MOLECULAR AND CELLULAR BIOLOGY, Sept. 1994, p. 5731–5740
0270-7306/94/$04.00+0

Vol. 14, No. 9

Ty1 In Vitro Integration: Effects of Mutations in *cis* and in *trans*

LELITA T. BRAITERMAN† AND JEF D. BOEKE*

Department of Molecular Biology & Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received 8 December 1993/Returned for modification 2 February 1994/Accepted 26 May 1994

Mutations within the *TYB* gene of Ty1 encoding integrase (IN) as well as alterations in its substrate, a linear DNA molecule, were examined for their effects on in vitro IN activity, using a recently developed physical assay. Five different codon-insertion mutations, two frameshift mutations, and one missense mutation, previously identified as transposition-deficient mutations, were tested. Virus-like particles, the source of IN, from two different protease mutants and a reverse transcriptase mutant exhibited near-normal to normal IN activity. Two frameshift mutations mapping within the phylogenetically variable C-terminal domain of IN resulted in significant in vitro IN activity. In contrast, three mutations within the amino-terminal conserved domain of IN completely abolished IN activity. When the substrate termini were mutated, we found that substrates with as few as 4 bp of Ty1 termini were capable of efficiently generating integration products. Surprisingly, certain substrates that lacked obvious similarity to Ty1 termini were also readily integrated into both linear and circular targets, whereas others were not used as substrates at all. Termini rich in adenosine residues were among the more active substrates; however, certain substrates lacking terminal adenosine residues can form small quantities of integration products, including complete integration reactions.

The retrotransposon Ty1 of the yeast *Saccharomyces cerevisiae* reinvades the genome of its host by a process similar to the retroviral life cycle (6). Ty1 resembles retroviral proviruses in its long terminal repeat (LTR) structure and in the presence of short target site duplications. Ty1 mRNA is packaged into virus-like particles (VLPs) formed by the Ty1-encoded gene products TYA, the functional analog of retroviral Gag protein, and TYA/TYB, analogous to the retroviral Gag/Pol polyprotein (26, 33). The Ty1 protease (PR) processes TYA/TYB to release PR, integrase (IN), and reverse transcriptase (RT) which are embedded within the TYA/TYB polyprotein (27). The RT domain generates a double-stranded (ds) cDNA copy of Ty1 RNA, and then the IN protein inserts this cDNA into the host genome. The process of retrovirus and retrotransposon integration has been studied by using in vitro reactions which are dependent on three macromolecular components: (i) target DNA; (ii) dsDNA with specific terminal sequences, termed substrate DNA; and (iii) IN, either purified or supplied as a nucleoprotein complex (7, 11, 12, 16, 18, 24, 35).

When retrotransposon IN and retroviral IN sequence alignments are examined, an N-terminal conserved domain and a C-terminal variable domain are observed (29, 30). Mutations within the conserved N-terminal half of IN inactivate both retroviral infectivity (20, 37) and Ty1 transposition (10, 21). Two features that distinguish retrotransposon Ty1 IN from retroviral IN are (i) the apparent lack of 3′ processing activity (it is predicted from the sequence that the ends of the Ty1 dsDNA copy terminate with the conserved 5′TG. . .CA3′ short inverted repeat, suggesting a lack of end processing [22]) and (ii) a large C-terminal domain. Thus, the activities shared by retroviral and retrotransposon INs, i.e., DNA substrate end recognition and the DNA joining reactions, are proposed to map to the N-terminal region of IN that contains two amino acid sequence motifs that are highly conserved phylogenetically. On the basis of results of recent functional studies of retroviral integration and the reverse (disintegration) reactions, the conserved N terminus of IN has been further subdivided into an N-terminal zinc finger-like domain and a central catalytic core domain (13, 15, 23, 32, 39). In vitro assays for 3′ processing, strand transfer, and disintegration have been used to examine the activities of various retroviral IN mutations within these motifs. A highly conserved zinc-binding motif at the N terminus of IN (13, 15, 23), consisting of two histidine residues followed by two cysteine residues, designated HHCC, may be involved in substrate recognition (15, 32, 39). The other highly conserved motif, characterized by two invariant aspartate (D) residues and a glutamate (E) residue, referred to as the D, D(35)E region (30), was also mutagenized, and the invariant residues were essential for in vitro integration and disintegration activities (23, 31, 39).

We describe in the accompanying paper a very sensitive and quantitative physical assay for Ty1 in vitro integration whereby we are able to identify both single-end insertions and complete insertions of one DNA substrate into target DNA (9). Briefly, VLPs, the source of native Ty1 IN, are incubated with a radiolabeled substrate bearing Ty1 termini and either linear or circular target DNA. Integration products are radioactive and after agarose gel electrophoresis have a slower mobility than either substrate or target DNAs. We have characterized several different products formed with these target DNAs. The L1/L2 product represents substrate joined to linear target which is composed of a predominant product of single-end joining and a less abundant product of complete insertions. When circular target DNA is supplied, three products, C1, formed from the coordinated insertion of two substrates; C2.1, formed by a complete insertion into the circular target DNA; and C2.2, a single-end joining of substrate to the circular target DNA, are then observed (9). Although it is possible that one or more cellular factors copurifying with the VLPs contribute to this activity, we have shown that product formation is completely dependent on Ty1 IN.

Transposition in vivo can be easily assayed by using pGTy1 plasmids consisting of the regulated yeast *GAL1* promoter

* Corresponding author. Mailing address: Department of Molecular Biology & Genetics, Johns Hopkins University School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205. Phone: (410) 955-2481. Fax: (410) 955-0831. Electronic mail address: jef_boeke@qmail.bs.jhu.edu.

† Present address: Department of Pediatrics, Center for Medical Genetics, Johns Hopkins Medical Institutions, Baltimore, MD 21205.

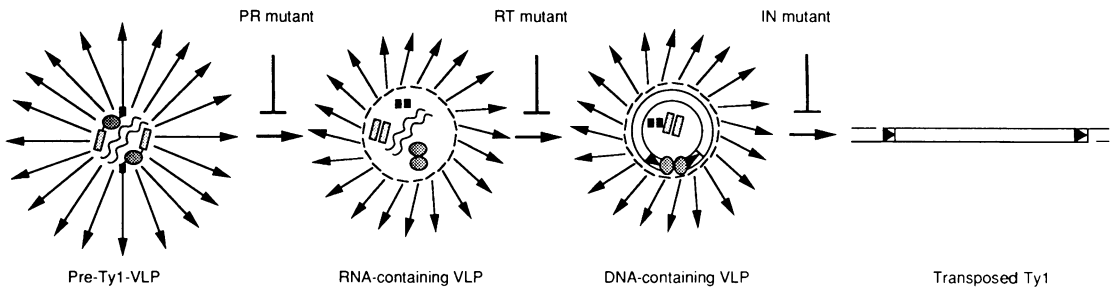

FIG. 1. Model for VLP maturation. A hypothetical model for the sequence of events leading to mature DNA-containing Ty1 VLPs and transposition is shown. There is no direct evidence for a discrete population of each of these VLPs during a normal transposition event. However, mutations within each functional domain of *TYB* lead to an accumulation of the indicated type of VLP. Mutations in PR block processing, and pre-VLPs accumulate; mutations in RT block reverse transcription, and RNA-containing VLPs accumulate; and mutations within IN block integration, and DNA-containing VLPs accumulate. The arrowheads represent N-termini of TYA and TYA/TYB; long arrows, unprocessed TYA; long arrows with attachments, unprocessed TYA/TYB; short arrows with bar, processed TYA; filled-in squares, PR dimer; ovals, IN dimer; rectangles, RT dimer; wavy lines, Ty1 RNA (two copies); open box, Ty1 cDNA; boxed triangles, LTRs; double lines, target DNA.

fused to a functional Ty1 element (Ty1-H3) (5). The *GAL1* sequences are not transcribed; thus, the transcripts from the pGTy1 plasmids are identical in structure to those from native chromosomal Ty1 elements. Upon galactose induction, not only is transposition easily assayed, but also large quantities of VLPs can be isolated from yeast extracts, providing a reliable source of IN.

A model for the Ty1-VLP maturation pathway is depicted in Fig. 1. TYA and TYA/TYB precursor proteins presumably self-assemble with Ty1 RNA to form an unprocessed pre-VLP (1, 14, 36, 42). Upon proteolytic processing, RNA-containing VLPs mature to DNA-containing VLPs after reverse transcription of the Ty1 RNA. In this report, we compared IN activity of wild-type VLPs and those VLPs derived from mutated pGTy1 plasmids in order to examine the effects of alterations to the virion environment. VLPs derived from five different Ty1 IN mutants as well as VLPs with PR and RT defects were assayed. All mutations studied inactivated transposition in vivo (25, 34). Most of them were generated by insertion of oligonucleotides semirandomly throughout a pGTy1 plasmid. These mutations result in alterations of a single domain of the TYA/TYB polyprotein (34). Two of the IN mutations within the N-terminal conserved domain, pGM66 and pGM119, were previously shown to be inactive in an in vitro integration assay which recovered complete insertions genetically (21). The pGM66 mutation was also shown to be inactive by using a physical in vitro integration assay developed for Ty1 (9). All *TYB* mutations referred to in this paper are summarized in Table 3.

The terminal 335 nucleotides of each end of the Ty1 chromosomal element and the dsDNA copy compose an intact LTR. The Ty1 terminal inverted repeats consist of only two nucleotides, TG. .CA, as opposed to the more extensive inverted repeats found in retroviruses (for a review, see reference 3). Therefore, it might be expected that only a small number of terminal bases would be required for Ty1 integration. Indeed, we previously showed that substrates with 12-bp Ty1 termini were sufficient for integration in vitro (22). Thus, we have begun an investigation of the sequence requirements of the substrate termini by using both the physical and the genetic assays described in the accompanying paper (9). Initial studies using various substrates with non-Ty1 ends showed surprisingly high levels of IN activity; however, this assay does not readily distinguish single-end insertions from complete insertions. In an effort to define the terminal sequences that are required for efficient insertion of a DNA substrate, we have developed a new strategy for readily generating large amounts of substrate DNAs with blunt-ended termini of any desired sequence.

MATERIALS AND METHODS

Yeast strains and plasmids. The *S. cerevisiae* strains and the transformants used in this study are listed in Tables 1 and 2. For rapid induction in the presence of galactose, YH8 haploid strains (which bear the *trp1Δ1* deletion that also removes the upstream activating sequences of *GAL3* that are required for rapid galactose induction [2]) were mated with strain YH82, which is *GAL3*$^+$. The media used in this work, SC−ura glucose, SC−ura galactose, SD, SGAL, and YPD, have been described elsewhere (38). Plasmids used to make substrates for IN reactions are pLB343 and pLB5 (see Fig. 2); plasmids used as target DNAs are pLB523 and pJEF2276 (see reference 9 for descriptions).

VLP isolation and characterization. Yeast cells bearing wild-type pGTy plasmid and RT$^-$ and IN$^-$ mutants (Table 3) were harvested after 24 h of galactose induction; VLPs were prepared from a 20 to 70% linear sucrose gradient as previously described (10). Gradient fractions (1.2 ml) were analyzed for RT activity (26) and IN protein as previously described (10). Briefly, pools of pairs of gradient fractions were analyzed by immunoblot analysis with anti-IN monoclonal antibody 8B11 (21); the IN-containing fractions were identified and 5-μl aliquots were used directly as the source of IN activity (9).

Yeast cells bearing pGTy1 plasmids with PR$^-$ mutations (Table 3) and the companion wild-type plasmid in isogenic strains were prepared by three different methods. PR$^-$ mutant VLPs were inactive when purified as described above. Therefore, PR$^-$ mutant VLPs were harvested from the 30/70 interface of a 20%/30%/70% sucrose step gradient as previously described (21). The three fractions with the greatest RT activities were diluted threefold with buffer B-EDTA (10 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [HEPES]-

TABLE 1. Yeast strains used in this study

Strain	Genotype	Reference
YH8	*MATα ura3-167 his3Δ200 leu2Δ1 trp1Δ1*	41
GRF167	*MATα ura3-167 his3Δ200*	5
YH82	*MATa ura3-52 lys2-801 his4-539 trp1Δ63 leu2Δ1*	41

TABLE 2. Transformants used in this study

Strain	Host	Plasmid	Description (reference)
JB224	GRF167	pJEF724	Wild-type unmarked pGTy1 (5)
LTB185	GRF167 × YH82	pJEF724	Wild-type unmarked pGTy1 (5)
LTB191	GRF167 × YH82	pJEF1266	PR mutant unmarked pGTy1 (34)
LTB151	YH8 × YH82	pJEF1562	Wild-type *lacO*-marked pGTy1 (34)
LTB167	YH8 × YH82	pGM100	IN mutant *lacO*-marked pGTy1 (34)
LTB169	YH8 × YH82	pGM119	IN mutant *lacO*-marked pGTy1 (34)
LTB171	YH8 × YH82	pGM66	IN mutant *lacO*-marked pGTy1 (34)
LTB175	YH8 × YH82	pGM129	IN mutant *lacO*-marked pGTy1 (34)
LTB177	YH8 × YH82	pGM95	IN mutant *lacO*-marked pGTy1 (34)
LTB111	YH8 × YH82	pX3	Wild-type *TRP1*-marked pGTy1 (41)
LTB120	YH8 × YH82	pAG149	RT mutant *TRP1*-marked pGTy1 (25)
LTB114	YH8 × YH82	pGM315	IN mutant (pGM66) *TRP1*-marked pGTy1 (34)
LTB128	YH8 × YH82	pGM313	PR mutant (pGM17) *TRP1*-marked pGTy1 (34)

KOH [pH 7.8], 15 mM KCl, 3 mM dithiothreitol) to reduce the sucrose concentration and then pelleted at 80,000 × *g* for 1 h at 4°C. Pellets were resuspended in 100 μl of TE (10 mM Tris-HCl [pH 7.4], 1 mM EDTA); 2 μl was used as the source of IN. In an effort to obtain more-purified preparations from strains bearing pGM17 and the wild-type pGTy1 plasmids, PR^- VLPs were isolated from two sequential sucrose gradients as previously described (21); 2 μl of pelleted and resuspended VLPs was used as the source of IN.

In vitro integration assay. IN assays were performed as previously described (9). The standard radiolabeled DNA substrate (936 bp) was prepared from plasmid pLB343 (Fig. 2). For later studies that required altered substrate ends, a PCR strategy was developed (see "PCR-generated substrates" below). Target DNA pJEF2276 was digested with *Xba*I for experiments using a linear target. The reaction mixtures contained 10 mM KCl, 10 mM Tris-HCl (pH 8.0), 15 mM $MgCl_2$, 1 mM dithiothreitol, 5% polyethylene glycol 8000, 5 ng of radiolabeled substrate (5×10^9 molecules), and 100 or 200 ng of linear or circular target DNA (3×10^{10} molecules or 6×10^{10}, respectively) giving molar ratios of 1:6 or 1:12, respectively; 5-μl volumes of VLP preparations are added last to the reaction mixture to give a total reaction volume of 20 μl. Reaction conditions were slightly modified in the experiment used to assay PR^- VLPs; 2 μl of either wild-type or PR^- pelleted VLPs was added to a reaction mixture that contained 10 mM $MgCl_2$, 10 mM Tris-HCl (pH 7.4), 10 mM KCl, 0.1 mM dithiothreitol, and 5% polyethylene glycol 8000, giving a final reaction mixture volume of 20 μl.

Alterations to the substrate ends by restriction digest. Six different substrate preparations were prepared from pLB343

TABLE 3. Mutations studied[a]

Domain and mutant	Type of mutation	Ty1 genome position[b]
PR		
pJEF1266	4-codon insertion	1703
pGM17	5-codon insertion	1873
IN		
pGM100	5-codon insertion	2164 (aa 43)
pGM119	5-codon insertion	2600 (aa 189)
pGM66	4-codon insertion	2724 (aa 229)
pGM129	Frameshift insertion	3280 (aa 415)
pGM95	Frameshift insertion	3796 (aa 587)
RT, pAG149	Missense mutation (FVDD→FVDE)	4578

[a] All mutations are described in further detail in references 10 and 34, except AG149 (25).

[b] The genome position refers to the Ty1-H3 sequence (4). The positions of the mutations in the IN protein sequence relative to the experimentally derived N terminus (8, 35) are given in parentheses. aa, amino acid.

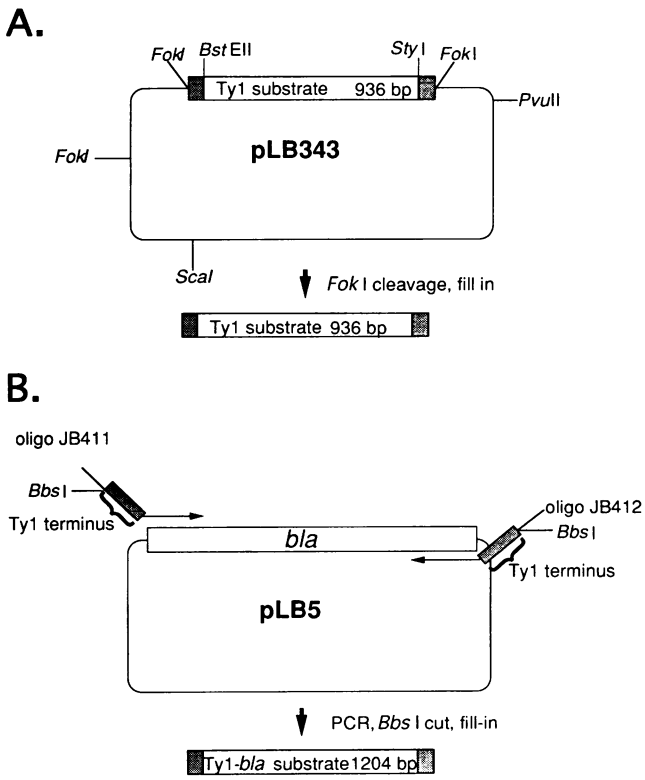

FIG. 2. IN substrates. (A) The structure of the wild-type DNA substrate derived from pLB343 by *Fok*I digestion of the 2,636-bp *Pvu*II-*Sca*I fragment is shown. Only restriction enzyme sites used to obtain the various substrates analyzed in Fig. 7 are shown. Shaded boxes designate the 12-bp Ty1 termini. (B) Linear DNA fragments bearing internal regions corresponding to Ty1 termini were obtained by PCR with primers JB411 and JB412 as shown. The blunt-ended DNA substrate with 4 bp of Ty1 termini was obtained by *Bbs*I digest of the PCR products and a subsequent fill-in reaction with deoxynucleoside triphosphates and PolIK. Shaded boxes designate 4 bp of Ty1 termini or regions of the substrate which were altered by changes to the PCR primers as described in Materials and Methods.

TABLE 4. Sequences of oligonucleotides used in this study

Oligonucleotide[a]	Sequence[b]
Right end	
JB411 (U5)	5′-CACAGAAGACCA**TGAG**GGGGTCTGACGCTCAGTG
JB421 (T)	5′-CACAGAAGACCA**AAAA**GGGGTCTGACGCTCAGTG
JB422 (A)	5′-CACAGAAGACCA**TAAA**GGGGTCTGACGCTCAGTG
JB481 (U3)	5′-CACAGAAGACCA**TGTT**GGGGTCTGACGCTCAGTG
Left end	
JB412 (U3)	5′-ACACGAAGACTG**TGTT**GTCCTCGTGATACGCC
JB423 (T)	5′-ACACGAAGACTC**AAAA**GACCTCGTGATACGCC
JB424 (A)	5′-ACACGAAGACTC**TAAA**GACCTCGTGATACGCC

[a] The symbols in parentheses correspond to the substrate terminus types in Fig. 8; U3, $AACA_{OH}$; U5, $CTCA_{OH}$; T, $TTTT_{OH}$; A, $TTTA_{OH}$.
[b] Sequences representing the 5′ termini of complete substrates are in boldface type. *Bbs*I sites are underlined.

by using various restriction enzyme combinations as indicated (see Fig. 7). For substrate lengths of 936, 690, and 201 bp, a 2,326-bp *Pvu*II-*Sca*I fragment was isolated as described elsewhere (9) and digested with *Fok*I. For all substrates, the ends were radiolabeled with [α-^{32}P]dATP, dCTP, dGTP, and dTTP and DNA polymerase I Klenow large fragment (PolIK). The fragments were separated on a 0.9% agarose gel, and bands were cut from the gel, electroeluted, and ethanol precipitated. The fragments were resuspended to give equimolar preparations.

PCR-generated substrates. Synthetic DNA oligonucleotides that incorporate the *Bbs*I restriction site and four nucleotides derived from native Ty1 termini (or variations of this region) were designed to hybridize to pLB5 or pUC19. Following PCR amplification with *Taq* polymerase (Cetus) and under the standard conditions recommended by the manufacturer, a 1,227-bp fragment containing two *Bbs*I sites and Ty1 termini flanking the *bla* gene is generated. Upon *Bbs*I digestion the Ty1 termini are exposed and can be radiolabeled as described for the 936-bp substrate (9), yielding a 1,208-bp blunt-ended substrate (Fig. 2). The sequences of the wild-type oligonucleotides used are as follows (*Bbs*I sites are underlined, and Ty1 termini are in boldface type): JB411, 5′-CACAGAAGAC CA**TGAG**GGGGTCTGACGCTCAGTG; JB412, 5′-ACAC GAAGACTG**TGTT**GTCCTCGTGATACGCC. All oligonucleotides used for generation of substrates are shown in Table 4.

RESULTS

Pre-VLPs: effect of a mutation within PR. Pre-VLPs are unprocessed and contain Ty1 RNA; a mutation within PR inhibits maturation of VLPs, resulting in accumulation of pre-VLPs (1, 36, 42). It is unknown whether mature and precursor polyproteins are significantly different in structure or function; however, PR defects result in complete inhibition of transposition in vivo (34, 42). When RT is embedded within the TYA/TYB polyprotein, PR$^-$ mutant VLPs exhibit high levels of RT activity using exogenous primer and templates; however, no dsDNA is found in these VLPs. One explanation for the lack of RT activity in vivo may be the physical constraint imposed by a PR$^-$ mutant, since processing untethers the enzymatic activities from the polyprotein providing the mobility to perform various functions. Alternatively, processing may enable RT to properly recognize or assemble a tRNA primer-template complex.

Previous studies of RT activity in PR$^-$ mutants indicate that precursor polyproteins have enzymatic activity in vitro (42).

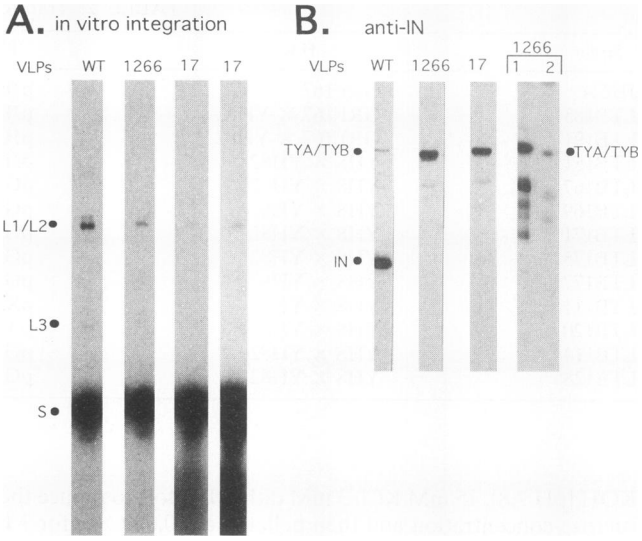

FIG. 3. Ty1 pre-VLPs have IN activity. Pre-VLPs were prepared by the sucrose step gradient method (21) from two PR mutant strains, LTB191 harboring plasmid pJEF1266, and LTB128 harboring a derivative of pGM17 (Table 2). The wild-type strain LTB185 harbors the pGTy1 plasmid pJEF724 (5). (A) In vitro IN activity of PR-VLPs. In vitro IN activity was determined under conditions optimized for pelleted VLPs as described in Materials and Methods. Three radioactive products, L1, L2, and L3, as well as the radioactive substrate (S) are detected by autoradiography of a dried agarose gel in the wild-type (WT) and pJEF1266 (1266) lanes. A small amount of the L1/L2 product is detected in the pGM17 (17) lane. Note that radioactive breakdown products are more abundant in the 17 lane; variable amounts of endogenous nuclease activity (seen as a smear below the S band) are detected in VLP preparations. We have also included a long exposure of the pGM17 lane to more clearly demonstrate the presence of product present. (B) Immunoblot analysis of PR$^-$ VLPs. By immunoblot analysis, both processed IN (lane WT) and unprocessed TYA/TYB form of IN (lanes 1266 and 17) are detected. VLPs from the 1266 PR$^-$ mutant were processed for immunoblot analysis after trichloroacetic acid precipitation. Unprocessed TYA/TYB and nonspecific degradation products are detected (1). An aliquot of the same VLP preparation was processed for immunoblot analysis after trichloroacetic acid precipitation from an in vitro integration assay (2).

Thus, we have examined transposition-deficient PR$^-$ mutant-derived pre-VLPs for IN activity. The TYA/TYB polyprotein of VLP preparations from two PR$^-$ mutants, pGM17 and pJEF1266 (Table 3), exhibited IN activity; the pJEF1266 mutation is more active than the pGM17 mutation in this preparation (Fig. 3A). By immunoblot analysis of PR$^-$ mutant pre-VLPs with anti-IN, the majority of IN protein detected is in the unprocessed TYA/TYB polyprotein form (Fig. 3B). Although some faster-migrating bands are observed, they are most likely the result of nonspecific proteolysis occurring during VLP purification, since these species do not comigrate with the processing intermediates observed for wild-type VLPs in vivo. Furthermore, their mobilities and abundance are not reproducible from preparation to preparation. We did not observe any significant alteration to the PR$^-$ VLP protein pattern after an in vitro IN reaction as assessed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and subsequent immunoblot analysis with anti-IN (Fig. 3B).

Another preparation of mutant pre-VLPs derived from the pGM17 mutant and the companion wild-type pGTy plasmid were purified from two sucrose gradients in an effort to reduce

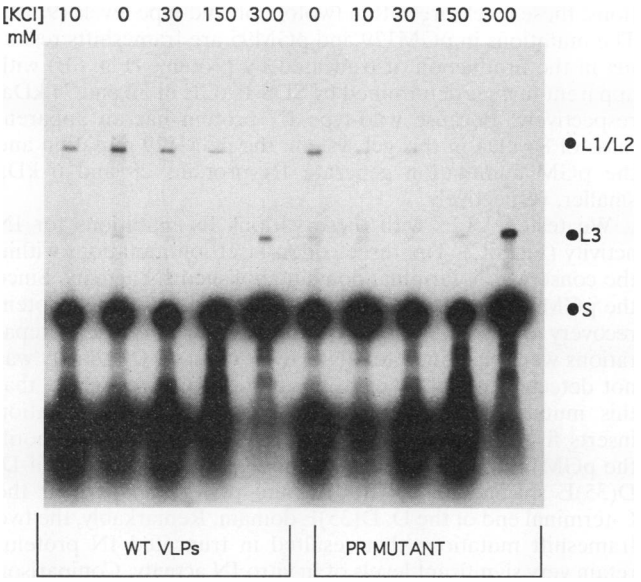

FIG. 4. Effects of salt concentration on in vitro IN activity of pre-VLPs. Numbers designate the concentration of KCl added to in vitro IN reaction mixtures. Equivalent amounts of immunoreactive IN protein were supplied to the in vitro IN assay under conditions established for pelleted VLPs. L1/L2 product is designated. Low quantities of L3 are observed except when 300 mM KCl was supplied in the reaction mixture. This product was not reproducibly detected at high levels in high salt. Wild-type (WT) VLPs, strain LTB111; PR MUTANT, strain LTB128. These VLPs represent preparations independent from those used for Fig. 3.

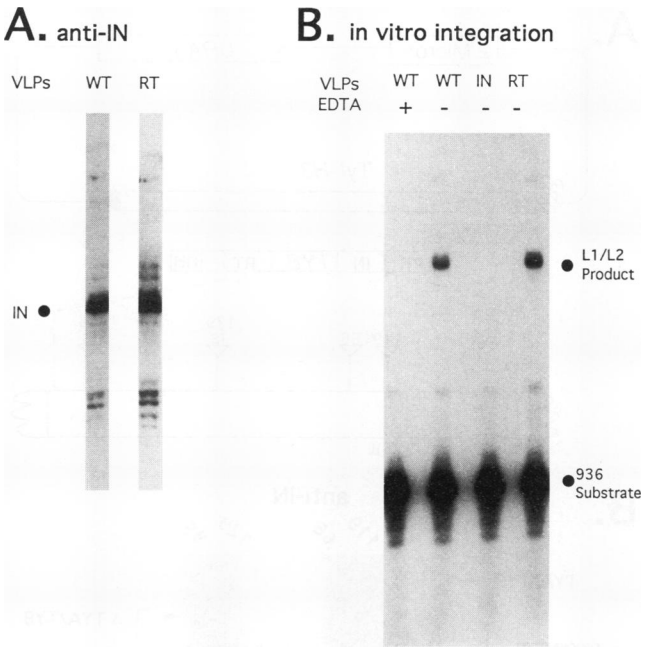

FIG. 5. RNA-containing VLPs have IN activity. VLPs from the RT$^-$ mutant, strain LTB120 harboring pAG149, and wild-type strain LTB111 harboring pX3 were harvested from a 20 to 70% linear sucrose gradient as described in Materials and Methods. (A) Immunoblot analysis of RT$^-$ VLPs. From immunoblot analysis of pooled gradient fractions, the IN-containing fractions were identified and used directly for in vitro IN assays. Mature IN accumulates in both VLP preparations. (B) In vitro IN activity of RT$^-$ VLPs. A total of 5 μl from the pooled fraction shown in panel A was assayed directly for IN activity. Equivalent amounts of L1/L2 product from wild-type and RT$^-$ VLPs are obtained. Only low amounts of L3 are formed in this experiment. IN$^-$ designates VLPs isolated from LTB114 harboring pGM66, an IN mutant. These VLPs were previously shown to be inactive in this assay (9) and the genetic assay (21) (see also Fig. 6C). EDTA was included at 50 mM where indicated (+).

the nuclease activity that copurifies with the VLPs in these strains. Reduced amounts of RT activity and protein relative to those of the wild-type preparation were recovered, indicating that the PR$^-$ VLPs are unstable.

However, upon reducing the wild-type VLP protein input to match that of the PR$^-$ VLPs, the IN activity was similar to that observed for wild-type VLPs (Fig. 4). High concentrations of salt had an inhibitory effect on IN activity from both preparations, although the PR$^-$ mutant VLPs were somewhat more sensitive (Fig. 4, compare the 150 mM KCl lanes). Thus, the virion environment provided by pre-VLPs can support the IN activity of the TYA/TYB polyprotein. We can only isolate active pre-VLPs from gradients which have been centrifuged for short periods. When PR$^-$ VLPs were isolated with longer centrifugation times and a 20 to 70% linear sucrose gradient, higher levels of breakdown products are observed by immunoblot analysis, and IN activity is not detected (data not shown).

Protease action does not appear to be required to form an IN capable of functioning in our in vitro reaction. Although it is formally possible that wild-type PR could be provided in *trans* from chromosomal Ty1 elements, we think this cannot explain our results because these assays are done under conditions in which the quantity of reaction products is linearly related to the amount of input VLPs (data not shown). Thus, we conclude that precursor TYB protein is active for in vitro integration.

RNA-containing VLPs: effect of a mutation within RT. VLPs derived from Ty1 elements with mutations within RT or lacking RT altogether contain some Ty1 RNA (14, 42). We have studied a mutation converting the second D of the highly conserved Y/FXDD box of Ty1 RT (Table 3), the putative active site, from FVDD to FVDE, which eliminates transposition in vivo (25); however, RT$^-$ mutant VLPs have quantities of IN comparable to those of wild-type VLPs (Fig. 5A). When in vitro IN activity was assayed, the amount of L1/L2 product is comparable to that of wild-type VLPs (Fig. 5B). Since the majority (~80%) of these events represent single-end insertions, we have only assessed the ability of IN within RNA-containing VLPs to mediate single-end insertions and not directly assessed the ability of these VLPs to mediate complete insertions; we have not carried out the two-dimensional analysis described in the accompanying paper (9).

DNA-containing VLPs: effects of mutations within IN. Five linker-insertion mutations were previously mapped to the IN domain, and their positions relative to each other are shown in Fig. 6A and Table 3 (10, 34). The pGM100 mutation is nearest the N terminus, located 43 residues from the putative PR cleavage site. Immunoblot analysis of IN in this mutant shows amounts of processed IN varying from very low amounts to near-wild-type levels in VLP preparations (Fig. 6B). In one experiment, VLP IN protein was reduced ~20-fold; however, RT activity was reduced only 3-fold (data not shown). Thus, this mutation has pleiotropic effects on VLP processing and stability. Levels of mature IN detected in homogenates from pGM100 are also variable (10). In contrast, the levels of mature IN comparable to that of the wild type are routinely recovered from pGM119 and pGM66 mutant VLP prepara-

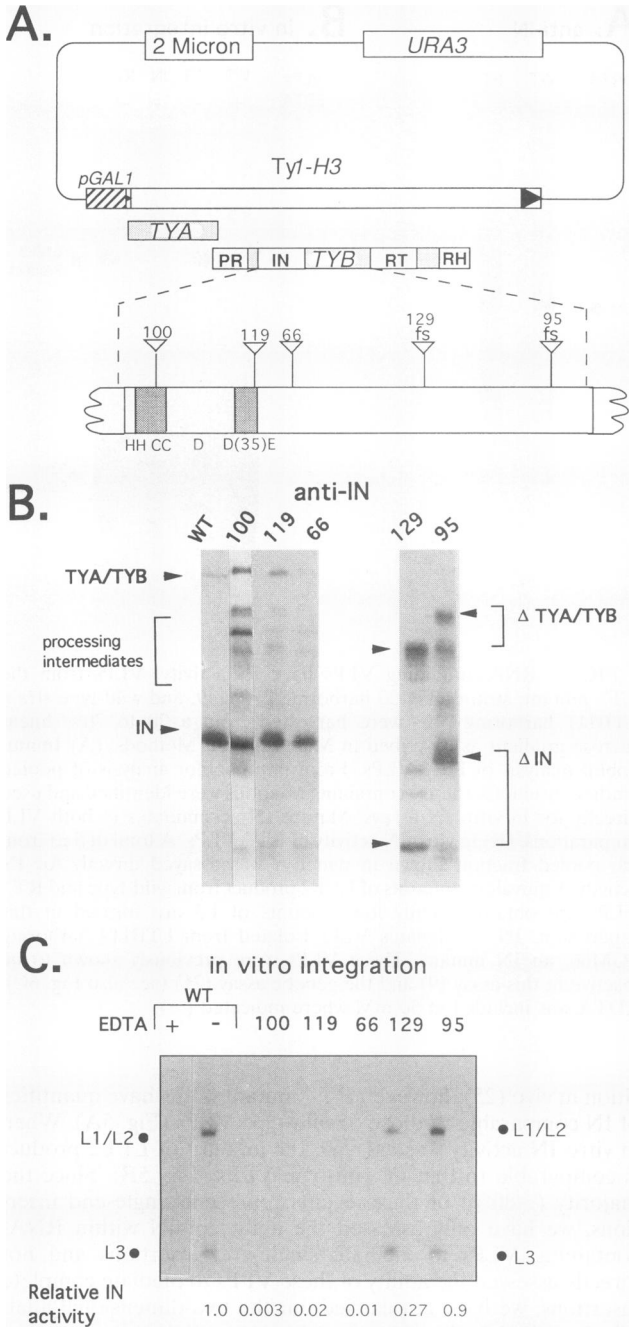

FIG. 6. Analysis of IN⁻ DNA-containing Ty1 VLPs. (A) Conserved regions within Ty1 IN. The pGTy1 plasmid structure and the functional organization of Ty1 are shown. The Ty1 IN domain is flanked by PR on the left and RT on the right. Thus, two proteolytic processing events must occur to accumulate mature IN protein, p90, consisting of a highly conserved N-terminal half and a variable C-terminal half of unknown function. An enlargement of the IN domain demonstrates the position of two highly conserved motifs, a zinc-binding site and the D, D(35)E motif, designated by shaded boxes. Positions of the five linker-insertion mutants (100, 119, 66, 129, and 95) relative to each other are shown. fs, frameshift mutation. (B) Immunoblot analysis of IN⁻ VLPs. IN⁻ mutant VLPs were isolated from a 20 to 70% linear sucrose gradient as described in Materials and Methods from five strains, LTB167, LTB169, LTB171, LTB175, and LTB177, harboring pGM100, pGM119, pGM66, pGM129, and pGM95, respectively. IN-containing fractions were determined by immunoblot analysis of pairs of pooled gradient fractions. An immunoblot of the pooled fraction with the greatest quantity of IN from each mutant is shown (the pGM100 lane is exposed 10 times longer). Two frameshift mutations, pGM129 and pGM95, resulted in truncated TYA/TYB (ΔTYA/TYB) and IN (ΔIN). Processing intermediates were also detected. (C) In vitro IN activity of IN⁻ VLPs. A 5-μl aliquot from each pooled gradient fraction shown in panel B was assayed directly for IN activity. The L1/L2 product is designated and was quantitated with a PhosphorImager (Molecular Dynamics). The presence of 50 mM EDTA is designated (+).

tions; these levels are within twofold of wild-type levels (9, 10). The mutations in pGM129 and pGM95 are frameshifts resulting in the production of truncated IN proteins (Fig. 6B) with apparent masses determined by SDS-PAGE of 52 and 74 kDa, respectively. Because wild-type IN protein has an apparent mass of 80 kDa in this gel system, the pGM129 mutation and the pGM95 mutation generate IN proteins 28 and 6 kDa smaller, respectively.

We tested VLPs with these various IN mutations for IN activity (Fig. 6C). The three codon-insertion mutations within the conserved N-terminal domain abolished IN activity. Since the pGM100 VLPs were most variable in terms of IN protein recovery and processing, three different pGM100 VLP preparations were examined for IN activity in vitro. IN activity was not detected in any preparation; therefore, we conclude that this mutation also inactivates IN. The pGM100 mutation inserts five codons within the N-terminal zinc-binding motif, the pGM119 mutation changes the highly conserved motif D, D(35)E spacing to D, D(40)E, and pGM66 maps near the C-terminal end of the D, D(35)E domain. Remarkably, the two frameshift mutations that resulted in truncated IN proteins retain very significant levels of in vitro IN activity. Comparison of the levels of L1/L2 and L3 products made by the various mutants suggests that the more severe frameshift mutant, pGM129, selectively affects L1/L2 formation. Because L1/L2 consists of substrate joined to target and L3 consists of two substrates joined together, this observation is consistent with a role for the C terminus of IN in target DNA recognition.

Substrate specificity: effects of altered substrate termini. A distinguishing feature of retroviral and retrotransposon LTR termini is the conserved dinucleotide CA at each 3′ end. To investigate the sequence requirements for Ty1 LTR termini, we also used the physical assay to detect in vitro IN activity. Initially, the termini were modified by use of various restriction enzymes to generate mutant substrate ends and later by the use of PCR to generate substrates with mutated end sequences. Various restriction enzyme digests of the donor substrate plasmid pLB343 generated either one or two ends lacking normal Ty1 termini (Fig. 7). Relative IN activity was obtained for each mutant substrate. Many substrates lacking Ty1 terminal sequence were used with an efficiency similar to that of the parental substrate bearing 12 bp of Ty1 terminal sequences at each end. Three trends were observed: (i) one terminal adenosine appears to be sufficient for recognition and efficient joining by IN (Fig. 7, substrate 924-U5); (ii) a fragment lacking adenosine termini, but with two internal adenosine residues near the ends, can also function as a substrate for IN; and (iii) two fragments lacking adenosine termini and with only one internal adenosine near an end did not function as an efficient substrate for IN. These results suggested two possibilities. (i) Ty1 IN either processes the termini of the substrate DNA to expose a terminal adenosine, or (ii) Ty1 IN will join other terminal substrate nucleotides to target DNA. Since this assay system detects mostly single-end joinings of substrate to the target DNA, it was not possible to define from this experiment

Substrate length	Description	Left end	Right end	Relative IN activity
936 bp	U3+U5 wild type	TGTTGGAATAGA ACAACCTTATCT	ACCCAATTCTCA TGGGTTAACAGT	1.0
924 bp	U5+ *Bst* EII	GTCACCTGTCCA CAGTGGACAGGT	ACCCAATTCTCA TGGGTTAACAGT	0.6
924 bp	U3+ *Sty* I	TGTTGGAATAGA ACAACCTTATCT	TCCGGATCCAAG AGGCCTAGGTTC	2.3
586 bp	*FOK*I+ *Sca*I	TGTCTGCTCCCG ACAGACGAGGGC	ACTTGGTTGAGT TGAACCAACTCA	1.35
912 bp	*Bst* EII+ *Sty* I	GTCACCTGTCCA CAGTGGACAGGT	TCCGGATCCAAG AGGCCTAGGTTC	0.7
680 bp	*FOK*I	CTCATGCATAAA GAGTACGTATTT	GACGGGCTTGTC CTGCCCGAACAG	<0.025
201 bp	*FOK*I+ *Pvu* II	AACATGCATAAA TTGTACGTATTT	CTGTCGTGCCAG GACAGCACGGTC	<0.05
1204 bp	*Bbs*I	TGTTGTCCTCGT AACACAGGAGCA	GTCTGGGGCTCA CAGACCCCGAGT	0.7

FIG. 7. IN activity with altered substrates. Radiolabeled substrates with various end sequences were prepared as described in Materials and Methods and used for in vitro IN reactions. The substrate length, description of the ends, and the sequences at both ends are listed. The shaded bases designate the 12 terminal Ty1 sequences at each end of the wild-type substrate as well as those bases within the altered substrates that match those in the wild-type sequence. In vitro integration reactions were carried out under standard reaction conditions; 5 μl of VLPs prepared from strain JB224 harboring pJEF724, using a 20 to 70% linear sucrose gradient, was supplied to the reaction mixture. Only the amount of L1/L2 product formed was quantitated as described above. The IN activity of the various substrates is relative to the IN activity of the 936-bp wild-type substrate.

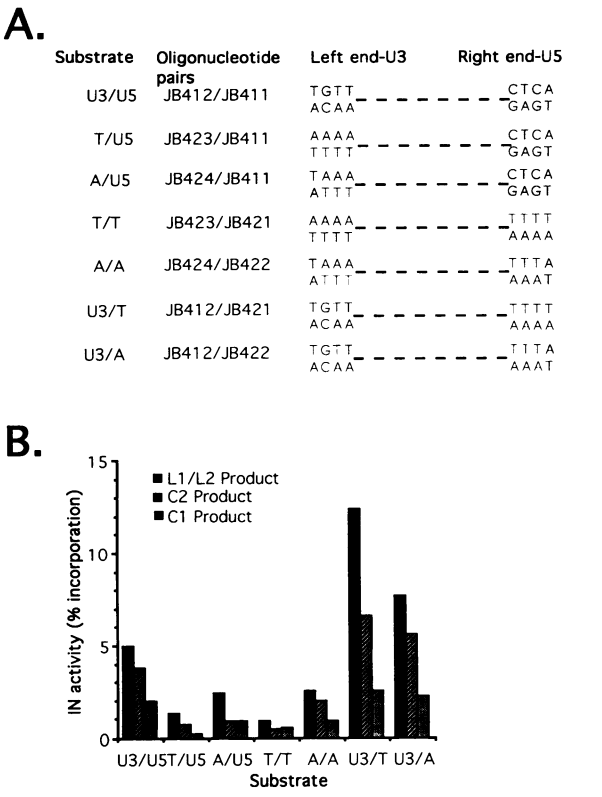

FIG. 8. In vitro integration activity with matched substrates differing only at their termini. (A) Oligonucleotide pairs that generate the PCR substrates. The various substrate designations, the oligonucleotide pairs that are used for PCR, and the structures of termini following *Bbs*I digestion and fill-in reactions are shown. The left end of the substrate designates the U3 terminus, and the right end designates the U5 terminus. (B) In vitro integration activity with linear and circular target DNA. Radiolabeled substrates were prepared as described in Materials and Methods, supplied to a standard in vitro integration reaction with 5 μl of VLPs prepared from strain JB224 bearing pJEF724, and then isolated on a 20 to 70% sucrose gradient as previously described. When the target DNA was linear, the percentage of radiolabeled substrate converted to L1/L2 product was then determined. When the target DNA was circular, the percentage of each radiolabeled substrate converted to the C1 product and C2 product was then determined. These products are described and characterized in the accompanying paper (9).

which base or substrate end participated in the strand exchange. Furthermore, in the above experiment, terminal sequences, internal sequences, and fragment length were all varying, making it difficult to draw firm conclusions. A new strategy that would allow detection of products by both the physical assay and by genetic selection to recover complete insertions was developed; thus, we could investigate the sequence requirements at both ends of the substrate that allow efficient coordinated insertion of one substrate as well as the detailed structure of the products. Importantly, these substrates would differ from each other only in the sequences of the four terminal base pairs at each end.

DNA substrates were generated such that the termini could be radiolabeled and the terminal sequences could be readily specified by the investigator (see Fig. 8 for strategy, Table 4 for oligonucleotides used as primers, and Materials and Methods for experimental detail). When such a DNA bearing 4 bp of Ty1 termini was supplied as a substrate to the in vitro integration reaction mixture, the amount of L1/L2 product formed was similar to that observed by using the standard 936-bp substrate (Fig. 7). Therefore, we used this strategy to generate substrates with various end sequences (Fig. 8A). Both linear target DNA and circular target DNA were used to detect IN activity, and percentages of radioactive substrate converted to the various products are shown (Fig. 8B). The absence of the U3 sequences dramatically reduced the amount of L1/L2, C1, and C2 products formed. Surprisingly, the T/T substrate (3′ terminal sequence $TTTT_{OH}$) which lacks 3′ terminal adenosine residues could function as a substrate for IN, although at a relatively low efficiency. The A/A substrate (3′ terminal sequence $TTTA_{OH}$) differing from the T/T substrate only in the presence of a single 3′ adenosine at each end generates products at higher efficiencies. However, the amount of A/A substrate converted to product is only half that obtained with the wild-type U3/U5 substrate. The substrates with a U3 end but lacking a U5 end dramatically increase the amount of integration products, suggesting that U5 actually is inhibitory. The U3 end is apparently preferred as a substrate by IN.

A potential criticism of the experiments with T termini is that untemplated A's could have been added to the 3′ end during the filling-in reaction (17) and that it was these molecules which might have undergone joining. We repeated the experiment using the T/T substrates, generated in the presence or absence of dATP, and observed the same low-level joining activity in both preparations (data not shown). Moreover, we showed that terminal T residues in the substrate can be directly joined to the target in a complete integration reaction (see below).

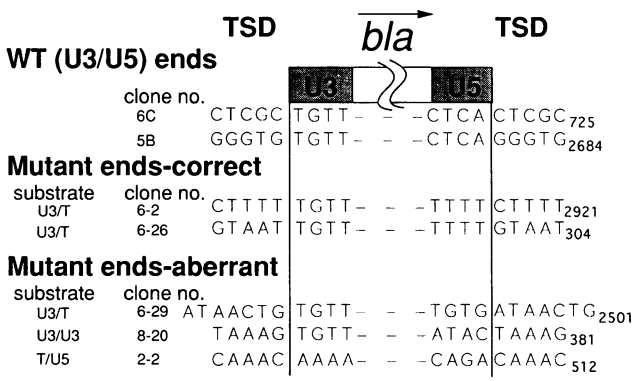

FIG. 9. Structures of junctions made by substrates with wild-type and mutant termini. In vitro integration into pLB523 target was carried out as described in the accompanying paper (9). The junction sequences of several wild-type (WT) and mutant substrate complete integration events are summarized; only the top strand is shown for simplicity. The subscript indicates the position of the insertion in the target plasmid. For the mutant substrates, the type of mutant substrate used is indicated (substrate nomenclature is as in Fig. 8). TSD, target site duplication. The inserted DNA in the first four recombinants corresponds precisely to the predicted structure of the substrates used. In the last three (aberrant) recombinants, the left junctions are normal in structure but the right junctions are not.

Complete integration reactions, using substrates with mutant ends. Most of the products observed with the physical assay represent single-end joinings, but the substrates described in the above section can also be used to investigate the essential components of substrate termini for complete insertions. Because the mutant termini flanked an active β-lactamase gene, we could use the above molecules as substrates together with a chloramphenicol-resistant target plasmid in a biological assay for complete integration. As described in the legend to Fig. 10 of the accompanying publication (9), the recombinant products can be selected by combined chloramphenicol and carbenicillin selection. There is a background of doubly drug-resistant colonies that makes quantitative comparison of the efficiencies of wild-type and mutant termini difficult with this system. However, we were able to recover insertions from the wild type and three of the mutant substrates. We have cloned and sequenced seven complete insertion events from a variety of these substrates (Fig. 9). The two wild-type substrates and two of the U3/T mutant substrates gave rise to integrations that precisely resemble a normal in vivo integration event (except of course, that the two mutant insertions end in four T residues rather than with a normal U5 terminus). These insertions end precisely at the predicted substrate termini and are flanked by 5-bp target site duplications. No extra bases are present at the junctions nor is there any evidence for processing of terminal dinucleotides by Ty1 IN. In addition, three insertions that differed from an in vivo integration in one or more ways were cloned. All three of these had target site duplications, although one of them (number 6-29) has a 7-bp duplication. All three aberrant events contained normal junctions at the (arbitrarily defined) left end (two of these left end junctions were U3 ends and one was a T end) but had aberrant structures at their right ends. In number 6-29, derived from a U3/T substrate, the right end junction sequence indicated that during the preparation of the substrate, the restriction enzyme *Bbs*I had failed to cleave off the last few base pairs of the PCR product; apparently IN joined the PCR-generated 3′ end to target DNA. In number 8-20, derived from a U3/U3 substrate, the last 123 bp of the right terminus were missing from the final product. In number 2-2, derived from a T/U5 substrate, the last 8 bp of the right terminus were missing.

DISCUSSION

By using a new in vitro Ty1 integration assay (9), we have assayed both mutant VLPs and mutant IN substrates to examine their effects on integration. The products formed in this assay system are predominantly the result of single-end joining of substrate DNA to target DNA; thus, we have examined here primarily IN activity and, to a lesser extent, the requirements for the coordinated insertion of both ends of one substrate.

Various alterations to the VLP protein components do not inhibit IN activity in vitro. Unprocessed TYA/TYB polyproteins are isolated in the form of pre-VLPs which accumulate in the yeast cytoplasm when mutations lie within the PR domain of inducible pGALTy1 plasmids. Near-wild-type levels of IN activity are detected in some pre-VLP preparations, indicating that the polyprotein is able to accomplish the substrate recognition and strand-transfer reactions carried out by mature IN. RNA-containing VLPs which accumulate when pGALTy1 plasmids bear RT mutations contain wild-type levels of both mature IN protein and IN activity. IN molecules within these RNA-containing VLPs are not associated with Ty1 dsDNA within the VLP. Therefore, association with and utilization of an exogenous substrate by IN does not require reverse transcription to have occurred within the VLP.

The amino terminus of Ty1 IN is essential for joining substrate to target. The Ty1 IN protein is less well characterized than many retroviral IN proteins. Two PR cleavage events are required to yield a mature IN protein of 71 kDa (8, 27, 35). Thus, Ty1 IN is much larger than retroviral IN proteins. In this study, three different mutations within the conserved N-terminal domain of native IN abolished IN activity in vitro. These findings are consistent with recent studies of retroviral INs; that is, the central core region of these proteins plays a key role in strand transfer.

In contrast, VLPs harboring C-terminally truncated IN proteins retained significant IN activity. This region of Ty1 IN is therefore not required for DNA joining or Ty1 end recognition reactions. A possible role for this domain of IN might be in vivo target site selection. Recently, Ty1 integration hot spots were identified upstream of tRNA genes on chromosome III. These same sequences do not function as hot spots in vitro (19, 28). We speculate that this IN domain, together with cellular factors, plays a role in guiding these nonrandom in vivo insertion events.

Terminal adenosines increase integration efficiency but are not essential for IN activity. Previous experiments have suggested that the LTR termini of linear viral DNAs are important for efficient integration. In vitro integration studies which used artificial substrates bearing 12 bp of Ty1 termini provide a sufficient number of bases for efficient integration; certain alterations to the structure or sequence of these termini apparently eliminated activity (22). In this study, the terminal sequences of artificial Ty1 substrates were altered and a surprising number of substrates bearing little or no resemblance to native Ty1 terminal sequence were found to serve as substrates. Surprisingly, IN activity was detected in some cases even when adenosine bases were absent from the 3′ ends of the substrate. However, there were substrates which were completely unable to serve as substrates, so there is sequence specificity for the substrate. A complicating factor in interpret-

ing these experiments is that the physical assay used in this study detects a large fraction of single-end insertions (80%) and a small fraction of complete insertions.

A second conclusion from these studies of transposon terminal sequence requirements was that U3-ended substrates yield more recombinant product than U5-ended substrates. It has been noted previously that the U3 end is more highly conserved among Ty1 elements than is the U5 end; 10 bp of U3 are identical to the U3 sequence of the distantly related *Drosophila copia* element (3, 40). Also, we note that both U3 and a number of the termini of substrates that work well for integration are rich in adenosine on the joining strand. The correlation between the number of adenosines within the last 4 bp of each terminus and integration activity determined by using the substrates in Fig. 7 is remarkably good ($r^2 = 0.731$), suggesting that the number of adenosines in the joining strand is an important determinant of how good the substrate will be. However, it is possible that this association is artifactual, as it is based on a small number of substrates that differ from each other in a variety of ways.

A procedure was developed to efficiently generate a family of closely related substrates with defined alterations to their ends that also contained the *bla* gene. The physical assay demonstrated that 4 or 5 bp of Ty1 termini were sufficient to yield near-wild-type levels of product. Surprisingly, substrates with T sequences on one end were relatively good substrates, and those with two T ends still showed some detectable activity. However, replacement of the terminal T nucleotides with A's significantly increased its ability to act as a substrate. These results were difficult to interpret because most of the products were single-end joinings and we could not easily determine whether one end was being preferentially joined.

Therefore, these substrates were also used in a minor variation of the genetic assay described in the accompanying paper (9), in which the substrate, which specifies carbenicillin resistance, inserts into a chloramphenicol-resistant target plasmid. Although a high background of uninformative transformants was obtained, by screening we were able to identify several insertion events. Of a total of 14 junctions analyzed, 11 were normal in structure and the other 3 had unusual substrate termini. These aberrant junctions may have arisen by joining of a structurally aberrant terminus in an otherwise normal integration reaction. Alternatively, two of the aberrant insertions could be explained by precise deletions at the right substrate terminus associated with the insertion event; we find this latter possibility less likely because it is more complex. Furthermore, all seven events analyzed had target site duplications, and six of these were of the appropriate length. Of the 11 normal termini, 6 derived from U3 ends, 3 were from T ends, and 2 were from U5 ends. The three aberrant ends derived from one each of a U3, U5, and T terminus. We suspect that the aberrant termini result from the complex, multistep method of substrate preparation used (PCR, restriction enzyme cleavage, and filling in) rather than an aberrant reaction by IN. From these observations, we conclude that Ty1 IN can utilize substrate termini that are very different from normal Ty1 termini in an otherwise normal reaction. These data show that any one of the four bases can act as the 3′ terminal nucleotide on the substrate for the integration reaction. It is interesting to note that although all of the substrates in Fig. 8 were tested for complete integration, we recovered complete integrations for only some of them, and most of the complete integration events were derived from the U3/T substrate, which also gave the strongest signal in the physical assay.

Although this assay yielded interesting data on the Ty1 IN joining reaction, there was a very high background of irrelevant transformants that made it impractical to apply the system quantitatively or to recover large numbers of insertion clones. Also, it appeared that the method used to generate these substrates (PCR amplification followed by *Bbs*I cleavage) sometimes generated products with inappropriate termini. We have recently developed an improved method for generating integration substrates and recovering the in vitro integration products that circumvents these problems (19).

ACKNOWLEDGMENTS

We are grateful to E. Caputo for technical assistance, A. Gabriel for supplying the RT mutant, and S. Devine for assistance with some experiments. We thank N. Craig, P. Englund, T. Kelly, and P. Hieter for helpful discussions; N. Craig and S. Devine for comments on the manuscript; and Roxanne Ashworth for assistance with DNA sequencing.

L.T.B. was supported by NSF predoctoral training grant RCD-9154644. This research was supported in part by NIH grant GM36481 and an American Cancer Society Faculty Research Award to J.D.B.

REFERENCES

1. **Adams, S. E., J. Mellor, K. Gull, R. B. Sim, M. F. Tuite, S. M. Kingsman, and A. J. Kingsman.** 1987. The functions and relationships of Ty-VLP proteins in yeast reflect those of mammalian proteins. Cell **49**:111–119.
2. **Bajwa, W., T. E. Torchia, and J. E. Hopper.** 1988. Yeast regulatory gene *GAL3*: carbon regulation; UAS_{GAL} elements in common with *GAL1*, *GAL2*, *GAL7*, *GAL10*, *GAL80*, and *MEL1*; encoded protein strikingly similar to yeast and *Escherichia coli* galactokinases. Mol. Cell. Biol. **8**:3439–3447.
3. **Boeke, J. D.** 1988. Retrotransposons, p. 59–103. *In* E. Domingo, J. J. Holland, and P. Ahlquist (ed.), RNA genetics, vol. II. Retroviruses, viroids, and RNA recombination. CRC Press, Boca Raton, Fla.
4. **Boeke, J. D., D. J. Eichinger, D. Castrillon, and G. R. Fink.** 1988. The *Saccharomyces cerevisiae* genome contains functional and nonfunctional copies of transposon Ty1. Mol. Cell. Biol. **8**:1432–1442.
5. **Boeke, J. D., D. J. Garfinkel, C. A. Styles, and G. R. Fink.** 1985. Ty elements transpose through an RNA intermediate. Cell **40**:491–500.
6. **Boeke, J. D., and S. B. Sandmeyer.** 1991. Yeast transposable elements, p. 193–261. *In* J. Broach, E. Jones, and J. Pringle (ed.), The molecular and cellular biology of the yeast *Saccharomyces*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
7. **Bowerman, B., P. O. Brown, J. M. Bishop, and H. E. Varmus.** 1989. A nucleoprotein complex initiates the integration of retroviral DNA. Genes Dev. **3**:467–478.
8. **Braiterman, L. T.** 1993. Integration of retrotransposon Ty1. Ph.D. thesis. The Johns Hopkins University School of Medicine, Baltimore, Md.
9. **Braiterman, L. T., and J. D. Boeke.** 1994. In vitro integration retrotransposon Ty1: a direct physical assay. Mol. Cell. Biol. **14**:5719–5730.
10. **Braiterman, L. T., G. M. Monokian, D. J. Eichinger, S. L. Merbs, A. Gabriel, and J. D. Boeke.** 1994. In-frame linker insertion mutagenesis of yeast transposon Ty*1*: phenotypic analysis. Gene **139**:19–26.
11. **Brown, P. O., B. Bowerman, H. E. Varmus, and J. M. Bishop.** 1987. Correct integration of retroviral DNA in vitro. Cell **49**:347–356.
12. **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.
13. **Burke, C. J., G. Sanyal, M. W. Bruner, J. A. Ryan, R. L. LaFemina, H. L. Robbins, A. S. Zeft, C. R. Middaugh, and M. G. Cordingly.** 1992. Structural implications of spectroscopic characterization of a putative zinc-finger peptide from HIV-1 integrase. J. Biol. Chem. **267**:9639–9644.
14. **Burns, N. R., H. R. Saibil, N. S. White, J. F. Pardon, P. A. Timmins, S. M. H. Richardson, B. M. Richards, S. E. Adams,**

S. M. Kingsman, and A. J. Kingsman. 1992. Symmetry, flexibility and permeability in the structure of yeast retrotransposon virus-like particles. EMBO J. **11:**1155–1164.
15. **Bushman, F. D., A. Engleman, 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.
16. **Bushman, F. D., T. Fujiwara, and R. Craigie.** 1990. Retroviral DNA integration directed by HIV integration protein in vitro. Science **249:**1555–1558.
17. **Clark, J. M.** 1988. Novel non-templated nucleotide addition reactions catalyzed by prokaryotic and eukaryotic DNA polymerases. Nucleic Acids Res. **16:**9677–9686.
18. **Craigie, R., 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 **62:**829–837.
19. **Devine, S., and J. D. Boeke.** Efficient integration of artificial transposons into plasmid targets *in vitro*: a useful DNA sequencing tool. Nucleic Acids Res., in press.
20. **Donehower, L., and H. E. Varmus.** 1984. A mutant murine leukemia virus with a single missense mutation in pol is defective in a function affecting integration. Proc. Natl. Acad. Sci. USA **80:**6461–6465.
21. **Eichinger, D. J., and J. D. Boeke.** 1988. The DNA intermediate in yeast Ty1 element transposition copurifies with virus-like particles: cell-free Ty1 transposition. Cell **54:**955–966.
22. **Eichinger, D. J., and J. D. Boeke.** 1990. A specific terminal structure is required for Ty1 transposition. Genes Dev. **4:**324–330.
23. **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.
24. **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:**3065–3069.
25. **Gabriel, A., and J. D. Boeke.** 1994. Unpublished data.
26. **Garfinkel, D. J., J. D. Boeke, and G. R. Fink.** 1985. Ty element transposition: reverse transcriptase and virus-like particles. Cell **42:**507–517.
27. **Garfinkel, D. J., A.-M. Hedge, S. D. Youngren, and T. D. Copeland.** 1991. Proteolytic processing of *pol-TYB* proteins from the yeast retrotransposon Ty1. J. Virol. **65:**4573–4581.
28. **Ji, H., D. P. Moore, M. A. Blomberg, L. B. Braiterman, D. F. Voytas, G. Natsoulis, and J. D. Boeke.** 1993. Hotspots for unselected Ty1 transposition events on yeast chromosome *III* are near tRNA genes and LTR sequences. Cell **73:**1007–1018.
29. **Johnson, M. S., M. A. McClure, D. F. Feng, J. Gray, and R. F. Doolittle.** 1986. Computer analysis of retroviral pol genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. Proc. Natl. Acad. Sci. USA **83:**7648–7652.
30. **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.
31. **Kulkosky, J., K. S. Jones, R. A. Katz, J. P. G. Mack, and A. M. Skalka.** 1992. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. Mol. Cell. Biol. **12:**2331–2338.
32. **Leavitt, A. D., R. B. Rose, and H. E. Varmus.** 1992. Both substrate and target oligonucleotide sequences affect in vitro integration mediated by human immunodeficiency virus type 1 integrase protein produced in *Saccharomyces cerevisiae*. J. Virol. **66:**2359–2368.
33. **Mellor, J., M. H. Malim, K. Gull, M. F. Tuite, S. McCready, T. Dibbayawan, S. M. Kingsman, and A. J. Kingsman.** 1985. Reverse transcriptase activity and Ty RNA are associated with virus-like particles in yeast. Nature (London) **318:**583–586.
34. **Monokian, G. M., L. T. Braiterman, and J. D. Boeke.** 1994. In-frame linker insertion mutagenesis of yeast transposon Ty*1*: mutations, transposition and dominance. Gene **139:**9–18.
35. **Moore, S. P., and D. G. Garfinkel.** 1994. Expression and partial purification of enzymatically active recombinant Ty1 integrase in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **91:**1843–1847.
36. **Müller, F., K. H. Bruhl, K. Freidel, K. V. Kowallik, and M. Ciriacy.** 1987. Processing of Ty1 proteins and formation of Ty1 virus-like particles in *Saccharomyces cerevisiae*. Mol. Gen. Genet. **207:**421–429.
37. **Panganiban, A. T., and H. M. Temin.** 1984. The retrovirus *pol* gene encodes a protein required for DNA integration: identification of a retroviral *int* locus. Proc. Natl. Acad. Sci. USA **81:**7885–7889.
38. **Rose, M. D., F. Winston, and P. Hieter.** 1990. Methods in yeast genetics: a laboratory course manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
39. **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.
40. **Williamson, V. M.** 1983. Transposable elements in yeast. Int. Rev. Cytol. **83:**1–25.
41. **Xu, H., and J. D. Boeke.** 1990. Host genes that influence transposition in yeast: the abundance of a rare tRNA regulates Ty1 transposition frequency. Proc. Natl. Acad. Sci. USA **87:**8360–8364.
42. **Youngren, S. D., J. D. Boeke, N. J. Sanders, and D. J. Garfinkel.** 1988. Functional organization of the retrotransposon Ty from *Saccharomyces cerevisiae*: the Ty protease is required for transposition. Mol. Cell. Biol. **8:**1421–1431.