In Vitro Integration of Retrotransposon Ty1: a Direct Physical Assay

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

Retrotransposon Ty1 of *Saccharomyces cerevisiae* inserts a double-stranded Ty1 cDNA into the yeast genome by a reaction analogous to the integration mechanism used by retroviruses. A quantitative in vitro integration assay that directly detects integrative recombination products was developed for Ty1. Blunt-ended artificial radioactive substrates bearing Ty1 termini integrate into circular or linear target DNAs. The reaction is specific for native integrase isolated in the form of virus-like particles; virus-like particles prepared from integrase mutants were completely inactive in this assay. The products are radioactive, allowing direct detection after gel electrophoresis by autoradiography. Using this simple and amenable system, we characterized the biochemical requirements of the system and the structures of the major integration products. Two classes of products were detected: those that were the result of bona fide complete integration events (concerted reactions) and single-end joinings of substrate to target (half-reactions). Additionally, we used a genetic selection scheme to identify and characterize target sites of complete integration events into a circular target plasmid; a 5-bp target site duplication flanking the inserted DNA resembling the duplication characteristic of in vivo integration was observed.

Genetic and biochemical studies of the yeast retrotransposon Ty1 have revealed many parallels with the life cycle of retroviruses (4, 6). These include the presence of open reading frames analogous to retroviral *gag* and *pol*, referred to as *TYA* and *TYB*, respectively; long terminal repeats (LTRs); and short target site duplications. Furthermore, retrotransposition of Ty1 involves assembly of a virus-like particle (VLP) containing Ty1 RNA and proteins in which reverse transcription produces a cDNA copy of Ty1 RNA. This DNA is then integrated into host DNA by a Ty1-encoded integrase (IN). The DNA integration step of the retrovirus and retrotransposon life cycle has been the focus of many recent studies (for reviews of retroviral integration, see references 11, 23, and 39). A variety of cell-free assay systems have been devised to show that both retroviral and retrotransposon nucleoprotein complexes (12, 13, 16, 21, 33) and purified IN (14, 15, 25, 30) can accomplish specific DNA cutting and joining reactions. Two commonly used approaches include (i) a labor-intensive λ bacteriophage-based assay that allows the products of a simple concerted integration to be recovered genetically (12, 16) and (ii) a half-reaction system in which radiolabeled double-stranded oligonucleotides serve both as one-ended substrate and target DNA, allowing direct detection of products generated (15, 26, 30).

The major difference between integration by retroviruses and Ty1 is an end-processing reaction observed in retroviral systems but not thought to occur in Ty1 (17). Ty1 cDNA is thought to be blunt ended, terminating in the conserved inverted dinucleotide 5'TG...CA3'. By using a phage λ -based assay, concerted integration was detected for the endogenous Ty1 cDNA within VLPs (16). Subsequent work

indicated that the VLPs could very efficiently mediate the complete integration event of blunt-ended 5'TG...CA3' artificial substrates into λ target DNA. By using this system, 12 bp of Ty1 DNA at each substrate end was shown to be sufficient for integration. Finally, it was shown that, as is the case for retroviruses and other transposons, the 3'-OH of the Ty1 substrate was necessary for integration (17). The mechanism for the retroviral IN-mediated strand transfer has been shown to occur via a one-step transesterification reaction (19, 28). Each strand of target DNA is attacked in a staggered fashion by the 3'-OH of the retroviral cDNA such that the 5'-PO₄ of target DNA is transesterified to the 3'-OH of retroviral cDNA.

Briefly, the genetic assay for Ty1 in vitro integration developed by Eichinger and Boeke (16) involves selection for recombinant λ phage in which the *supF*-marked Ty1 cDNA has been inserted into nonessential regions of the λ gTWES/ λ B phage target DNA (a strain with amber mutations in three essential genes). Following in vitro packaging in λ phage heads, the recombinant phage are able to form plaques on nonsuppressing bacterial host strains. The π an7 plasmid (bearing the bacterial *supF* gene) was inserted into a nonessential region of Ty1, resulting in a marked pGTy1-H3 plasmid which allowed efficient transposition in vivo and in vitro. VLPs isolated from cells expressing this plasmid contain a small amount of Ty1-H3- π an7 reverse transcripts, largely in the form of double-stranded linear DNA molecules. These VLPs provide the substrate DNA and IN. Alternatively, substrate DNA consisting of π an7 flanked by Ty1 termini can be provided in *trans* to unmarked Ty1 VLPs (17). In either case, the VLPs are incubated with λ gtWES/ λ B target DNA and Mg²⁺, and then the reaction mixture is treated to remove the protein components and the nucleic acid mixture is packaged into λ phage heads. Although this assay allowed recovery of complete integration events of both Ty1 DNA and exogenous DNA substrates in vitro, there are several drawbacks to this method: (i) the procedure is labor intensive, making it difficult to examine a large number of parameters that may affect integra-

* 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@gmail.bs.jhu.edu.

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

FIG. 1. Direct physical assay for Ty1 in vitro integration—reaction flow chart. The 936-bp substrate was obtained from pLB241 or pLB343 and then radiolabeled and used as indicated. The reaction is Mg²⁺ dependent; inhibition is observed in the presence of EDTA. Shaded areas indicate the sequences and positions of Ty1 termini. Positions of key enzyme sites used in this study are indicated. Ty1 minisubstrate and plasmid are not drawn to scale.

tion; (ii) the detection system is indirect, thus only complete integration events that produce viable phage will be recovered; and (iii) a gap flanking the substrate DNA results after insertion, and the gapped intermediate may be unstable or inefficiently packaged. Therefore, estimates of efficiency obtained with this system may underrepresent the integration frequency.

The genetic assay systems were developed to recover complete integration events, and the oligonucleotide assay was developed to evaluate purified IN; will IN mediate concerted reactions and half-reactions under the same conditions? Is IN more efficient at generating one type of product? Are the products generated in vitro representative of the products generated in vivo? Dissection of these issues requires a sensitive, quantitative in vitro assay in which all products can be analyzed without selection.

In this paper, we describe the development of a quantitative direct physical assay (Fig. 1) for integration of Ty1 in vitro which can be manipulated with ease. Artificial minisubstrates with 12 bp of Ty1 *cis*-acting sequences at each end, similar to those previously reported (17), are radiolabeled and mixed with either linear or circular plasmid target DNA and IN. In these experiments, the Ty1 IN protein is supplied in native form; VLPs isolated from *Saccharomyces cerevisiae* provide the IN activity. After incubation under appropriate conditions, the recombinant product molecules are separated from the substrate and target molecules by agarose gel electrophoresis and the dried gel is exposed to film for autoradiography. Joining of the radioactive substrate, typically about 1,000 bp in length, to a cold target typically ~3 kb in length, results in a variety of radioactive products migrating more slowly than both the substrate and the target DNAs in agarose gels.

We report the biochemical requirements of the reaction and analysis of the structures of the reaction products and the

target insertion sites. From studies with two-dimensional (2-D) agarose gel electrophoresis and electron microscopy, the various products characterized belong to two classes of integration events: (i) complete integration events reflecting concerted integration products and (ii) one-ended insertions reflecting half-reactions. The predominant class, one-ended insertions that result from the joining of one end of the substrate DNA to target DNA, composes ~80% of the products. An unusual type of product, a linear DNA target with a substrate joined at each end, is observed when circular target DNA is supplied. This product apparently resulted from one end each of two substrate molecules being joined to the circular target in a concerted fashion.

MATERIALS AND METHODS

Construction of LTR substrate and target plasmids. Most of the experiments in this paper made use of *FokI* digestion to prepare the transposition substrates. Therefore, only DNA fragments lacking *FokI* sites could be used as substrates. The mini-LTR substrate plasmid, pLB343, was constructed from M138 π 1U3 Δ 12/U5 Δ 12 (17) as follows. Deletion of the π an7 sequences without alteration of the Ty1 terminal sequences was accomplished by digestion with *BstEII* and *StyI*, which cut just inside the Ty1 termini; a *BspMI*-to-*BstYI* 869-bp fragment from plasmid pACYC184, chosen because it lacked *FokI* restriction sites, was then cloned into these sites, resulting in the regeneration of the *BstEII* and *StyI* sites, to give M13LB82. The *SalI* fragment from M13LB82 was cloned into the *SalI* site of pSP65, resulting in pLB241. The *StyI* site within the 869-bp insert fragment (from pACYC184) was removed by partial digestion of pLB241 with *StyI* followed by filling in with the DNA polymerase I Klenow large fragment (PolIK) and religation to give pLB343. All restriction fragments were gel purified with low-melting-point agarose.

A second mini-LTR plasmid, pJEF2343, is the source of the Ty1-*bla* integration substrate which confers carbenicillin resistance in *Escherichia coli*. It is derived from plasmid pTZ18U via a complex series of steps (information available upon request) and consists of the *bla* gene from pACYC177 inserted into a synthetically derived sequence representing the 12-bp Ty1 termini, separated by a polylinker and flanked by *NsiI* sites (see Fig. 10A). In this case, the transposition substrate is prepared by *NsiI* digestion rather than *FokI* digestion. This cassette is inserted into a pTZ18U derivative that is kanamycin resistant and ampicillin sensitive.

Target plasmid pJEF2276 (3,157 bp) was constructed by inserting the 1.2-kb *cat* gene-containing *XmnI*-*Tth111I* fragment (filled in with PolIK) from pACYC184 into *DraI*-digested pUC18. This chloramphenicol-resistant, ampicillin-sensitive plasmid served as either a linear target or a circular target in both the physical and the genetic assays described below. Although this plasmid is sensitive to ampicillin, it still contains a segment of the *bla* gene, and therefore it has some homology to the Ty1-*bla* substrate. Digest of pJEF2276 with *FspI* and *SmaI* resulted in the deletion of a 156-bp region of the polylinker to generate pLB523 (3,001 bp). This region was removed to facilitate 2-D analysis of the products by restriction digest.

Yeast strains and GALTy1 plasmids. The *S. cerevisiae* strains and yeast plasmids used in this study are listed in Tables 1 and 2. The media used in this study, SC—ura glucose, SC—ura galactose, SD, SGAL, and YPD, have been described elsewhere (34), except that SC medium was modified by the omission of *para*-aminobenzoic acid. Galactose-containing me-

TABLE 1. Yeast strains used in this study

Strain ^a	Genotype	Reference
YH8	<i>MATα ura3-167 his3Δ200 leu2Δ1 trp1Δ1</i>	40
GRF167	<i>MATα ura3-167 his3Δ200</i>	5
YH82	<i>MATα ura3-52 lys2-801 his4-539 trp1Δ63 leu2Δ1</i>	40

^a GRF167 and YH8 are isogenic strains.

dia contain 2% (wt/vol) galactose (prepared as a 20% filter-sterilized stock).

VLP isolation and characterization. Yeast cells bearing various pGAL-Ty1 plasmids (Table 2) were harvested after 24 h of galactose induction, and VLPs were prepared as described elsewhere (9). Briefly, a yeast extract was obtained by glass bead lysis, clarified at 10,000 rpm in a Sorvall RC-6 centrifuge, and then layered over a preformed 20 to 70% linear sucrose gradient. After an 18- to 24-h ultracentrifuge spin in an SW28 rotor at 25,000 rpm, 1.2-ml fractions were collected and then 5 μ l of each fraction was directly analyzed for reverse transcriptase activity (22). For certain experiments, the presence of IN protein was determined by using a previously described immunoblotting method (9). Briefly, trichloroacetic acid-precipitated 100- μ l aliquots of pools of two fractions were prepared for sodium dodecyl sulfate (SDS)-gel electrophoresis, transferred to Immobilon P, and then probed with monoclonal antibody 8B11, an IN-specific monoclonal antibody, and ¹²⁵I-protein A (16). IN activity was measured by using either 2 or 5 μ l directly from two pooled gradient fractions that were shown to contain either reverse transcriptase activity or IN protein.

Preparation of LTR substrates and target DNA. A two-step fragment purification procedure was necessary to produce a homogeneous population of labeled substrate molecules. A 2,426-bp *ScaI-PvuII* fragment from pLB343 was purified on a 0.9% agarose-TBE (89 mM Tris, 89 mM boric acid, and 0.2 mM EDTA) or -TTE (89 mM Tris, 28.5 mM taurine, and 0.5 mM EDTA) gel, and then the DNA was electroeluted from a gel band in NaCl-TE buffer (5 mM NaCl, 10 mM Tris-HCl, 1 mM EDTA) by using 150 V for 45 min. The ethanol-precipitated DNA was digested with *FokI*. Following this, blunt ends were generated by filling in with PolIK, using [α -³²P]dATP and a 0.25 mM concentration of dGTP, dCTP, and dTTP at room temperature. To ensure that most substrate molecules of the preparation had blunt ends, 1 mM dATP was added for a 10-min chase following a 10-min labeling reaction. The 936-bp radiolabeled fragment was then electrophoretically separated on a second 0.9% agarose gel and isolated by electroelution as described above.

The cloning experiments required the use of unlabeled, selectable Ty1-*bla* substrate derived from plasmid pJEF2343. The plasmid was digested with *Nsi* and then treated with T4

DNA polymerase, and unlabeled deoxynucleoside triphosphates (dNTPs) were used to generate a blunt-ended substrate as previously described (1).

Plasmids bearing the substrate fragments and the target plasmids, pJEF2276 and pLB523, were purified from CsCl-ethidium bromide gradients. Target plasmids were used directly in circular form or linearized with *XbaI*. The target plasmid was also linearized with *SmaI* or *SacI* in some experiments, as indicated.

Assay for Ty1 in vitro integration. A cocktail containing 10 mM KCl, 10 mM Tris-HCl (pH 8.0), 15 mM MgCl₂, 1 mM dithiothreitol, 5% polyethylene glycol 8000, 5 ng of radiolabeled substrate DNA (5×10^9 molecules), and 100 or 200 ng of target DNA pJEF2276 or pLB523 (3×10^{10} and 6×10^{10} molecules, respectively) was prepared in a 15- μ l volume. Although 10 mM KCl is used in the standard reaction mixture, we have subsequently found that higher activity is obtained with no KCl. For the standard reaction mixture, the molar ratio of target to substrate is 6:1; 5 μ l (unless otherwise indicated) of two pooled peak VLP gradient fractions is directly added to the reaction mixture and incubated at 25 to 30°C for 30 min unless otherwise stated. When standard reaction conditions were altered, that component omitted from the cocktail was then added separately before addition of the VLPs. For all reaction mixtures, the VLPs were added last. Reactions were stopped by the addition of a mixture of proteinase K and EDTA to a final concentration of 50 μ g/ml and 25 mM, respectively, and incubated at room temperature for a minimum of 30 min. Reaction mixtures were further incubated at 65°C for 15 min in 0.1% SDS and 1 $\times$ gel sample buffer (25 mM EDTA, 5% glycerol, bromophenol blue) and then electrophoretically separated on a 0.6% agarose gel containing 0.5 μ g of ethidium bromide per ml. TBE or TTE gel running buffer was used, and electrophoresis was carried out in an IBI slab gel unit (8 cm by 8 cm) at 50 V for 14 to 16 h. The gels were photographed to evaluate the recovery of input DNA and dried onto DE81 chromatography paper. The dried gels were exposed to XAR-5 film for autoradiography or exposed to a phosphor screen for quantitation with a PhosphorImager (Molecular Dynamics).

2-D agarose gel electrophoresis. 2-D agarose gel analysis was performed as described by Brewer and Fangman (10). For the first dimension, 500 ng of *HindIII*-digested λ DNA was added to each sample. The DNAs were separated on a 0.4% agarose gel at 1 V/cm for 16 h at 4°C in either TBE or TTE gel running buffer. After being stained in 1 μ g of ethidium bromide per ml, the individual gel lanes were cut from the gel, placed along the horizontal of a gel tray, and cast into a 1% agarose gel containing 0.5 μ g of ethidium bromide per ml. After electrophoresis at a higher voltage, 3.3 V/cm, for 4 h at room temperature in TBE or TTE gel running buffer, the gels were photographed, dried, and exposed as described for the integration assays.

Characterization of Ty1 integration products. A 10 $\times$ integration reaction mixture (a 200- μ l volume composed of 180 μ l of reaction mixture and 20 μ l of VLPs) was performed to obtain sufficient labeled material for product analysis. The target plasmid pLB523 was used for these reactions; otherwise, standard reaction conditions were used as described above. After the reaction and heat treatment were stopped, the products were loaded into four lanes (13 mm wide), separated on a 1% vertical agarose-TTE gel (1.5 mm thick), and run at 10 V/cm for 5 to 6 h at room temperature. The gel was wrapped in Saran Wrap and exposed to XAR-5 film for 16 h at room temperature. The desired bands were cut from the gel after being lined up with the autoradiogram. The DNA was electro-

TABLE 2. Transformants used in this study

Strain	Host	Plasmid	Description (reference[s])
JB224	GRF167	pJEF724	Wild-type unmarked pGAL-Ty1 (5)
LTB185	GRF167 $\times$ YH82	pJEF724	Wild-type unmarked pGAL-Ty1 (5)
LTB111	YH8 $\times$ YH82	pX3	Wild-type <i>TRP1</i> -marked pGAL-Ty1 (40)
LTB114	YH8 $\times$ YH82	pGM315	IN ⁻ mutant (pGM66) <i>TRP1</i> -marked pGAL-Ty1 (9, 29)

eluted at 150 V for 45 min and then ethanol precipitated and resuspended in 40 μ l of TE (10 mM Tris-HCl [pH 8.0], 1 mM EDTA); 10 μ l was used for restriction digests. Aliquots of isolated undigested product and digested product were analyzed by using 2-D agarose gels as described above. Nicked plasmid DNA pJEF2343 (4 kb) was added to the C2 sample as a control for the migration of a 4-kb relaxed circle in the 2-D gel system.

Electron microscopy. DNA from integration assays was purified by phenol-chloroform extraction and ethanol precipitated, and then the resuspended DNA was spread on grids as described by Perez-Morga and Englund (31). Determination of the lengths of the DNA fragments was done by using an UltraViolet Products camera and imaging system and the Image 1.42 computer program by Wayne Rasband, furnished by the NIH Research Services Branch, National Institute of Mental Health.

Alterations to target DNA ends. Two strategies were used to alter the ends of the linear target DNA. First, restriction digestion of pJEF2276 with three different enzymes, *Xba*I (5' overhang), *Sac*I (3' overhang), and *Sma*I (blunt) was used to obtain linear target DNA preparations with different end sequences. Second, the ends of *Xba*I-digested pLB523 were treated as follows to obtain four different target preparations: (i) phosphatase treatment, resulting in removal of the 5' phosphate; (ii) incubation with ddCTP and PolIK, resulting in addition of a nucleotide without a 3'-OH; (iii) phosphatase treatment and subsequent addition of ddCTP; and (iv) ligation of the self-complementary oligonucleotide JB484 (5'-CTAGC CGCGGGGATTTTCCCCGCGG) that forms stem-loop structures at the *Xba*I termini. To test the efficiencies of the various treatments, an aliquot of each 3-kb target DNA preparation (and control untreated 3-kb linear target DNA) was incubated with T4 DNA ligase in the presence of ATP and then analyzed on an agarose gel. Only the control untreated target shifted in mobility; all other target DNA preparations migrated at 3 kb, indicating that the functional end groups were absent.

Genetic assay for in vitro integration. Integration into a circular plasmid target was also assessed by a genetic assay. In this assay, integration of a Ty1 substrate fragment bearing the *bla* gene (conferring carbenicillin resistance) into a plasmid bearing the *cat* gene (conferring chloramphenicol resistance) was monitored genetically. This assay allows the direct and detailed structural analysis of individual complete integration events. Integration assays were performed with unlabeled or labeled Ty1-*bla* substrate DNA prepared in parallel. The integration efficiency of the unlabeled substrate was judged by integration assays, using the radiolabeled Ty1-*bla* substrate DNA as described above. Phenol-chloroform-extracted, ethanol-precipitated DNA from the integration assays was digested with *Xho*I in order to linearize any residual carbenicillin-resistant pGALTy1 plasmid DNA that copurified with the VLPs (no *Xho*I sites are found in the integration recombinant). Ten microliters of the 20- μ l digest was transformed into *E. coli* MH5 which was plated onto Luria broth-carbenicillin, and then the carbenicillin-resistant colonies were replica plated to Luria broth-chloramphenicol plates. Miniprep DNA, isolated from colonies which were both carbenicillin and chloramphenicol resistant, was digested with *Bgl*I and *Kpn*I to identify recombinant plasmids. Each enzyme cut once in the target plasmid and once within the insert; thus, two bands were detected when an integration event occurred. Recombinant plasmids from *E. coli* DH5 α were isolated and sequenced by using oligonucleotides JB193 (5'-CCGCGCCACATAGCA GAAC) and JB157 (5'-TAAGATGATCTTCTTGA) which

FIG. 2. Native IN preparations (VLPs). VLPs (strain LTB185) were recovered from a preformed 20 to 70% linear sucrose gradient centrifuged for 24 h. Aliquots (2 μ l) from fractions were assayed for reverse transcriptase activity (percent incorporation), and 100- μ l aliquots from pairs of fractions were pooled and analyzed for IN protein by immunoblot analysis with monoclonal antibody 8B11 (16) and ^{125} I-protein A. p190 TYA/B is the precursor polypeptide processed to yield IN protein p90. The bands migrating between p190 and p90 are processing intermediates detected in various amounts with each VLP preparation.

prime outward from the right and left termini of the Ty1-*bla* inserts, respectively.

RESULTS

Physical assay for Ty1 integration; product formation requires functional IN. A quantitative in vitro integration assay that directly detects integrative recombinant products was developed for Ty1. Integration assays were performed with VLPs previously characterized by the presence of reverse transcriptase activity and IN protein (Fig. 2). Greater than 97% of the yeast protein loaded on the gradient (supernatant from a 10,000-rpm centrifugation step) fractionates in the upper part of the gradient, whereas the VLP fractions represent less than 3% of the protein load. By immunoblot analysis (Fig. 3A), equivalent amounts of IN protein from VLP gradient fractions of wild-type (WT-LTB111) and mutant integrase (IN-LTB114) were assayed for integration activity by using linearized target plasmid DNA and radiolabeled artificial minisubstrates (Fig. 3B). The sequences at each end of the artificial substrates consist of 12 bp derived from the left end of the Ty1 LTR, designated U3, and 12 bp derived from the right

FIG. 3. IN activity. (A) Immunoblot analysis of wild-type (WT) and IN⁻ VLPs. VLPs were isolated as previously described (9) from two strains: LTB111, harboring a wild-type inducible pGTy1 plasmid, and LTB114, harboring the identical plasmid with a mutation within IN (pGM66) (see references 8 and 9 for a full description of this mutant). By immunoblot analysis, comparable amounts of mature IN protein were recovered from the two strains. More processing intermediates are observed in the IN⁻ VLP fractions. (B) In vitro IN activity of wild-type and IN⁻ VLPs. Aliquots (2 μ l) from pairs of pooled fractions were assayed directly for in vitro IN activity. The fractions that contained wild-type IN by immunoblot analysis also exhibited IN activity (upper panel). Three radioactive products, L1, L2, and L3, and the radioactive substrate (S) are detected by autoradiography of a dried agarose gel. The gradient fractions that contained the IN⁻ protein by immunoblot analysis did not form radioactive products in this assay system, just as no products were detected when IN⁻ VLPs were used for the genetic assay (16). S, radioactive substrate; *, products that are not reproducibly observed. EDTA was included at 50 mM as indicated (+).

end of the Ty1 LTR, designated U5. Only fractions containing wild-type VLPs result in product formation.

Three major products are detected and numbered in order of increasing mobility. L1 and L2 appear as a broad band sometimes resembling a doublet migrating at 4 kb, and L3 migrates at 2 kb when integration assays are analyzed on a 0.6% agarose gel. When the products were analyzed on 1% agarose gels, the broad 4-kb band resolved into two bands, L1 and L2, composing 30 and 70% of the 4-kb product, respectively. (These results suggest that some products are complete integration events, but other products may result from a single-end joining of the substrate to the target [see "Product analysis: integration into linear targets" below].) The 2-kb L3 product presumably consists of one substrate joined to another substrate.

Under optimal conditions, 5 to 17% of the substrate DNA is converted to 4-kb product. The conversion of 17% of the substrate molecules to product represents 8.5×10^8 integra-

tion events. The amount of VLPs used per integration assay represents approximately 10^7 cell equivalents. Thus, a one-cell equivalent gives rise to 85 integration events. This number of integration events per cell equivalent appears to be 85-fold higher than that observed with the previously described λ system (17). The difference in efficiency can be attributed to at least four factors, as will be shown below. (i) In the current assay, a significant number of products that are formed are incomplete integration events and would not score as an integration event in the λ assay. (ii) The current assay was far easier to optimize, and the biochemical conditions used are somewhat different. (iii) The substrate and target molecules are different in sequence and size. (iv) The previous detection system was indirect. Products were selected such that only insertions that could be efficiently packaged into phage λ heads in vitro and give rise to viable phage λ were detected.

Biochemical analysis of Ty1 in vitro integration. Previous studies (16) established that the biochemical requirements for concerted integration by Ty1 integrase were simple and did not require an exogenous energy source. To optimize the above reaction by using a linear target, we examined the requirement for divalent cations, salt, pH, time, and temperature (Fig. 4). This reaction shows a strong preference for Mg^{2+} as the divalent cation. When higher concentrations of Mg^{2+} were tested, a nonspecific nuclease activity that copurified with the VLPs was also activated; therefore, 15 mM Mg^{2+} was chosen for the standard reaction condition. The substitution of Mn^{2+} for Mg^{2+} did not support the integration reaction to a significant extent. Low concentrations of KCl or NaCl were preferred, and high salt strongly inhibited the reaction. When a broad pH range was examined by using various buffers, surprisingly, the reaction was apparently optimal at two specific pHs, pH 6.0 with citrate-phosphate buffer and pH 8.0 with Tris-HCl buffer. The L1 and L2 products appeared as a discrete doublet with equal intensities when citrate-phosphate buffer, pH 6.0, was used in the reaction mixture, whereas L1 is less abundant than L2 under all other reaction conditions examined. Of course, the identity of the buffer system itself may also influence integration efficiency; different buffers had to be used to span the pH range of interest effectively.

The time course of this reaction is similar to that previously reported for Ty1 integration (16); products are observed as soon as 1 min after the addition of IN and appear to level off at 20 to 30 min. The reaction is efficient at various temperatures (22 to 30°C). Surprisingly, IN activity was readily detected at 4°C (the cocktail was precooled before the addition of IN). Addition of polyethylene glycol stimulates the reaction 16-fold. The standard reaction mixture contains 10 mM Tris-HCl (pH 8.0), 10 mM KCl, 15 mM $MgCl_2$, 1.0 mM dithiothreitol, 5% polyethylene glycol, 100 ng of plasmid DNA, 5 ng of substrate DNA, and 2 μ l of VLPs in a 20- μ l total reaction volume. The resulting molar ratio of substrate to target DNA is 1:6. The integration products are stable in 0.1% SDS at 65°C for 15 min.

Integration products: linear and circular target DNAs. Both linear (L) and circular (C) plasmid DNA can serve as target DNAs (Fig. 5). Three major radioactive products are observed when circular DNA serves as the target. These products are numbered in order of increasing electrophoretic mobility (Fig. 5B, lane 2). These are: (i) an unexpected, abundant product with an apparent molecular size of 5 kb, C1 (see "Product analysis: integration into circular targets" below); (ii) the expected relaxed circular 4-kb product, C2; and (iii) a 2-kb product, C3. L3 and C3 appear to be identical, namely, one substrate inserted into another substrate. In this experiment, the concentration of substrate DNA was decreased, changing

FIG. 4. Optimization of in vitro reaction conditions. In vitro integration assays were performed with 2 μ l of two pooled wild-type VLP (strain JB224) gradient fractions exhibiting the highest amounts of IN protein. The standard reaction conditions described in Materials and Methods were used, except for the component examined in each panel. Effects of NaCl and KCl concentration, pH, time, Mg^{2+} , Mn^{2+} , and temperature on IN activity were examined. Mg-1 and Mg-2 are two independent experiments. Only the amount of L1/L2 product formed was quantitated; each value was adjusted to relative IN activity based on the condition exhibiting the highest level of activity.

the molar ratio of substrate to either linear or circular target DNA. The fraction of radioactive substrate converted to the various 4-kb products, i.e., the L1/L2 or C2 products, remained relatively constant despite dramatic decreases in the amount of substrate supplied. In contrast, the L3 product generated decreased in parallel with the input of the radioactive substrate DNA. Furthermore, when circular target DNA was supplied to the reaction mixture, the unexpected C1 product and C3 product (Fig. 5B) also decreased in parallel with the input of the radioactive substrate DNA, suggesting that the latter requires two molecules of substrate whereas the former requires one molecule of substrate.

Thus, we detect different products in the two types of reactions, except that C3 is indistinguishable from L3. Product L1 migrates close to L2 on many gels, is often diffuse compared with L2, and sometimes appears as a shadow. L2 has the electrophoretic mobility expected for a simple, concerted insertion into a linear target. These results suggest that L1 might represent a population of single-end joined molecules whereas L2 might represent complete integration events. Models of the structures for each product are shown in Fig. 6. Product C1 has the mobility expected for a linear target attached to two molecules of substrate, and C2 has the mobility expected for a nicked circle resulting from a complete integration event of a substrate into a circular target. Because of the complexity of the reaction products, we have also analyzed them by 2-D gel electrophoresis (Fig. 7) and restriction enzyme digestion (Fig. 8).

Product analysis: integration into linear targets. We used 2-D agarose gel electrophoresis (2) to examine the reaction products (Fig. 7). Conditions were similar to those reported for the analysis of replication forks (10). Migration in the first

FIG. 5. Integration products formed with linear and circular target DNA. In vitro integration assays were performed with 2 μ l of wild-type VLP (strain LTB111) fractions as in Fig. 4. (A) Products formed with linear target DNA and decreasing amounts of substrate DNA. The L1/L2 and L3 radioactive products and radioactive substrate (S) are designated. (B) Products formed with circular target DNA and decreasing amounts of substrate DNA. The C1, C2, and C3 radioactive products and radioactive substrate (S) are designated. (A and B) Lanes 1, EDTA was included during the reaction. In order to visualize the products with reduced amounts of radioactive substrate, lanes 4, 5, and 6 were exposed 24 times longer than lanes 1, 2, and 3. No products were observed in lanes 1 upon longer exposure.

FIG. 6. Models of Ty1 in vitro integration products. (A) L1 and L2 products are recovered when linear DNA serves as a target. L1 is shown as the result of one end of the substrate joined to the target DNA, and L2 is shown as the result of a complete integration event into the target DNA. Furthermore, the ends of the linear target serve as a preferred insertion site (see Fig. 8A). (B) C1 and C2 products were recovered when circular DNA serves as a target. C1 is shown as the result of one end of two substrate DNAs joined to each strand of a circular target DNA in a concerted fashion. C2 is shown as composed of two products, C2.1, the result of a complete integration event into a circular target, and C2.2, the result of one end of the substrate joined to the target DNA. Shaded boxes, Ty1 termini; open rectangle, non-Ty1 sequence; double lines, double-stranded DNA target; A, *AvaI* site.

dimension was dependent on molecular weight. Thus, all conformations of the same molecular weight (other than supercoiled plasmid DNA) have the same mobility. Separation in the second dimension is then dependent on structure. Linear species form a smooth arc, and the arrowheads in Fig. 7 indicate the λ *HindIII* DNA fragments included as internal standards in each reaction mixture; supercoiled plasmid migrates faster than corresponding linear DNAs, and nicked

FIG. 7. 2-D agarose gel electrophoresis: analysis of products. In vitro integration assays using either linear target DNA (A) or circular target DNA (B) were performed as for Fig. 4, except that VLPs were from strain JB224. The positions of the products L1, L3, C2, and C3 are designated with brackets. L2 and C1 are designated with arrows. The positions of bacteriophage λ *HindIII* fragments of 23.1, 9.4, 6.5, 4.3, 2.3, and 2.0 kb are designated by long arrows. 0.95 indicates the mobility of the 936-bp radiolabeled substrate in this gel system. Very long arrows designate the direction of electrophoresis.

FIG. 8. Characterization of isolated products. IN reactions were carried out as described in Materials and Methods (VLPs were from strain LTB111). Each product was electroeluted from an excised gel band after separation on a 1% agarose gel. (A) Linear target: analysis of purified products. The mobility of *AvaI*-digested L1 product shifted to the right in the first dimension relative to the undigested L1; both undigested and digested are positioned off the linear arc, thus mobility in the second dimension is retarded. The mobility of *AvaI*-digested L2 shifted to the right in the first dimension relative to undigested L2; however, both undigested and digested products migrate on the linear arc. (B) Circular target: analysis of purified products. Both *AvaI*-digested and undigested C1 are positioned on the linear arc; however, C1 migrated faster after digestion. C2 is a discrete spot positioned off of the linear arc; after digestion, two products are observed and designated C2.1, positioned on the linear arc, and C2.2, positioned off of the linear arc. Long arrows designate the direction of electrophoresis.

circular plasmid DNA is retarded relative to the corresponding linear DNA. 2-D agarose gel electrophoresis of integration reactions supplied with linear target DNA (Fig. 7A) revealed three major species: a nonlinear species, L1, which appears as a vertical streak clearly migrating off the arc of linear mole-

TABLE 3. Summary of in vitro reaction products detected

Major products	Linear target ^a (species [% of total])	Circular target ^b	
		Species (% of total ^c)	% Total C2 products ^d
Complete insertions (or mixture) ^e	L2 (70 ^c)	C2.1 (8.4)	18 ^f
Single-end joining			
One substrate	L1 (28)	C2.2 (29.6)	82
Two substrates	ND ^g	C1 (62)	NA ^h

^a The values given were determined by PhosphorImager analysis of Fig. 7.

^b The values were obtained from the experiment shown in Fig. 8B.

^c The value for C1 is calculated as follows: % C1 = [C1/(C1 + C2.1 + C2.2)] × 100.

^d This class represents the formation of products dependent on the joining of one substrate to one target; thus, these values can be compared with the values obtained for L1 and L2.

^e This species is composed of a small fraction of complete integration events and a major fraction of insertions near the ends of the target DNA (see text). *Ava*I digestion results in fragmentation of the product, resulting in a smear; therefore, direct quantitation of the fraction of complete integration events is not possible. A total of 61% of L2 is recovered as a 3.5-kb species; since both ends of the substrate were labeled, we estimate that half of the radioactive material should have been lost. Thus, there is a slight overrecovery of the material in this experiment.

^f The value for C2.1 is calculated (from PhosphorImager analysis of Fig. 7B) as follows: % C2.1 = [C2.1/2(C2.2) + C2.1] × 100.

^g ND, not determined. When two substrates are joined in a concerted fashion to a linear target, radiolabeled fragments of various lengths are generated; therefore, this product cannot be quantitated as a single species.

^h NA, not applicable.

cules; a linear 4-kb band, L2; and a linear 2-kb band, L3. L1 appears to represent a continuous collection of products with exactly the same molecular weight as the expected 4-kb L2 linear species. L3 migrates on the arc of linear species, consistent with the prediction of one substrate inserted into another.

The identity of the major spots was made by electroeluting individual L1 to L3 bands from one-dimensional gels, restriction digestion of each extracted DNA, and then by 2-D gel analysis. Each band was electroeluted from a 1% vertical agarose gel and digested with *Ava*I, which cuts in the center of the substrate molecule and does not cut the target. Aliquots of undigested and digested products were then analyzed by 2-D agarose gel electrophoresis. The mobility of L1 shifted from a streak at 4 kb to a more compact 3.5-kb streak upon digestion (Fig. 8A), indicating that half of a substrate molecule was removed by restriction enzyme digestion (the released 450-bp fragment is observed only when it does not run off the gel). Thus, the L1 product represents a collection of single-end joinings of the substrate DNA along the length of the linear target (Fig. 6A).

Unexpectedly, digestion of the 4-kb band L2 with *Ava*I did not generate a smear of product as would be predicted if L2 represented complete integration events of substrate randomly distributed along the length of target DNA. Rather, greater than 61% of the L2 products are recovered as a 3.5-kb band. The presence of this band indicates that these products are formed by either complete integration events very near a target end or single-end joinings of substrate very near an end of the linearized target to form species that migrate very similarly to a linear one. Thus, 30% of the integration products are single-end insertions (L1) along the length of the target DNA, and the L2 fraction (70%) of the integration products represents a mixture of complete integration events and insertions near the ends (Table 3). Digestion of the L3 product with *Ava*I did not abolish the product as would be expected if one

substrate inserted into another. Instead, a major product recovered migrated as a 900-bp linear fragment (data not shown), the length of one substrate, suggesting that many of these molecules are also integrated near one end.

We used electron microscopy to confirm the presence of single-end joinings. The DNA from an entire reaction mixture using linear target was extracted with phenol-chloroform and then spread on grids by a new technique that allows the use of small quantities of DNA (31). Single-end insertions visualized as branched structures are detected when linear DNA served as a target (Fig. 9). Measurement of the three arms of the branched molecules confirmed that one arm was equivalent in length to that of the substrate and the other two added up to the length of the target. Among the relatively small number of events observed, the distribution of insertion sites appeared to be random. However, single-end insertions very near the ends of the target probably resembled linear products and therefore might have been missed in this assay. We did not observe any branched structures when EDTA was included in a control reaction.

Product analysis: integration into circular targets. 2-D agarose gel electrophoresis of the products recovered when circular DNA serves as a target is shown in Fig. 7B. Three major species are observed, and only one spot, corresponding to C2, is positioned off the linear arc, and this species comigrates with nicked 4-kb control plasmid DNA. C3 and the unexpected 5-kb band, C1, which is the size of two substrates plus one target DNA, migrate on the linear arc, indicating that both these species are linear. Since we only identified one nonlinear product spot, either single-end joinings to circular targets do not occur or, alternatively, these products have the same mobility as concerted integration products, i.e., C2 is a mixture of 4-kb "panhandles" and open circles (Fig. 6B). We do not have an electrophoresis standard with the panhandle structure expected if a substrate molecule is joined at one end to a circular target. Because of their complex nature, these products were also isolated from gels and then subjected to restriction enzyme digestion.

The strategy outlined for L1, L2, and L3 was used to characterize the DNA extracted from bands C1, C2, and C3. Digests of each product with *Ava*I were then analyzed by 2-D agarose gel electrophoresis (Fig. 8B). The shift in mobility of C1 from 5 to 4 kb is consistent with the removal of one substrate length. If this product results from a concerted insertion of two substrate DNAs into the circular target DNA, then a linearized product would be formed and each end of the target DNA would be joined to a substrate DNA. *Ava*I digestion would remove ~450 bp of substrate DNA from each end of the 5-kb product, resulting in a new 4-kb product, compatible with the observed results and the model shown in Fig. 6B. By using a PCR detection method, small amounts of a product of this type were previously isolated and cloned from a reaction mixture in which purified Moloney murine leukemia virus IN was the source of integration activity (15).

Digestion of C2 with *Ava*I results in two major products, a 4-kb linear product designated C2.1 and a nonlinear product of 3.5 kb designated C2.2; the nonlinear product must represent single-end joinings to the circular target, thus identifying the circular counterpart of L1. Digestion of the 4-kb circular product generated by a complete integration event results in the linear 4-kb species, C2.1. Thus, the results of product analysis are compatible with the models shown in Fig. 6B. Quantitation of the products recovered showed that 18% of the digested material is in the C2.1 form and 82% of the digested material is in the C2.2 form (Table 3). Since both ends of the substrate are radiolabeled, removal of the ~450-bp

FIG. 9. Identification of branched products by electron microscopy. Linear target DNA (A), substrate DNA (B), or DNA purified after an in vitro integration reaction (C to F) which was performed under the standard reaction conditions with VLPs from strain JB224. Arrowheads indicate branched structures.

substrate fragment results in an estimated loss of half of the radioactive C2.2 material. Thus, only 70% of the input radioactive material is recovered (Table 3). The results from the 2-D agarose gel analysis of undigested and digested C3 were indistinguishable from those of L3, verifying their identity (data not shown).

By PhosphorImager analysis, 2% of the substrate (5×10^9 molecules) is converted to C2 (1×10^8 molecules). If 20% of C2 is recovered as C2.1, representing complete integration events, then 2×10^7 concerted integrations are recovered. Thus, this system detects a 10-fold increase in the number of complete integration events (1.5×10^6) reported by Eichinger and Boeke (17). These results suggest that the major increase in sensitivity observed with the physical assay is the result of a previously undetected single-end joining activity whereby either one substrate or two substrates are joined to the target DNA.

Linear targets give rise to true integration products and not end-ligation events. The products from the linear reaction are dependent on IN (Fig. 3) and do not require a high-energy cofactor for their formation. The efficient use of circular target DNA shows that recombinant product formation does not require an end-joining reaction. Nevertheless, a remarkably large fraction of the linear products appear to have the Ty1 sequences joined near the end of the linear targets. An end-to-end joining activity was identified in yeast nuclear extracts when an *in vitro* homologous recombination assay was developed by L. S. Symington (35). The experiments we have described thus far do not explicitly rule out a ligation mechanism for the generation of recombinant products with linear targets. Thus, we performed a series of manipulations of the linear target molecule to determine whether special structures were required at the target DNA ends for efficient recombinant product formation with linear targets, as would be expected if any ligation-like activity were present.

First, input target DNA previously linearized with enzymes which produced different termini, i.e., blunt ends (*Sma*I), 5' overhangs (*Xba*I), and 3' overhangs (*Sac*I), did not alter the nature of the 4-kb L1/L2 product. Second, the nature or quantity of L1/L2 was not affected when the ends of the linearized target were treated so that the 5'-PO₄ and/or the 3'-OH was removed. Third, an oligonucleotide was hybridized to the ends of the linear target DNA, forming terminal stem-loop structures that were ligated to the ends, also without affecting L1/L2 product formation (7). Therefore, the functional groups at the target ends are not directly involved in the product formation, and the majority of products formed when linearized target DNA is supplied to the reaction mixture must consist of either single-end joinings near the ends of the target which behave as a 4-kb linear product or complete integration events near the ends.

Five-base-pair target site duplication. To assay for complete integration event events, we developed a simple selection-based cloning strategy. Circular target DNA bearing a *cat* gene and a new minisubstrate DNA engineered such that the Ty1 termini flank the *bla* gene (Fig. 10A) are incubated with VLPs. The reaction products were extracted with phenol-chloroform, transformed into *E. coli*, and grown in the presence of carbenicillin and chloramphenicol. The recovered clones were sequenced, enabling the identification of target site duplications. In five of six cases in which the target DNA was sequenced on both ends, a 5-bp duplication of the target plasmid was identified (Fig. 10B). In the sixth case, clone 2-3, the insertion site was in a region of the plasmid where homology existed between the *bla* gene of the substrate and sequences of the target. The substrate U3 sequences were unaltered; however,

FIG. 10. Cloning of *in vitro* integration products. (A) Plasmid pJEF2343 was constructed so that the Ty1 termini flank the *bla* gene and will be exposed upon digestion with *Nsi*I. The *Nsi*I ends were made blunt by treatment with T4 DNA polymerase in the presence of dNTPs. Shaded areas designate Ty1 sequences. (B) Insertion sites and target site duplications. *In vitro* integration assays were performed with circular target pJEF2276, the substrate DNA described above, and wild-type VLPs (from strain JB224) in the presence or absence of EDTA. Colonies resistant to both carbenicillin and chloramphenicol were examined for recombinant products by restriction digestion. A total of 21 colonies were recovered from reactions without EDTA, and 4 colonies were recovered with EDTA added. Only colonies recovered from reactions without EDTA contained recombinant plasmids of the expected structure. The plasmids from the latter colonies were sequenced; the insertion sites and target site duplications are shown. Essential regions of the target plasmid pJEF2276 are designated with shaded arrows. Shaded nucleotides designate the inferred target half-site (not confirmed by sequencing). Plasmids are not drawn to scale.

the U5 sequences as well as 97 bases of the target DNA were deleted. This unusual product may have arisen from a single-end joining of substrate to a region of the target with homology to substrate. Recovery of this event may have been partially mediated by homologous recombination in *E. coli*. Five additional insertion sites were sequenced at one end. As expected, the insertion sites appear throughout nonessential regions of the plasmid, although 3 of 11 insertions clustered near the plasmid *ori* sequence.

DISCUSSION

Direct detection of Ty1 integration intermediates. The development of *in vitro* assay systems for IN has allowed much

progress in understanding the mechanistic aspects of integration (11, 19). This progress has been made principally with the use of two types of assay systems, namely, (i) genetic assays that allow selection of complete integration events and (ii) physical assays that allow detection of half-reactions developed as a means to assay purified IN activity. Thus, two types of products have been recovered, but not under the same conditions or in the same assay. In fact, the purified IN and IN isolated in the form of nucleoprotein complexes behave quite differently. Integration by purified Moloney murine leukemia virus IN and human immunodeficiency virus type 1 IN requires Mn^{2+} (15, 37) for optimal activity, whereas IN in the context of nucleoprotein complexes prefers Mg^{2+} (12, 18, 20). Each assay system has its drawbacks. The genetic assays are (i) indirect—recovery of the products is dependent on the efficiency of the detection system, (ii) selective—only complete integration events are detected, and (iii) labor-intensive—thus, less amenable to multiple manipulations. The oligonucleotide-based assays are limited in that they detect only single-end insertions and thus reflect the *in vivo* integration process incompletely.

In this report, we have provided exogenous minisubstrate and target DNAs to study unselected Ty1 integration events *in vitro* by using a physical assay. The assay is simple and amenable to many manipulations. Since the products are radioactive, they are easy to identify, isolate, and characterize. Additionally, the complete integration event products can be cloned without use of an amplification procedure.

Two major classes of recombinant products. Using this system, we have identified two classes of recombinant products: (i) complete integration events, the result of a concerted joining of both ends of one substrate DNA to a target DNA, and (ii) single-end insertions, the result of one end of a substrate DNA joining to the target. An unusual type of single-end insertion product, C1, was also identified. In this product, two substrate DNAs were joined simultaneously to circular targets, generating a linear species. The C1 product was identified only when circular target DNA was supplied to the integration reaction mixture and was a major fraction of the events detected by using the standard conditions employed in this study; 62% of the products detected are C1. Fragmentation of a linear target would result if this type of product was generated when linear DNA was supplied as the target. There are two observations that suggest this type of reaction occurs when linear target is supplied. (i) A background smear of radioactivity is present in the gel lanes of linear reactions that is not present in the gel lanes of circular reactions when these reactions are done in parallel. (ii) Quantitation of the lengths of DNA of the electron micrographs revealed fragment lengths of various sizes smaller than 4 kb. These fragments were not observed in the EDTA control; however, an IN^- control has not been examined for these fragments, so the latter result may simply be the result of nuclease contamination.

A major product, detected when both linear and circular targets were supplied, was the single-end insertion. Under the same conditions both complete integration events and single-end insertions are detected. From the experiments with circular targets, we directly determined the fraction of complete integration events by quantitating the amount of product C2.1 recovered from C2. The C2.2 product, representing single-end joinings, was about four times more abundant than complete integration events.

The finding that a large fraction of reaction products recovered with this assay system are single-end insertions has several implications. (i) One end of the substrate may be preferred for joining by IN, accounting for the high frequency of single-end insertions. There is some indication that U3 ends

may be more active than U5 ends (see accompanying paper [8]). (ii) The various reaction products may reflect the lack of properly structured higher-order IN complexes on the exogenous substrate. Presumably, VLPs have greater numbers of IN molecules per particle than would be required to insert one Ty1 double-stranded DNA. The nature of IN within the particles is unknown; however, multimeric IN complexes are hypothesized to carry out the integration reaction (24). (iii) Packaging of substrate DNA may be required for efficient complete integration events. Basic proteins or DNA-bending proteins might facilitate the positioning of both ends of one substrate DNA to promote the coordinated insertion of the two ends. Small basic proteins such as HU protein and histone H1 stimulated the recovery of complete integration events mediated by purified human immunodeficiency virus IN by 10-fold (14), whereas these basic proteins failed to stimulate product formation in oligonucleotide-based assays. (iv) Finally, the *in vitro* conditions may change the rates of integration such that these various products are the result of the artificial *in vitro* assay conditions.

Termini of linear targets appear to be favored integration sites *in vitro*. The ends of linear targets were shown to be favored locations for insertion events. The insertions near the ends occurred with various forms of linear target DNA as well as when the insertion was into a second substrate molecule. Since no homology exists between the ends of the substrate and the linear target, these results suggest that the sequences near the ends of the target do not play a role in the selection of these integration sites. We were unable to distinguish whether these near-terminus integration reactions were complete integration events or single-end joinings. We suspect that the L1/L2 products that migrate as complete insertion products represent mostly single-end joinings near the termini (rather than complete insertions near the termini) because in this case the ratio of single-end joinings to complete reaction products in the linear target reactions would resemble the ratio observed with circular targets. In either case, we showed that the terminal functional groups of the target DNA are not required for the joining reaction. These *in vitro* findings support the idea that IN might prefer to insert in cases where ends are exposed. Such ends may occur (rarely) *in vivo*; various types of DNA damage can introduce single-strand breaks into the chromosome (for a review, see reference 32). Several lines of evidence indicate that double-strand breaks of the chromosome can occur even in mitotic cells (35, 36). Furthermore, DNA repair of lesions sustained by the chromosome may introduce transient gaps so that 3'-OH groups are exposed. The HeT-A and TART sequences of *Drosophila melanogaster*, which have retrotransposon-like features, are known to "heal" broken chromosome ends (3, 27); a Ty1-like element, Ty5, from *S. cerevisiae* is also telomere associated (38); perhaps Ty1 can also perform a one-ended integration reaction *in vivo* at or near telomeres.

ACKNOWLEDGMENTS

We are grateful to P. Englund, V. Klein, and M. Delannoy for help with the electron microscopy; E. Caputo for technical assistance; and L. Lin for providing purified *FokI* enzyme. We thank N. Craig, P. Brown, M. Heck, L. Symington, and T. Kelly for helpful discussions and N. Craig and S. Devine for comments on the manuscript.

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

- Ausubel, F., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl (ed.). 1992. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
- Bell, L., and B. Byers. 1983. Separation of branched from linear DNA by two-dimensional gel electrophoresis. *Anal. Biochem.* **130**:527-535.
- Biessmann, H., K. Valgeirsdottir, A. Lofsky, C. Chin, B. Ginther, R. W. Levis, and M.-L. Pardue. 1992. HeT-A, a transposable element specifically involved in "healing" broken chromosome ends in *Drosophila melanogaster*. *Mol. Cell. Biol.* **12**:3910-3918.
- Boeke, J. D., and D. J. Garfinkel. 1988. Yeast Ty elements as retroviruses, p. 15-39. In Y. Koltin and M. J. Leibowitz (ed.), *Viruses of fungi and simple eukaryotes*. Marcel Dekker Publications, Inc., New York.
- 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.
- 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 Press, Cold Spring Harbor, N.Y.
- Braiterman, L. T. 1993. Integration of retrotransposon Ty1. Ph.D. thesis. The Johns Hopkins University School of Medicine, Baltimore, Md.
- Braiterman, L. T., and J. D. Boeke. 1994. Ty1 in vitro integration: effects of mutations in *cis* and *trans*. *Mol. Cell. Biol.* **14**:5731-5740.
- 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 Ty1: phenotypic analysis. *Gene* **139**:19-26.
- Brewer, B., and W. L. Fangman. 1987. The localization of replication origins on ARS plasmids in *S. cerevisiae*. *Cell* **51**:463-471.
- Brown, P. O. 1990. Integration of retroviral DNA. *Curr. Top. Microbiol. Immunol.* **157**:19-48.
- Brown, P. O., B. Bowerman, H. E. Varmus, and J. M. Bishop. 1987. Correct integration of retroviral DNA in vitro. *Cell* **49**:347-357.
- 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.
- Bushman, F. D., T. Fujiwara, and R. Craigie. 1990. Retroviral DNA integration directed by HIV integration protein in vitro. *Science* **249**:1555-1558.
- 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.
- 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.
- Eichinger, D. J., and J. D. Boeke. 1990. A specific terminal structure is required for Ty1 transposition. *Genes Dev.* **4**:324-330.
- Ellison, V., H. Abrams, T. Roe, J. Lifson, and P. O. Brown. 1990. Human immunodeficiency virus integration in a cell-free system. *J. Virol.* **64**:2711-2715.
- 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.
- Farnet, C. M., and W. A. Haseltine. 1990. Integration of human immunodeficiency virus type 1 DNA in vitro. *Proc. Natl. Acad. Sci. USA* **87**:4164-4168.
- 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.
- Garfinkel, D. J., J. D. Boeke, and G. R. Fink. 1985. Ty element transposition: reverse transcriptase and virus-like particles. *Cell* **42**:507-517.
- Grandgenett, D. P., and S. R. Mumm. 1990. Unraveling retrovirus integration. *Cell* **60**:3-4.
- 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.
- Jonsson, C. B., G. A. Donzella, and M. J. Roth. 1993. Characterization of the forward and reverse integration reactions of the Moloney murine leukemia virus integrase protein purified from *E. coli*. *J. Biol. Chem.* **268**:1462-1469.
- Katz, R. A., G. Merkel, J. Kulkosky, J. Leis, and A. M. Skalka. 1990. The avian retroviral IN protein is both necessary and sufficient for integration recombination in vitro. *Cell* **63**:87-95.
- Levis, R. W., R. Ganesan, K. Houchens, L. A. Tolar, and F.-M. Sheen. 1993. Transposons in place of telomeric repeats at a *Drosophila* telomere. *Cell* **75**:1083-1093.
- Mizuuchi, K. 1992. Transpositional recombination: mechanistic insights from studies of Mu and other elements. *Annu. Rev. Biochem.* **61**:1011-1051.
- Monokian, G. M., L. T. Braiterman, and J. D. Boeke. 1994. In-frame linker insertion mutagenesis of yeast transposon Ty1: mutations, transposition and dominance. *Gene* **139**:19-26.
- 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.
- Perez-Morga, D. L., and P. T. Englund. 1993. Microtechnique for electron microscopy of DNA. *Nucleic Acids Res.* **21**:1327-1328.
- Petes, T. D., R. E. Malone, and L. S. Symington. 1991. Recombination in yeast, p. 407. In J. Broach, E. Jones, and J. Pringle (ed.), *The molecular and cellular biology of the yeast Saccharomyces*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Pryciak, P. M., A. Sil, and H. E. Varmus. 1992. Retroviral integration into minichromosomes in vitro. *EMBO J.* **11**:291-303.
- 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.
- Symington, L. S. 1991. Double-strand break repair and recombination catalyzed by a nuclear extract of *Saccharomyces cerevisiae*. *EMBO J.* **10**:987-996.
- Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein, and F. W. Stahl. 1983. The double-strand-break repair model for recombination. *Cell* **33**:25-35.
- 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.
- Voytas, D. F., and J. D. Boeke. 1992. Yeast retrotransposon revealed. *Nature (London)* **358**:717.
- Whitcomb, J. M., and S. H. Hughes. 1992. Retroviral reverse transcriptase and integration: progress and problems. *Annu. Rev. Cell Biol.* **8**:275-306.
- Xu, H., and J. D. Boeke. 1990. Localization of sequences required in *cis* for yeast Ty1 element transposition near the long terminal repeats: analysis of mini-Ty1 elements. *Mol. Cell. Biol.* **10**:2695-2702.