.4), 10 mM dithiothreitol, 10 mM MnCl_2 , 10 mM 3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate (CHAPS), 0.01% Nonidet P-40, 10 mM NaCl, and 5% ethylene glycol. When the crossbone substrate was composed of two independent oligonucleotides, each labeled substrate half was present in the reaction mixture at 33.5 nM. Standard reaction mixtures were incubated at 37°C for 2 h, terminated, and electrophoresed through 20% denaturing polyacrylamide gels as previously described (11). The sensitivity of reaction products to *E. coli* exonuclease III was determined by stopping reactions at 70°C for 15 min and subsequent proteolytic digestion with 20 μ g of proteinase K at 37°C for 30 min. Following a second heat inactivation at 70°C for 15 min, half of the reaction volume was removed, adjusted to 30 μ l with 50 mM Tris-HCl (pH 7.6)–1 mM MgCl_2 –10 mM β -mercaptoethanol, and digested with 100 U of *E. coli* exonuclease III for 0.5 to 1 h at room temperature, followed by 30 min at 37°C.

For time course studies, reactions were scaled up fivefold and aliquots were removed and terminated at different intervals. Time course reactions which were subjected to heparin trapping of initially formed IN-DNA complexes were performed as follows. The tailed crossbone substrate oligonucleotides were incubated for 5 min at 37°C with 15 pmol of WT M-MuLV IN in the presence of MnCl_2 . Following this brief incubation, 0.2 μ g of heparin was added and aliquots of the reaction mixture were terminated at different time intervals.

Complementation assays were performed by mixing each M-MuLV IN mutant at 0.5 μ M (7.5 pmol of each) at room temperature or 37°C for 5 to 15 min in reaction buffer without DNA substrates. The reaction volume was then adjusted to 15 μ l by substrate addition and incubated at 37°C for 2 h. For these reactions, NaCl was present at 50 mM. Similar complementation results were obtained under multiple assay conditions. These included assays performed in the absence of NaCl, preincubation of the M-MuLV IN mutant proteins with MnCl_2 , or brief preincubation of the mutants with DNA components prior to addition of the divalent cation (data not shown). For the component half assays, reactions contained either tailed LTR half (5467), untail LTR half (5527), or target half (4985) oligonucleotides at 33.5 nM and either crossbone oligonucleotide 4166, 4167, 5354, or 5355 at 33.5 nM. For competition reactions, reaction conditions were modified as follows. The tailed competition reactions contained all reaction buffer components and each tailed half crossbone (oligonucleotides 4166 and 4167) at 33.5 nM. The different component halves (oligonucleotides 5467, 5527, and 4985) were then titrated into each reaction. The reaction mixtures were briefly preincubated at 37°C and initiated by addition of 1 μ M M-MuLV IN. Both the component half and competition reaction mixtures were incubated at 37°C for 1 h. Untailed crossbone competition experiments were similarly performed with each untail LTR half crossbone (oligonucleotides 5354 and 5355) at 33.5 nM.

RESULTS

M-MuLV IN coordinated-disintegration reactions (crossbones). To more accurately reconstruct the integration intermediate structure, a disintegration assay which is dependent on coordination of separate component oligonucleotides was utilized (Fig. 1A). The substrate in the M-MuLV coordinated-disintegration assay was a 3'-end-labeled DNA hairpin that contains self-complementary LTR and target DNA sequences which are separated by an ss region (half crossbone). Coordination of two half-crossbone oligonucleotides by M-MuLV IN generates an integration intermediate DNA structure (crossbones) and promotes activity at one or two 3' ends of the substrate. Since the assay temperature is well above the temperature required for efficient annealing of the ss region, co-

FIG. 1. M-MuLV IN coordinated disintegration. (A) Substrate determinants and basis of the assay. Each half-crossbone oligonucleotide contains a self-complementary 5-bp LTR and a 5-bp target DNA domain. The 3'-end label on

ordination of the crossbone substrate may be mediated through IN-DNA and IN-IN interactions between bound half crossbones. The coordinated-disintegration assay may therefore reflect not only the bimolecular nature of integration but also aspects of the multimeric protein-protein interactions intrinsic to IN function.

Incubation of the half-crossbone oligonucleotides with M-MuLV IN yielded several distinct reaction products (Fig. 1A and B). Substrate coordination resulted in intermolecular transesterifications at one or both 3' ends of the crossbone structure and are respectively referred to as single and double disintegration products (Fig. 1B, lane 2). A faster-migrating product of the reaction, termed foldback, results from an intramolecular nucleophilic attack of an oligonucleotide's 3'-OH group on its own LTR-target DNA junction. The foldback reaction yields two distinct product species because of intramolecular activity on the two different half crossbones; their migration differences are likely due to the effect of their sequence differences on product structure (see Fig. 6B, foldback products in lanes 2 and 3). Lastly, a 16-nucleotide (nt) endonucleolytic cleavage product which may correspond to hydrolysis of the LTR-target DNA junction was also generated and is referred to as an LTR hydrolysis product. Activity in the coordinated-disintegration assay was absolutely dependent on Mn^{2+} , and no activity was noted with Mg^{2+} (data not shown). The structures of the coordinated-disintegration reaction products were investigated by subjecting half of the reaction mixture to digestion with *E. coli* exonuclease III. DNA molecules with a 3'-OH present in a blunt or 3' recessed duplex structure are sensitive to this 3'→5' exonuclease. The unreacted substrate, single disintegration, and LTR hydrolysis products were sensitive to exonuclease III digestion. The double-disintegration and foldback reaction products were not sensitive to exonuclease III, indicating that these products are covalently closed circular molecules (Fig. 1B, lane 3). The nature of these reaction products was further confirmed by limited digestion with nuclease P1, which converted the circular double-disintegration and foldback reaction products to their expected linear sizes (32 and 16 nt, respectively; data not shown). Additionally, the single disintegration product comigrated with a 5'-end-labeled dumbbell oligonucleotide that mimics the structure of this product, indicating their structural similarity (data not shown).

A feature of the coordinated crossbone substrate is the 5' ss tails of the LTRs. These ss dinucleotide tails represent the 5' overhangs which remain after 3' processing of the viral DNA and are presumably removed in vivo during or after strand transfer by an uncharacterized repair process. An in vitro role for the 5' ss tails of LTR substrates in the maintenance of a stable IN-LTR complex has been proposed (17, 60). The influence of the LTR 5' ss tails on crossbone substrate coordination was addressed by using half-crossbone oligonucleotides which lacked the LTR 5' ss tails but had sequences otherwise identical to that of the standard crossbone substrate. Reactions with the untailed substrate demonstrated that the LTR 5' ss tails were not required for coordination of the oligonucleotides into the crossbone substrate structure: M-MuLV IN was capable of mediating the intermolecular single and double disinte-

each half crossbone is indicated by a filled circle. (B) Coordinated disintegration of crossbone substrates containing (lanes 1 to 3) or lacking (lanes 4 to 6) LTR 5' ss tails. Lanes: 1 and 4, control reaction buffer incubations; 2 and 5, incubation with M-MuLV IN; 3 and 6, *E. coli* exonuclease III digestion of reaction mixtures. Reaction product structures are indicated beside their respective migration positions in the gel.

gration reactions (Fig. 1B, lanes 5 and 6). However, elimination of the 5' tail caused a greater-than-10-fold reduction in generation of the foldback product in multiple independent assays. Additionally, the LTR hydrolysis reaction product was not produced in the absence of the LTR 5' tail. These results suggest a distinction between the intermolecular and intramolecular reactions that is delineated by the 5' ss tails of the crossbone LTR domain.

Coordinated-disintegration reaction time courses. Reactions with the tailed and untailed crossbone substrates were monitored by time course analysis. The time course reactions were performed with either one or both of the oligonucleotide components of the crossbone substrate to determine the effect of partial or full sequence complementarity of the ss region on substrate coordination. Coordination of two of the same half-crossbone oligonucleotide molecules by M-MuLV IN generated a crossbone structure in which only the central two nucleotides of the 4-nt ss region were complementary (Fig. 2A, left panel). Inclusion of two complementary half-crossbone oligonucleotides in the reaction time course allowed the coordination of a crossbone structure in which the ss region had full base-pairing capacity (Fig. 2A, right panel).

Reactions containing a single-tailed half crossbone indicated that M-MuLV IN was capable of crossbone substrate coordination in the absence of full ss region complementarity (Fig. 2B, left panel). However, partial complementarity resulted in 18-fold less efficient single disintegration than reactions that contained a complementary ss region and there was no double disintegration (Fig. 2B, right panel). This suggested that the rates of intermolecular single and double disintegration differed and were affected by the complementarity of the ss region sequence (data not shown). The yield of the intramolecular foldback product did not significantly vary, irrespective of ss region complementarity (Fig. 2B). In time course reactions containing a single untail oligonucleotide, the lack of complementarity in the ss region combined with the absence of the LTR 5' ss tails severely impaired crossbone substrate coordination (Fig. 2C, left panel). Efficient coordinated disintegration of the untail crossbone substrate required full ss region complementarity (Fig. 2C, right panel), suggesting that the absence of the LTR 5' ss tails further sensitized M-MuLV IN to mispairing of the ss region.

Although the reaction time course indicated that single and double disintegration were catalyzed at different rates, it could not readily be determined if double disintegration results from either processive or coordinated M-MuLV IN activity. To address this question more directly, the coordinated disintegration reaction was performed with the addition of a heparin trap. IN was briefly incubated with the tailed crossbone substrate in the presence of a divalent cation to allow the formation of IN-DNA complexes. After this brief reaction pulse, a time point was taken and heparin was added to trap free enzyme and block subsequent formation of new IN-DNA complexes (Fig. 2D). After heparin addition, single disintegration levels remained constant throughout the reaction's time course (Fig. 2D, top panel). Significantly, none of the single disintegration product was converted to a double disintegration product after addition of heparin (Fig. 2D, bottom panel). In the absence of the LTR 5' ss tails, coordinated disintegration of the untail crossbone substrate yielded similar results in the presence of a heparin trap (data not shown). This result indicates that double disintegration does not arise from processive activity of the M-MuLV IN protein bound in the single disintegration complex. Similar results were obtained when initial IN-DNA complexes were formed under noncatalytic conditions with Mg^{2+} , followed by addition of the heparin trap and

initiation with Mn^{2+} (data not shown). Extension of the time of initial complex formation with Mg^{2+} prior to addition of the heparin trap (5 to 60 min) did not enhance conversion of the single disintegration product to a double disintegration product (data not shown). No activity was observed when the divalent cation was omitted until after addition of the heparin trap, indicating that stable coordination of the crossbone substrate required metal-dependent IN-DNA interactions (data not shown).

Effect of crossbone substrate LTR positioning on coordinated disintegration. M-MuLV integration in vivo results in the duplication of host target sequences flanking the integrated provirus (1, 2, 30). The virus-specific duplication of host target sequences correlates with a 4-bp staggered attack by the LTRs on the target DNA. This precise coordination of the viral LTRs and target DNA may be the result of specific protein-protein interactions important for assembly of the preintegration complex and concerted integration (40). Analogously, coordination of the crossbone substrate oligonucleotides recreates the site of M-MuLV integration, with the two LTR domains staggered by 4 bp with respect to the target DNA region of the substrate. Half-crossbone oligonucleotides with altered ss region lengths were used to examine the possible dependence of coordinated disintegration on crossbone LTR position. Substrates were designed such that the LTRs would be staggered by 1 (X1), 6 (X6), or 9 (X9) bp in the coordinated crossbone structure (Fig. 3A). Additionally, a substrate was designed with the 4-nt ss region present as a 3' overhang instead of being recessed within the LTR-target DNA junction (X^{-4}). The coordination of the X^{-4} oligonucleotides would therefore invert the LTR positions of the integration site with respect to the standard crossbone substrate. The influence of the sequence within the ss region was also addressed by using oligonucleotides that have ss regions containing all thymidylate (X4T) or adenylate (X4A) residues.

Each of these substrates was assayed under identical conditions and analyzed for sensitivity to *E. coli* exonuclease III to facilitate identification of circular products. Various LTR hydrolysis product yields were obtained with each of the X substrate variants, making their definitive identification difficult. M-MuLV IN was able to coordinate the X1 and X6 substrates and generate the intermolecular single (product B) and double disintegration products (product C) (Fig. 3B, X1, lanes 1 to 3, and X6, lanes 1 to 3), with efficiencies reproducibly similar to that observed with the standard crossbone substrate (Fig. 3B, X4, lanes 1 to 3). Although separation of the crossbone LTRs by 9 bp did not affect single disintegration, double disintegration of the X9 substrate was not detected (Fig. 3B, X9, lanes 1 to 3). For the X^{-4} substrate, inversion of the LTR positions in the crossbone structure resulted in no double disintegration and reduced levels of single disintegration; however, M-MuLV IN was still capable of mediating the intramolecular foldback reaction (Fig. 3B, X^{-4} , lanes 1 to 3). Little variation in the yield of the intramolecular foldback product was detected by varying the length of the ss region (X1, X6, X9, and X^{-4} substrates; Fig. 3B). These results indicate that double disintegration is more sensitive to LTR position than are single disintegration and intramolecular foldback reactions.

The effect of the ss region sequence and mispairing was further investigated with the X4T and X4A half-crossbone oligonucleotides (Fig. 3B, X4T/X4A, lanes 1 to 7). Reaction mixtures that contained only the X4T (Fig. 3B, lanes 2 and 5) or X4A (Fig. 3B, lanes 3 and 6) half crossbone demonstrated that the intermolecular single disintegration and intramolecular foldback products were relatively unaffected by the mispaired ss region; however, double disintegration levels were

reduced. Reaction mixtures containing both the X4A and X4T half crossbones demonstrated only modest enhancement of single and double disintegration (Fig. 3B, lanes 4 and 7) compared with reaction mixtures with either half crossbone alone (Fig. 3B, lanes 2 and 3). Foldback reaction levels were unaffected. These results indicate that the sequence, as well as the complementarity, of the ss region influenced intermolecular disintegration.

Coordinated-disintegration activity of M-MuLV IN mutants. By using a panel of purified M-MuLV IN mutants (Fig. 4A), the requirement of specific domains of IN for coordinated disintegration was addressed. M-MuLV HHCC mutants include *in5247*, which contains a linker insertion that changes the spacing of the conserved histidines, and C97Y (*bi5349*), which contains a single point mutation of the Zn²⁺ finger-like domain (50). NΔ105 and NΔ176 contain N-terminal deletions of the first 105 and 176 amino acids of M-MuLV IN and thus lack the HHCC finger domain. The NΔ105 mutant retains an intact DD(X₃₅)E motif; however, the N terminus of NΔ176 initiates downstream of the first aspartate of the active site and thus contains only the D(35)E region of the catalytic triad. K227A contains a mutation of a lysine residue that is conserved in all known retroviral IN proteins (19). C209A is a mutant with a point mutation within the active site which eliminates the third and sole non-HHCC cysteine residue present in M-MuLV IN. CΔ22, CΔ28, CΔ40, CΔ45, and CΔ86 are named for the number of amino acids deleted from the C terminus of M-MuLV IN. Studies of these C-terminal deletion mutants in vivo demonstrated that the terminal 28 amino acids were dispensable for establishment of the M-MuLV provirus (48). The CΔ232 and CΔ144 IN mutants contain large deletions of the C terminus and terminate within and just beyond the active site, respectively.

The mutant M-MuLV IN proteins were assayed for coordinated disintegration of the tailed crossbone substrate (Fig. 4B). *in5247* and C97Y (Fig. 4B, lanes 2 and 3) yielded low levels of intermolecular disintegration products and no intramolecular foldback product. In contrast, NΔ105 generated intermolecular disintegration reaction products but was severely compromised for the foldback reaction (Fig. 4B, lane 4). CΔ232, CΔ144, and NΔ176 were inactive (Fig. 4B, lanes 5, 6, and 7, respectively). The K227A mutant was deficient for coordinated disintegration, demonstrating only very limited levels of single disintegration (Fig. 4B, lane 8). Although the C209A point mutant demonstrated single and double disintegration levels similar to those of the WT, foldback was reduced two- to threefold (Fig. 4B, lane 9). M-MuLV IN mutants with the C terminus truncated by 22 (CΔ22, Fig. 4B, lane 10) or 28 (CΔ28, Fig. 4B, lane 11) amino acids were able to catalyze both intermolecular and intramolecular reactions. Further C-terminal deletions abolished all activity (CΔ40, CΔ45, and CΔ86; Fig. 4B, lanes 12 to 14).

The partial or full deficiency of some of the M-MuLV IN mutants for coordinated disintegration may result from the lack of domains required for IN-DNA and IN-IN interactions. Alternatively, these mutations may affect proper folding of the enzyme. These possibilities were investigated by complementation analysis with mixed pairs of M-MuLV IN mutants. Domains of IN required for multimeric IN-IN interactions can be identified by the ability of defective IN mutants to complement each other in *trans*, yielding an active enzyme (9, 37, 57). For these complementation reactions, the concentration of NΔ105 typically assayed (15 pmol) was halved and mixed with an equimolar amount of a second, inactive mutant. At a limiting protein (7.5 pmol), NΔ105 displayed reduced levels of intermolecular single disintegration; double disintegration was re-

duced to a much larger extent (Fig. 4B, lane 15; see also below). The poor ability of NΔ105 to mediate the intramolecular foldback reaction and efficient double disintegration was remedied by providing the HHCC domain in *trans* from the inactive C-terminal deletion mutant CΔ38, CΔ41, CΔ86, CΔ232, or CΔ144 (Fig. 4B, lanes 16 to 20). When reactions were performed with the standard concentration of NΔ105 (15 pmol), addition of any of these proteins resulted in only slight stimulation of the basal levels of intermolecular disintegration and foldback (data not shown). In contrast, neither NΔ176 nor any of the C-terminal deletion mutants could be complemented by a broad titration range of these proteins (data not shown). The HHCC finger mutants C97Y and *in5247* also could not be complemented by the M-MuLV IN C-terminal deletion proteins, implying that a mutant HHCC domain can dominantly interfere with multimeric IN-IN interactions (data not shown).

The complementation of NΔ105 by the HHCC domain was further investigated. Titration of NΔ105 alone in the coordinated-disintegration assay was performed by using the tailed crossbone substrate (Fig. 4C, left panel). At very low enzyme concentrations, NΔ105 generated low single disintegration product levels and virtually no double disintegration product; however, with higher enzyme concentrations, the double disintegration product yield increased (Fig. 4C, lanes 1 to 7). To determine the effect of the HHCC domain on limiting NΔ105 activity, complementations were performed by titrating the CΔ232 protein into these reactions (Fig. 4C, center and right panels). When the concentration of CΔ232 was varied against a fixed, low level of NΔ105 (2.5 pmol; Fig. 4C, center panel, lane 8), 2-fold maximal enhancement of single disintegration and, more dramatically, up to 55-fold maximal stimulation of double disintegration were observed (Fig. 4C, center panel lanes 9 to 13). When titrated against an intermediate concentration of NΔ105 (7.5 pmol; Fig. 4C, right panel, lane 14), CΔ232 enhanced levels of single disintegration only 1.2-fold yet maximally stimulated double disintegration 8.5-fold (Fig. 4C, right panel, lanes 15 to 19). Interestingly, less dramatic stimulation of NΔ105 by the CΔ232 protein was seen when preparations of NΔ105 which were refolded at concentrations higher than 2 to 5 μM were used; possibly because of an increased aggregation state. In support of this observation, a titration of dithiothreitol with more-concentrated preparations of NΔ105 resulted in stimulation of integration activity (26; data not shown). The impaired activity of NΔ105 at limiting concentrations could thus be overcome either by saturating enzyme concentrations or provision of the HHCC domain in *trans*. These effects may be a result of enhanced protein-protein interactions and/or stabilized protein-DNA interactions.

By using the untailed half crossbone, the influence of the LTR 5' ss tails on the coordinated disintegration activity of the M-MuLV IN mutants was investigated (Fig. 4D). In reactions with the untailed substrate, the HHCC finger mutants (Fig. 4D, lanes 3 and 4), as well as the fingerless NΔ105 mutant (Fig. 4D, lane 5), were incapable of substrate coordination, further emphasizing the importance of the HHCC domain. It was possible that the inability of the fingerless mutant to discriminate the untailed crossbone's LTR and target DNA domains accounted for this result. However, an untailed crossbone substrate composed entirely of LTR sequences was not efficiently coordinated by M-MuLV IN (data not shown). With the untailed crossbone substrate, the yields of the intermolecular single and double disintegration products obtained with the C209A mutant were fourfold lower than those observed with the tailed crossbone substrate (Fig. 4D, lane 10, and B, lane 9). C-terminal deletion mutants CΔ22 and CΔ28 were capable of

FIG. 3. Effect of LTR position on coordinated disintegration. (A) The standard crossbone substrate (4166-4167) is indicated as X4. The X1 and X6 substrates are both derived from oligonucleotides having self-complementary ss regions (4609 and 4608, respectively). The X9 substrate is derived from oligonucleotides 4896 and 4897; X-4 is derived from oligonucleotides 4893 and 4898, and the X4T/X4A substrate (not depicted) is derived from oligonucleotides 4892 and 5699. (B) Activity of M-MuLV IN with each of the crossbone substrate variants. Substrates and product bands: a, input substrate; b, single disintegration reaction product; c, double disintegration reaction product; d, foldback reaction product. Lanes: 1, control buffer incubation; 2, incubation with M-MuLV IN; 3, *E. coli* exonuclease III digestion of the terminated reaction. For the X4T/X4A substrate, lane 1 is the control, lanes 2 and 3 are independent incubations of the X4T and X4A oligonucleotides, and lane 4 is coinubation of X4T and X4A with M-MuLV IN; lanes 5 to 8, respectively, are *E. coli* exonuclease III digestion of the terminated X4T, X4A, and X4T/X4A reaction mixtures. Double disintegration of the X4A/X4T substrate yielded three product species: one each resulting from disintegration of a mispaired X4T or X4A crossbone substrate and one from the X4T/X4A coordinated substrate.

coordinated disintegration with the untaild crossbone substrate (Fig. 4D, lanes 11 and 12); however, two- to fourfold lower levels of single and double disintegration were observed relative to the tailed crossbone reactions. CA232 (Fig. 4D, lane 6), CA144 (Fig. 4D, lane 7), NA176 (Fig. 4D, lane 8), K227A (Fig. 4D, lane 9), CA40 (Fig. 4D, lane 13), CA45 (Fig. 4D, lane 14), and CA86 (Fig. 4D, lane 15) were inactive for untaild coordinated disintegration. The inability of NA105 to coordinate the untaild crossbone substrate was remedied by complementation with an inactive CA40, CA45, CA86, CA232, or CA144 deletion mutant (Fig. 4D, lanes 17 to 21). Thus, in the absence of the LTR 5' ss tails, substrate coordination demonstrated a marked dependence on the HHCC domain of IN. Collectively, these data suggest that the HHCC domain and LTR 5' ss tails possess similar functions in promoting effective protein-DNA and protein-protein interactions.

Influence of LTR and target DNA determinants on substrate coordination. To investigate the reliance of M-MuLV IN on LTR and target DNA determinants in mediating substrate coordination, reactions were performed with a half-crossbone oligonucleotide and oligonucleotides that represent these regions of the crossbone substrate's target or LTR domain. Reactions containing these component half oligonucleotides can result in the bimolecular coordination of a partially mispaired crossbone substrate, as shown in Fig. 2A (left panel), or coordination of a single half crossbone and a component half oligonucleotide (Fig. 5A). The coordinated substrate in the target component half reactions (Fig. 5A, left) would contain two target half sites and a single LTR region. In the LTR component half reactions, coordination would produce a substrate with two LTR domains and a single target half site (Fig. 5A, right). These partial substrate reactions potentiate disintegration at a single 3' end, yielding labeled products that are 4 nt larger than the input substrate. The roles of the LTR 5' ss tails and HHCC domain of IN in these component half reactions were also determined by using the fingerless NA105 mutant with the tailed and untaild substrate components. The different substrate combinations assayed and the activity of the WT and fingerless NA105 M-MuLV IN proteins for these component half reactions are summarized in Fig. 5B.

In the absence of component half oligonucleotides, WT M-MuLV IN did not efficiently mediate coordinated disintegration with either of the tailed half crossbones (Fig. 5C, lanes 3 and 4) because of the lack of ss region complementarity. Similarly, the fingerless NA105 mutant was inactive with either tailed half crossbone by itself (Fig. 5C, lanes 5 and 6). These controls indicated that the generation of a product 4 nt larger than the input substrate would signify component half coordination.

WT M-MuLV IN and the fingerless NA105 mutant were incubated with either a complementary (even lanes) or non-complementary tailed half crossbone (odd lanes) and each of the component halves (Fig. 5B, substrates A, B, and C, and Fig. 5C). WT M-MuLV IN properly coordinated the tailed crossbone with the untaild LTR (substrate A; Fig. 5C, lane 8), tailed LTR (substrate B; Fig. 5C, lane 12), and target DNA (substrate C; Fig. 5C, lane 16) component halves. In the absence of the HHCC finger, NA105 was able to coordinate the untaild LTR (substrate A; Fig. 5C, lane 10), tailed LTR (substrate B; Fig. 5C, lane 14), and target DNA (substrate C; Fig. 5C, lane 18) component halves with a complementary tailed half crossbone. Complementarity of the ss region was an absolute requirement for efficient component half coordination for the WT (Fig. 5C, lanes 7, 11, and 15) and fingerless NA105 (Fig. 5C, lanes 9, 13, and 17) M-MuLV IN proteins.

Notably, NA105 coordinated the tailed half crossbone with

FIG. 4. Activity of mutant M-MuLV IN and complementation assays. (A) Schematic map of WT and mutant M-MuLV IN proteins assayed. (B) Mutant activities with the tailed crossbone substrate (oligonucleotides 4166 and 4167). The M-MuLV IN protein assayed is indicated above each reaction lane (lanes 1 to 14); lanes 15 to 20 are N Δ 105 complementations with the indicated mutants. (C) Stimulation of N Δ 105 double disintegration by saturating enzyme concentration and C Δ 232. Lanes: 1, control; 2 to 7, titration of N Δ 105; 8 to 13, titration of C Δ 232 versus 2.5 pmol of N Δ 105; 14 to 19, C Δ 232 titration versus 7.5 pmol of N Δ 105. (D) Mutant activities with the untail crossbone substrate (oligonucleotides 5354 and 5355). Lane 1 is a control buffer incubation. Mutants assayed are indicated above lanes 2 to 15, and lanes 16 to 21 are complementations of N Δ 105 with the mutants indicated.

the untailed LTR or target DNA component half (Fig. 5B, substrates A and C) with slightly lower efficiencies than did WT IN. However, these reactions were not severely impaired by the absence of the HHCC finger, indicating that the presence of a single LTR 5' ss tail obviated the necessity of the HHCC domain in the component half reactions. This suggested that the fingerless NΔ105 mutant would be distinctly dependent on the presence of at least one LTR 5' ss tail for component half coordination. This possibility was directly addressed in the component half reactions by using the untailed half crossbone (Fig. 5B, substrates D, E, and F, and Fig. 5D). WT M-MuLV IN was able to coordinate the untailed LTR (substrate D; Fig. 5D, lane 8), tailed LTR (substrate E; Fig. 5D, lane 12), and target DNA (substrate F; Fig. 5D, lane 16) component halves with the untailed half crossbone. For each of these reactions, only the component halves complementary to the untailed half crossbone ss region were coordinated (Fig. 5D, odd-numbered lanes). In the absence of the HHCC domain, NΔ105 efficiently coordinated the complementary tailed LTR component half (substrate E; Fig. 5D, lane 14) because of the presence of a single LTR 5' ss tail in the reaction. In the absence of any LTR 5' ss tail, NΔ105 was incapable of coordinating the untailed LTR (substrate D; Fig. 5D, lane 10) or target DNA (substrate F; Fig. 5D, lane 18) component half. Collectively, these data imply that the LTR 5' ss tails and the HHCC domain of IN similarly affect IN-IN interactions and stable intermolecular coordination of LTR and target DNA determinants.

Substrate determinant competition experiments. Since coordination of a mispaired crossbone substrate is inefficient, the influence of the LTR and target DNA component half reactions on coordinated disintegration of the complete crossbone substrates that have or lack the LTR 5' ss tails was addressed (Fig. 6A). The component half coordination products in this assay were the same as those shown in Fig. 5A. Mispaired coordination of either tailed half crossbone was less efficient than coordination of the fully complementary tailed crossbone substrate (Fig. 6B, compare lanes 2 and 3 with lane 4). This demonstrated the basal level of coordinated disintegration in the absence of any component half reaction. With the target component half included in the assay, decreased levels of single and double disintegration were observed, while a lesser effect on LTR hydrolysis and the foldback reaction was noted (Fig. 6B, lanes 5 to 10). Sharper reductions of single and double disintegration, LTR hydrolysis, and foldback reactions were observed when the tailed LTR component half was included in the reaction mixture (Fig. 6B, lanes 11 to 16). With an untailed LTR component half, the levels of single and double disintegration decreased as the concentration of the untailed LTR component half competitor increased (Fig. 6B, lanes 17 to 22). The untailed LTR component half reaction had a lesser effect on the intramolecular foldback reaction of tailed coordinated disintegration (Fig. 6B, lanes 17 to 22) than did the tailed LTR component half reaction (Fig. 6B, lanes 11 to 16). Significantly, the tailed LTR component half reaction reached a plateau immediately, whereas the target component and untailed LTR component half products reached a plateau at higher molar concentrations (bottom panel of Fig. 6B).

Coordinated disintegration of the tailed crossbone substrate was never completely eliminated by competition with any of the component half reactions. This suggested that the LTR 5' tails of the crossbone substrate contribute to a stable IN-DNA complex and decrease the access of component half competitor DNA. To address this possibility, the effect of the component half reactions on coordinated disintegration of the untailed crossbone substrate was investigated (Fig. 6C). WT M-MuLV IN inefficiently mediated mispaired coordination of either un-

tailed half crossbone (Fig. 6C, lanes 2 and 3) and efficiently coordinated the two complementary half-crossbone oligonucleotides to form the untailed crossbone substrate (Fig. 6C, lane 4). In the presence of increasing concentrations of the target component half, decreased levels of single and double disintegration were observed (Fig. 6C, lanes 5 to 10). Compared with tailed crossbone competition, double disintegration of the untailed crossbone substrate was more strongly affected by the target component half titration (compare Fig. 6B and C, lanes 5 to 10). The tailed LTR component half reaction affected both single and double disintegration of the untailed crossbone substrate (Fig. 6C, lanes 11 to 16). The untailed LTR component half was least efficient in reducing single and double disintegrations of the untailed crossbone substrate (Fig. 6C, lanes 17 to 22). These data suggest a subtle influence of both LTR and target DNA determinants on stable coordination of the crossbone substrate by M-MuLV IN.

DISCUSSION

In these studies, we investigated aspects of substrate recognition and multimeric M-MuLV IN functions by using a bimolecular, coordinated-disintegration assay. Nuclease sensitivity of the coordinated-disintegration reaction products defined two prominent product classes: intermolecular transesterification products resultant of disintegration at either or both LTR-target DNA junctions (single and double disintegrations) and an intramolecular transesterification product (foldback) that results from nucleophilic attack of a substrate oligonucleotide's 3'-OH on its own LTR-target DNA junction. The distinctions between these intermolecular and intramolecular reactions were demonstrated in several experiments. First, substrates lacking LTR 5' ss tails yielded less intramolecular foldback reaction product, whereas the intermolecular reactions were insensitive to absence of the 5' ss tails. This result suggests that these intermolecular and intramolecular reactions result from recognition and coordination of alternative substrate conformations. Second, partial, as well as full, mispairing of the crossbone ss region diminished the efficiency of intermolecular disintegration but had a lesser effect on the intramolecular foldback reaction. Third, a mutant with the HHCC domain deleted (NΔ105) was fully capable of mediating the intermolecular disintegration reactions but was deficient for the intramolecular foldback reaction. Lastly, altered length of the ss region had little effect on the intramolecular foldback reaction.

The dependence of foldback on the 5' ss tails of the crossbone LTRs and the HHCC domain of IN is similar to their requirement for efficient 3' processing and strand transfer. Retroviral IN proteins can mediate 3' processing by endonucleolytic cleavage of a 3'-terminal dinucleotide or by an intramolecular transesterification reaction using the terminal 3'-OH group of the LTR DNA (14, 21, 34, 39, 61). The intramolecular nature of the foldback reaction suggests that it is analogous to the formation of the cyclic dinucleotide generated by *in vitro* 3' processing. The data presented here suggest that the intermolecular and intramolecular reactions of coordinated disintegration represent different substrate recognition mechanisms. Whether foldback is dependent on or independent of bimolecular coordination is not clear. One proposed model suggests that the intramolecular foldback reaction results from noncoordinated binding of a single half-crossbone oligonucleotide (9, 37, 55). In this respect, the foldback reaction is similar to disintegration in being unimolecular. Alternatively, foldback might equally result from coordination of two half crossbones in a different configuration that is independent of the ss region sequence or conformation. For exam-

FIG. 6. Competition of coordinated disintegration by component half reactions. (A) Schematic basis of the assay. (B) Competition of coordinated disintegration of the tailed crossbone substrate. The structure and concentration of each component half in the reaction mixture are indicated above the lanes. Structures of the reaction products are indicated on the left. The LTR hydrolysis product is indicated by the letter L. The lower panel is a light exposure showing the competition products. Lanes: 1, control; 2, M-MuLV IN incubation with half crossbone 4166; 3, M-MuLV IN incubation with half crossbone 4167; 4, coin-cubation of half crossbones 4166 and 4167 with IN; 5 to 10, target half competition; 11 to 16, tailed LTR half competition; 17 to 22, untailed LTR half competition. (C) Competition of coordinated disintegration of the untailed crossbone substrate. The lower panel is a light exposure showing the competition products. Lanes are as shown in panel B, except that lane 1 is a control buffer incubation, lanes 2 to 4 are independent incubations of half crossbones 5354 and 5355 and coin-cubation of 5354 and 5355 with M-MuLV IN, respectively.

ple, binding of the LTR domain and target DNA domain of a crossbone in *cis* by one protomer independently or within a multimeric complex may promote the intramolecular foldback reaction, whereas when it is bound in *trans* by different protomers, intermolecular activity is potentiated. This concept is not without precedent in the family of polynucleotidyl transferases. For F1p recombination, *cis* and *trans* modes of DNA half-site binding mediated by distinct F1p monomers within a complex were respectively correlated with intramolecular and intermolecular activity (7). Given that M-MuLV IN requires coordination of two functional viral termini to process either in

vivo (40), the dependence of foldback on bimolecular coordination is therefore plausible. Similarly, asymmetric recognition of the retroviral termini and target DNA by distinct HIV-1 IN protomers has been proposed (8).

For coordinated disintegration, M-MuLV IN was surprisingly tolerant of altered crossbone LTR positions. Such flexibility in the mechanism of crossbone substrate coordination might be accounted for by maintenance of proper IN-IN interactions through deformation of the ss region. It is possible that in the absence of specific cellular factors (28, 36, 41–46), proper substrate coordination relies solely on the structural flexibility of naked DNA (23, 53) and the ability of IN to exploit such a feature. Indeed, Pryciak et al. have demonstrated that the exposed face of nucleosome-assembled DNA and native chromatin may serve as preferred integration targets (44, 46). An extension of that work detected enhanced integration events near regions of local DNA distortion potentiated by nucleosome cores on DNA structure (42, 43).

The reduced efficiency of substrate coordination caused by mispairing of the crossbone ss region was demonstrated by

several experiments. The time course of coordinated disintegration, altered sequence and length of the ss region, and the component half reactions all demonstrated reduced product yields when the ss region was mispaired. Absence of the LTR 5' ss tails also made M-MuLV IN more vulnerable to mispairing of the ss region. Similar effects of mispairing on bimolecular disintegration were noted in several other studies (37, 55). Although one study found that Y substrate disintegration was insensitive to a single base mismatch in the integration site (8), another study found that complete mispairing of the integration site resulted in significant reduction of activity (37). This suggests that integration site mispairing can diminish productive IN-DNA interactions, even in a unimolecular context. Previous work demonstrated that chemical modification or mutations in one domain of IN can make M-MuLV IN susceptible to structural changes in the disintegration substrate (27). Given these data, ss region mispairing in the crossbone substrate may similarly destabilize coordinated IN-IN interactions.

The sensitivity of double disintegration to the absence of the HHCC domain (Fig. 4C) and ss region mispairing suggests a dependence on coordinated IN-IN interactions. Heparin-trapped time course reactions demonstrated that initial IN-DNA complex formation required a divalent cation (Mn^{2+} or Mg^{2+}) and indicated that the M-MuLV IN protein in the coordinated complex does not processively execute double disintegration. Processive IN activity was also not observed with heparin-trapped complexes formed in Mg^{2+} for longer time periods. These results suggest that double disintegration requires IN-IN interactions within a multimeric IN-DNA complex or independent binding of IN to a single disintegration product that is free in solution.

Evidence of a functional role for the HHCC domain in influencing IN-IN and IN-DNA interactions was suggested by the nonproportional levels of single and double disintegrations demonstrated by titration of $\Delta 105$. One interpretation of this result is that at limiting enzyme concentrations, $\Delta 105$ IN-IN interactions are impaired by the lack of the HHCC domain. This could be overcome by increasing the enzyme concentration; similarly, titration of the HHCC domain ($\Delta 232$) could assist in $\Delta 105$ interactions. A second model suggested by these data is that the $\Delta 105$ -DNA complex may be stabilized *in trans* by the presence of the HHCC domain, thereby promoting coordination of a second IN protomer for double disintegration. Furthermore, $\Delta 105$ was incapable of coordinating the crossbone substrate lacking the LTR 5' ss tails. Summarily, bimolecular substrate coordination displayed a marked dependence on the presence of either the HHCC domain or LTR 5' ss tail, suggesting their possible functional redundancy within a multimeric context.

The coordination of the LTR and target DNA component halves indicated that features of both determinants are recognized by M-MuLV IN for assembly of the crossbone substrate. The elimination of the LTR 5' ss tails from these component half reactions indicated a critical requirement of the HHCC domain for stable substrate coordination. Although required for integration, the HHCC domain and LTR 5' ss tail are dispensable for unimolecular disintegration (8, 10, 17, 58). The component half reactions reduced the levels of coordinated disintegration of the optimal, fully complementary crossbone substrate. This was more evident when untailed coordinated disintegration was assayed: in the absence of the crossbone substrate LTR 5' ss tails, the component half determinants affected double disintegration more strongly than single disintegration. These results are consistent with a possible role of the 5' LTR tail in the maintenance of a stable IN-DNA com-

plex (17, 60) and also suggest that ss region mispairing can reduce complex stability, thereby allowing the partial component half reactions to predominate. It has been proposed that the 5' ss tails present on a 3'-processed LTR may contribute to processivity of HIV-1 IN by stabilizing the IN-DNA complex (17, 60). It has also been suggested that residue C-56 of HIV-1 IN may affect complex stability in a like manner (18). For M-MuLV IN, possibly similar effects on IN-IN and IN-DNA interactions have been found to be influenced by residue C-209 (12, 26).

The ability of IN to function catalytically within a multimeric complex has been proposed (24), and the ability of mixtures of defective IN mutants to complement one another is generally accepted to reflect aspects of those multimeric interactions (9, 37, 57). Analysis of HIV-2 IN mutants highlighted the critical requirement of the primary sequence extending from the region proximal to the HHCC finger and terminating 50 residues upstream of the C terminus for efficient complementation in integration assays (57). For coordinated disintegration, our complementation studies with $\Delta 105$ and $\Delta 176$ suggest that the region common to $\Delta 105$ and the HHCC finger protein is required *in cis* with the active site and C terminus for complementation of M-MuLV IN protomers. This possibility was inferred from the finding that the noncomplementing pair $\Delta 144$ and $\Delta 176$ lacks this region of overlap and could not be complemented by any protein mixture. The region of M-MuLV IN from amino acids 105 to 176 corresponds to the region comprising $\beta 1$, $\beta 2$, $\beta 3$, and $\alpha 1$ in the HIV-1 IN crystal structure (16). If M-MuLV IN folds similarly to HIV-1 IN, the complementation of $\Delta 232$ and $\Delta 105$ could allow an intermolecular $\alpha 1$ - $\alpha 5$ interaction. Further work to define the minimal region of the M-MuLV IN HHCC domain required for complementation is under way.

ACKNOWLEDGMENTS

This work was supported by American Cancer Society grant VM 134 and by National Science Foundation grant DMB-9105091. M.J.R. is a Stohlmann Scholar of the Leukemia Society of America, Inc. G.A.D. was supported by NIH training grant 5T32AI07043-02, and C.B.J. was supported by NIH postdoctoral fellowship grant IF32CA09242.

REFERENCES

1. Brown, P. O., B. Bowerman, H. E. Varmus, and J. M. Bishop. 1989. Retroviral integration: structure of the initial covalent product and its precursor and a role for the viral IN protein. *Proc. Natl. Acad. Sci. USA* **86**:2525-2529.
2. Brown, P. O., and H. E. Varmus. 1989. Retroviruses, p. 53-108. *In* D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
3. Burke, C. J., G. Sanya, M. W. Bruner, J. A. Ryan, R. L. LaFemina, H. L. Robbins, A. S. Zeff, C. R. Middaugh, and M. G. Cordingley. 1992. Structural implications of spectroscopic characterization of a putative zinc finger peptide from HIV-1 integrase. *J. Biol. Chem.* **267**:9639-9646.
4. Bushman, F. D., and R. Craigie. 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.
5. Bushman, F. D., A. Engelman, I. Palmer, P. Wingfield, and R. Craigie. 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.
6. Bushman, F. D., and B. Wang. 1994. Rous sarcoma virus integrase protein: mapping functions for catalysis and substrate binding. *J. Virol.* **68**:2215-2223.
7. Chen, J.-W., J. Lee, and M. Jayaram. 1992. DNA cleavage *in trans* by the active site tyrosine during Flp recombination: switching protein partners before exchanging strands. *Cell* **69**:647-658.
8. Chow, S. A., and P. O. Brown. 1994. 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.
9. Chow, S. A., and P. O. Brown. 1994. 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.
10. Chow, S. A., K. A. Vincent, V. Ellison, and P. O. Brown. 1992. Reversal of

- integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science* **255**:723–726.
11. Donzella, G. A., C. B. Jonsson, and M. J. Roth. 1993. Influence of substrate structure on disintegration activity of Moloney murine leukemia virus integrase. *J. Virol.* **67**:7077–7087.
 12. Donzella, G. A., O. Leon, and M. J. Roth. Unpublished data.
 13. Doolittle, R. F., D. F. Feng, M. S. Johnson, and M. M. McClure. 1989. Origins and evolutionary relationships of retroviruses. *Q. Rev. Biol.* **64**:1–30.
 14. Dotan, I. B., B. P. Scottoline, T. S. Heuer, and P. O. Brown. 1995. Characterization of recombinant murine leukemia virus integrase. *J. Virol.* **69**:456–468.
 15. Drellich, M., R. Wilhelm, and J. Mous. 1992. Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein in vitro. *Virology* **188**:459–468.
 16. Dyda, F., A. B. Hickman, T. M. Jenkins, A. Engelman, R. Craigie, and D. R. Davies. 1994. Crystal structure of the catalytic domain of HIV integrase: similarity to other polynucleotidyl transferases. *Science* **266**:1981–1986.
 17. Ellison, V., and P. O. Brown. 1994. A stable complex between integrase and viral DNA ends mediates human immunodeficiency virus integration in vitro. *Proc. Natl. Acad. Sci. USA* **91**:7316–7320.
 18. Ellison, V., J. Gerton, K. A. Vincent, and P. O. Brown. 1995. An essential interaction between distinct domains of HIV-1 integrase mediates assembly of the active multimer. *J. Biol. Chem.* **270**:3320–3326.
 19. Engelman, A., and R. Craigie. 1992. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J. Virol.* **66**:6361–6369.
 20. Engelman, A., G. Englund, J. M. Orenstein, M. A. Martin, and R. Craigie. 1995. Multiple effects of mutations in human immunodeficiency virus type 1 integrase in viral replication. *J. Virol.* **69**:2729–2736.
 21. Engelman, A., K. Mizuuchi, and R. Craigie. 1991. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell* **67**:1211–1221.
 22. Fujiwara, T., and K. Mizuuchi. 1988. Retroviral DNA integration: structure of an integration intermediate. *Cell* **54**:497–504.
 23. Hagerman, P. J. 1988. DNA flexibility. *Annu. Rev. Biophys. Biophys. Chem.* **17**:265–286.
 24. Jones, K. S., J. Coleman, G. W. Merkel, T. M. Laue, and A. M. Skalka. 1992. Retroviral integrase functions as a multimer and can turn over catalytically. *J. Biol. Chem.* **267**:16037–16040.
 25. Jonsson, C. B., G. A. Donzella, and M. J. Roth. 1993. Characterization of the forward and reverse integration reactions of the Moloney murine leukemia virus integrase protein purified from *Escherichia coli*. *J. Biol. Chem.* **268**:1462–1469.
 26. Jonsson, C. B., G. A. Donzella, E. Gaucan, C. M. Smith, and M. J. Roth. Functional domains of M-MuLV integrase defined by mutation and complementation analysis. Submitted for publication.
 27. Jonsson, C. B., and M. J. Roth. 1993. Role of the His-Cys finger of murine leukemia virus integrase protein in integration and disintegration. *J. Virol.* **67**:5562–5571.
 28. Kalpana, G. V., S. Marmon, W. Wang, G. R. Crabtree, and S. P. Goff. 1994. Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. *Science* **266**:2002–2006.
 29. Katz, R. A., G. Merkel, J. Kulkosky, J. Leis, and A. M. Skalka. 1990. The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. *Cell* **63**:87–95.
 30. Katz, R. A., and A. M. Skalka. 1994. The retroviral enzymes. *Annu. Rev. Biochem.* **63**:133–173.
 31. Katzman, M., R. A. Katz, A. M. Skalka, and J. Leis. 1989. The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the in vitro sites of integration. *J. Virol.* **63**:5319–5327.
 32. Khan, E., J. P. G. Mack, R. A. Katz, J. Kulkosky, and A. M. Skalka. 1991. Retroviral integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Res.* **19**:851–860.
 33. Kulkosky, J., K. S. Jones, R. A. Katz, J. P. G. Mack, and A. M. Skalka. 1992. Residues required 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.
 34. Kulkosky, J., and A. M. Skalka. 1990. HIV DNA integration: observations and inferences. *J. Acquired Immune Defic. Syndr.* **3**:839–851.
 35. Leavitt, A. D., L. Shiue, and H. E. Varmus. 1993. Site directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions in vitro. *J. Biol. Chem.* **268**:2113–2119.
 36. Lee, M. S., and R. Craigie. 1994. Protection of retroviral DNA from auto-integration: involvement of a cellular factor. *Proc. Natl. Acad. Sci. USA* **91**:9823–9827.
 37. Mazumder, A., A. Engelman, R. Craigie, M. Fesen, and Y. Pommier. 1994. Intermolecular disintegration and intramolecular strand transfer activities of wild-type and mutant HIV-1 integrase. *Nucleic Acids Res.* **22**:1037–1043.
 38. McEuen, A. R., B. Edwards, K. A. Koepke, A. E. Ball, B. A. Jennings, A. J. Wolstenholme, M. J. Danson, and D. W. Hough. 1992. Zinc binding by retroviral integrase. *Biochem. Biophys. Res. Commun.* **189**:813–818.
 39. Mizuuchi, K. 1992. Polynucleotidyl transfer reactions in transpositional recombination. *J. Biol. Chem.* **267**:21273–21276.
 40. Murphy, J. E., and S. P. Goff. 1992. A mutation at one end of Moloney murine leukemia virus DNA blocks cleavage of both ends by viral integrase in vivo. *J. Virol.* **66**:5092–5095.
 41. Panganiban, A. T., and K. J. Talbot. 1993. Efficient insertion from an internal long terminal repeat (LTR)-LTR sequence on a reticuloendotheliosis virus vector is imprecise and cell specific. *J. Virol.* **67**:1564–1571.
 42. Pruss, D., F. D. Bushman, and A. P. Wolffe. 1994. Human immunodeficiency virus integrase directs integration to sites of severe distortion within the nucleosome core. *Proc. Natl. Acad. Sci. USA* **91**:5913–5917.
 43. Pruss, D., R. Reeves, F. D. Bushman, and A. P. Wolffe. 1994. The influence of DNA and nucleosome structure on integration events directed by HIV integrase. *J. Biol. Chem.* **269**:25031–25041.
 44. Pryciak, P. M., A. Sil, and H. E. Varmus. 1992. Retroviral integration into minichromosomes in vitro. *EMBO J.* **11**:291–303.
 45. Pryciak, P. M., and H. E. Varmus. 1992. *Fv-I* restriction and its effects on murine leukemia virus integration in vivo and in vitro. *J. Virol.* **67**:5959–5966.
 46. Pryciak, P. M., and H. E. Varmus. 1992. Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell* **69**:769–780.
 47. Puras Lutzke, R. A., C. Vink, and R. H. A. Plasterk. 1994. Characterization of the minimal DNA binding domain of the HIV integrase protein. *Nucleic Acids Res.* **22**:4125–4131.
 48. Roth, M. J. 1991. Mutational analysis of the carboxyl terminus of the Moloney murine leukemia virus integration protein. *J. Virol.* **65**:2141–2145.
 49. Roth, M. J., P. Schwartzberg, and S. P. Goff. 1989. Structure of the termini of DNA intermediates in the integration of retroviral DNA: dependence on IN function and terminal DNA sequences. *Cell* **58**:47–54.
 50. Roth, M. J., P. Schwartzberg, N. Tanese, and S. P. Goff. 1990. Analysis of mutations in the integration function of the Moloney murine leukemia virus: effects on DNA binding and cutting. *J. Virol.* **64**:4709–4717.
 51. Schauer, M., and A. Billich. 1992. The N-terminal region of HIV-1 integrase is required for integration activity, but not for DNA binding. *Biochem. Biophys. Res. Commun.* **185**:874–880.
 52. Shin, C., B. Taddeo, W. A. Haseltine, and C. M. Farnet. 1994. Genetic analysis of the human immunodeficiency virus type 1 integrase protein. *J. Virol.* **68**:1633–1642.
 53. Straus, J. K., and L. J. Mauer III. 1995. DNA bending by asymmetric phosphate neutralization. *Science* **266**:1829–1833.
 54. Taddeo, B., W. A. Haseltine, and C. M. Farnet. 1994. Integrase mutants of human immunodeficiency virus type 1 with a specific defect in integration. *J. Virol.* **68**:8401–8405.
 55. van Den Ent, F. M. I., C. Vink, and R. H. A. Plasterk. 1994. DNA substrate requirements for different activities of the human immunodeficiency virus type 1 integrase protein. *J. Virol.* **68**:7825–7832.
 56. van Gent, D. C., A. A. M. Oude Groeneger, and R. H. A. Plasterk. 1992. Mutational analysis of the integrase protein of human immunodeficiency virus type 2. *Proc. Natl. Acad. Sci. USA* **89**:9598–9602.
 57. van Gent, D. C., C. Vink, A. A. M. Oude Groeneger, and R. H. A. Plasterk. 1993. Complementation between HIV integrase proteins mutated in different domains. *EMBO J.* **12**:3261–3267.
 58. Vincent, K. A., V. Ellison, S. A. Chow, and P. O. Brown. 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.
 59. Vink, C., A. A. M. Oude Groeneger, and R. H. A. Plasterk. 1993. Identification of the catalytic and DNA binding region of the human immunodeficiency virus type 1 integrase protein. *Nucleic Acids Res.* **21**:1419–1425.
 60. Vink, C., R. A. Puras Lutzke, and R. H. A. Plasterk. 1994. Formation of a stable complex between the human immunodeficiency virus integrase protein and viral DNA. *Nucleic Acids Res.* **22**:4103–4110.
 61. Vink, C., E. Yeheskiely, G. A. van der Marel, J. H. van Boom, and R. H. A. Plasterk. 1991. Site specific hydrolysis and alcoholysis of human immunodeficiency virus DNA termini mediated by the viral integrase protein. *Nucleic Acids Res.* **19**:6691–6698.
 62. Vora, A. C., M. L. Fitzgerald, and D. P. Grandgenett. 1990. Removal of 3'-OH terminal nucleotides from blunt-ended long terminal repeat termini by the avian retrovirus integration protein. *J. Virol.* **64**:5656–5659.
 63. Wiskerchen, M., and M. A. Muesing. 1995. Human immunodeficiency virus type 1 integrase: effects of mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA templates, and sustain viral propagation in primary cells. *J. Virol.* **69**:376–386.
 64. Woerner, A. M., M. Klutch, J. G. Levin, and C. J. Marcus-Sekura. 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.
 65. Woerner, A. M., and C. J. Marcus-Sekura. 1993. Characterization of a DNA binding domain in the C-terminus of HIV-1 integrase by deletion mutagenesis. *Nucleic Acids Res.* **21**:3507–3511.