

Characterization of a Transpositionally Active Ty3 Element and Identification of the Ty3 Integrase Protein

LORI J. HANSEN AND SUZANNE B. SANDMEYER*

Department of Microbiology and Molecular Genetics, College of Medicine, University of California, Irvine, California 92717

Received 28 November 1989/Accepted 1 March 1990

Ty3 is a *Saccharomyces cerevisiae* retrotransposon associated with tRNA genes. Two Ty3 elements have been cloned and characterized. The complete nucleotide sequence for one element, Ty3-2, was reported previously (L. J. Hansen, D. L. Chalker, and S. B. Sandmeyer, *Mol. Cell. Biol.* 9:5245–5256, 1988). However, this element is incapable of autonomous transposition. The complete DNA sequence of a transpositionally competent Ty3 element, Ty3-1, is presented here. Its sequence translates into two overlapping open reading frames, *TYA3-1* and *TYB3-1*, which encode proteins with homology to the proteins specified by the retroviral *gag* and *pol* genes, respectively. Comparison of the Ty3-1 nucleotide sequence to Ty3-2 suggests that the *TYB3-2* open reading frame of Ty3-2 is truncated by the deletion of a single nucleotide, which causes a frameshift mutation. Restoration of the reading frame with insertion of a single adenine by site-directed mutagenesis converted Ty3-2 into a transpositionally active element, Ty3-2(+A). Western blot analysis with antibodies made against synthetic peptides identified integrase (IN) proteins in viruslike particle preparations from cells expressing Ty3 elements. Cells expressing Ty3-1 and Ty3-2(+A) produce antibody-reactive proteins with approximate molecular masses of 61 and 58 kilodaltons (kDa), while cells expressing Ty3-2 produce reactive proteins of approximately 52 and 49 kDa. Together, these data show that the 61- or 58-kDa protein, or both, provides the integrase function of Ty3.

Retrotransposons from many eucaryotic systems have been characterized. These elements have long terminal repeats (LTRs), encode proteins with homology to retroviral proteins, and replicate through reverse transcription. They are distinguished from retroviruses primarily by the absence of an obligatory extracellular phase (2, 3, 15, 57). Four retrotransposon families have been characterized for *Saccharomyces cerevisiae*: Ty1 (9), Ty2 (26, 60), Ty3 (10, 25), and Ty4 (55). Ty1 and Ty2 are similar at the nucleotide and amino acid sequence level, and delta elements comprise their LTRs. They are present in approximately 30 to 50 copies per haploid genome and are related to the *Drosophila* copia (41), the *Nicotiana tabacum* Tnt1 (23), and the *Arabidopsis thaliana* Ta1 (59) retrotransposons. This group of elements is distinguished primarily by the fact that the catalytic proteins are encoded in the following order (5' to 3'): protease (PR), integrase (IN), and reverse transcriptase (RT). Ty4 has been recently characterized. Tau elements make up its LTRs, and the complete element is present in low copy number. Ty3 is present in only one to four copies per genome. Its LTR sequences are sigma elements. Ty3 is related to the *Drosophila* gypsylike elements (29, 38, 45, 56, 61) and the *Lilium henryi* del (54) and *Schizosaccharomyces pombe* Tf1 (H. Levin and J. Boeke, personal communication) retrotransposons and only distantly related to Ty1 and Ty2. Comparisons of the polymerase sequences from the latter group of elements show that they resemble retroviral proteins more closely than do those encoded by the former group of retrotransposons. The order in which they express the catalytic proteins is the same as in retroviruses: PR, RT, and IN.

Ty3 is distinguished from other characterized retrotransposons and retroviruses by its specificity for integration next to tRNA genes. Ty3 insertions from cells expressing high

levels of Ty3 RNA have been characterized. They are immediately upstream of the 5' ends of tRNA genes and are flanked by 5-base-pair (bp) direct repeats of the target sequence (D. Chalker and S. B. Sandmeyer, unpublished data). A newly discovered repeated sequence in *Dictyostelium discoideum* is also associated with tRNA genes, although the positions of characterized insertions are distinct from sigma element insertion sites (39). Sigma elements have been found exclusively at positions -19 to -16 with respect to the 5' ends of tRNA-coding sequences (8, 12, 46, 47). The two Ty3 elements present in *S. cerevisiae* AB950 have been cloned and designated Ty3-1 and Ty3-2. Ty3-1 occurs 16 bp upstream of the 5' end of a tRNA^{Cys} gene, while Ty3-2 is 17 bp upstream of a tRNA^{Ile} gene (10, 25).

The nucleotide sequence of Ty3-2 was reported previously (25). Computer translation of the nucleotide sequence revealed two long open reading frames (ORFs), *TYA3-2* and *TYB3-2*, and a short ORF, called ORF3. *TYA3-2* encodes a protein domain that has homology to the retroviral nucleocapsid protein (NC). *TYB3-2* encodes proteins with homology to retroviral PR, RT, and IN proteins.

Transcription of copia (18, 52) and Ty1 (19, 40, 42) retrotransposons results in production of intracellular nucleoprotein complexes. These are referred to as viruslike particles (VLPs) because they resemble the nucleocapsid cores of retroviruses. In the case of Ty1, production of high levels of RNA and VLPs has been linked to replication and transposition of the element (5, 17). The *GAL1-10* upstream activating sequence (32) was used to induce Ty3 transcription, which resulted in expression of Ty3 proteins. VLPs for Ty3-1 and Ty3-2 have been isolated from cells expressing high levels of Ty3 transcripts by sucrose gradient fractionation. The fractions containing VLPs were identified on the basis of the activity of reverse transcriptase and the presence of other Ty3-encoded proteins (unpublished data). In a previous study, it was found that cells expressing high levels

* Corresponding author.

of Ty3-1 transcripts undergo transposition, whereas cells expressing Ty3-2 do not (25). Ty3-2 is therefore capable of forming VLPs and has a functional reverse transcriptase but is blocked at a step before integration.

In order to understand the disparity in the abilities of Ty3-1 and Ty3-2 to transpose, the nucleotide sequence of a functional Ty3 element, Ty3-1, was determined. The effects of sequence differences between Ty3-1 and Ty3-2 were examined by constructing hybrid Ty3 elements and testing them for transposition. Differences in the proteins predicted from Ty3-1 and Ty3-2 nucleotide sequences correlated with different mobilities of proteins encoded by the IN regions of the two elements, resulting in identification of the proteins required for Ty3 integration. These studies are presented here.

MATERIALS AND METHODS

Nucleotide-sequencing strategy. The nucleotide sequences of the Ty3-1 sigma elements and the tRNA^{Cys} gene associated with Ty3-1 were reported previously (10). The dideoxy-chain termination method (48) was used to determine the nucleotide sequence of the Ty3-1 internal domain. All sequence analyses used the enzyme Sequenase (United States Biochemical Corp.) and ³⁵S-dATP (1,000 Ci/mmol; Amersham Corp.). Subclones from the plasmid pTy3-1 (25) were constructed with Ty3-1 restriction fragments cloned into pIBI20 or pIBI21 (International Biotechnologies, Inc.). Small-scale preparations of these plasmids were obtained by the boiling method of Holmes and Quigley (27). The sequences of the insertions were determined from the double-stranded plasmids (24) in polymerase reactions primed with the M13 universal primer and separately with a reverse sequencing primer (International Biotechnologies). These two primers hybridize adjacent to opposite ends of the insert DNA. Synthetic oligonucleotides (Operon Technologies, Inc.) based on the Ty3-2 sequence were used as sequencing primers for analysis of regions outside and overlapping the ends of cloned fragments. The nucleotide sequence was compiled, edited, and translated by using the MS-DOS Sequence Analysis Program (SnAP) written by A. Goldin and G. Gutman (University of California, Irvine).

Construction of chimeric Ty3 elements. Chimeric Ty3 elements were designed in which the 5' and 3' regions of the Ty3-1 and Ty3-2 internal domains were exchanged. These were constructed in the pEGTy3 expression plasmid so that they were galactose inducible (25). This plasmid contains the *GAL1-10* upstream activating sequence fused to a Ty3 element upstream of the putative TATAA sequence in the 5' sigma element. Ty3 transcription is induced when transformed cells are grown in galactose-containing medium. pEGTy3 contains the yeast *URA3* gene, which permits positive and negative selection of cells containing the plasmid. Yeast cells transformed with the plasmid are selected by growth on synthetic medium minus uracil, while cells which have lost the plasmid can be selected by growth in medium containing 5-fluoro-orotic acid (5-FOA), which is toxic to cells expressing the *URA3* gene (6).

A *Sall* restriction site within the internal domain and *XhoI* sites present in each sigma element were used to construct the chimeric plasmids. To create pEGTy3-2/1, the 5' 2.9-kilobase-pair *XhoI-Sall* restriction fragment of Ty3-2 was ligated to pEGTy3-1, which had been digested with *Sall* and partially digested with *XhoI*, thereby replacing the corresponding Ty3-1 sequence. Similarly, for pEGTy3-1/2, the 3' 2.2-kilobase-pair *Sall-XhoI* restriction fragment from Ty3-2

was substituted for the downstream portion of the Ty3-1 internal domain.

Site-directed mutagenesis of Ty3-2. A pIBI21 plasmid containing the 2.2-kilobase-pair *Sall-XhoI* fragment from the 3' portion of Ty3-2 (25) was constructed in order to produce the coding strand upon superinfection with the helper phage M13K07 (International Biotechnologies). This plasmid was transformed into *Escherichia coli* RZ1032 [*lysA* (61-62) *thi-1* *relA1* *spoT1* *dut-1* *ung-1* (Tet^r) *supE44*], and single-stranded DNA was isolated. The mutagenic 41-mer oligonucleotide (5'GTTATTAATGTACAATTCCTGAAAAAGTTTGTATACCGTCC3'), which inserts an adenine (shown in bold-face type) into the sense strand, was used to prime the polymerization reaction (34), and the resulting DNA was transformed into *E. coli* HB101 [F⁻ *hsd-20* (r_B⁻ m_B⁻) *recA13* *leuB6* *ara-14* *proA2* *lacY1* *galK2* *rpsL20* (Sm^r) *xyl-5* *mtl-1* *supE44* k⁻]. Plasmid DNA isolated from transformants was screened by sequence analysis. The *Sall-XhoI* restriction fragment containing the mutated sequence was then substituted into pEGTy3-2 by replacing the 3' *Sall-XhoI* fragment. This plasmid differs from pEGTy3-2 by a single nucleotide and is designated pEGTy3-2(+A).

Transposition assays. Yeast strains were cultured by standard methods (51). *S. cerevisiae* yVB110 was derived from yVB109 (*MATa* Δ *trp1-901* *ura3-52* *his3-200* *ade2-101* *lys2-1* *leu1-12* *can1-100*) by serially deleting the three endogenous Ty3 elements (V. Bilanchone, personal communication). The strains are otherwise isogenic. The Ty3 null strain, yVB110, was transformed separately with the plasmids pEGTy3-1, pEGTy3-2, and pEGTy3-2(+A) and the chimeric plasmids pEGTy3-1/2 and pEGTy3-2/1 by a modification of the technique of Ito et al. (30), and transformants were selected on synthetic medium minus uracil. The assay for transposition of the chimeric Ty3 elements was performed as described previously (25). Briefly, cells containing pEGTy3-1/2 and pEGTy3-2/1 were grown on galactose-containing synthetic medium minus uracil to induce Ty3 transcription and transposition. After selection against cells retaining the donor plasmid by growth on 5-FOA-containing medium, in situ hybridization was used to screen for colonies which had acquired Ty3 sequences through transposition. This assay was modified to a more sensitive DNA slot blot hybridization analysis of the pEGTy3-1, pEGTy3-2, and pEGTy3-2(+A) plasmid-containing cells. After induction of transposition on galactose-containing synthetic medium minus uracil at 30°C for 7 days, 10 colonies per type of transformant were transferred to YPD medium (1% yeast extract, 2% peptone, 2% glucose) and incubated at 30°C for 1 day. These colonies, in which multiple independent transposition events had occurred, were patched onto medium containing 5-FOA, and 50 clonally homogeneous colonies were isolated from each of the original colonies. In these colonies, each transposition was represented equivalently. Cells representing the 500 colonies from each of the three transformants were grown in pools of 10 in YPD broth. Genomic DNA was extracted from the 150 pooled cultures by the method of Boeke et al. (5) and prepared for slot blot analysis by the method of Adams and Hatfield (1). The DNA was denatured with basic 15× SSPE (100 ml of 20× SSPE [3 M NaCl, 0.166 M Na₂HPO₄, 0.022 M EDTA, pH 7.0] plus 30 ml of 2 N NaOH) at 80°C, neutralized with acidic Tris hydrochloride (100 ml of 2 M Tris base plus 21 ml of concentrated HCl), loaded onto nitrocellulose filters (Schleicher & Schuell, Inc.) contained in a slot blot apparatus, and washed with 10× SSPE (pH 7.0). The nitrocellulose filters were baked at 80°C for 1 h. The DNA was hybridized to a ³²P-labeled Ty3

internal domain probe as previously described (10), and the filters were exposed to Kodak XAR-5 film in the presence of a Quanta III intensifying screen (Du Pont Co.).

VLP preparation. VLPs were isolated essentially by the method of Eichinger and Boeke (17). Cultures were grown in galactose-containing medium at 30°C to an optical density at 600 nm of 0.9 to 1.0; the cells were then broken and fractionated as described previously (17). Sucrose step gradient fractions were analyzed for reverse transcriptase activity (19). Peak fractions at the 70%-30% interface were pooled and concentrated by ultracentrifugation at 38,000 rpm for 1 h at 4°C in a Beckman Ti50 rotor. Concentrated VLPs were suspended in 150 µl of H₂O per liter of starting culture. VLPs were mixed 1:1 with 2× sodium dodecyl sulfate-polyacrylamide gel electrophoresis loading buffer (100 mM Tris hydrochloride [pH 6.8], 50% glycerol, 0.1% sodium dodecyl sulfate, 5% 2-mercaptoethanol, 0.001% bromophenol blue), boiled for 2 min, and stored at -70°C. For the null strain, protein from the equivalent sucrose fractions was pooled and treated in a similar manner.

Western blot (immunoblot) analysis. A peptide with the amino acid sequence CHAQIEMETNNQRR (University of California, San Diego, Protein Facility), was designed from the predicted integrase sequence (nucleotides 4,503 to 4,544; see Fig. 1). The amino-terminal cysteine is not Ty3 encoded but was added in order to couple the peptide to bovine serum albumin (Sigma Chemical Co.) with the bifunctional reagent *m*-maleimidazolyl-*N*-hydroxysuccinimide ester (33, 37). Two rabbits were injected subcutaneously with approximately 0.3 mg of bovine serum albumin-conjugated peptide (4 to 6 peptides per bovine serum albumin molecule) in Freund complete adjuvant, followed by a boost 2 weeks later with 0.3 mg of conjugated peptide in Freund incomplete adjuvant. Two additional boosts at 2-week intervals contained 0.5 mg of free peptide in Freund incomplete adjuvant.

The anti-IN antibodies were affinity purified for use in the Western blot analysis. The immunoglobulin G fraction was isolated by passing the antiserum over a protein A column (Immunopure Immobilized Protein A gel; Pierce Chemical Co. [22]), followed by elution of the bound immunoglobulin G with 0.1 M glycine, pH 2.8. The peak fractions were identified by optical density at 280 nm, adjusted to neutral pH, and dialyzed against phosphate-buffered saline (PBS). The immunoglobulin G fraction was then passed over a Sulfolink Coupling Gel (Pierce) column to which the IN peptide (CHAQIEMETNNQRR) had been conjugated through the cysteine residue (20). The anti-IN antibodies were eluted with 0.1 M glycine, pH 2.8. Peak fractions were identified by optical density at 280 nm, pooled, and dialyzed against PBS. Bovine serum albumin was added to a concentration of 1% (wt/vol) to stabilize the immunoglobulin G, and samples were stored at -20°C.

Concentrated VLP proteins and the equivalent preparation from the null strain (approximately 20 µg of total protein) were fractionated on a Laemmli sodium dodecyl sulfate-10% polyacrylamide gel (36) and transferred to a NitroScreen West membrane (Du Pont, NEN Research Products) by the semidry electrophoretic transfer method of Kyhse-Andersen (35). The membrane was incubated with blocking buffer (PBS containing 5% nonfat dry milk, 0.02% sodium azide, and 0.2% Tween 20) for 1 h at room temperature, incubated with the affinity-purified anti-IN antibody (diluted 1:50 in blocking buffer) for 1 h at room temperature, washed with PBS-0.1% Tween 20, and incubated for 1 h at room temperature with ¹²⁵I-protein A (ICN Pharmaceuticals Inc.) in blocking buffer (0.1 mCi of ¹²⁵I-protein A per ml

blocking buffer). The membrane was then washed extensively with PBS-0.1% Tween 20 and exposed to Kodak XAR-5 film.

RESULTS

Nucleotide sequence of Ty3-1. The Ty3-1 retrotransposon lies at position -16 relative to the 5' end of a tRNA^{Cys} gene. Ty3-1 is 5,351 bp long and consists of two identical 340-bp sigma elements flanking an internal domain of 4,671 bp (Fig. 1). The four sigma elements from Ty3-1 and Ty3-2 are identical. The internal domains of Ty3-1 and Ty3-2 are highly similar (99.2%), but there are a few notable differences.

TYA3 is analogous to the retroviral *gag* gene. Starting at the first methionine codon at nucleotide positions 416 to 418, *TYA3-1* and *TYA3-2* are both 290 codons long. There are eight nucleotide differences between the two elements in this region. These result in only two amino acid differences: glycine in Ty3-1 versus arginine in Ty3-2 and proline in Ty3-1 versus alanine in Ty3-2. These residues are 37 and 10 codons, respectively, upstream of the C-X₂-C-X₄-H-X₄-C motif at the carboxyl-terminal end of the predicted *TYA3* protein.

The *TYB3* sequence predicts proteins with homology to the retroviral *pol* proteins PR, RT, and IN in that order (25). As is the case in Ty3-2, the *TYB3* ORF overlaps the last 13 codons of the *TYA3* ORF and is in the +1 reading frame with respect to *TYA3*. In Ty3-2, there is nearly perfect duplication (76 of 78 nucleotides [25]) of the sequence from nucleotide position 1776 to position 1854 in Ty3-1. In Ty3-1, this region contains a 60-bp sequence flanked by direct 18-bp repeats. Therefore, the pattern for Ty3-1 is 18 bp-60 bp-18 bp, while for Ty3-2 it is 18 bp-60 bp-18 bp-60 bp-18 bp (Fig. 1). Because the repeats do not interrupt the *TYB3* reading frame, there is a 6-amino-acid repeat in the predicted *TYB3-1* protein, and the *TYB3-2* protein sequence is 26 amino acids longer than that of *TYB3-1* in this region. The duplication in the Ty3-2 sequence occurs between the regions that contain homology to the PR and RT proteins. This kind of structural polymorphism has also been seen in long interspersed nuclear elements. Mouse L1 elements contain two or three copies of a 42-bp sequence that preserves the reading frame (50).

Another significant difference between Ty3-1 and Ty3-2 is the presence of an additional nucleotide at Ty3-1 position 4744, compared with the homologous region of Ty3-2. The consequence of this difference is that the carboxyl terminus of *TYB3-2* is truncated by 75 codons with respect to *TYB3-1*, and the last 30 codons of *TYB3-2* are within a different reading frame from those specified in *TYB3-1* (Fig. 2). *TYB3-1* is 1,270 codons long and extends into the 3' sigma element for 15 codons. This contrasts with *TYB3-2*, which is only 1,221 codons in length (including the 26-amino-acid repeat) and terminates upstream of the 3' sigma element. The overall difference in the size of *TYB3-1* and *TYB3-2* is 49 codons. There are 30 nucleotide differences between *TYB3-1* and *TYB3-2* which result in six differences in amino acid sequence. One of these involves a charge difference in the RT region. *TYB3-1* residue 594 is an aspartate, while the equivalent *TYB3-2* residue is an alanine. The nucleotide and amino acid differences are indicated in Fig. 1.

Transposition assays of the chimeric and mutant Ty3 elements. In order to localize the region responsible for the transposition defect of Ty3-2, chimeric Ty3 elements that combined different regions of Ty3-1 and Ty3-2 were constructed. These were made by using the *Xho*I restriction site

AACTTCATGGAGGACACCTAGTAAATAAAGCTCGACTCAGGATC -71
 GAACATAGGACCAACAGATTGGCAATCTGCTGGCTACCACTCGCCATACAGGCTGATTTTCTGAAG -1
 TGTGTATCTCAAAATGAGATATGCTAGTACAACTGCTGACCTGACCGTTCATAAACAACATATGA 70
 AACACCTTATAAAGGACGACACATGAGCAAAACCGACCTCCCTAGCTGAATCCCAAGATAT 140
 AAATGCTGAACTTAGTATTAGATCCGAGATTCGGCTCCCACTTAGTATGATTCATATTTATA 210
 TAATATATAAGTAACTTCCGTAATTAATCTGATAAAGTGTGTAACAATGTTGACAACTGGTACTCCCTA 280
 AGAGCTTTATATTAGGATGTCAGACACTCCGATTAATCTGACGGCCGTAATACACCTGGTAGCC 350
 TTAAGGTTACTAATGTCGAAGCAACCTCGAAAGCCGACCTAGCTACACCCAGCCAGTATGAG 420
 L F K R T I E K P H L A T P H P S M S 2
 CTTTATGGATCAATCCAGGAGGAGAAATATCCAAACTCCCACTAGAACTGCTTCTAATCTCCG 490
 F M D Q I P G G G N Y P K L P V E C L P N F P 25
 ATCCAACATCTTTGACCTTCAGAGTAGAATGACTCCATAAAGTCAAAACTTTATCTCCGAATAA 560
 I Q P S L T P A G C R N D S H K L K R N F I S E I H 49
 C TGTAAACATGCTATGATATCTGGCCAGTATGCGACTGCTATGCTACTGACAGAGACATTTAT 630
 L N H S M I S W P M D A S R I V Y C R R H L L 72
 AAACCCGCTCGTCAAGGATGACTTGTGACAGCAAGCTATCTGAAATAACATTCGACACA 700
 N P A A Q M G D F V O E G I L E I T F D T 95
 TFCATACAGGATATATGACATTTCTAAGCCACAGATCAATAAATCTTAACTGCAATCAGCC 770
 F I O G L Y Q H F Y K P P D I N K R I F N A I T Q 119
 AACTTCCGAGCTAACTGGTATGAGCTCTCAACCAAGCTTCAAGAAATTTGGACAGAAATGCC 840
 L S E A K L G I E R L M Q R F R K I W D R M P 142
 ACCGACTTCAGCCCAAAAGTGGCTATGACATATGACTAGCTATGACAAAGAAACCTATATAT 910
 P D F H T E K A A I N H Y T T R L L T K E Y H 165
 ATTGTCAGATGCAAAACGAGACATTAAGAGCCGCTGGAAGAGGCTTACGACAGCACTGACTAA 980
 I V R M H K P T L R D A M E E A Y O T T A L E T 189
 CTGAAGACTTCCGAGATTCGAATGATGCTGACGACACTATCTGCTGACCAACCACTTAT 1050
 E R F P P G F L E D A D G D T T I I G A T T H L 212
 ACAGAAAGATGACTGACTGATGATTCAGAAAGATTCGACCCAGATGACTGCTATCCGTA 1120
 Q E E Y D S D Y D S E D N L F Q H C Y V H T V 235
 AGGCAAGAGATCTTACATAAACAATGCAATCAATCGAAACAGGAGAAATAACAACCTCTAGAG 1190
 R T R R S Y H K P M S H H R M R R H M H P S R E 259
 AAGATGTATAAAAATGCGCTATGCTTCTATGTGAAGAGAGGAGCACTGCTGCAAGTATGAGCC 1260
 E C I K N R L L P Y K K R E G R L H E R A J 282
 T H V E H 5
 ACGTAAGCGATCTTACCGATTTGAACCTGAACTCAAAAGACCAAACTCTTTATCAAAACTT 1330
 R K A S S M R S 290
 V R R V L T D L E L E S K D O O T P F I K T L 28
 ACCAATGTACATATATGCCATCCCGAGATGCAATCCCGGAAACCAATAAAATACAAAC 1400
 P I V H I G E P E N D N T A E K T I K I Q H 51
 ACCAAATGAAACCTGTTTGAACGATGATCCCACTGCTTATCCGAGAGATATTAGCAACTC 1470
 T (R V K T L F) (S P T S F I R) R D I V E L L 75
 TCAAAATGCAATGACAGACCCCTCACTCCCTTTAGAGATTCGACCCCAAACTCCGCTTAC 1540
 K Y E I Y E T P L R F R G F V A T K S A V T 98
 ATCCGAGCAGTACCACTGACTCAAACTCAAGTCTGATTAATCTTTAGCCGCTACATAGT 1610
 S E A V T I D L R I N D L H I T L A A Y I L D 121
 AAGATGATACCAATGTTAATGGAATCCACTTACCCGCTACCGGAACTTCCGACAGTAC 1680
 W M D Y Q L L I G N P I L R R Y P K I L H T V L 145
 TGAATACGAGAGAGCCGCTTAAAGCCAGACTTATCCGCTCCGAAACCGTAAATAACGTTAG 1750
 H T R E S P D S L K P T Y R S E T V M H V R 168
 AACCTACTCCGCTGATGCTGATGACCCGCAAACTAAAGTCTGTTTCCGCAACTCTCGAA 1820
 T Y S A C H R G B P H I K L S F A P T J L E 191
 GCACTGACCCGAAATCCGCTGGTAACTGCTGACTCGACGCAAAACCTGCTCTGCAACCACTA 1890
 A T D P K S A G M R G D S R T K T L S L A T T T 215
 CTCTCGACAAATGACCCGCTTACCGCTTGAATACCCGAGTACTCAAGTACATTTCCGCAAT 1960
 F A A I D P L T L D N F G S T A Q S T F A G F 238
 CCGACTACTGAAAGAGGAGCTTCCGAGAGGATGAAATCTCCAGCTGTTCTCAACCTTACG 2030
 P I P E A S I L E E D G R Y S H V V S T I Q 262
 AGTGTAGAACCTAATCTACTGATGACGAAATAGGACACCTTTGCACTTTGCCGATTTGGTACAAC 2100
 S V E P H A T D H S M R D T F C T L P V M L Q Q 285
 AGAATATAGAGATGACTAGTAACTGCTCCACCAAGACTGCGGACATTAATCAATCCCGGAA 2170
 K Y R I I R M D L P R P A D I H M I P V R 308
 ACATGATATGAAATTAACCTGGCCAGACTACTCTGACTACGCCATACCATGTTACAGAAAGAAC 2240
 H D I E I R P G A R L P R [D Q P Y H V T E K H 332
 GAACAAGAAATCAACAATACTCAAAACTGCTGATAACAAGTCTATTCTCCCTCAAAGTCCCTT 2310
 E Q E I W K I V Q K L L D M K F I V P S K S P C 355
 CGACTCCCTGACTCTGCTCGGAAAGAGCGGTACTCCGACTGCTGCTGATTACCGACCT 2380
 S S P V V L V P K K D G T F R L C V D Y R T L 378
 GAACAAGCTACCTCCGACCATCCATTAACCAAGTCCGACACTTATGAGCGTATGGAAAT 2450
 H K A T I S D P F P L R I D N L L S R I G N 402
 GCCGAGATTTTACCGCTAGATTTGATAGTGGTACCACAGATCCGATGAAACCAAGACCGCT 2520
 A Q I F T T L D L H S G Y H Q I P H E F K D R Y 425
 ACAAAACCGCTTGTGACACCACTCCGTAAGTAAATATACCTCATGCCATTTGGCTTACTCAATG 2590
 K T A F V T P S G K Y E I V N H P G L V M A 448
 ACCTAGTACCTCCGAGATGACTGCTGATGATTTAGACCTGAGATGCTCAATGTTACTGAT 2660
 P S T P A R Y N A D T F R D L R F V H V Y L 472
 GATATATTAATCTCCGAACTCCGAAACATTTAGAACCTTATGACAGCTAGTAAAGATTA 2730
 I L I F S E S P E E H M K H L D T V L E R L K 495

AGAACGGAACCTTATTAGAGAAAAATGTAATTTGCATCTGAAAGAACTGACTTTTGGCTA 2800
 N E M L I V K R R K C R F A S E T F L G 518
 TACTATTGGAATCCAGAAATGCTCCACTACAGCAAAATGCGACCAATCCGAGACTTTCCGACGCT 2870
 S I G I O K I A P L O H K C A A I R D F P T P 542
 AAAACGTAAGCAAGCAGAGATTTTGAAGATGATTAATCTACAGACGATTTCTGAAATGCT 2940
 K T V K Q A D R F L C N I N Y V R F I P C S 565
 CCAAGTTCCACAGCAATCCACTGTTTATTGTGCAAAAGTCAATGGACAAAACAGCAAGC 3010
 K I A Q O P I Q L F I C D K S Q M T E K Q D K A 588
 AATTGATAAACTAAAGAGCGCTTGTACTCCCGCTCTACCTACCAACAAGAACTCAACTC 3080
 I D R L R D A L C N S P V L V P F N R I K A N Y 612
 CGACTTACACAGACCGCTCAAAAGCGGCTTGGCTGCTGACAAAGTGCAGCAAGCAAAAC 3150
 R L T T A S K D G I G A V L E E V D N K H R L 635
 TTGTGCTGCTGCTGCTTACTCTAAATCTTACAGAGTCCCGAAGAAATCTCTGCTGGCAAT 3220
 V G V V G Y P S K S L E S A Q K N N Y P A G E L 658
 AGAAGCTTGGAAATTAAGCAACTCCACCTTCCGATATATGCTTACGGAAGACTTTCAGT 3290
 L L G I I K A H H F R Y M L H G K H P T L 682
 AGAACAGACACATGTTGTTATCATTACAAACAGCAAGCCGACCGCTGCAACCGCTGCT 3360
 R T H I S L L S L D H K H E P A R R V O R M L 705
 TAGTAGCTAGCCACATGACTTACCTAGTAACTAGCTGACCCAGCAAGCTTCTGCGAGATC 3430
 D D L A T Y D F T L E Y L A G P R V V A A 728
 CATATCCCTGCTGATATGACTAATCCCGGAAACCTCCGACTATGACAGAAAGTGGAAATC 3500
 I S R A V Y T I T P E T S R P I D G E S W K S 752
 TACTACAATGACAGCCATATGAGTCTGCTTAACTATGAAAGAAATGACACCAAGCAAGTCA 3570
 Y K S D P L C S A V L I H H K E L T Q H M V T 775
 CACTGAGATATGACCGCTTCCGATGACGAGAAACTGCAACTACAGACACTTCCGAAAGAA 3640
 P E D H S A F R S Y O R K L E L S E T F R K H 798
 TTATCCCTGAGAGCAATGATCTATTAACAGACCGACTAGTACCAATAAAACAGAGACCA 3710
 Y S L E D E H Y Y V Q D R L V V P I K Q R A 822
 GTTAGAGACTATATGACCACTTATTGGAGACATTTGGTGTACAGTACCGCTTGGGAAA 3780
 V H R L Y H D T L F G G F G V T P L A K I 845
 TCAAGCAATTTACTATGGCAAAATGACCAATTCGATACATACATCAGGACTCCGTAAGT 3850
 S P I Y M P K L Q H S I I O Y I R T V O 868
 TCACTAATAAATGACAGCCGCTTACAGGACTTACACCCACTCCCTACAGCAAGGAGAA 3920
 Q D I R S H R P A L H G L L O P L F I A E G R 892
 TGCTGATATCAATGATTTGTGACAGGATTTCCCGGCACTCAAAATGATGATGATCTCG 3990
 W L D I S H B F V T G L P P T S M H I L V 915
 TCGTAGTGTCTTTTGAAGCGGCTCACTCATAGTACAGGAAACCTTAGCCCAACACT 4060
 V V D R F S K R A H F I A T R K T L D A T Q L 938
 AATAGACTACTTTCGATGATTTTCATATGATGTTTCCGCAAGCAATACCGCTGATAGAG 4130
 I D L F R Y I P S Y H C F P R T I T S R D 962
 CTCCTATGACCCGCAAAATGACAGAACTCAGCAAAAGACTAGGATAAATGACAGACTCTCCG 4200
 V R H T A D K Y Q E L T K R L G I K S T M S S A 985
 CGAACCCCAACAGATGAGCAATCCGACGAGATGACAGACTTAAACAGGTTACTAAGAGCTA 4270
 N H P O T D G O S R T I O T L N L R A Y 1008
 TGCTCAACCAATTTGAGATGATATTTACCAAAATGCAATTTTACAAATCTACACT 4340
 A S T M I Q M H M H V Y L P Q I E F V Y M S T P 1032
 ACTGAGCACTTGGAAATCACCATTGAAATGATTTAGGATATTACCAATACCCCTGCTATTAAG 4410
 T R T L G K S P F E I D L G L P P T P A I K S 1055
 CAGATGACAGTCAACGAGAGTCTTACTGCTGAGACTGCCAAACCTCAAGGCTTACCAT 4480
 D D E V H A R S P T A V E L A R H L K A L T I 1070
 CCAAGCAGGAGCAGCTAGACAGCTCAAACTGGAATGGAATCAATCAACTCAAGACTGAAACCC 4550
 Q T K E Q L E H A Q I E H E T H M H O R R K P 1101
 TTATGTTAAATAGGAGTACCTATTGATGATGAGATGATCAATCAAGAAAGTCTTATGAT 4620
 L L L H I G D H V L V H R D A Y F K R G A Y M K 1125
 AAGTACAATAATACCTCGGACTTTCGAGTTGCTCAAGAAATAAAGCAATACCCCTACGACTAGA 4690
 V Q Q I Y V G P F R V V K K I N D M A Y E L D 1149
 TTTAACTCTCAAGAAAGCAGAGTATTAATGTAATCTCAAAATGCTGAAAGTCTTATACCTCCA 4760
 L H S H K R R H R V I H V O F L K R F V Y R P 1172
 GACCGTACCAAGAAATAACCAATCAGCTCCTCAAGAAATTAAGAGGACAGCAAGTACTGCAC 4830
 D A Y P K M K P I S S T E R I K R A H E V T A L 1195
 TCTAGGATAGTACTACACAAACTTACTTATGCTACAGTCAAGACTTACAGCAACTCTGCT 4900
 I G I D T H R Y L C H N Q D V D F T L S V 1219
 AGAATCTCAGAGCTGAAATTTGCCAAATCCGAAAGCAAGCAAGTCAATATAGCAACTTAGA 4970
 E Y S E A E F C Q I P E R T R S I L A N F R 1242
 CACTTACGAAACAGCAACCTGAGAGAGGAGATGTTGTTCAAAATGAGATGTCAGT 5040
 Q L Y E T Q D N P E R E E D V S Q T H E I C Q Y 1265
 ATGCAATGCTACCTGACCTGATCAAAACATATGAAACACTTATAAACAAGCAACATG 5110
 D N T S P 1270
 AGACAAACCCGACTCCCTAGCTGACTACCAAGTATAAATGCTGAACTAGTTAGATCCGA 5180
 GATTCCGCTCCACACTAGTATGATGATTTATATAATATAAGATAAGTAACTCCGTA 5250
 ATTAATCTGAAACTGTTTACAGCTGCTTCCCTAGACTGTTATATAGGATGTCAGACA 5320
 CTCGGTATTACTCGAGCCGTAATACAGAACTTCCATTTGATGCTTATATGGAATATGA 5390

<i>TYB3-1</i>	1161	V	Q	F	L	K	K	F	V	Y	R	P	D	A	Y	P	K	N	K	P	I	1180	
<i>TYB3-2</i>		V	Q	F	L	K	S	L	Y	T	V	Q	T	R	T	Q	R	I	N	Q	S	1206	
	1187																						
<i>TYB3-1</i>		S	S	T	E	R	I	K	R	A	H	E	V	T	A	L	I	G	I	D	T	1200	
<i>TYB3-2</i>		A	P	L	R	E	L	R	E	H	T	K	L	L	H	S	*						
																						1221	
<i>TYB3-1</i>		T	H	K	T	Y	L	C	H	M	Q	D	V	D	P	T	L	S	V	E	Y	1220	
<i>TYB3-1</i>		S	E	A	E	F	C	Q	I	P	E	R	T	R	R	S	I	L	A	N	F	1240	
<i>TYB3-1</i>		R	Q	L	Y	E	T	Q	D	N	P	E	R	E	E	D	V	V	S	Q	N	1260	
<i>TYB3-1</i>		E	I	C	Q	Y	D	N	T	S	P	*										1270	

FIG. 2. Comparison of *TYB3-1* and *TYB3-2* carboxyl termini. The deduced carboxyl-terminal amino acid sequence of *TYB3-1* (residues 1161 to 1270) is shown above the carboxyl terminus of *TYB3-2* (residues 1187 to 1221). There is a -1 frameshift in the reading frame of *TYB3-2* (after residue 1191) with respect to *TYB3-1* due to a single-nucleotide difference in this region (indicated by an arrowhead). The last 15 *TYB3-1* residues are encoded by the 3' sigma element and are underlined. Stop codons are indicated by asterisks (*).

present in each sigma element and the unique *Sall* site present in the internal domain (Fig. 3). The *Sall* site lies downstream of the 78-bp duplication (Fig. 1) and is 576 bp upstream of the region encoding amino acid homology to retroviral IN proteins. One chimeric element, Ty3-2/1, contained the Ty3-2 sequence upstream and the Ty3-1 sequence downstream of the *Sall* site, and the other, Ty3-1/2, contained the reverse order of the Ty3 domains. These chimeric elements were created in the high-copy pEGTy3 (25) vector so that expression of the Ty3 elements was galactose inducible. The two chimeric Ty3 elements were grown under inducing conditions and tested by colony hybridization for evidence of transposition (see Materials and Methods). It was previously shown that transformants grown under non-inducing conditions (glucose-containing medium) do not undergo transposition (25). The hybridization assay showed that replacement of the 3' Ty3-1 *Sall*-*Xho*I restriction fragment with the equivalent Ty3-2 sequence rendered Ty3-1 nonfunctional (Ty3-1/2) (Table 1, assay A). Conversely, replacing the 3' Ty3-2 sequence with the Ty3-1 sequence allowed Ty3-2 to transpose at the about the same frequency observed for Ty3-1 (25; this report). This demonstrates that the transposition defect of Ty3-2 lies in the downstream region of *TYB3-2* and that there is no significant effect of the duplication in Ty3-2 on transposition.

In order to determine whether the truncation of *TYB3-2* is solely responsible for the transposition defect, a single

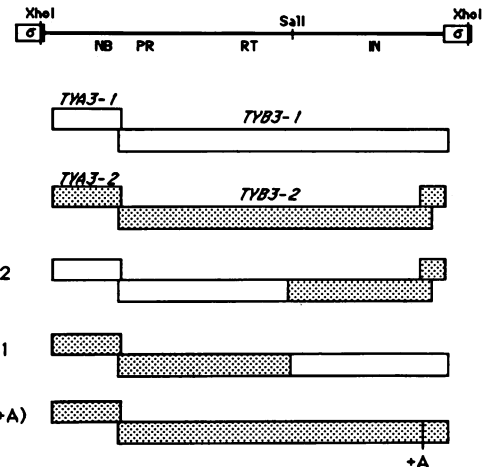


FIG. 3. Comparison of structures of wild-type, chimeric, and mutant Ty3 elements. The overall structure of a Ty3 element is diagrammed on the top line. The sigma element LTRs are boxed, and the heavy line between them represents the internal domain. The *Xho*I and *Sall* sites used in the construction of the chimeric Ty3 elements are indicated. Regions that contain homology to retroviral proteins are indicated below the internal domain. The reading frames for the various Ty3 elements are shown by boxes. The open boxes represent Ty3-1 sequence, while the stippled boxes represent Ty3-2 sequence. Addition of a single nucleotide is indicated for Ty3-2(+A).

adenine was introduced at Ty3-2 nucleotide position 4821 (equivalent to Ty3-1 position 4744) by site-directed mutagenesis. Addition of this nucleotide restores the reading frame to that of Ty3-1. The mutated Ty3-2 nucleotide sequence, designated Ty3-2(+A), was cloned into the galactose expression plasmid and designated pEGTy3-2(+A). Ty3-2(+A) was tested for its ability to transpose. yVB110 cells transformed separately with pEGTy3-1, pEGTy3-2, and pEGTy3-2(+A) were grown on galactose-containing synthetic medium minus uracil to induce Ty3 transcription, cured of the plasmid by growth on 5-FOA-containing medium, and streaked for individual colonies. DNA was extracted from pools of individual colonies, transferred to nitrocellulose in a slot blot apparatus, and hybridized to a radiolabeled probe specific for the Ty3 internal domain. The autoradiograph resulting from this experiment is shown in Fig. 4. As expected, Ty3-1 transposed at a relatively high frequency. At least one transposition occurred per 50 isolates of the original 10 pEGTy3-1-transformed colonies (Fig. 4A). The inability of Ty3-2 to transpose is illustrated in Fig. 4B; none of the 500 colonies derived from pEGTy3-2 transformants had acquired a genomic copy of Ty3 (the lower right slots that show hybridization are control samples of DNA from

FIG. 1. Nucleotide sequence of Ty3-1 and the associated tRNA gene and deduced amino acid sequence of Ty3-1 proteins. The Ty3-1 nucleotide sequence is shown above the predicted amino acid sequence. Numbering of the nucleotide sequence is relative to the first base of the 5' sigma element. Arrows above the 8-bp inverted repeats delineate the ends of the 5' and 3' sigma elements (nucleotides 1 to 340 and 5012 to 5351, respectively). The tRNA^{Cys}-coding sequence in lines 1 and 2 (-90 to -16 relative to the 5' sigma element) is overlined with an arrow indicating the direction of transcription. The 5-bp flanking direct repeat of the target sequence is underlined. The downward-pointing arrow at nucleotide position 223 denotes the major transcriptional start site. Numbering of the amino acids begins with the putative initiation methionine codon (shaded) for *TYA3-1* and with the first amino acid encoded in the *TYB3-1* ORF. Ty3-2 nucleotide and amino acid sequence differences are shown above and below the Ty3-1 sequences, respectively. Brackets indicate the amino acid sequences used in protein alignments to show homology to the retroviral NC, PR, RT, RH, and IN proteins (25). Highly conserved amino acid residues (31) are shaded. Predicted minus- and plus-strand primer regions are overlined. Arrows over the internal domain indicate the 18-bp direct repeats and the 60-bp sequence that is discussed in Results. The asterisk (*) above nucleotide 4744 indicates the presence of an adenine in Ty3-1 that is absent in Ty3-2. The boxed sequence (*TYB3-1* residues 1086 to 1099) indicates the peptide used for antibody production.

TABLE 1. Ty3 mobilization after induction with galactose

Assay and GAL-Ty3 fusion	Tpn/10 ^a	Tpn/500 ^b (% of total)
Assay A		
pEGTy3-1/2	0	0 (0)
pEGTy3-2/1	10	45 (9.0)
Assay B		
pEGTy3-1	10	54 (10.8)
pEGTy3-2	0	0 (0)
pEGTy3-2(+A)	7	44 (8.8)

^a Number of colonies which were positive for Ty3 transposition (Tpn) per 10 original colonies.

^b Estimated number of Ty3 transposition-positive isolates per 500 clonal isolates of 10 original colonies.

cells having a single chromosomal copy of Ty3). Because this result was obtained by a relatively insensitive assay (at least one transposition per 500 colonies is required for detection), it is possible that Ty3-2 transposes, but at a greatly reduced frequency. The result of Ty3-2(+A) is seen in Fig. 4C. The addition of a single adenine into the sense strand conferred transposition activity on Ty3-2. Clonal derivatives of 7 of the 10 original pEGTy3-2(+A) transformants hybridized to the Ty3 probe, indicating that transposition had occurred. The results of this analysis are presented in Table 1 (assay B). Estimates of the number of transposed Ty3 elements per pool of 10 DNA samples were determined by densitometric scanning of the autoradiogram shown in Fig. 4.

Characterization of the Ty3 IN by Western blot analysis.

Antibodies were produced in rabbits against a synthetic peptide corresponding to a sequence in the predicted TYB3 protein which is close to the region that has homology to retroviral IN proteins. This region of the protein is predicted

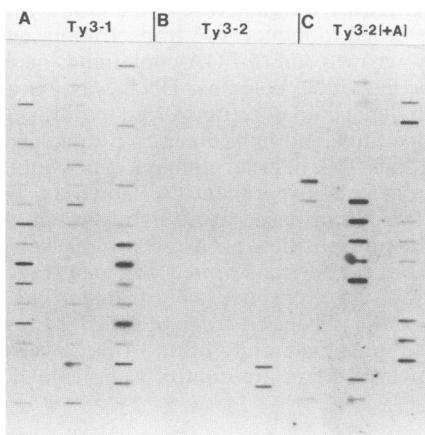


FIG. 4. Slot blot analysis of Ty3 transposition. Cells containing the plasmids pEGTy3-1 (A), pEGTy3-2 (B), and pEGTy3-2(+A) (C) were induced for Ty3 transcription (see Materials and Methods). DNA was isolated from pools of individual colonies, applied to nitrocellulose in a slot blot apparatus, and hybridized to a Ty3 internal domain-specific probe. For each of the 10 original induced colonies, 50 colonies were isolated; each slot reflects hybridization of 20 μ g of DNA from 10 isolated colonies. The bottom right region of each panel contains three positive controls followed by a negative control (yVB110) (one of the positive Ty3-2 controls did not hybridize well); the samples in the rightmost column of panel C are shifted up by one slot.

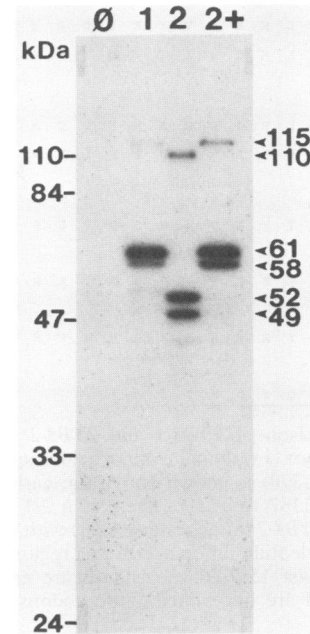


FIG. 5. Western blot analysis of Ty3 VLPs with anti-IN rabbit antiserum. VLPs were isolated from galactose-induced yeast cells containing pEGTy3-1 (1), pEGTy3-2 (2), or pEGTy3-2(+A) (2+), and an extract was prepared from the null strain alone (ϕ), as described in Materials and Methods. VLP proteins and the null-strain proteins were fractionated on a 10% sodium dodecyl sulfate-polyacrylamide gel, transferred to NitroScreen West membrane (Du Pont, NEN Research Products), and reacted with affinity-purified anti-IN antiserum. The antibody-protein complexes were visualized by incubation with ¹²⁵I-protein A followed by exposure to X-ray film.

to be antigenic on the basis of hydrophobicity plots (28; data not shown). When Ty3 is induced under control of the *GALI-10* upstream activating sequence, Ty3 VLPs are produced. Because the integrase protein is present in VLPs of other retrotransposons and in retroviral nucleocapsid cores (7, 17), concentrated VLPs prepared as described in Materials and Methods were utilized to identify the Ty3 IN protein by Western blot analysis. The result of this experiment carried out with Ty3-1- and Ty3-2-derived VLPs is shown in Fig. 5. For each element, there were two proteins against which the antipeptide serum reacted most strongly. These proteins were not found in similar extracts of Ty3 null cells (yVB110). The Ty3-1 IN proteins had molecular masses of approximately 61 and 58 kilodaltons (kDa), while the Ty3-2 IN proteins were approximately 52 and 49 kDa in size. Antiserum was raised against a second peptide, CQSER TIQTLNRLRAY, which was designed from a region in Ty3-1 (nucleotides 4224 to 4270) that encodes a protein with homology to retroviral IN proteins. This serum reacted to the same proteins identified with the first antibody but showed a higher level of nonspecific reactivity with proteins present in the null strain. The size difference of 9 kDa between the Ty3-1 and Ty3-2 proteins agrees with the computer-predicted 8.7-kDa truncation of the TYB3-2-encoded protein. As expected, addition of an adenine to Ty3-2 in the Ty3-2(+A) mutant results in production of IN proteins of the same size as those seen in Ty3-1 VLPs. The autoradiograph also revealed the presence of reactive proteins of approximately 115 and 110 kDa for Ty3-1 and Ty3-2, respectively. The difference in size between these species corre-

sponds roughly with the difference in length between the *TYB3-1* and *TYB3-2* ORFs. The position of the amino terminus of proteins of this size in the amino acid sequence predicted from *TYB3* is near the beginning of RT homology.

DISCUSSION

In this study we have determined the nucleotide sequence of an active Ty3 element, Ty3-1, and have identified the Ty3 IN protein by mutational and Western blot analyses. A comparison of the nucleotide sequence of Ty3-1 with that of Ty3-2, which is unable to transpose, revealed several nucleotide differences and a nucleotide deletion in the defective element. The presence of the region containing this deletion correlated with transpositional inactivity in chimeric Ty3 elements. Mutations in retroviral and other retrotransposon integrases have previously been shown to block integration (4, 13, 14, 43, 49). The single-nucleotide deletion in Ty3-2 was shown to be specifically responsible for the transposition defect by reversion of the phenotype with insertion mutagenesis of Ty3-2.

The mutation in Ty3-2 was useful in identification of the IN protein encoded by Ty3. Western blot analysis identified Ty3 IN proteins that are present in Ty3-1 and Ty3-2 VLPs. Comparison of the sequences of the proteins predicted to be encoded by *TYB3* in the two elements, together with Western blot analysis, showed that the deletion mutation in Ty3-2 results in a truncated IN protein. The size of the wild-type and mutant Ty3 IN proteins allowed us to infer which region encodes Ty3 IN. Ty3-1 and Ty3-2 each produced pairs of proteins: 61 and 58 kDa and 52 and 49 kDa, respectively. The predicted amino-terminal half of the Ty3 IN protein has homology to retroviral IN proteins. Characterized retroviral IN proteins are smaller than the Ty3 IN proteins. For example, the molecular masses of the Rous sarcoma virus, human immunodeficiency virus, and murine leukemia virus proteins are 32, 34, and 46 kDa, respectively (57). This difference in size appears to be largely accounted for by the longer carboxyl-terminal region of the Ty3 IN protein, which we have shown is required for activity.

The second ORF of the defective Ty3 predicts a protein shorter than the protein encoded by the active element by about 9 kDa. The fact that the predicted difference at the carboxyl-terminal end corresponds closely to the difference in molecular weight of the observed IN proteins from the two elements suggests that carboxyl-terminal processing of the wild-type element, if it occurs, does not result in a significant change in molecular weight. In addition, the IN amino-terminal coding regions of the two elements are highly similar. The fact that both Ty3 elements studied encode a pair of IN proteins differing by 3 kDa suggests that in each case the smaller protein is derived from the larger by removal of a 3-kDa peptide from the amino terminus. The presence of a 115-kDa IN-related protein in the Ty3-1 VLP suggests that the Ty3 IN also exists as part of a polyprotein that includes RT, which is encoded upstream of IN. This protein could be a precursor form of IN, or it could be analogous to the Rous sarcoma virus β component of the $\alpha\beta$ heterodimer form of RT (11, 16, 21). If the latter possibility is the case, it would indicate that the size of the Ty3 RT domain is about 55 kDa. In the Western blot analysis, the 115-kDa protein reacted much less intensely to the antisera than did IN, showing that it is present in lower amounts than is IN in the VLP. The function of such a composite protein remains to be elucidated.

Preferred positions have been described for retroviral

integration in regions which contain DNase I-hypersensitive sites (44, 58) and at specific sequences which may be associated with nuclear matrix proteins (53). Nevertheless, the specific features of these regions which cause preferential insertion have not been distinguished from among a plethora of possibilities. While it has been shown that the retroviral IN protein is required for integration, it has not been specifically shown that this protein is required for target cleavage. Similarly, a central role for this protein in target site selection has not been demonstrated. In contrast to the diverse features of retrovirus insertion sites, tRNA genes are common to all characterized Ty3 insertions. This specificity should simplify *in vivo* identification of Ty3 integration target determinants. Furthermore, if IN cleaves the target, a simple model would predict that Ty3 IN protein is also capable of specific interaction with tRNA genes. Testing the hypothesis that Ty3 IN nicking of the target directs position-specific insertion should contribute to a more general understanding of the role of the IN protein in retrotransposition.

ACKNOWLEDGMENTS

We thank D. Chalker and C. Connolly for assistance with the experiments.

This work was supported by Public Health Service grant GM33281 from the National Institutes of Health. L.J.H. was supported by predoctoral training grant GM07134 from the National Institutes of Health.

LITERATURE CITED

- Adams, C. W., and G. W. Hatfield. 1984. Effects of promoter strengths and growth conditions on copy number of transcription-fusion vectors. *J. Biol. Chem.* **259**:7399-7403.
- Bingham, P. M., and Z. Zachar. 1989. Retrotransposons and the FB transposon from *Drosophila melanogaster*, p. 485-502. In D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
- Boeke, J. D. 1989. Transposable elements in *Saccharomyces cerevisiae*, p. 335-374. In D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
- Boeke, J. D., D. 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.
- 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., F. LaCroute, and G. R. Fink. 1984. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol. Gen. Genet.* **197**:345-346.
- Bowerman, B., P. O. Brown, J. M. Bishop, and H. E. Varmus. 1989. A nucleoprotein complex mediates the integration of retroviral DNA. *Genes Dev.* **3**:469-478.
- Brodeur, G. M., S. B. Sandmeyer, and M. V. Olson. 1983. Consistent association between sigma elements and tRNA genes in yeast. *Proc. Natl. Acad. Sci. USA* **80**:3292-3296.
- Clare, J., and P. Farabaugh. 1985. Nucleotide sequence of a yeast Ty element: evidence for an unusual mechanism of gene expression. *Proc. Natl. Acad. Sci. USA* **82**:2829-2833.
- Clark, D. J., V. W. Bilanchone, L. J. Haywood, S. L. Dildine, and S. B. Sandmeyer. 1988. A yeast sigma composite element, TY3, has properties of a retrotransposon. *J. Biol. Chem.* **263**:1413-1423.
- Copeland, T. D., D. P. Grandgenett, and S. Oroszlan. 1980. Amino acid sequence analysis of reverse transcriptase subunits from avian myeloblastosis virus. *J. Virol.* **36**:115-119.
- Del Rey, F. J., T. F. Donahue, and G. R. Fink. 1982. Sigma, a repetitive element found adjacent to tRNA genes of yeast. *Proc.*

- Natl. Acad. Sci. USA 79:4138-4142.
13. Donehower, L. A. 1988. Analysis of mutant Moloney murine leukemia viruses containing linker insertion mutations in the 3' region of *pol*. J. Virol. 62:3958-3964.
 14. Donehower, L. A., and H. E. Varmus. 1984. A mutant murine leukemia virus with a single missense codon in *pol* is defective in a function affecting integration. Proc. Natl. Acad. Sci. USA 81:6461-6465.
 15. Doolittle, R. F., D.-F. Feng, M. S. Johnson, and M. A. McClure. 1989. Origins and evolutionary relationships of retroviruses. Q. Rev. Biol. 64:1-30.
 16. Duyk, G., J. Leis, M. Longiaru, and A. M. Skalka. 1983. Selective cleavage in the avian retroviral long terminal repeat sequence by the endonuclease associated with the $\alpha\beta$ form of avian reverse transcriptase. Proc. Natl. Acad. Sci. USA 80:6745-6749.
 17. 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.
 18. Emori, Y., S. Shiba, S. Inouye, S. Yuki, and K. Saigo. 1985. The nucleotide sequences of copia and copia-related RNA in *Drosophila* virus-like particles. Nature (London) 315:773-776.
 19. Garfinkel, D. J., J. D. Boeke, and G. R. Fink. 1985. Ty element transposition: reverse transcriptase and virus-like particles. Cell 42:507-517.
 20. Gentry, L. E., L. R. Rohrschneider, J. E. Casnellie, and E. G. Krebs. 1983. Antibodies to a defined region of pp60^{src} neutralize the tyrosine-specific kinase activity. J. Biol. Chem. 258:11219-11228.
 21. Gibson, W., and I. Verma. 1974. Studies on the reverse transcriptase RNA tumor viruses. Structural relatedness of two subunits of avian RNA tumor viruses. Proc. Natl. Acad. Sci. USA 71:4991-4994.
 22. Goding, J. W. 1978. Use of staphylococcal protein A as an immunological reagent. J. Immunol. Methods 20:241-253.
 23. Grandbastien, M. A., A. Spielmann, and M. Caboche. 1989. Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature (London) 337:376-380.
 24. Haltiner, M., T. Kempe, and R. Tjian. 1985. A novel strategy for constructing clustered point mutations. Nucleic Acids Res. 13:1015-1025.
 25. Hansen, L. J., D. L. Chalker, and S. B. Sandmeyer. 1988. Ty3, a yeast retrotransposon associated with tRNA genes, has homology to animal retroviruses. Mol. Cell. Biol. 8:5245-5256.
 26. Hauber, J., P. Nelbock-Hochstetter, and H. Feldmann. 1985. Nucleotide sequence and characteristics of a Ty element from yeast. Nucleic Acids Res. 13:2745-2758.
 27. Holmes, D. S., and M. Quigley. 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114:193-197.
 28. Hopp, T. P., and K. R. Woods. 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA 78:3824-3828.
 29. Inouye, S., S. Yuki, and K. Saigo. 1986. Complete nucleotide sequence and genome organization of a *Drosophila* transposable genetic element, 297. Eur. J. Biochem. 154:417-425.
 30. Ito, H., Y. Fukuda, K. Murata, and A. Kimura. 1983. Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. 153:163-168.
 31. 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.
 32. Johnston, M., and R. W. Davis. 1984. Sequences that regulate the divergent *GALI-GAL10* promoter in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 4:1440-1448.
 33. Kitagawa, T., and T. Aikawa. 1976. Enzyme coupled immunoassay of insulin using a novel coupling reagent. J. Biochem. 79:233-236.
 34. Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA 82:488-492.
 35. Kyhse-Andersen, J. 1984. Electrophoretic blotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. J. Biochem. Biophys. Methods 10:203-209.
 36. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680-685.
 37. Liu, F.-T., M. Zinnecker, T. Hamakoa, and D. H. Katz. 1979. New procedures for preparation and isolation of conjugates of proteins and a synthetic copolymer of D-amino acids and immunochemical characterization of such conjugates. Biochemistry 18:690-697.
 38. Marlor, R. L., S. M. Parkhurst, and V. G. Corces. 1986. The *Drosophila melanogaster* gypsy transposable element encodes putative gene products homologous to retroviral proteins. Mol. Cell. Biol. 6:1129-1134.
 39. Marschalek, R., T. Brechner, E. Amon-Bohm, and T. Dinger-mann. 1989. Transfer RNA genes: landmarks for integration of mobile genetic elements in *Dictyostelium discoideum*. Science 244:1493-1496.
 40. Mellor, J., M. H. Malim, K. Gull, M. F. Tuite, S. McCreedy, 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.
 41. Mount, S. M., and G. M. Rubin. 1985. Complete nucleotide sequence of the *Drosophila* transposable element copia: homology between copia and retroviral proteins. Mol. Cell. Biol. 5:1630-1638.
 42. Muller, 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.
 43. Panganiban, A. T., and H. M. Temin. 1984. The retrovirus *pol* gene encodes a product required for DNA integration: identification of a retrovirus *int* locus. Proc. Natl. Acad. Sci. USA 81:7885-7889.
 44. Robinson, H. L., and G. C. Gagnon. 1986. Patterns of proviral insertion and deletion in avian leukosis virus-induced lymphomas. J. Virol. 57:28-36.
 45. Saigo, K., W. Kugimiya, Y. Matsuo, S. Inouye, K. Yoshioka, and S. Yuki. 1984. Identification of the coding sequence for a reverse transcriptase-like enzyme in a transposable genetic element in *Drosophila melanogaster*. Nature (London) 312:659-661.
 46. Sandmeyer, S. B., V. W. Bilanchone, D. J. Clark, P. Morcos, G. F. Carle, and G. M. Brodeur. 1988. Sigma elements are position-specific for many different yeast tRNA genes. Nucleic Acids Res. 16:1499-1515.
 47. Sandmeyer, S. B., and M. V. Olson. 1982. Insertion of a repetitive element at the same position in the 5'-flanking regions of two dissimilar yeast tRNA genes. Proc. Natl. Acad. Sci. USA 79:7674-7678.
 48. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
 49. Schwartzberg, P., J. Colicelli, and S. P. Goff. 1984. Construction and analysis of deletion mutations in the *pol* gene of murine leukemia virus: a new viral function required for productive infection. Cell 37:1043-1052.
 50. Shehee, W. R., S.-F. Chao, D. D. Loeb, M. B. Comer, C. A. Hutchison III, and M. H. Edgell. 1987. Determination of a functional ancestral sequence and definition of the 5' end of A-type mouse L1 elements. J. Mol. Biol. 196:757-767.
 51. Sherman, F., G. R. Fink, and J. B. Hicks. 1986. Methods in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 52. Shiba, T., and K. Saigo. 1983. Retrovirus-like particles containing RNA homologous to the transposable element copia in *Drosophila melanogaster*. Nature (London) 302:119-124.
 53. Shih, C.-C., J. P. Stoye, and J. M. Coffin. 1988. Highly preferred targets for retrovirus integration. Cell 53:531-537.
 54. Smyth, D. R., P. Kalitsis, J. L. Joseph, and J. W. SENTRY. 1989. Plant retrotransposon from *Lilium henryi* is related to Ty3 of yeast and the gypsy group of *Drosophila*. Proc. Natl. Acad. Sci.

- USA 86:5015-5019.
55. **Stucka, R., H. Lochmuller, and H. Feldmann.** 1989. Ty4, a novel low-copy number element in *Saccharomyces cerevisiae*: one copy is located in a cluster of Ty elements and tRNA genes. *Nucleic Acids Res.* 17:4993-5001.
 56. **Toh, H., R. Kikuno, H. Hayashida, T. Miyata, W. Kugimiya, S. Inouye, S. Yuki, and K. Saigo.** 1985. Close structural resemblance between putative polymerase of a *Drosophila* transposable genetic element 17.6 and *pol* gene product of Moloney murine leukaemia virus. *EMBO J.* 4:1267-1272.
 57. **Varmus, H., and P. Brown.** 1989. Retroviruses, p. 53-108. *In* D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
 58. **Vijaya, S., D. L. Steffen, and H. L. Robinson.** 1986. Acceptor sites for retroviral integrations map near DNase I-hypersensitive sites in chromatin. *J. Virol.* 60:683-692.
 59. **Voytas, D. F., and F. M. Ausubel.** 1988. A copia-like transposable element family in *Arabidopsis thaliana*. *Nature (London)* 336:242-244.
 60. **Warmington, J. R., R. B. Waring, C. S. Newlon, K. J. Indge, and S. G. Oliver.** 1985. Nucleotide sequence characterization of Ty 1-17, a class II transposon from yeast. *Nucleic Acids Res.* 13:6679-6693.
 61. **Yuki, S., S. Inouye, S. Ishimaru, and K. Saigo.** 1986. Nucleotide sequence characterization of a *Drosophila* retrotransposon, 412. *Eur. J. Biochem.* 158:403-410.