JOURNAL OF VIROLOGY, June 1996, p. 3571–3580 Vol. 70, No. 6
0022-538X/96/$04.00+0

Concerted Integration of Linear Retroviral DNA by the Avian Sarcoma Virus Integrase In Vitro: Dependence on Both Long Terminal Repeat Termini

ASHOK AIYAR,[1]† PATRICK HINDMARSH,[1] ANNA MARIE SKALKA,[2] AND JONATHAN LEIS[1]*

Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106,[1] *and Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111*

Received 19 September 1995/Accepted 27 February 1996

We have reconstituted integration reactions in vitro with specially designed donor DNAs, a supercoiled plasmid acceptor, purified bacterium-derived Rous sarcoma virus integrase (IN), and a host cell DNA-bending protein, HMG1. The duplex donor DNAs are approximately 300 deoxynucleotides in length and contain only 15 bp of the RSV U3 and U5 termini at the respective ends. The donor has blunt U3 and U5 termini which end with the sequence 5′CATT. Joining of the donor DNA to the acceptor DNA is detected by using a simple biochemical assay. Integration was found to be dependent on both U3 and U5 termini; mutations in either result in a significant decrease in the level of integration in vitro. Restriction digestion of the products is consistent with most integrants representing a concerted integration in which both long terminal repeat termini come from the same donor molecule. The U5 and U3 sequences in the substrate flank a *supF* tRNA gene, permitting biological selection of integrants. Many integrants have been sequenced, and have all of the hallmarks of authentic viral integration, including the removal of a terminal TT dinucleotide from each donor DNA end, and duplication of acceptor sequences at the integration site without introducing deletions into the acceptor. Target site selection in the acceptor plasmid was random except that the orientation of integrants selected was apparently influenced by *supF* transcription. Mutations which substituted the conserved CA dinucleotide with a GA pair led to a decreased rate of integration. In 2 of 14 mutant integrants sequenced, deoxynucleotides were deleted from either the U5 or U3 terminus. In one instance, an internal CA dinucleotide was used, which resulted in a 10-bp U5 donor deletion. In the other, an internal GA dinucleotide was used, which produced a 5-bp U3 donor deletion. Both of these integrants provide further evidence that concerted integration in this reconstituted system requires interactions between IN and the U3 and U5 termini from the same donor molecule.

Integration of viral DNA into the host chromosome is an obligatory step in retroviral replication. Shortly after infection, viral RNA is reverse transcribed into a linear double-stranded DNA flanked by long terminal repeat (LTR) sequences. The ends of the LTRs are recognized by the virus-encoded integrase (IN), which catalyzes the integration of the DNA in a two-step reaction. In the first step, termed processing, a TT dinucleotide is removed from each viral 3′ end catalyzed by IN (3, 17, 18). The newly processed ends are then joined to the host DNA by IN in a single transesterification reaction, termed joining (12).

The standard assay for reconstituted integration systems uses short duplex oligodeoxynucleotide substrates which resemble the terminal 10 to 15 bp of the virus LTR U3 or U5 sequence. These short duplex substrates have facilitated investigation of substrate sequence requirements for both processing and joining reactions (10, 17, 18) and examination of the effects of mutations in either IN or the IN recognition sequences on the process of integration. However, one limitation of this assay is that it examines the action of IN on only a single viral terminus. Biological evidence strongly suggests that both viral DNA ends are coordinately recognized by IN. Murine leukemia viruses with a mutation in the U3 terminus failed to correctly process the wild-type U5 end (23). Thus, in vivo, both ends are likely held and processed together by a multimeric form of the viral IN (28). In addition, we have described mutations in avian leukosis-sarcoma virus U5 sequences that are severely defective when tested for integration in vitro by using oligodeoxynucleotides that represent only the U5 terminus. The effects of these mutations are not as severe in the context of an infectious virus, in which integration into the host genome is clearly detected (6). In the second step of integration, the processed viral ends are always joined to host DNA at the same site in a staggered manner which results in a short duplication of the host DNA. The size of the duplication is dependent on the virus and not the host cell which is infected. This observation also suggests that the viral DNA ends, which are several kilobases apart on the linear viral DNA sequence, are juxtaposed by IN during integration (8).

Experimental systems designed to examine concerted integration have been described for murine leukemia virus, avian leukosis-sarcoma virus, and human immunodeficiency virus. These assays have relied on the construction of plasmid vectors with a U3-U5 junction that has been modified to include the cleavage site of the restriction endonuclease *Nde*I. Digestion of the plasmid with *Nde*I linearizes the DNA such that it has U3 and U5 sequences at the respective ends. However, the *Nde*I ends differ from wild type in both sequence and structure. While viral ends are blunt, the *Nde*I ends resemble processed viral ends in that they have a two deoxynucleotide 5′ overhang. Incubation of these linear DNAs with IN and λ DNA as an integration acceptor results in the insertion of the viral DNA. The efficiency of the reaction is low and requires biological

* Corresponding author. Mailing address: Department of Biochemistry, Case Western Reserve University, 2119 Abington Rd., Cleveland, OH 44106. Phone: (216) 368-3360. Fax: (216) 368-4544. Electronic mail address: jxl8@biochemistry.cwru.edu.

† Present address: McArdle Center for Cancer Research, University of Wisconsin, Madison, WI 53706.

selection for detection (5, 10, 13–15, 17). A different experimental system to biochemically detect concerted integration has been described by Vora et al. (30). However, most of the integrations in these reactions result from the insertion of two ends from different donors into an acceptor, and as a consequence, large deletions between insertion sites are often observed. More recently, Kukolj and Skalka (20) have reported enhanced and coordinated processing of two viral DNA ends in vitro by both avian sarcoma virus and human immunodeficiency virus type 1 IN. Their assay uses substrates that consist of short duplex oligodeoxynucleotides that represent U3 and U5 ends that are covalently linked across the termini by short single-stranded deoxynucleotide sequences. With this assay, it has been possible to biochemically examine the effects of mutations at one viral end on reactions of the other.

We describe herein another approach to developing a simple assay to monitor concerted integration events. The substrate in this assay is a very short linear DNA which is flanked by only 15 bp of the natural U5 and U3 LTRs. The donors are blunt ended and contain the wild-type viral sequences at both ends. Integration can be monitored biochemically, and integrants can be cloned biologically. In contrast to the single-end duplex oligodeoxynucleotide assay, integration with this system requires the presence of viral sequences from both ends on the same integration donor molecule for maximum efficiency. Integration was also stimulated by the addition of a ubiquitous nuclear DNA-bending protein, high-mobility-group protein 1 (HMG1) (24).

MATERIALS AND METHODS

Bacterial strains and growth conditions. *Escherichia coli* DH5α (Bethesda Research Laboratories, Bethesda, Md.) and MC1061/P3 (Invitrogen) were used for these studies. MC1061/P3 is a derivative of MC1061 with the male episome P3 which can be selected for by the presence of an encoded Kan^r gene. In addition, P3 possesses *amp*(Am) and *tet*(Am) genes, the expression of which can be rescued by the *supF* amber suppressor tRNA. Under these conditions, MC1061/P3 can be selected for ampicillin, tetracycline, and kanamycin resistance.

Reagents. Avian sarcoma virus IN was prepared as described by Jones et al. (16). HMG1 was purified as described by Chow et al. (7). Proteinase K (30 U/mg) and glycogen were from Boehringer Mannheim Biochemical. Vent DNA polymerase (2 U/μl) and *Bam*HI (20 U/μl) were from New England Biolabs (Beverly, Mass.). Oligodeoxynucleotides were synthesized on an Applied Biosystems synthesizer or purchased from Midland Certified Reagent Company (Midland, Tex.). Oligodeoxynucleotides were purified by denaturing polyacrylamide gel electrophoresis followed by reverse-phase chromatography as previously described (1). The following oligodeoxynucleotides were used in this study: U5(WT), 5′AATGAAGCCTTCTGCTGGGCGGAGCCTATG3′; U5(CTTC→GAAG), 5′AATCTTCCCTTCTGCTGGGCGGAGCCTATG3′; U5(CA→GG), 5′AACCAAGCCTTCTGCTGGGCGGAGCCTATG3′; U3(WT), 5′AATGTAGTCTTATGCGTTGCCCGGATCCGG3′; ΔU3, 5′AGCAATGGCAACAACGTTGCCCGGATCCGG3′; ΔU5, 5′AGCACTCGCAACAACTGGGCGGAGCCTATG3′; U5seq, 5′TTCAAAAGTCCGAAA3′; and U3seq, 5′AGAATTCGGCGTTGC3′.

The U5(WT) and U3(WT) oligodeoxynucleotides were used to prepare the wild-type donor DNA substrate. The U5(CTTC→GAAG), U5(CA→GG), ΔU5, and ΔU3 oligodeoxynucleotides were used to prepare donor substrates with mutations in the U5 terminus sequence or lacking the U3 terminus. The U5seq and U3seq oligodeoxynucleotides were used as sequencing primers. The U3seq primer is complementary to plasmid πvx nucleotides 326 to 312, and the U5seq primer is complementary to plasmid πvx nucleotides 116 to 130.

Plasmid constructions and preparations. Plasmid πvx (GenBank accession number VB0010) was used as a template to amplify the *supF* tRNA gene. This plasmid was propagated in *E. coli* MC1061/P3 under the conditions described above. The integration target used was plasmid pBCSK$^+$ (Stratagene, La Jolla, Calif.), which was propagated in *E. coli* DH5α. For other experiments, a derivative of pBCSK$^+$ which lacks the multiple cloning site was used. This target DNA was constructed by digesting pBCSK$^+$ with the enzyme *Bss*HII, whose restriction sites flank the multiple cloning sequence. The resultant large fragment, lacking the multiple cloning site, was gel purified and self-ligated to produce plasmid pBCΔMCS. pBCΔMCS lacks a *Bam*HI restriction enzyme site present within the integration donor and pBCSK$^+$. Plasmids were purified by using Qiaprep columns (Qiagen, Chatsworth, Calif.) or pZ523 columns (5′→3′, Boulder, Colo.) according to manufacturer's instructions. The growth of DH5α containing either pBCSK$^+$ or pBCΔMCS was selected for by the addition of 35 μg of chloramphenicol per ml.

PCR amplification of integration donors and purification of integration donors. Integration donors were amplified by using the thermostable Vent DNA polymerase (New England Biolabs) and primers listed above. Twenty-five picomoles of each primer and 10 ng of πvx DNA as the template were used during each PCR. Vent DNA polymerase was used according to manufacturer's instructions. A total of 20 rounds of amplification were performed in each PCR. The amplification conditions were 94°C for 2 min, 50°C for 1 min, and 72°C for 45 s for three rounds. This was followed by amplification conditions of 94°C for 2 min, 57°C for 1 min, and 72°C for 45 s for 17 additional rounds. The resultant product donor DNA was purified by gel electrophoresis on 2% agarose gels run on 0.5× Tris-borate-EDTA (26). The purified DNA (600 ng) was recovered by electroelution onto DE-81 paper as described by Sambrook et al. (26) or by using Qiaex-II resin (Qiagen) and then precipitated with ethanol. The recovered DNA was washed with 70% ethanol, dried, and suspended in either TE (10 mM Tris-HCl [pH 8.0], 1 mM EDTA) or deionized distilled water.

The integration donors, which were approximately 300 bp in length, were internally labeled during the PCR by the inclusion of [α-^{32}P]dCTP (3,000 Ci/mmol, 10 mCi/ml; New England Nuclear, Boston, Mass.) during the amplification. The final concentrations of deoxynucleoside triphosphates during amplification reactions were 0.25 mM each unlabeled dATP, dGTP, and TTP. The final dCTP concentration was 0.0502 mM (12 Ci/mmol, 0.6 mCi/ml).

Integration reactions. The integration reaction conditions were similar to those described by Fitzgerald et al. (14). Briefly, 30 ng (~0.3 pmol of ends) of donor was mixed with 50 ng of acceptor (~0.02 pmol) and 180 ng of avian leukosis-sarcoma virus IN (~6 pmol) in a 5-μl preincubation reaction mixture. The IN was diluted in a buffer containing 30% glycerol, 0.5 M NaCl, 50 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES; pH 8.0), 1 mM dithiothreitol, and 0.1 mM EDTA. Where specified, HMG1 was added to the reaction mixtures in the concentrations indicated. The preincubation reaction mixtures were placed on ice for at least 5 h or, more typically, overnight. The volume of each preincubation mixture was then increased to 30 μl with a final concentrations of 20 mM Tris-HCl (pH 7.5), 166 mM NaCl, 5 mM dithiothreitol, 10% dimethyl sulfoxide, 0.05% Nonidet P-40, 1% glycerol, 1.6 mM HEPES (pH 8.0), 3.3 μM EDTA, and 6.7 mM $MgCl_2$, and the integration assay mixture was incubated at 37°C for 90 min. Where indicated, the IN was added only during the integration reaction. The integration reactions were stopped by increasing the volume to 160 μl with the addition of EDTA (final concentration of 4.25 mM), sodium dodecyl sulfate (final concentration of 0.44%), and proteinase K (final concentration of 0.06 mg/ml). After digestion for 60 min at 37°C, the reaction mixtures were extracted with phenol followed by phenol-chloroform-isoamyl alcohol (25:24:1 mixture). Seventeen microliters of 3 M sodium acetate (pH 5.2) was added along with 1 μl of glycogen (10-mg/ml stock solution). The reaction products were precipitated by the addition of 400 μl of 100% ethanol and washed twice with 70% ethanol prior to electrophoresis and autoradiography. The reaction products were separated on a 1% agarose gel run in 0.5× Tris-borate-EDTA at 10 V/cm for 2 h. Gels were stained with ethidium bromide, and bands detected by exposure to UV light were used as internal markers for each experiment (see Fig. 1B). Following electrophoresis, gels were submerged in 5% trichloroacetic acid for 20 min or until the bromophenol blue dye turned bright yellow. After being washed with water, the gels were dried on DE-81 paper (Whatman) in a Bio-Rad slab gel dryer at 80°C for approximately 2 h under vacuum. The dried gels were exposed to autoradiographic film overnight at −80°C in a film cassette with GAFMED TA-3 midspeed screens.

Cloning and sequencing of integrants. In one experiment, the nicked circle integrant band was recovered from a 1% agarose gel by using a Qiaex-II DNA extraction kit (Qiagen). The efficiency of recovery was improved by the addition of 400 ng of wheat germ tRNA to the agarose slice prior to solubilization with the QX-1 buffer provided in the kit. In all subsequent experiments, the integration products were used directly for transformation of bacteria without purification of the integrant band. Integration products were introduced into *E. coli* MC1061/P3 by electroporation, using a Bio-Rad electroporator with 0.1-cm electroporation cuvettes, 1.8-kV voltage, 25-μF capacitance, and 200-ohm resistance. The P3 episome is maintained at a low copy number. Therefore, drug selection typically was done with only 20 μg of ampicillin, 10 μg of kanamycin, and 5 μg of tetracycline per ml. Under these conditions, we did not detect colonies after *supF* selection when the donor, acceptor, or donor and acceptor in the absence of IN were electroporated into cells. Plasmid DNAs were recovered from individual clones, and integration junctions were sequenced by using primers U3seq (for sequencing the U3 junction) and U5seq (for sequencing the U5 junction). Sequencing was performed with a Sequenase kit as instructed by the manufacturer (U.S. Biochemical, Cleveland, Ohio).

Restriction digestion conditions. The enzyme *Bam*HI was used under conditions recommended by the manufacturer.

RESULTS

Design of integration donors and acceptors. The integration donor DNAs used in this study were constructed by PCR amplification of the *supF* gene from plasmid πvx. The oligodeoxynucleotides used to construct the donors, as well as donors with mutations in either the U5 or U3 LTR terminus, are listed in Materials and Methods. A typical donor consists of the ter-

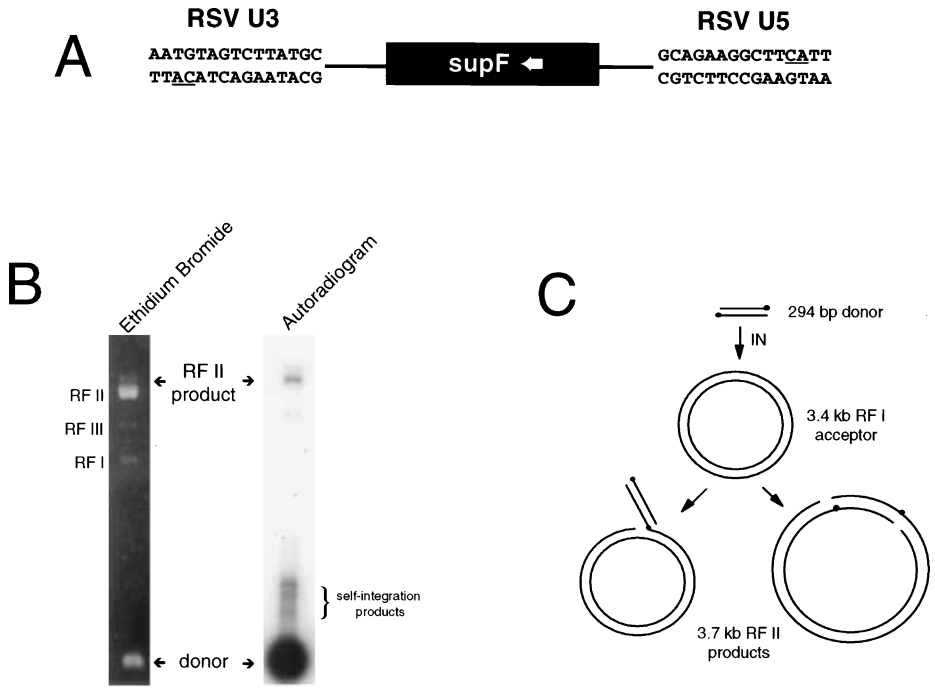


FIG. 1. Reconstitution of integration in vitro. (A) Diagrammatic representation of an integration donor DNA prepared by PCR as described in Materials and Methods. The 15 base-paired deoxynucleotides shown at each end represent the U3 and U5 preintegration termini of Rous sarcoma virus (RSV). The conserved CA dinucleotides used during joining to host sequences are underlined. The U3 and U5 sequences flank a *supF* gene, indicated by the solid rectangle. The arrow denotes direction of *supF* transcription. (B) Gel electrophoresis separation of products from an integration reaction. An ethidium bromide stain and an autoradiogram of the gel-separated products are shown. The positions of the supercoiled, nicked, and linearized acceptor DNAs are marked as RF I, RF II, and RF III, respectively. The migration positions of the donor and RFII integration products are marked by arrows. Self-integration products represent donor DNAs that have integrated into other donor DNAs. (C) Schematic representation of concerted and unconcerted reaction products, both of which are predicted to migrate as RFII circles. The donor is shown as a pair of short lines with a solid circle at each end.

minal 15 bp of the Rous sarcoma virus U3 and U5 sequences flanking the *supF* tRNA gene (Fig. 1A). The donors were approximately 300 bp long. To avoid the addition of 3′ nontemplate nucleotides, we used Vent thermostable DNA polymerase in the PCRs. Plasmid pBCSK$^+$ (Stratagene), which is 3.4 kb in length, was used as the integration acceptor. Integration of the donor into the supercoiled form of pBCSK$^+$ results in the formation of a nicked circular DNA which is 3.7 kb in length and can be separated from the starting plasmid by using gel electrophoresis conditions described in Materials and Methods. The integration product migrates slightly more slowly than the nicked pBCSK$^+$. Integration into a linearized acceptor results in a linear DNA product which also migrates more slowly than the starting linear acceptor DNA. For some experiments, a derivative of pBCSK$^+$ that lacks the multiple cloning site (pBCΔ MCS) was used. This plasmid, which is 3.2 kb in length, lacks a *Bam*HI restriction site that is present in the integration donor.

Analysis of integration in vitro. Integration reactions were carried out as described in Materials and Methods. The products of a typical integration reaction with supercoiled acceptor DNA were separated by agarose gel electrophoresis and detected by autoradiography (Fig. 1). The products of the integration reaction in which one or both ends of the donor DNA are inserted into a supercoiled target are shown schematically in Fig. 1C. The ethidium bromide-stained gel and corresponding autoradiogram are shown in Fig. 1B. The unintegrated donor DNA, which is in 15-fold molar excess relative to the target, can be seen as a well-stained rapidly migrating band at the bottom of the gel (donor). A small amount of the starting replicative form I (RFI) acceptor DNA still remains at the end of reaction and is seen as a lightly stained band at the top of the gel. Most of the starting acceptor DNA has been nicked through the action of the DNA endonuclease activity associated with IN and migrates at the position of RFII DNA. Approximately 10% of the starting acceptor DNA has IN-catalyzed insertions of donor DNA. This results in a nicked circle product that is 3.7 kb in size which migrates slightly more slowly than the 3.4-kb nicked circle (RFII product). This larger nicked circle could contain both one-ended (nonconcerted integration) and two-ended (concerted integration) products, as shown in Fig. 1C. If IN is omitted from the reaction mixture, only the starting donor and the RFI acceptor DNA are observed (data not shown).

The donor DNA was internally labeled with [α-^{32}P]dCTP during the PCRs so that integration products could also be detected by the appearance of radioactivity at the position of the acceptor DNA. The agarose gel was dried, and an autoradiogram was prepared (Fig. 1B). The most intensely labeled band is the unused donor DNA migrating at the bottom of the gel. Two labeled bands that migrate slightly more slowly than the unreacted donor DNA represent self-integration products, as they are detected in reaction mixtures that lack the acceptor plasmid DNA. When the acceptor DNA is included in the reaction, a labeled band that comigrates with the 3.7-kb nicked circle band detected by ethidium bromide staining is observed (RFII product). There are minor labeled DNA products that migrate above and below the RFII product. The band above has an estimated molecular weight that would be consistent with integration of two donor DNAs into an acceptor. The band below may represent integration of a donor into an acceptor DNA that was converted to linear DNA by IN prior to the integration event. The migration of these minor bands relative to ethidium bromide-stained markers is consistent with this interpretation.

Extended preincubation and the addition of HMG1 increase the efficiency of integration. In our initial experiments, donor, acceptor, and IN were preincubated on ice for 30 min as

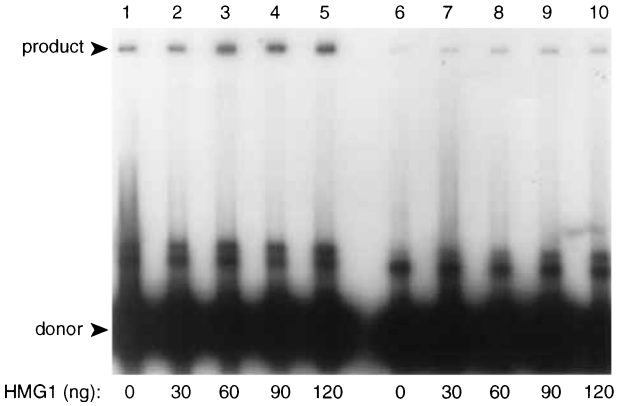


FIG. 2. Extended preincubation and HMG1 increase integration efficiency. Integration reactions with the wild-type donor, acceptor DNA, and recombinant IN were preincubated for 30 min (lanes 6 to 10) or overnight on ice (lanes 1 to 5), using buffer conditions described in Materials and Methods. Lanes 1 and 6 lacked the HMG1 protein. Lanes 2 to 5 and 7 to 10 included various concentrations of HMG1 (as indicated). Reaction mixtures were then incubated at 37°C for 90 min as described in Materials and Methods, and the products were separated by agarose gel electrophoresis. The labeled donor and RFII integration product are indicated by arrowheads.

recommended by Fitzgerald et al. (14). Following this preincubation period, the required divalent cation was added and the reaction mixtures were incubated at 37°C for 90 min (Fig. 2). While optimizing the reaction conditions, we noticed that an extended overnight preincubation period increased the yield of integrants by two- to threefold (Fig. 2, compare lanes 1 and 6). It seemed possible that the extended preincubation period permitted the formation of an IN multimer which brought both ends of the donor DNA together into a noncovalent complex that facilitated the integration reaction. The formation of such a complex would require the short donor DNA to circularize. Purified rat HMG1 has recently been shown to bend DNA, inducing circularization of duplex molecules as short as 60 bp (24). We therefore added the HMG1 protein to the preincubation mixture. This resulted in a further three- to fourfold increase in the efficiency of integration (Fig. 2; compare lanes 2 through 5 to lane 1).

The effect of HMG1 is on the integration donor and not the acceptor. An alternative explanation for how HMG1 stimulates integration is that it introduces bends into the acceptor DNA, which renders it more susceptible to nucleophilic attack by the integration donor. To distinguish between these two possibilities, we performed the staggered incubation experiments shown in Fig. 3. Donor or acceptor DNA was preincubated on ice with either HMG1 or IN. Then all of the remaining components were added to the integration reaction mixture, and a second incubation was carried out at 37°C. Preincubation of the donor DNA with HMG1 led to the largest increase in efficiency of integration (Fig. 3, lane 5). In contrast, preincubation of HMG1 with the acceptor DNA had no detectable effect on the efficiency of the integration (Fig. 3, lane 4), even if the donor DNA was also preincubated with IN. These results indicate that HMG1 acts on the donor and not the acceptor DNA. Furthermore, if the donor DNA is preincubated with HMG1, it is no longer necessary to preincubate the donor with IN to obtain the increase in efficiency of integration. These results strongly suggest that HMG1 facilitates the bending of the donor DNA such that both ends can be coordinately recognized by an IN multimer.

Linear DNA is a more efficient acceptor than supercoiled DNA. To examine whether the topology of the acceptor DNA affected the efficiency of integration, we compared integration reactions by using either linear or supercoiled acceptor DNA. As shown in Fig. 4, the linear DNA was a better acceptor than supercoiled DNA. A similar observation has been reported for the yeast Ty1 integration system (2). Nevertheless, we have used the supercoiled DNA of the acceptor in our standard assay conditions because it (the open circular product) is more efficient in transforming bacteria, thereby permitting analysis of integrants in vivo.

Mutations at either the U3 or U5 end severely decrease the efficiency of integration. To establish whether the integration products detected in our reactions were largely the result of uncoordinated integration events or a combination of both uncoordinated and coordinated integration, the U3 sequences in the donor were substituted with πvx sequences (Fig. 5A). The U5 sequences were unchanged. If the efficiency of integration of this donor with a single LTR end was half that observed with the wild type, we would conclude that the majority of the integration events represented insertions of single donor ends into the acceptor DNA. However, if the efficiency of integration was decreased by more than 50%, we would conclude that coordinated integration requiring both LTR ends likely predominated during the reaction. As shown in Fig. 5B, the presence of only one LTR terminus reduced the observed integration by 75 to 85% when a supercoiled acceptor DNA was used (compare lanes 1 and 3). The minor bands observed above and below the RFII product are as shown in Fig. 1. The ethidium bromide-stained bands served as an internal standard in each gel. A similar result was obtained when the integration of the wild-type and ΔU3 donors into linear DNA were compared, as shown in Fig. 5C (compare lanes 1 and 2). We interpret these data to imply that the majority of

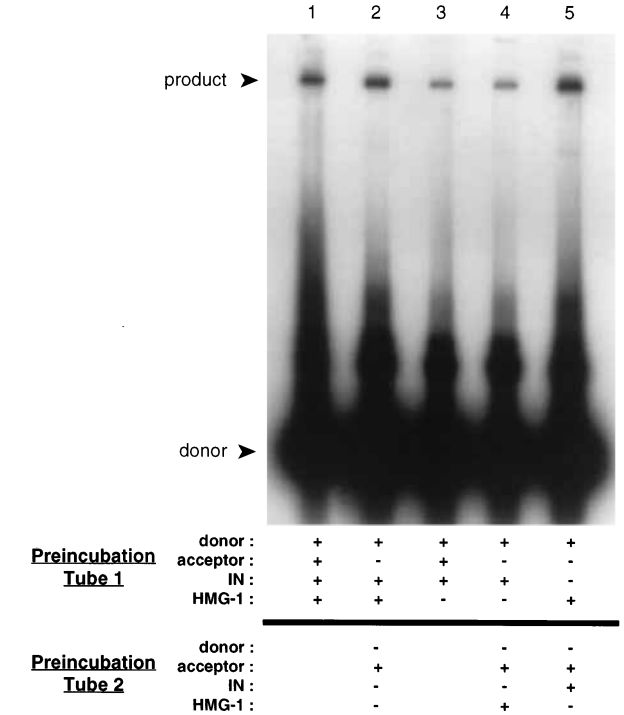


FIG. 3. The HMG1 protein acts on the donor DNA. Donor DNA, acceptor DNA, HMG1, and IN were preincubated together (lane 1) or in two separate incubations (lanes 2, 4, and 5) as indicated. The components added to each incubation are listed at the bottom. Preincubations were performed overnight on ice, and integration reactions were performed at 37°C for 90 min as described in Materials and Methods. Components omitted from the preincubation mixture were added to the integration reaction as indicated. The labeled donor and RFII products are indicated by arrowheads.

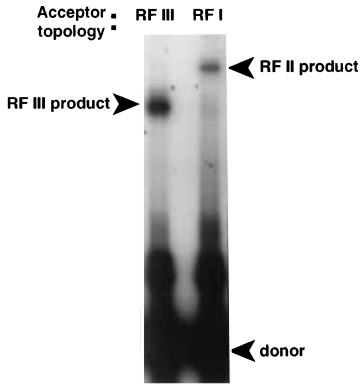


FIG. 4. Influence of acceptor DNA topology on integration. Integration reactions were performed as described in Materials and Methods, using either supercoiled (RFI) or linear (RFIII) acceptor DNA. Autoradiograms of separated reaction products are shown in lane 1 (RFIII acceptor) and lane 2 (RFI acceptor). The donor, RFII, and RFIII products are marked with arrowheads.

integrants observed with the wild-type donors result from the coordinate insertion of both ends of the same donor into the acceptor DNA. Results similar to those shown in Fig. 5B and C were obtained if the U5 rather than the U3 LTR terminus was deleted from the donor substrate (Fig. 5D). Therefore, removal of either LTR terminus in this system results in significant loss of detectable integration products.

We have previously examined the effects of mutations in U5 on integration in vivo and in vitro (6). One of these mutations changed the U5 sequence at the 3′ end of the plus strand from CTTCATT to GAAGATT (Fig. 5A). This mutation changed the highly conserved CA at the terminus to a GA dinucleotide sequence (indicated in boldface). Three additional nucleotides immediately 5′ to the CA were also changed. When Mg^{2+} was used as the divalent cation during the reaction, this mutation abrogated the ability of the mutated U5 sequence to participate in the processing reaction using duplex oligodeoxynucleotide substrates (6). In contrast, a virus containing the same U5 mutation could integrate its DNA into host chromosomal DNA, albeit at lower levels compared with the wild type (6). We therefore examined the effect of this same U5 mutation in the context of a donor substrate with a wild-type U3 terminus as described in Materials and Methods. The effect of a second mutation that changed only the highly conserved CA dinucleotide at the terminus in the sequence CTTCATT to CTTGGTT was also evaluated (Fig. 5A). As shown in Fig. 5B (compare lanes 2 and 4 with lane 1), donor DNAs containing either of these mutated U5 sequences showed integration products but at reduced levels compared with wild-type donor DNAs. The degree of defectiveness of the four-nucleotide-substitution mutation, referred to as S4, was not as severe as the complete omission of one of the two LTR ends from the donor but was of the same magnitude as previously reported in vivo (6). This result implies that the presence of the wild-type U3 end can partly alleviate the effect of the U5 mutation during integration.

Restriction digestion of the integration products. The results from the experiments described above, particularly those that replace the viral ends with unrelated sequences, led us to conclude that a substantial proportion of products resulted from integration dependent on both termini. The following experiment was carried out to independently confirm this conclusion. Integration products were digested with *Bam*HI, which cleaves the integration donor but not the acceptor DNA (pBCΔMCS). As depicted schematically in Fig. 6A, restriction digestion of products derived from single-ended integration events should continue to migrate as nicked DNA circles. In contrast, digestion of products derived from two-ended concerted integration reactions should migrate as linear DNA. When the products of an integration reaction using supercoiled pBCΔMCS acceptor were subjected to exhaustive digestion with *Bam*HI and separated by gel electrophoresis, a prominent band that migrated to a position slightly higher than that of linear pBCΔMCS DNA was

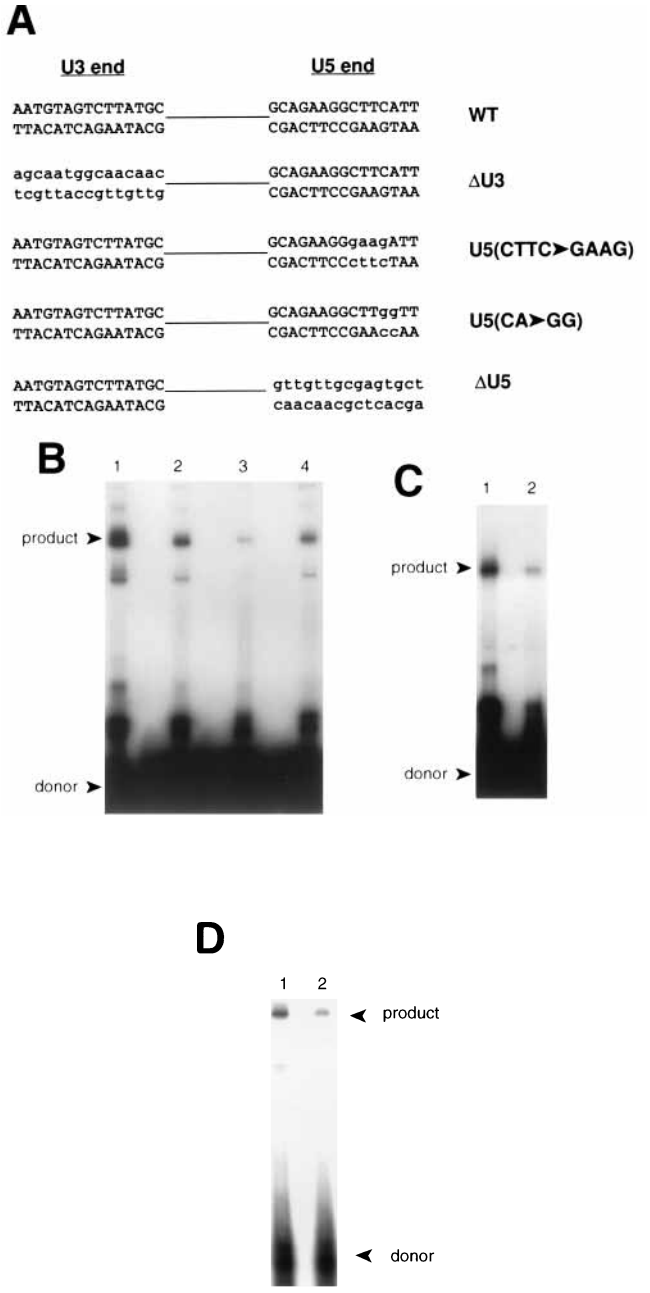


FIG. 5. Effects of mutations in either of the two viral termini on integration. (A) Schematic representation of wild-type (WT) and mutant donor DNAs. The ΔU3, ΔU5, and U5(CTTC→GAAG) mutant donors are described in the text. Mutated sequences are shown in lowercase letters. (B) Autoradiogram of reaction products from integration reactions performed with the donors described in panel A. Lane 1, wild-type donor; lane 2, U5(CTTC→GAAG) donor; lane 3, ΔU3 donor; lane 4, U5(CA→GG) donor. The labeled donor and RFII products are indicated by arrowheads. (C) Autoradiogram of reaction products from integration reactions performed using the wild-type donor (lane 1) and ΔU3 donor (lane 2) with a linear acceptor DNA. The labeled donor and RFIII products are indicated by arrowheads. (D) Autoradiogram of reaction products performed with donors described in panel A. Lane 1, wild-type donor; lane 2, ΔU5 donor.

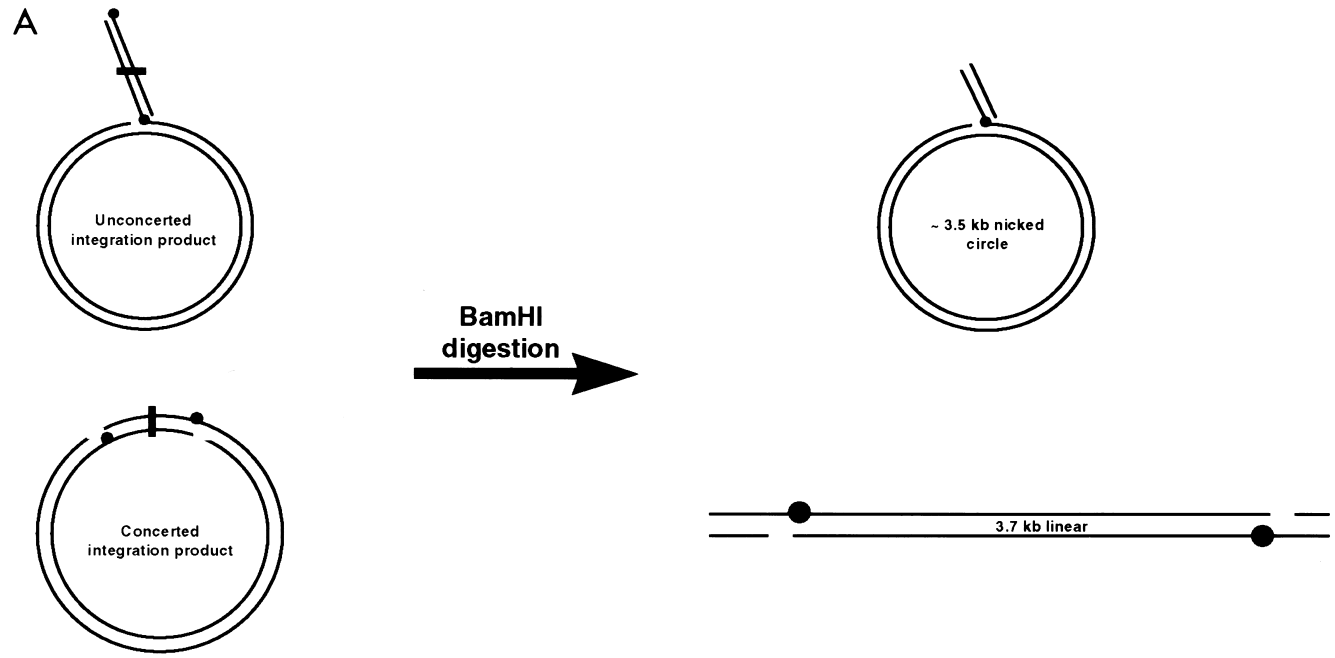
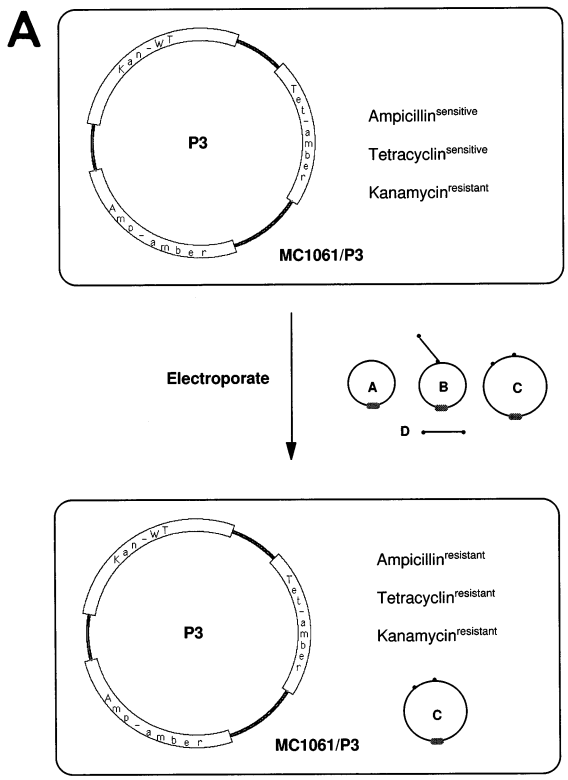


FIG. 6. Restriction digestion analysis of integration products. (A) Schematic representation of the expected DNAs upon restriction digestion of concerted and unconcerted integration products. A unique *Bam*HI restriction site present in the donor is indicated by the solid rectangular bar. Digestion products of unconcerted integration will migrate as RFII circles, while digestion products of concerted integration will migrate as linear DNA. (B) Integration reactions were performed with the wild-type donor DNA and pBCΔMCS acceptor DNA as described in Materials and Methods. An autoradiogram of reaction products separated by gel electrophoresis after digestion with *Bam*HI (lane +) or without digestion (lane −) is shown. The labeled donor, RFII integration product, and major restriction enzyme-digested products are marked with arrowheads. The major *Bam*HI restriction enzyme-digested product migrates as a linear DNA of 3.5 kb.

observed (Fig. 6B). The size of this product, as judged from ethidium bromide-stained bands in the gel, is consistent with it being larger than the linear pBCΔMCS DNA by about 0.3 kb. We did not detect a band the size of an RFII DNA product with a shortened donor which would represent one-ended integration events. We conclude from this result and those obtained by using donors with mutations in either the U3 or U5 terminus sequence that the majority of the integration products reflect the concerted integration of both ends of a single donor DNA into an acceptor.

Biological selection for integrants. To examine the junction sequences of individual integration events, we used the selection scheme as outlined in Fig. 7A. Products from the reconstituted integration reaction using the wild-type donor DNA, when introduced into MC1061/P3, produced colonies on plates containing ampicillin (30 μg/ml), tetracycline (7.5 μg/ml), and kanamycin (10 μg/ml). In contrast to results obtained with *E. coli* CA244 [*trpA49*(Am) *lacZ125*(Am)], control transformations of donor, acceptor, or acceptor and donor in the absence of integration reactions into MC1061/P3 resulted in no ampicillin- and tetracycline-resistant colonies (1a). Plasmid DNAs were recovered from several colonies and sequenced by using oligodeoxynucleotides that prime DNA synthesis in opposite directions on the different strands, thereby providing the sequence of the donor-acceptor DNA junctions. Results from these analyses are summarized in Table 1. All of the integrants sequenced showed the hallmarks of authentic retroviral integration, including the loss of the 3′ TT dinucleotide from the LTR termini and a short duplication of acceptor DNA at the site of integration. The size of the duplications varied between 4 to 7 nucleotides, with the largest number of duplications being 6 nucleotides in length. Moreover, the integration events occurred over the entire length of the target DNA (Fig. 7B). Thus, the integration products formed in this reconstituted system are similar to those observed with in vivo reactions. Table 1 also lists the size of the duplication of acceptor DNA at each site of integration. Most of the 4-, 5-, and 7-bp duplications of the acceptor DNA were clustered approximately between acceptor plasmid positions 750 to 1650. Outside of this region, the duplications were 6 bp, more typical of an avian retrovirus (8).

Curiously, all of first 12 integrants isolated from the initial drug selection conditions integrated in the same orientation which we designated as left. For all integrants in this orientation, the direction of *supF* transcription was toward the plasmid origin of replication. As this appeared unusual, we repeated the selection of integrants under less stringent drug conditions, with ampicillin at 20 μg/ml and tetracycline at 5 μg/ml. Under these conditions, we obtained a significantly larger number of resistant colonies. To ensure that these colonies represented authentic integration events, we examined more than 200 in a colony PCR assay using primers that would detect the integrated donor. More than 95% of the colonies



contained an integrated donor DNA of the correct size. The presence of the insert was confirmed by restriction digestion analysis. A number of these clones were sequenced, and it was found that many again integrated in the same orientation. However, several were integrated in the opposite or right orientation, in which the direction of *supF* transcription is away from the plasmid origin of replication (Table 1). A total of 25 wild-type clones have now been sequenced. While the integration into the target DNA in vitro appears to be random, approximately 15% of the clones selected in our biological screen represented integration events at the exact same site, and so only 21 different integration sites are listed in Table 1. In one instance, we detected integration into the same region but in opposite orientations (integrants at positions 211 to 216 and 208 to 213 in Table 1).

Sequence analysis of integrants obtained with the U5(CTTC→GAAG) donor. Integration products from using the U5(CTTC→GAAG) donor were introduced into *E. coli*, and individual clones were isolated and sequenced. The less stringent drug selection conditions were used. The yield of recombinant clones was less than the wild-type yield, as expected (Fig. 5B), and the S4 mutation was detected in individual clones. The sites of integration for 15 such clones are presented in Table 2. A large proportion of the integration events analyzed resembled wild-type events. Integration was distributed throughout the plasmid DNA. As with the wild type, 2 bp from each end of the viral DNA were deleted. There also was a duplication of the acceptor DNA at the site of integration. The frequency of

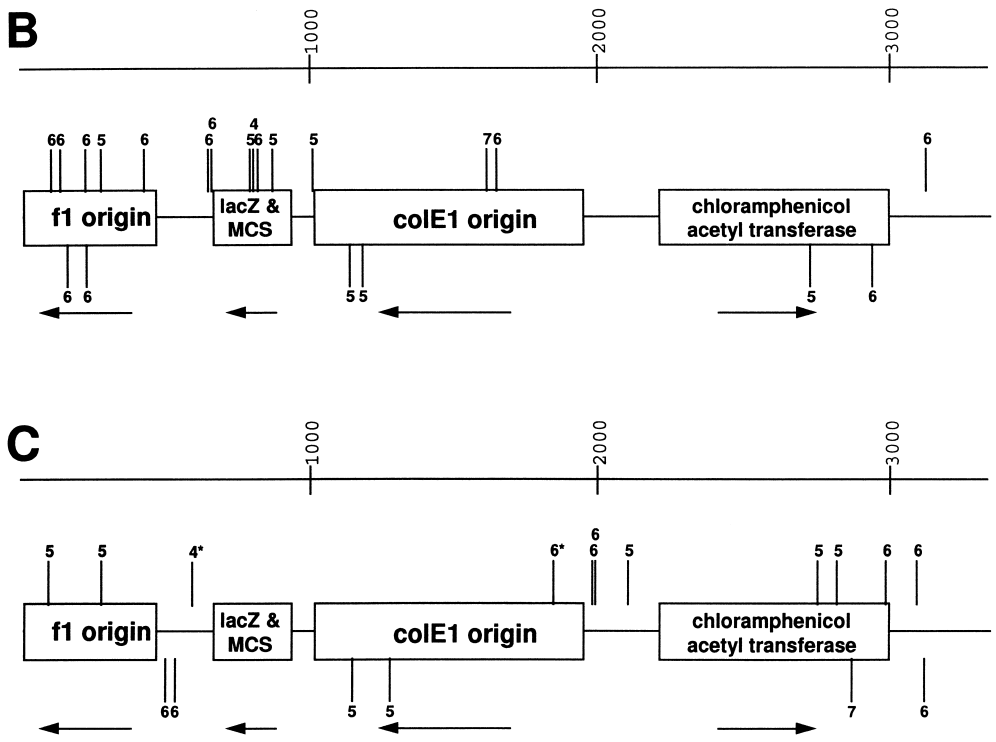


FIG. 7. Biological selection of integrants. (A) Diagrammatic representation of the selection scheme for integrants. The parent *E. coli* strain MC1061/P3 is shown with the P3 episome that confers kanamycin resistance. Products of the integration reaction include nicked acceptor DNA (labeled A), nicked circles resulting from nonconcerted integration (labeled B), nicked circles resulting from concerted integration (labled C), and unused or self-integrated donor DNA (labeled D). Of these four DNAs, only the products of concerted integration possess both the *supF* tRNA gene and a plasmid origin of replication (indicated by the stippled oval), necessary to confer ampicillin and tetracycline resistance to the transformed bacteria. WT, wild type. (B) Location, orientation, and duplication size of cell DNA for integrants, using a wild-type donor DNA as described in Table 1. The data are presented with a linear representation of the acceptor DNA drawn to scale. Each vertical bar represents an integration position. Bars above the plasmid represent integration in the left orientation, while bars below the plasmid represent integrations in the right orientation. The number above or below each bar represents the size of the duplication of the acceptor DNA for that integrant. MCS, multiple cloning site. (C) Location, orientation, and size of duplication of cell DNA for integrants of U5(CTTC→GAAG) mutant donor DNAs listed in Tables 2 and 3 are schematically shown as described in panel B. Asterisks denote integrants associated with a deletion in donor sequence listed in Table 3.

TABLE 1. Sites of integration of the wild-type donor into acceptor DNA

Sequence of donor-acceptor junction[a]	Size (bp) of duplication of cell DNA	Plasmid position of integration	Orientation of integration[b]
taagggTGTAG···CTTCAtaaggg	6	94–99	Left
ccctatTGTAG···CTTCAccctat	6	122–127	Left
gtttggTGTAG···CTTCAgtttgg	6	143–148	Right
ggcgatTGTAG···CTTCAggcgat	6	211–216	Right
ccctgaTGTAG···CTTCAccctga	6	208–213	Left
ccggcTGTAG···CTTCAccggc	5	328–332	Left
ccaacaTGTAG···CTTCAccaaca	6	485–490	Left
ccctatTGTAG···CTTCAccctat	6	640–645	Left
cccggtTGTAG···CTTCAcccggt	6	656–661	Left
ccctcTGTAG···CTTCAccctc	5	781–785	Left
cgccTGTAG···CTTCAcgcc	4	800–803	Left
catgatTGTAG···CTTCAcatgat	6	806–811	Left
ccacaTGTAG···CTTCAccaca	5	854–859	Left
ccccgTGTAG···CTTCAccccg	5	1003–1007	Left
ccacaTGTAG···CTTCAccaca	5	1123–1127	Right
ccagcTGTAG···CTTCAccagc	5	1169–1173	Right
ctctgtaTGTAG···CTTCActctgta	7	1610–1616	Left
ctagccTGTAG···CTTCActagcc	6	1637–1642	Left
ccctgTGTAG···CTTCAccctg	5	2769–2773	Right
actaccTGTAG···CTTCAactacc	6	2954–2959	Right
acagcgTGTAG···CTTCAacagcg	6	3117–3122	Left

[a] Deoxynucleotide sequence of the junction of the donor DNA integrated into the target DNA. The sequence of only one strand of the duplex is shown. Lowercase letters denote duplication of the cell DNA; uppercase letters indicate the processed viral DNA sequences which has lost 2 bp from each end.
[b] Orientation of the integrated donor DNA relative to the ColE1 origin of replication.

non-6-bp duplication, particularly 5-bp duplication, was slightly greater than observed with the wild-type donor. In contrast to the wild type, in two clones (Table 3), a deletion in either the U5 or the U3 terminus was detected. The deletion in the U5 terminus was 10 bp and ended at the first internal CA dinucleotide. The U3 terminus deletion was more unusual. In this case, there was a 5-bp deletion that ended at the first internal GA dinucleotide (Table 3). This was the only integrant that we have detected so far that did not end with a CA dinucleotide at either donor terminus after integration.

DISCUSSION

To facilitate analysis of concerted integration and sequence requirements for this reaction, we designed small integration donors that possess only the terminal 15 bp derived from viral U5 and U3 LTR termini in avian leukosis-sarcoma virus DNA. Both ends have blunt termini, so that integration requires IN to process the ends as well as join the processed ends into an acceptor DNA. While the efficiency of integration is low, with approximately 0.5% of the donor DNA becoming integrated into the acceptor plasmid over the course of the reaction, the amount of donor plus plasmid product is sufficient to detect by ethidium bromide staining of agarose gels. Removal of either the U3 or U5 terminal IN recognition sequence results in significant decreases in integrants. Restriction endonuclease analysis of the donor plus plasmid products also reveals the presence of concerted integration events. Thus, under conditions described here, both ends are joined to the acceptor plasmid in a coordinated and concerted reaction.

Vora et al. (30) have recently described results with a similar, but longer, donor substrate which has preprocessed ends. Coordinated integration products were reported. However, the integration events produced in this system appear to reflect the reac-

TABLE 2. Sites of integration of the mutant donor into acceptor DNA

Sequence of donor-acceptor junction[a]	Size (bp) of duplication of cell DNA	Plasmid position of integration	Orientation of integration[b]
tgccgTGTAG···GAAGAtgccg	5	85–89	Left
accccTGTAG···GAAGAacccc	5	256–260	Right
tcgctaTGTAG···GAAGAtcgcta	6	515–520	Right
cacatcTGTAG···GAAGAcacatc	6	543–548	Right
gaaagTGTAG···GAAGAgaaag	5	1147–1151	Right
ggtggTGTAG···GAAGAggtgg	5	1269–1274	Right
gcccagTGTAG···GAAGAgcccag	6	1992–1997	Left
cccggtTGTAG···GAAGAcccggt	6	1981–1986	Left
gctccTGTAG···GAAGAgctcc	5	2107–2111	Left
ccctgTGTAG···GAAGAccctg	5	2769–2773	Left
aagggTGTAG···GAAGAaaggg	5	2807–2811	Left
gttcaggTGTAG···GAAGAgttcagg	7	2895–2901	Right
agttttTGTAG···GAAGAagtttt	6	2987–2992	Left
aatgccTGTAG···GAAGAaatgcc	6	3097–3012	Left
aggatcTGTAG···GAAGAaggatc	6	3107–3112	Right

[a] Deoxynucleotide sequence of the junction of the donor DNA integrated into the target DNA. The sequence of only one strand of the duplex is shown. Lowercase letters denote the duplication of the cell DNA; uppercase letters indicate the processed viral DNA sequences which has lost 2 bp from each end.
[b] Orientation of the integrated donor DNA relative to the ColE1 origin of replication.

TABLE 3. Sites of integration of the mutant donor into acceptor DNA with LTR deletions

Sequence of donor-acceptor junction[a]	Size (bp) of duplication of cell DNA	Plasmid position of integration[b]
tgccgTGTAGTCTTA···AGCAGAAGGGAAGAtgccg	5	85–89[c]
actgTGTAGTCTTA···AGCAactg	4	587–590
cagtggTCTTA···AGCAGAAGGGAAGAcagtgg	6	1849–1854

[a] Deoxynucleotide sequence of the junction of the donor DNA integrated into the target DNA. The sequence of only one strand of the duplex is shown. Lowercase letters denote the duplication of the cell DNA; uppercase letters indicate the processed viral DNA sequences from each end.
[b] Both integrated deletion donor DNAs clones were in the left orientation relative to the ColE1 origin of replication.
[c] The integrant was taken from Table 2 for reference.

tion of separate donor DNA molecules with the acceptor DNA, in which only one LTR terminus per donor is used. The difference in results with their system and the one reported here could reflect a difference in affinity of IN for blunt versus preprocessed viral DNA ends and/or differences in lengths of the two donor DNAs used. The latter interpretation would be consistent with our observation that addition of a ubiquitous DNA-bending protein, HMG1, to the integration reaction increased efficiency by approximately fourfold. This protein has been shown to bend DNA as short as 60 bp into a circular structure whose ends can be covalently linked by DNA ligase (24). The addition of the bacterial DNA-bending protein, HU, also stimulated the integration reaction, although to a smaller extent (data not shown). Since preincubation of HMG1 with the donor DNA alone was sufficient to observe the full effects of stimulation by HMG1, we believe that this protein bends the donor into a conformation that facilitates the concerted recognition of U5 and U3 termini from the same molecule by an IN multimer. However, it is also possible that HMG1 causes bends near the end of a linear donor, which improves its efficiency of nucleophilic attack on the acceptor DNA by some unknown mechanism.

We have used integration donors that encode the *supF* tRNA gene to facilitate biological selection and cloning of integrants. Sequence analyses of several donor-acceptor junctions confirmed that the biochemical integration events resulting from the reconstituted system exhibited hallmarks of authentic retroviral integration in vivo. This included the removal of 2 bp from each LTR terminus and the duplication of acceptor DNA at the site of integration. We do not observe deletions of the acceptor DNA at the insertion sites. The sites of integration also appear to be random. Integrants were found in the region of the ColE1 origin of replication, in the *lacZ* gene, in the chloramphenicol acetyltransferase gene, and in the region of the F1 origin of replication. As others have reported (8), we observed instances of multiple integrations into the same site on the acceptor plasmid. In one instance, we observed integration at the same region but in opposite orientations (Table 1). When the bacteria transformed with integration reaction products were grown under conditions recommended for selection of *supF*-containing plasmids such as πvx, all of the recovered recombinant plasmids represented products of integration in which *supF* transcription was oriented toward the ColE1 origin of replication. When the selection was repeated under less stringent conditions, integrants in both transcription orientation were obtained. Examination of the target DNA sequence indicates that there is a strong rho-independent consensus transcription termination signal that would be functional for donors integrated in the leftward direction (21). Thus, donors integrated in this orientation are expected to produce more copies of *supF* tRNA and may be able to grow better in the presence of the antibiotic. This is consistent with our finding that integration in the rightward orientation was detected only when the concentration of the antibiotic applied during the selection was reduced. These results suggest that the pool of recovered integrants is influenced by transcription orientation within the acceptor plasmid. For this reason, we cannot conclude that we have detected hot spots of integration per se within an acceptor DNA as previously reported (19, 27), even though we have detected the exact same integration events in a few independently isolated clones from different transformations.

The sizes of the duplications of the acceptor DNA at the sites of donor DNA integration ranged from 4 to 7 bp, with the 6-bp duplication being found in the largest number of integrants (57% of the total). Integration of avian retroviral DNA in vivo usually occurs with a 6-bp duplication of the cell DNA (8, 28). While some of our integrants display duplications that are slightly smaller or larger than those observed in vivo, we note that the frequency of the expected duplication detected in this system is higher than that described by others using preprocessed donors, in which case significantly larger duplications as well as deletions within acceptor sequences have been detected (5, 10, 14, 17). Neither of these events has been observed to occur during avian retroviral integration in vivo. With the present reconstituted system, the size of the duplication of the acceptor DNA appears to be influenced by the site of integration. Initial statistical analysis indicates that there may be clustering of events which produce six-nucleotide duplications with the wild-type donor. This bias may be a result of plasmid supercoiling, and this is presently being tested by using a covalently closed relaxed circular acceptor DNA.

We have used the present integration system to examine the effect of mutations in the U5 IN recognition sequences. A particularly interesting mutation changes four nucleotides near the U5 terminus of the donor including the highly conserved CA dinucleotide to GA. This mutation, referred to as S4, was previously analyzed in vitro with a duplex oligodeoxynucleotide assay and in vivo by using a PCR-based assay (6). Paradoxically, while this mutation abrogated processing in the duplex oligodeoxynucleotide assay, viruses with this mutation integrated into host chromosome in vivo with efficiencies that were only two- to threefold lower than that of the wild type (6). This discrepancy might suggest that a mutation in the U5 terminus could be partially compensated for by the presence of a wild-type U3 terminus on the same donor DNA. There are several other examples of substrate alterations that cause severe defects when analyzed as single-LTR donors but cause mild defects when analyzed in vivo (4, 11, 22, 25). Using the present purified reconstituted concerted integration system, we have found that a donor DNA containing a U5 S4 mutation integrates into the acceptor DNA randomly but with an efficiency far greater than predicted from analysis of the same mutation in the context of a single-LTR donor. In one of the integrants sequenced, we observed abnormal processing of the wild-type U3 end of a donor that included a mutant U5 sequence. This might indicate only if both ends of the donor were coordinately recognized and used during integration. Tang et al. (29) have shown that the deleterious effects of mutations within one of the inverted repeats of the transposon Tn*7* can be partially repressed by sequence changes within the other inverted repeat of the same transposon. This finding suggests that coordinated recognition

of both donor ends has been conserved between the bacterial insertion sequences and the retroviruses, which are otherwise evolutionarily distinct. In the present reconstituted system, IN catalyzes a cooperative and concerted integration reaction, independent of any cell accessory factors, in which LTR sequences at one end of a donor can influence IN recognition of LTR sequences at the other. Similar results have been reported by Kukolj and Skalka (20). The ability of IN to catalyze some integration of DNA in which one of the two LTRs has base changes may be advantageous to a virus which utilizes an error prone reverse transcriptase to replicate its RNA.

One additional difference between wild-type and S4 integrants is worth noting. Two of the S4 integrants were found to have a deletion of the U5 or U3 sequence, bringing an internal CA or GA sequence, respectively, to the virus-acceptor DNA junction (Table 3). Previously, Colicelli and Goff (9) demonstrated that mutations which altered the highly conserved dinucleotide terminal sequence resulted in pseudorevertants that integrated by using internal alternate CA dinucleotides. These results, taken together, indicate that the present in vitro system displays all of the subtleties known for in vivo integration and should therefore be very useful for further analysis of the mechanism of concerted retroviral integration.

ACKNOWLEDGMENTS

We thank Stephen Lippard, MIT, for generous gifts of the HMG1 protein and Ken Marians, Sloan Kettering Institute, for the bacterial HU protein.

This work was supported in part by Public Health Service grants CA38046 (J.L.), CA47486 (A.M.S.), and CA06927 (A.M.S.) from the National Cancer Institute and by Cancer Research Center grant P30 CA43703. A.A. was supported by in part by funds from AIDS Institutional Training Grant AI07381 from the NIH.

REFERENCES

1. **Aiyar, A., Z. Ge, and J. Leis.** 1994. A specific orientation of RNA secondary structures is required for initiation of reverse transcription. J. Virol. **68:**611–618.
1a. **Aiyar, A., and J. Leis.** Unpublished data.
2. **Braiterman, L. T., and J. D. Boeke.** 1994. In vitro integration of retrotransposon Ty1: a direct physical assay. Mol. Cell. Biol. **14:**5719–5730.
3. **Brown, P., B. Bowerman, H. Varmus, and J. M. Bishop.** 1987. Correct integration of retroviral DNA *in vitro*. Cell **49:**347–356.
4. **Bushman, F. D., and R. Craigie.** 1990. Sequence requirements for integration of Moloney murine leukemia virus DNA in vitro. J. Virol. **64:**5645–5648.
5. **Bushman, F. D., T. Fujiwara, and R. Craigie.** 1990. Retroviral DNA integration directed by HIV integration protein in vitro. Science **249:**1555–1558.
6. **Cobrinik, D., A. Aiyar, Z. Ge, H. Huang, and J. Leis.** 1991. Overlapping retroviral U5 sequence elements are required for efficient integration and initiation of reverse transcription. J. Virol. **65:**3864–3872.
7. **Chow, C., C. Barnes, and S. Lippard.** 1995. A single HMG domain in high-mobility group 1 protein binds to DNAs as small as 20 base pairs containing the major cisplatin adduct. Biochemistry **34:**2956–2964.
8. **Coffin, J. M.** 1990. Retroviridae and their replication, p. 1437–1500. *In* B. N. Fields, D. M. Knipe, R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman (ed.), Virology, 2nd ed. Raven Press, Ltd., New York.
9. **Colicelli, J., and S. Goff.** 1985. Mutants and pseudorevertants of Moloney murine leukemia virus with alterations at the integration site. Cell **42:**573–580.
10. **Craigie, R., A.-T. Fujiwara, and F. Bushman.** 1990. The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. Cell **64:**829–837.
11. **Donzella, G. A., C. B. Jonsson, and M. J. Roth.** 1993. Influence of substrate structure on disintegration activity of Moloney murine virus integrase. J. Virol. **67:**7077–7087.
12. **Engleman, A., K. Mizuuchi, and R. Craigie.** 1991. HIV-1 DNA integration: mechanism of viral DNA cleavage and strand transfer. Cell **67:**1211–1221.
13. **Fitzgerald, M. L., and D. P. Grandgenett.** 1994. Retroviral integration: in vitro host site selection by avian integrase. J. Virol. **68:**4314–4321.
14. **Fitzgerald, M. L., A. C. Vora, W. G. Zeh, and D. P. Grandgenett.** 1992. Concerted integration of viral DNA termini by purified avian myeloblastosis virus integrase. J. Virol. **66:**6257–6263.
15. **Fujiwara, T., and R. Craigie.** 1989. Integration of mini-retroviral DNA: a cell free reaction for biochemical analysis of retroviral integration. Proc. Natl. Acad. Sci. USA **86:**3056–3069.
16. **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.
17. **Katz, R., G. Merkel, J. Kulkosky, J. Leis, and A. Skalka.** 1990. The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. Cell **63:**87–95.
18. **Katzman, M., R. Katz, A. M. Skalka, and J. Leis.** 1989. The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the in vivo sites of integration. J. Virol. **63:**5319–5327.
19. **Kitamura, Y., Y. M. Lee, and J. M. Coffin.** 1992. Nonrandom integration of retroviral DNA *in vitro*: effect of CpG methylation. Proc. Natl. Acad. Sci. USA **89:**5532–5536.
20. **Kukolj, G., and A. M. Skalka.** 1995. Enhanced and coordinated processing of synapsed viral DNA ends by retroviral integrases *in vitro*. Genes Dev. **9:**2556–2567.
21. **Lewin, B.** 1990. Genes IV. Oxford University Press, Oxford.
22. **Murphy, J. E., T. de Los Santos, and S. P. Goff.** 1993. Mutational analysis of the sequences at the termini of the Moloney murine leukemia virus DNA required for integration. Virology **195:**432–440.
23. **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 the viral integrase in vivo. J. Virol. **66:**5092–5095.
24. **Pil, P. M., C. S. Chow, and S. J. Lippard.** 1993. High-mobility-group 1 protein mediates DNA bending as determined by ring closures. Proc. Natl. Acad. Sci. USA **90:**9465–9469.
25. **Roth, M. J., P. L. Schwartzberg, and S. P. Goff.** 1990. Structure of the termini of DNA intermediates in the integration of retroviral DNA: dependence on IN function and terminal DNA sequence. Cell **58:**47–54.
26. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
27. **Shih, C. C., J. P. Stoye, and J. M. Coffin.** 1988. Highly preferred targets for retroviral integration. Cell **53:**531–537.
28. **Skalka, A. M.** 1993. Retroviral DNA integration: lessons for transposon shuffling. Gene **135:**175–182.
29. **Tang, Y., S. Cotterill, and C. P. Lichenstein.** 1995. Genetic analysis of the terminal 8-bp inverted repeats of transposon Tn7. Gene **162:**175–182.
30. **Vora, A. C., M. McCord, M. L. Fitzgerald, R. B. Inman, and D. P. Grandgenett.** 1994. Efficient concerted integration of retrovirus-like DNA in vitro by avian myeloblastosis virus integrase. Nucleic Acids Res. **224:**4454–4461.