Functional Organization of the Retrotransposon Ty from *Saccharomyces cerevisiae*: Ty Protease Is Required for Transposition

SUSAN D. YOUNGREN, JEF D. BOEKE, NANCY J. SANDERS, and DAVID J. GARFINKEL

Basic Research Program, Bioanetics Research, Inc., National Cancer Institute-Frederick Cancer Research Facility, Frederick, Maryland 21701, and Department of Molecular Biology and Genetics, Johns Hopkins School of Medicine, Baltimore, Maryland 21205

Received 25 August 1987/ Accepted 18 December 1987

We used several mutations generated in vitro to further characterize the functions of the products encoded by the *TyB* gene of the transpositionally active retrotransposon TyH3 from *Saccharomyces cerevisiae*. Mutations close to a core protein domain of *TyB*, which is homologous to retroviral proteases, have striking effects on Ty protein processing, the physiology of Ty viruslike particles, and transposition. The Ty protease is required for processing of both *TyA* and *TyB* proteins. Mutations in the protease resulted in the synthesis of morphologically and functionally aberrant Ty viruslike particles. The mutant particles displayed reverse transcriptase activity, but did not synthesize Ty DNA in vitro. Ty RNA was present in the mutant particles, but at very low levels. Transposition of a genetically tagged element ceased when the protease domain was mutated, demonstrating that Ty protease is essential for transposition. One of these mutations also defined a segment of *TyB* encoding an active reverse transcriptase. These results indicate that the Ty protease, like its retroviral counterpart, plays an important role in particle assembly, replication, and transposition of these elements.

Yeast Ty elements make up a family of retrotransposons found in *Saccharomyces cerevisiae* (18, 43). Most Ty elements are approximately 6 kilobases (kb) long and are similar to retrovirial proviruses in structure. Terminal repeated delta sequences, or long terminal repeats (LTRs), of 335 nucleotides bracket an internal region of 5.3 kb. Ty elements are transcribed from LTR to LTR, forming a 5.7-kb RNA that is terminally redundant for 45 nucleotides (15). This structure is similar to retroviral RNA. Ty elements contain a sequence just inside the 5′ LTR that can serve as a binding site for a tRNA^Met^ (14) and an oligopurine tract near the 3′ LTR. Both sequence motifs are important for priming retroviral reverse transcription (21). Moreover, the organization of Ty elements is similar to that of the gag-pol region of retroviruses (Fig. 1A). In Ty elements, there are two overlapping genes: *TyA*, which is equivalent to gag and specifies a protein that has homology to DNA-binding proteins (11, 24, 51); and *TyB*, which is equivalent to pol and specifies a protein that has limited homology to retroviral protease, integrase, and reverse transcriptase (11, 24, 47, 48, 51). The gene product of *TyB* is thought to be synthesized as a TyA-TyB fusion protein resulting from a specific frameshift event at the end of *TyA* that puts *TyB* in frame with *TyA* (11, 35, 36).

A system for studying Ty transposition has been developed in which a genetically tagged element, TyH3, is fused to the controllable GAL1 promoter of *S. cerevisiae* on a high-copy-number plasmid (referred to as pGTYH3) (51). Upon galactose induction, the frequency of transposition of both the marked element and genomic Ty elements dramatically increases (a condition referred to as transposition-induction). Using this system, we found that transposition of Ty elements proceeds via an RNA intermediate and resembles the process of retroviral reverse transcription and integration.

The evolutionary relatedness of Ty elements and retroviruses became even more evident when Ty viruslike particles (Ty-VLPs) accumulated in cells induced for transposition (20, 37). The cytoplasm of transposition-induced cells is filled with spherical particles of approximately 60 nm in diameter. The cofractionation of Ty-VLPs, reverse transcriptase activity, a TyB-encoded protein, and genome-length Ty RNA, as well as the ability of this complex to synthesize Ty DNA, suggests that reverse transcription takes place in the particle and that Ty-VLPs are transposition intermediates (20, 37). Recently, the major structural proteins of Ty-VLPs have been shown to be derived from the primary product of TyA by proteolytic cleavage (1, 39). At least some of these cleavages appear to be mediated by the product of the protease domain of TyB. Deletion of residues that are homologous to retroviral proteases results in the accumulation of unprocessed TyA, the putative TyATyB polypeptide, and morphologically altered Ty-VLPs.

In addition to Ty-encoded products, there are host genes that are important for Ty transposition. The best example is the SPT3 gene, which affects the length and abundance of Ty transcripts (52, 53). *sp3-101* mutants have a small amount of a truncated Ty transcript lacking approximately 800 nucleotides at the 5′ end. The *SPT3* gene is required for transposition of endogenous Ty elements, but transposition of GAL1-promoted Ty elements is unaffected (7). This mutation presumably blocks the transposition of endogenous elements by destroying the ability of Ty RNA to act as an mRNA and as a template for reverse transcription.

The ability to regulate transposition and to mark a specific Ty to monitor its movement makes possible the genetic and biochemical analysis of retrotransposition. To map the position and define the biological functions of proteins encoded by a transpositionally competent Ty, we have begun to analyze a series of linker mutations within TyH3. In the work described in this paper we investigated the effects of Ty protease domain mutations on transposition, reverse
transcription, and Ty-VLP physiology. To minimize the effect of possible genetic complementation between TyH3 mutants and the gene products from the ensemble of endogenous chromosomal Ty elements, we analyzed TyH3 mutants in an spt3 mutant background. The data indicate that correct Ty protein processing is essential for retrotransposition and for normal particle formation. Protease mutants still display a VLP-associated reverse transcriptase activity, but the particles cannot synthesize Ty DNA in vitro. We also have characterized a mutation that helps define the coding region for Ty reverse transcriptase.

**MATERIALS AND METHODS**

Yeast strains, general genetic methods, and media. The yeast strains used in this study are described below and in Table 1. Strain DG531 is an isogenic spt3-101 derivative of strain BWG1-7A (MATa ade1-100 ura3-52 leu2-3,112 his4-519 GAL+) (23) which was constructed by transformation of the spt3-101 integrating plasmid pFW33 (kindly provided by F. Winston) into strain BWG1-7A. spt3-101 is a frameshift mutation in the SPT3 gene that is generated in vitro (52). Transplacement of the spt3-101 mutation was selected from a pFW33 integrant on 5-fluoroorotic acid medium (6). spt3-101 mutants were detected by a mating test with SPT3 and spt3 tester strains and confirmed by measuring the size of Ty RNA. spt3 × spt3 crosses give rise to a low frequency of diploids, and the SPT3 gene is required for formation of a full-sized Ty transcript (52). All yeast media were those described by Sherman et al. (46).

Nucleic acid manipulations. The molecular linker insertions were generated by the methods of Boeke (3) and Barany (2). In general, plasmid pGTyH3, or a marker derivative, was partially digested with a restriction endonuclease under conditions that produced a single cleavage. In some cases, the ends were filled in by using the Klenow fragment of DNA polymerase, and the full-length linear DNA was isolated after electrophoresis on a 0.35% or 0.5% agarose gel. Unphosphorylated linkers were ligated to the purified linear DNAs, and the products were processed by using the linker-tailing technique of Lathe et al. (31). The following linkers were used: a SacI 8-mer (dCGAGCTTCG) (New England BioLabs, Inc., Beverly, Mass.), or a MluI 12-mer (dTGCACGCCTGCA) (Biopolymers Laboratory, The Johns Hopkins University School of Medicine, Baltimore, Md.). The Sau3a site (at nucleotide 3087) was converted to a SacI site by partially digesting pGTyH3 with Sau3a and then ligating the partial digest with a BamHI fragment that contains the Km’ cassette from pUC4-KIXX (2). Subsequent loss of the cassette by digestion with SacI and recircularization of pGTyH3 generated an in-frame 4-codon insertion. Plasmids containing linker insertions were initially mapped with restriction enzymes to check for deletions and to map the site of insertion. To verify the position of the linker, all but two of the insertions (SacI at nucleotide 3087 and SacI at nucleotide 3301) described in Table 2 were sequenced (44) and compared with the sequence of TyH3 (4). The set of pGTyH3 linker mutations marked with the NEO gene was constructed by inserting a restriction fragment containing NEO from pSFL105 (pGTyH3NEO) and is diagrammed in Fig. 1B. A more detailed characterization of the other linker mutations listed in Table 2 will appear separately (J. Boeke and G. Monikian, manuscript in preparation). The GAL1-lacZ fusion plasmid pCGS286 (pGTyGAL1-lacZ) was kindly provided by J. Mao. Rapid plasmid DNA isolation, plasmid transfor-
mation of *Escherichia coli*, restriction enzyme analysis, gel electrophoresis, and Southern hybridization analyses were as described by Maniatis et al. (33). Plasmids were introduced into *S. cerevisiae* by the lithium acetate transformation procedure of Ito et al. (26). Total RNA or DNA was prepared from yeast cells essentially as described by Sherman et al. (46), although in some instances yeast DNA was isolated by the method of Holm et al. (25). RNA was prepared from subcellular fractions by phenol extraction followed by ethanol precipitation. RNA was denatured with glyoxal (34), separated by agarose gel electrophoresis, and transferred to Hybond N as recommended by the supplier (Amersham Corp., Arlington Heights, Ill.). 32P-labeled hybridization probes were made by randomly primed DNA synthesis of purified restriction fragments (16, 17). Autoradiograms were scanned by using a Quick Scan Jr. densitometer (Helena Laboratories, Beaumont, Tex.). The EcoRI fragment that contains lacO was labeled by filling in with [α-32P]dATP and the Klenow fragment of DNA polymerase as described previously (5).

**Isolation of Ty-VLPs.** The VLPs were isolated essentially as described previously (20) with the following modifications. The yeast cells were grown overnight at 30°C in SC-Ura synthetic complete medium with 5% raffinose and then diluted into SC-Ura containing 2% galactose and grown for 18 h at 22°C. Yeast spheroplasts were made and homogenized, and the large-cell debris were removed by centrifugation as described previously. The S10 supernatant from this spin was layered onto a sucrose step gradient consisting of 20%, 30%, and 75% layers in a buffer composed of 50 mM KCl, 50 mM HEPES (N-2-hydroxyethylpiperazine-N’-2-ethanesulfonic acid; pH 7.8), 5 mM EDTA (pH 8), and 3 mM dithiotreitol (D. Eichinger, personal communication). The gradients were centrifuged in an SW28 rotor (Beckman Instruments, Inc., Fullerton, Calif.) for 3 h at 25,000 rpm and then fractionated in a fractionator (no. 640; Isco, Lincoln, Nebr.). The Ty-VLPs typically collected at the boundary between the 30% and 75% layers. In some experiments the Ty-VLPs present in the S10 supernatant were concentrated by centrifugation in a Ti50 rotor (Beckman) at 44,000 rpm for 1 h, resulting in a P40 pellet. Fractions were assayed for reverse transcriptase activity as described below. Protein concentrations were determined by the Bradford assay (8). All the fractions were stored on ice at 4°C or −20°C.

**Reverse transcriptase assay.** The exogenous and endogenous reactions were performed essentially as described by Garfinkel et al. (20), except that the endogenous assays were run with 3 mM dithiotreitol and [α-32P]dGTP (3,000 Ci/mmol; Amersham).

**Preparation of antisera.** Peptide 136 is derived from the N-terminal sequence of TyA (Fig. 1C) and contains the following amino acids (encoded by nucleotides 393 to 419): Asp-Val-Ser-Ala-Ser-Lys-Thr-Glu-Glu-Cys. Peptide 136 was synthesized by A. Huang and J. Young (Division of Infectious Disease, Children’s Hospital, Boston, Mass.) and coupled via the cysteine to keyhole limpet hemocyanin or bovine serum albumin by using *m*-maleimidobenzoyl-N-succinimide ester (32). Antipeptide antibodies were raised in male New Zealand White rabbits and purified by affinity chromatography.

### TABLE 1. Yeast strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Plasmid</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>JB183</td>
<td>MATα <em>ura3-52 his3Δ200 trp1-289 lys2 GAL</em></td>
<td>pAB100*</td>
<td>5</td>
</tr>
<tr>
<td>FW665</td>
<td>MATα lys2-173R his4-9176 *ura3-52 sp3-101</td>
<td>pGTyH3lacZ</td>
<td>This work</td>
</tr>
<tr>
<td>FW689</td>
<td>MATα lys2-173 his4-916 *ura3-52 sp3-101</td>
<td>pGTyH3lacO</td>
<td>This work</td>
</tr>
<tr>
<td>JB314</td>
<td>Mata adel-100 <em>ura3-52 leu2-3,112 his4-519 GAL</em></td>
<td>pGTyH3neoO</td>
<td>This work</td>
</tr>
<tr>
<td>DG531</td>
<td>Mata adel-100 <em>ura3-52 leu2-3,112 his4-519 sp3-101 GAL</em></td>
<td>pGTyH3neo::MluI-1876</td>
<td>This work</td>
</tr>
<tr>
<td>DG681</td>
<td>DG531</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DG692</td>
<td>DG531</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DG714</td>
<td>DG531</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DG716</td>
<td>DG531</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DG718</td>
<td>DG531</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DG720</td>
<td>DG531</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YH8</td>
<td>MATα his3Δ200 <em>ura3-167 leu2Δ1 trp1Δ1 GAL</em></td>
<td>pGTyH3lacO</td>
<td>H. Xu, unpublished</td>
</tr>
<tr>
<td>GM140</td>
<td>YH8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM36</td>
<td>YH8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* a pAB100 is a TRPl-based centromere plasmid containing a promoterless HIS3 gene (5).
* b pCGS286.
* c pJefl105.
* d pJefl269.
* e pGMO17.

### TABLE 2. Effects of linker insertions on TyA and TyB proteins

<table>
<thead>
<tr>
<th>Insertion</th>
<th>Original restriction site</th>
<th>Ty-VLPs (TyA)*</th>
<th>Peptide 136 (TyA)*</th>
<th>TyB2 (TyB)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>M142-1210 (TyA)</td>
<td>SspI</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>M142-1443 (TyA)</td>
<td>AluI</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>M142-1615 (TyB)</td>
<td>Hinfl</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>SacI-1702 (TyB)</td>
<td>BglII</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M16-1876 (TyB)</td>
<td>Ddel</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M16-2975 (TyB)</td>
<td>Hinfl</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SacI-3087 (TyB)</td>
<td>Sau3a</td>
<td>NT*</td>
<td>NT</td>
<td>+</td>
</tr>
<tr>
<td>M16-3241 (TyB)</td>
<td>SspI</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>SacI-3301 (TyB)</td>
<td>BglII</td>
<td>NT</td>
<td>NT</td>
<td>+</td>
</tr>
<tr>
<td>M16-3800 (TyB)</td>
<td>Hinfl</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pGTyH3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

* a Total yeast proteins were prepared from transposition-induced cells and immunoblotted as described in the legend to Fig. 2 and Materials and Methods.
* b The molecular linker used to generate the mutation is designated to the left of the dash, and its nucleotide position is designated to the right. The nucleotide position refers to that of the original restriction site. Further details are described in the text. All mutations were made in pGTyH3.
* c The protein patterns obtained from the Western blots were scored (+) if they resembled those of the wild type and (−) if they did not. Antibodies were directed specifically against TyA (Ty-VLPs and peptide 136) or TyB (TyB2) proteins.
* d All of the linker insertions were verified by DNA sequencing, except for the SacI-3087 and SacI-3301 insertions.
* e NT. Not tested.
chromatography, using established techniques (22). Anti-VLP antibodies raised against Ty1-15-encoded particles were a generous gift from A. J. and S. M. Kingsman. The anti-TyB2 antibody was raised in rabbits to a fusion protein isolated from *E. coli*. Briefly, the region from the EcoRV site at nucleotide 2779 to the PvuII site at nucleotide 3944 of TyH3 (diagrammed in Fig. 1C) was cloned into the Path 2 expression vector (13), and the resulting *trp*E-*TyB* fusion protein was induced with indoleacrylic acid. Protein extracts from the induced cells were separated in sodium dodecyl sulfate (SDS)-10% polyacrylamide gels (30), the gels were stained with Coomasie blue, and the resulting fusion protein band was excised from each gel. The gel slices were used to immunize rabbits, and the sera were obtained by standard techniques (50). The antibodies were purified from whole sera by affinity chromatography over a bed of protein A-agarose.

**Western blot analysis.** Western immunoblots were run by the procedures developed by Towbin et al. (49). Total yeast proteins and Ty-VLPs were separated on SDS-polyacrylamide gels, with acrylamide concentrations ranging from 7.5% to 12.5%. The proteins were transferred to nitrocellulose filters and screened with the antisera described above. Antibodies were added in excess, as determined by titration experiments. Antibody-antigen complexes were detected with 125I-labeled protein A (Amersham). Total yeast cell proteins were isolated by trichloroacetic acid extraction as described by Ohashi et al. (40).

**Electron microscopy.** Standard procedures for preparing thin sections of yeast cells for electron microscopy were used (10).

**Transposition assay.** The transposition assay involving the use of TyH3 marked with the *NEO* gene will be described elsewhere in greater detail (7a). Briefly, yeast cells containing the relevant *URA3*-based plasmid were induced for transposition by growth for 5 days at 22°C on SC-Ura plates containing galactose at a final concentration of 2%. Single colonies were restreaked on SC-Ura plates containing glucose (the *GAL1* promoter is strongly catabolite repressed in the presence of glucose) and then grown nonselectively on YPD plates to allow for loss of the plasmid. Ura− segregants were identified by replica plating to SC-Ura plates or to medium containing 5-fluoroorotic acid (6). Transposition events were detected by the ability of the Ura− segregants to grow on YPD plates that contain the antibiotic G418 ( Gibco Laboratories, Grand Island, N.Y.) at a final concentration of 200 μg/ml. This concentration of G418 was chosen because it is just above the minimum concentration required to inhibit the growth of strain DG531. Transpositions were scored after incubation for 2 days at 30°C. The transposition efficiency is defined as the number of G418− Ura− segregants divided by the total number of Ura− segregants. The transposition assay involving the use of TyH3 marker with the *lacO* sequence was very similar to the one described above, except that the Ura− segregants were analyzed by Southern filter hybridizations to detect transposition events (5). The transposition efficiency is defined as the number of Ura− segregants whose genomes contain additional sequences that hybridize with *lacO*− or Ty−specific probes divided by total number of Ura− segregants.

**RESULTS**

**Mutational analysis of TyH3.** We exploited two recent advances in the study of Ty elements to understand the biochemistry and genetics of retrotransposition. The first is a major refinement of the transposition assay described previously (5). Functional Ty elements tagged with selectable genes also retain their ability to transpose normally (7a). One such construction contains the *GAL1*-promoted TyH3 marked with the aminoglycoside phosphotransferase gene from Tn903 (41) and is referred to as pGTyH3NEO. The *NEO* gene confers dominant resistance to the antibiotic G418 in *S. cerevisiae* (27) and to neomycin and kanamycin in *E. coli*. Transposition of TyH3NEO results in *S. cerevisiae* cells that remain resistant to G418 after the loss of the pGTyH3NEO plasmid. When the *GAL1* promoter is induced, TyH3NEO transposes at levels comparable to elements marked with other sequences. Large numbers of cells which suffered transposition events can now be analyzed by a plate assay. This assay does not estimate the actual number of marked transpositions per cell, because phenotypic resistance to G418 can result from a single TyH3NEO transposition. Therefore, cells containing single or multiple transpositions will both be scored as positive. Expression of the *NEO* gene within a Ty is independent of orientation, and the *NEO* gene is expressed when placed in several different Ty elements (M. Curcio and D. Garfinkel, unpublished results).

The second advance is the analysis of Ty elements in an *spt3-101* mutant background. This minimizes the effect of possible genetic complementation between TyH3 mutants and the gene products from the ensemble of chromosomal endogenous elements. Small amounts of Ty*A* proteins were detected in *SPT3* yeast strains (1, 39), but no Ty*A* proteins were detected in the *spt3-101* mutants (Fig. 2). As expected, *GAL1*-promoted expression of TyH3 in an *spt*3 mutant background restored the synthesis of Ty proteins to a level that was comparable to that of identical constructions in *SPT3* strains. Shut-off of *GAL1* expression on glucose also occurred normally in the *spt3-101* mutant.

**TyA and TyB proteins encoded by TyH3.** Total protein extracts or subcellular fractions enriched for Ty-VLPs were prepared from transposition-induced strain DG692 expressing pGTyH3NEO and analyzed by Western blot procedures to detect Ty*A* and Ty*B* proteins. We used antibodies directed against specific regions of Ty*A* and Ty*B* (Fig. 1C; see Materials and Methods). Antibodies raised against purified Ty-VLPs reacted with several proteins encoded by the Ty*A* gene. Two polypeptides of 38 and 41 kilodaltons (kDa) (p38-*TyA* and p41-*TyA*) isolated from TyH3-VLPs strongly reacted with anti-Ty-VLP antibodies, whereas several other proteins, ranging in size from approximately 13 to 58 kDa, reacted with lower intensities (Fig. 3A and C, part i, panel VLP, lane 4). In contrast, p54-*TyA* and p58-*TyA* from total cell extracts strongly reacted with anti-Ty-VLP antibodies (Fig. 3B and C, part ii, panel VLP, lane 4). Longer autoradiography revealed weak signals for a few of the smaller proteins. Identical results were obtained when Ty-VLPs proteins and total cellular proteins were analyzed on the same gel (Fig. 3C, lanes 4). Antipeptide antibodies directed against the C terminus of Ty*A* also showed that although p54-*TyA* and p58-*TyA* are the major Ty*A* proteins present in total cell extracts, these larger proteins are not detected in TyH3-VLPs (data not shown). The difference in the extent of Ty*A* processing between total cell extracts and purified VLPs may result from the presence of a large pool of nonparticulate, unprocessed Ty*A* proteins.

To investigate the protein products of Ty*B*, we raised antibodies against a *trp*E-*TyB* fusion protein (TyB2) containing the Ty*B* coding sequence from an EcoRV site at nucleotide 2779 to a PvuII site at nucleotide 3944 (Fig. 1C). This
Two mutants were identified that had altered mobilities of both TyA and TyB polypeptides. Mutant strain DG716 contained a SacI linker inserted at a BglII site at nucleotide 1702. This site is 1 nucleotide 3' from the core protease homology domain mentioned above. Mutant strain DG714 contained an MluI linker inserted at a DdeI site at nucleotide 1876. These in-phase mutations resulted from the addition of 4 or 5 codons to the TyB open reading frame (Fig. 1B).

Total yeast cell protein and fractions enriched for Ty-VLPs from the protease mutant strain DG714 and DG716 were analyzed by using the same antiseras as described above. The GALI-promoted expression of Ty elements containing protease mutations decisively altered Ty protein processing (Fig. 3A and B, lanes 2 and 3) when compared with the wild-type TyH3 element (Fig. 3A and B, lanes 4). In each instance, larger proteins were observed. It was especially evident for TyA proteins present in total cell extracts. In protease mutants, p58-TyA strongly reacted with anti-VLP antibodies, whereas in the wild-type, p54-TyA was the dominant immunoreactive protein (Fig. 3B, panel VLP, lanes 2 to 4). For the VLP-associated TyA proteins, the general protein pattern was the same, but the proteins were larger (Fig. 3A, panel VLP, lanes 2 and 3), suggesting that several correct processing events occurred, but others were aberrant. There was also a 190-kDa protein in the protease mutants that weakly reacted with anti-VLP antibodies (Fig. 3A and B, panel VLP, lanes 2 and 3). For TyB, novel proteins of 110, 160, and 190 kDa accumulated, whereas the wild-type p90-TyB protein was not detected in the protease mutants (Fig. 3A and B, panel B2, lanes 2 and 3).

A novel 190-kDa protein accumulated in protease mutants. This protein is probably the TyA-TyB polypeptide because both TyA (Fig. 3A and B, panel VLP, lanes 2 and 3) and TyB (Fig. 3A and B, panel B2, lanes 2 and 3) specific antibodies reacted with a precursor of the size (193 kDa) predicted from the coding potential of TyH3.

The protein processing observed in the mutants was detected only in fractions enriched for Ty-VLPs (Fig. 3A, lanes 2 and 3). It was virtually absent in total protein extracts from transposition-induced cells (Fig. 3B, lanes 2 and 3). The processed material found in the protease mutants may result from residual Ty protease activity or the action of a cellular protease. The VLP proteins from the protease mutants may also contain other potentially aberrant post-translational modifications that could affect their electrophoretic mobility.

Protease mutations alter Ty-VLP morphology. To determine the role of Ty protease in the synthesis of Ty-VLPs, we compared the morphology of particles formed in transposition-induced strains DG716 and DG714 expressing pGTyH3NEO::SacI-1702 or pGTyH3NEO::MluI-1876 with the wild-type particles from strain DG692 by using electron microscopy (Fig. 4). In thin sections of the protease mutants, the VLPs had a distinctly altered morphology. Although almost all of the Ty-VLPs appeared partially open or cracked, they retained the same diameter (approximately 60 nm), abundance, and location within the cytoplasm. The host strains contained the spt3-101 mutation, precluding the accumulation of Ty-VLPs from the chromosomal elements. The spt3-101 mutation itself had no noticeable effect on the synthesis or morphology of wild-type TyH3-VLPs.

Protease mutations affect the synthesis of Ty DNA. Subcellular fractions enriched for Ty-VLPs contain a reverse transcriptase activity that can be detected by two different assays (20). The first assay involves the polymerization of [α-32P]GTP into the exogenous primer template oligo(dG)-poly(dC). Ty-VLPs also have the ability to catalyze the
primer-template-independent (endogenous) synthesis of Ty DNA when supplied with all four deoxyribonucleoside triphosphates with only one labeled with \(^{32}\)P. Particles prepared from the protease mutants demonstrated exogenous reverse transcriptase activity, but not endogenous activity (Table 3). The exogenous reverse transcriptase activity of the protease mutants was particulate and cosedimented in sucrose gradients with the bulk of the Ty proteins. No endogenous activity was found in any of the gradient fractions that had exogenous activity.

Considering the altered morphology of Ty-VLPs prepared from Ty protease mutants, the lack of endogenous activity may be the result of the absence of the Ty RNA template for reverse transcription. To determine whether these particles contained Ty RNA, we performed Northern filter hybridizations by using RNA from subcellular fractions that contained either mutant or wild-type TyH3-VLPs (Fig. 5). In both cases, we found genome-length 6.6-kb TyH3NEO RNA packaged in the VLPs, even though we made no attempt to

TABE 3. Reverse transcriptase activity of Ty protease mutants

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>10(^{-3}) Exogenous activity</th>
<th>Endogenous activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGAL1-lacZ</td>
<td>0.03</td>
<td>203</td>
</tr>
<tr>
<td>pGTYH3NEO</td>
<td>7</td>
<td>3,108</td>
</tr>
<tr>
<td>pGTYH3NEO::MluI-1876</td>
<td>7.5</td>
<td>204</td>
</tr>
<tr>
<td>pGTYH3NEO::SacI-1702</td>
<td>15</td>
<td>221</td>
</tr>
<tr>
<td>pGTYH3NEO-BglIIIΔ1702-3301</td>
<td>103</td>
<td>NT*</td>
</tr>
</tbody>
</table>

* The designated plasmids were introduced into strain DG531 by transformation. Ty-VLP-enriched fractions were isolated from transposition-induced cells and assayed for reverse transcriptase activity.

- Activity was expressed as nanomoles of dGTP incorporated into oligo(dG)-poly(dC) per hour per microgram of protein.
- Activity was expressed as counts per minute of \([\alpha-32P]dGTP\) incorporated per hour per microgram of protein.
- NT, Not tested.
FUNCTIONAL ORGANIZATION OF Ty ELEMENTS

FIG. 4. VLP morphology of a Ty protease mutant. Strains DG692 (pGTyH3NEO) and DG714 (pGTyH3NEO::MluI-1876) were transposition induced and prepared for electron microscopy. Both strains also carry the sp3-101 mutation. The mitochondria (M) and particles (P) are labeled in the ×27,360 magnification. The inset is a ×68,400 magnification of the particles.

protect RNA from degradation during the fractionation. However, the amount of Ty RNA present in the protease mutant VLPs was approximately 1/10 that found in wild-type particles. The decreased amount of template may account for the inability of the mutant particles to produce detectable levels of Ty DNA in vitro.

To monitor the integrity of a highly expressed RNA that should not be associated with VLPs, we compared the amount of GALL lacZ RNA present in total cellular RNA with the GALL lacZ RNA in fractionated material. The absence of GALL lacZ RNA in the P40 pellet (Fig. 5) or in sucrose gradient fractions (data not shown) suggests that unencapsidated RNAs are degraded or lost during fractionation. Identical results were previously obtained with the less abundant yeast URA3 transcript (20).

Ty protease is required for retrotransposition. We examined the effect of the protease mutations and a control in-frame linker insertion from outside the protease region (SacI-3301; Table 1) on transposition in an sp3 mutant background (Table 4). Transposition of TyH3NEO::SacI-3301 indicated that the presence of linker sequences per se does not affect transposition and that codon insertions in a Ty can be phenotypically silent. However, the protease mutations abolished transposition of TyH3NEO. GALL-promoted TyH3::MluI-1876 marked with a synthetic lacO sequence also failed to transpose or to activate endogenous Ty element transposition when assayed in a SPT3 background. Therefore, the Ty protease is required for transposition; no cellular activity can substitute for its function.

Identification of the Ty reverse transcriptase-coding sequence. Previous structural (11, 24, 38, 47, 51) and functional (5, 20, 37) studies indicate that the TyB gene encodes a reverse transcriptase. The Ty homology domain with retroviral reverse transcriptases begins near nucleotide 4000 of TyB and continues until the end of the open reading frame (Fig. 1A). To define the functional significance of this region, we characterized the reverse transcriptase activity of a Ty element containing an in-frame 1,599-nucleotide deletion that removes most of the protease domain, all of the integrase domain, and a significant portion of unassigned TyB coding sequence. The remaining TyB sequence includes a small amount of the protease region joined to sequences starting at nucleotide 3301 (Fig. 1B). Ty-VLPs prepared from transposition-induced strain DG718 expressing pGTyH3NEO-BglIIA1702-3301 contained 15-fold-higher reverse transcriptase activities than did Ty-VLPs from strain DG692 expressing wild-type pGTyH3NEO (Table 3). The increase in activity is accompanied by a comparable increase in the levels of several TyB proteins (data not shown). These results

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevant genotype</th>
<th>Transposition efficiency^a</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGTyH3NEO</td>
<td>sp3-101</td>
<td>35/40</td>
</tr>
<tr>
<td>pGTyH3NEO::MluI-1876</td>
<td>sp3-101</td>
<td>0/48</td>
</tr>
<tr>
<td>pGTyH3NEO::SacI-1702</td>
<td>sp3-101</td>
<td>0/48</td>
</tr>
<tr>
<td>pGTyH3NEO::BglIIA1702-3301</td>
<td>sp3-101</td>
<td>0/48</td>
</tr>
<tr>
<td>pGTyH3NEO::SacI-3301</td>
<td>sp3-101</td>
<td>0/48</td>
</tr>
<tr>
<td>pGTyH3lacO</td>
<td>SPT3</td>
<td>10/10</td>
</tr>
<tr>
<td>pGTyH3lacO::MluI-1876</td>
<td>SPT3</td>
<td>0/10</td>
</tr>
</tbody>
</table>

^a The designated plasmids were introduced into the sp3-101 mutant strain DG531 or the SPT3 strain YH8 by transformation. The transformants were transposition induced for 5 days at 20°C and then analyzed as described in Materials and Methods.

^b For Ty elements marked with the NEO gene, the transposition efficiency is the number of G418^R Ura^- segregants divided by the total number of Ura^- segregants analyzed. For Ty elements marked with the lacO sequence, the transposition efficiency is the number of Ura^- segregants that have transpositions as judged by Southern analysis with lacO- or Ty-specific probes divided by the total number of Ura^- segregants analyzed.
suggest that an active reverse transcriptase is encoded by sequences 3' of nucleotide 3301 in TyB. This region includes residues homologous to retroviral reverse transcriptases.

**DISCUSSION**

In the present work, the relatedness of protein domains between the TyB gene and the retroviral pol gene was used in combination with our ability to monitor Ty transposition to investigate Ty gene function. A 9-amino-acid domain homologous to retroviral proteases is located near the 5' end of TyB. An enzymatically active protease responsible for the cleavage of several retroviral proteins is also found in the corresponding region of murine leukemia virus (12, 28, 54). Mutations adjacent to the Ty protease homology domain profoundly affect both TyA and TyB protein processing. Codon insertions bracketing the protease that extend into TyA or into the integrase homology domain have no effect on protein processing. These results strongly suggest that the Ty protease mutants directly affect the activity of the Ty protease. However, we cannot eliminate the possibility that the Ty protease mutations cause the defects indirectly. The mutations might alter particle structure, rendering TyA and
TyB proteins resistant to cleavage, or the mutations might affect the activity of a cellular protease.

In the work described in this paper, we demonstrated that a Ty protease is required for several important steps in the retrotransposition of Ty elements. Protease mutants cannot synthesize Ty DNA in vitro, properly process TyA- or TyB-encoded proteins, or form morphologically normal Ty-VLPs. These biochemical defects have a dramatic effect on transposition. When induced for transposition, protease mutants cannot transpose or activate the transposition of chromosomal elements. This is the first demonstration that a Ty-encoded function, the Ty protease, is essential for transposition.

Recent studies have shown that TyA encodes the major capsid proteins of the Ty-VLP. The predominant full-length TyA proteins from Ty1-15 and Ty9C, called p1 and p2 (1) or pro-TyA and TyA (39), can form recognizable particles in the absence of most of TyB (1, 39). The TyH3-encoded p58-TyA and p54-TyA proteins identified in total cell extracts probably correspond to products encoded by the related elements Ty1-15 and Ty9C. TyH3, Ty1-15, and Ty9C are members of the Ty1 structural class of elements, which are present in approximately 25 to 35 copies per haploid genome in common laboratory yeast strains (9, 19, 29). Although these elements are closely related, microheterogeneity has been detected by restriction site polymorphisms and by comparison of DNA sequences.

Results of our analysis of TyA proteins from TyH3-VLPs markedly differ from the results obtained with Ty1-15 and Ty9C particles. When the major Ty-VLP antigens recognized by anti-VLP antibodies were compared, p38-TyA and p41-TyA were the dominant proteins, whereas the relatively unprocessed p2 or TyA protein in the 50- to 60-kDa range appeared to be a major component of the particles encoded by Ty1-15 or Ty9C. We detected p1- and p2-sized proteins in wild-type TyH3-VLPs, but these proteins made up a minor fraction of the immunoreactive protein found in TyH3 particles. These results probably do not reflect differences in antibody reactivities for several reasons. First, the anti-VLP antibodies used in our experiments were raised against VLPs from the related element Ty1-15. Second, an antibody raised against the C terminus of TyA reacted exclusively with a subset of smaller proteins from Ty-VLPs, yet it reacted only with full-length TyA proteins from total cell extracts. Thus, two independently derived antibodies showed similar patterns of reactivity toward Ty proteins. Preliminary results also suggest that a major TyH3-VLP protein comigrates with p38-TyA (S. Youngren, unpublished results). The functional significance of this difference in the extent of TyA processing is unclear because the transpositional competence of Ty1-15 and Ty9C remains to be determined.

A 190-kDa protein is detected in protease mutants created by the insertion of four or five extra codons. This protein reacts with both TyA- and TyB-specific antibodies, strongly suggesting that the 190-kDa protein is the TyA-TyB precursor. Since anti-TyB2 antibodies failed to detect p90-TyB in the protease mutants, we believe that a specific TyB protein, p90-TyB, is first synthesized as part of p190-TyA-TyB and is then processed to its final size. Recently, Muller et al. (39) found a 160-kDa VLP-associated protein immunologically related to TyA in cells expressing a Ty element that has a large deletion near the protease domain. These results are also best explained by the lack of proteolytic cleavage of the TyA-TyB precursor.

Mutations in the Ty protease domain affect reverse transcriptase activity. We obtained the same levels of exogenous reverse transcriptase activity from Ty-VLPs from a wild-type Ty element and the protease mutants. However, the protease mutants did not synthesize detectable levels of Ty DNA in vitro. Probably, the absence of the endogenous reverse transcriptase activity was caused by the significantly lower levels of Ty RNA template in the mutant VLPs. The mutant particles may be more fragile, or may package less Ty RNA. Alternatively, an incorrectly processed Ty reverse transcriptase may be active when provided with an exogenous primer-template, but may be incapable of effectively using the substrates and cofactors present in the endogenous reaction to synthesize Ty DNA. To further investigate these possibilities, the physical nature of the reverse transcriptase activity in protease mutants must be determined.

Mutations in Ty protease or in the retroviral murine leukemia virus protease have pleiotropic effects. Mutations in the murine leukemia virus protease region cause defects in protein processing, reverse transcription, and virus maturation (12, 28). The major viral precursor polyproteins (Pr65 gag, Pr200 gag-pol, and gPr80-env) are synthesized in cells carrying in-phase deletions in the protease domain. These cells release morphologically immature virions that contain mostly unprocessed gag and gag-pol polyproteins. The mutant virions apparently contain a functionally normal surface glycoprotein, but their specific infectivity is greatly reduced. The infection is blocked well before reverse transcription of the genome takes place, and it is unclear whether the mutant virions can enter the host cell. In addition, there are conflicting views on whether the murine gag-pol precursor is an enzymatically active reverse transcriptase (12, 42). Like Ty, retroviral protease mutants fail to transpose or to form the provirus. However, the major defect of Ty or murine leukemia virus protease mutants may well be the inability to mature prior to reverse transcription. Further understanding of the molecular events that cause Ty transposition should elucidate the similarities and differences between the Ty and retroviral life cycles.

ACKNOWLEDGMENTS

This research was sponsored by the National Cancer Institute under contract no. N01-C0-23909 with BioNetics Research, Inc. J.D.B. was supported by Public Health Service grant GM36481 from the National Institutes of Health.

We are grateful to G. Monkonik for technical assistance, M. Gonda and K. Nagashima (PRI, Inc., National Cancer Institute-Frederick Cancer Research Facility, Frederick, Md.) for help with electron microscopy, S. Showalter (PRI, Inc., National Cancer Institute-Frederick Cancer Research Facility) for help in generating antibodies, S. Lucas for preparation of the manuscript, and G. R. Fink for his support and encouragement in the beginning of this work. N. Copeland, J. Strathern, and especially M. Curcio provided helpful discussion and comments on this manuscript.

LITERATURE CITED


