MOLECULAR AND CELLULAR BIOLOGY, June 1995, p. 3310–3317 Vol. 15, No. 6
0270-7306/95/$04.00+0

A Novel Mechanism of Self-Primed Reverse Transcription Defines a New Family of Retroelements

HENRY L. LEVIN*

Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, Bethesda, Maryland 20892

Received 22 December 1994/Returned for modification 16 February 1995/Accepted 3 March 1995

Retroviruses and long terminal repeat (LTR)-containing retrotransposons initiate reverse transcription by using a specific tRNA primer that anneals to the primer-binding site of the retroelement transcript. Sequences from a large number of retroviruses and LTR-containing retrotransposons had indicated that the role of tRNAs in priming reverse transcription is universal among these LTR-containing retroelements. Data presented here strongly support the surprising conclusion that Tf1, a highly active LTR-containing retrotransposon isolated from *Schizosaccharomyces pombe*, undergoes a novel self-priming process that requires hybridization between the primer-binding site and the first 11 bases of the Tf1 transcript. Single-base mutations in these regions block transposition and reverse transcription, while compensatory mutations that reestablish complementarity rescue both defects. In addition, the sequence of the minus-strand RNA primer of reverse transcription was consistent with its being derived from the 5′ end of the Tf1 transcript. Evidence that this mechanism defines a new family of retroelements is presented.

The propagation of retroviruses and retrotransposons containing long terminal repeats (LTR) requires the insertion of a cDNA copy of their sequence into the genome of a host cell. Many features of the reverse transcription process are conserved among the broad family of LTR-containing elements (7, 24). The first steps of reverse transcription are universal among these elements and include the hybridization of a specific tRNA molecule to the retroelement mRNA at the primer-binding site (PBS), found just downstream of the 5′ LTR (1, 13). As demonstrated for both retroviruses (25) and retrotransposons (3, 11), a tRNA molecule is required for priming the synthesis of minus-strand strong-stop DNA, the first product of reverse transcription (Fig. 1A). The role of tRNA in the priming of reverse transcription is thought to be common to all LTR-containing elements since the retroviruses and LTR-containing retrotransposons that have been sequenced have a PBS that is complementary to known tRNA species.

Tf1 is an LTR-containing retrotransposon that was isolated from the fission yeast *Schizosaccharomyces pombe*. It has coding sequences similar to the protease, reverse transcriptase, and integrase (IN) proteins of retroviruses and retrotransposons (16). In vivo assays for transposition demonstrate that Tf1 is highly active, resulting in transposition frequencies as high as 20% (17, 23).

In light of the conservation of reverse transcription priming, it is surprising that no tRNA primer has been identified for Tf1. In addition to the absence of tRNA homology, Tf1 lacks the highly conserved UGG in the PBS of retroelements that hybridizes to the last 3 bases of the tRNA molecule. Evidence for the presence of an RNA primer that functions precisely at the conventional PBS location of Tf1 came from the analysis of DNA isolated from virus-like particles. This DNA was used as the template in primer extension analyses, and treatment with RNase demonstrated that there are at least 9 bases of RNA at the 5′ end of the minus-strand strong-stop DNA. The position of this RNA is immediately downstream of the 5′ LTR and defines the PBS for Tf1 (17). Instead of using an uncharacterized tRNA as the primer, data from this study strongly support the model that Tf1 mRNA undergoes a novel self-priming mechanism of reverse transcription. The first 11 bases of the Tf1 transcript are exactly complementary to the PBS location identified by the primer extension experiments. This observation raised the possibility that the first 11 bases of the transcript fold back and anneal to the PBS. If an RNA endonuclease cuts the first 11 bases from the transcript, the 3′ OH required to prime DNA synthesis would become available (Fig. 1B). Data are presented below showing that individual base changes in the PBS or the 5′ end of the transcript block transposition and reverse transcription while compensatory mutations that reestablish complementarity rescue both defects. In addition, the RNA primer of reverse transcription was identified by hybridization and does indeed include the sequence of the first 11 bases of the Tf1 transcript. Evidence that this mechanism defines a new family of retroelements is presented.

MATERIALS AND METHODS

Media. The *S. pombe* minimal liquid and plate media were composed of EMM (20). Selective plates contained EMM and 2 g of dropout mix per liter, a powder that contained all amino acids, adenine, and uracil except for nutrients that are absent as required for selection (21). Vitamin B_1 (thiamine; 10 μM) was added to EMM plates when indicated. 5-Fluoroorotic acid (5-FOA; PRC Inc., Gainesville, Fla.) plates were made by adding 1 g of 5-FOA per liter to EMM supplemented with 100 μg of uracil per ml. YES 5-FOA/G418 plates contained 5 g of yeast extract (Difco), 2 g of complete dropout mix, 1 g of 5-FOA, and 500 mg (corrected for purity) of Geneticin (G418; Gibco) per liter.

Strains and plasmid constructions. The yeast strains used in this paper are listed in Table 1. The oligonucleotides used are listed in Table 2. Plasmid pHL449-1 contained the *nmt1* promoter fused to Tf1-*neo* at the start of the transposon transcript and was identical to pHL414 (17), except that an artificial intron (GTAGGTGCTATTTTACTAGTCTAAGCTAATCAATAG) was inserted in reverse orientation into the *Nru*I site of *neo*. Although the intron was included in these plasmids with the intention of using a direct-transposition assay (4), none of the experiments described in this paper involved the intron, since transposition events were scored only after the plasmid copy of *neo* was evicted. pHL476-3 was identical to pHL449-1, except that a reading frameshift in the beginning of IN was created from a three-piece ligation of the 5-kb vector *Xho*I-*Bam*HI fragment of pHL414, the 2-kb *Apa*I-*Bam*HI 3′ Tf1 fragment of pHL449-1, and a 4-kb *Apa*I-*Xho*I fragment of Tf1 from pHL338-9 that had a frameshift mutation created at the beginning of IN at a *Bsp*HI site. All PCRs described below were performed with the *Pfu* enzyme (Stratagene). The point

* Phone: (301) 402-4281. Fax: (301) 496-0243. Electronic mail address: Henry_Levin@nih.gov.

TABLE 1. Yeast strains used in this study

Strain	Genotype or description	Source	Plasmid description
YHL912	*h⁻ ura4-294 leu1-32*	J. Boeke, 21X5A	
YHL1282	YHL912/pHL449-1	This study	Wild-type Tf1-*neo*AI
YHL1554	YHL912/pHL476-3	This study	Tf1-*neo*AI with frameshift in beginning of IN
YHL2738	YHL912/pHL793-1	This study	Tf1-*neo*AI with C-to-G mutation at 7th position of 5′ end
YHL2722	YHL912/pHL791-1	This study	Tf1-*neo*AI with G-to-C mutation at 7th position of PBS
YHL2786	YHL912/pHL799-1	This study	Tf1-*neo*AI with double mutation at 7th position
YHL2754	YHL912/pHL795-1	This study	Tf1-*neo*AI with G-to-C mutation at 5th position of 5′ end
YHL2706	YHL912/pHL789-1	This study	Tf1-*neo*AI with C-to-G mutation at 5th position of PBS
YHL2770	YHL912/pHL797-1	This study	Tf1-*neo*AI with double mutation at 5th position

mutations were all generated on 980-base PCR *Xho*I-*Avr*II fragments that contained the Tf1 mRNA leader and that were ligated into the unique *Xho*I and *Avr*II sites of pHL411-62. The plasmids with single point mutations in the PBS (pHL791-1 and pHL789-1) were constructed by fusion PCR with two overlapping approximately 30-base oligonucleotides with the mutant sequence and flanking PCR oligonucleotides that contained restriction sites unique to the Tf1 plasmid. The fusion PCR products were produced by mixing oligonucleotides that hybridize to the 5′ (HL38) and 3′ (HL39) ends of the fusion product with template that consisted of two half PCR fragments that were made with the overlapping mutation oligonucleotides and one of the two flanking oligonucleotides. The fusion products were ligated into pHL411-62 (identical to pHL449-1 except without the *neo*) as *Xho*I-*Avr*II fragments, and the mutations were finally ligated into pHL449-1 on a 4-kb *Bst*XI fragment. The specific oligonucleotides used in the construction of each mutation are shown in Table 2. Mutations made in the 5′ end of the transcript (pHL793-1 and pHL795-1) were made with single PCR products by incorporating the desired mutation within an oligonucleotide that overlapped the unique *Xho*I site used to insert the *Xho*I-*Avr*II PCR fragment into pHL411-62. The double-mutant plasmids pHL799-1 and pHL797-1 were created with DNA from the PBS mutant plasmids (pHL791-1 and pHL789-1) as the template with HL39 (wild-type 3′ PCR oligonucleotide) and either of the 5′-endpoint mutant oligonucleotides HL46 and HL47. The mutation sites of each plasmid were sequenced to verify the nature of the base changes. In addition, each PCR-generated construct was made in duplicate by using an independent PCR to control against the effects of untemplated PCR-generated mutations. Each duplicated plasmid gave results in all transposition assays equivalent to that of its partner. For the sake of simplicity, the results for only one member of each pair are reported below.

Transposition assay. Strains were first grown as patches on EMM to induce the *nmt1* promoter, which was fused to the *neo*-marked Tf1. The presence of the bacterial *neo* gene allows *S. pombe* to grow in 500 μg of G418 per ml. After 5 days of 32°C incubation, the plate was replica printed to EMM medium containing 5-FOA to select against cells containing the Tf1-*neo* plasmid (2). This plate was then replica printed to YES medium (20) containing G418 as well as 5-FOA and incubated at 32°C for 2 days to determine at what frequency Tf1-*neo* inserts into the genome (15, 17). Quantitative measurements of transposition frequencies were performed as follows. Strains were grown as patches of cells on EMM-uracil dropout agar (20) for 2 days at 32°C. These cells were then diluted and grown in liquid EMM-uracil dropout from an optical density at 600 nm of 0.1 to saturation. The cells were then spread onto 5-FOA plates, and the resultant colonies were printed to YES plates containing 500 μg of G418 per ml (17).

Particle preparation and extraction of nucleic acid. The preparations of large-scale yeast extracts and the subsequent analysis on sucrose gradients were based on previously published protocols (5, 6, 8). A 500-ml portion of EMM-uracil medium was inoculated at an optical density at 600 nm of 0.05 and grown to an optical density at 600 nm of 1.0. The cells were harvested, washed, and broken as previously described (17). Glass beads loaded to just above the meniscus greatly increased the efficiency of breakage and recovery of virus-like particles. Harvested cells (1 ml) were resuspended in 3.5 ml of buffer B/EDTA (15 mM KCl, 10 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid [HEPES] · KOHs [pH 7.8], 5 mM EDTA) containing 1 μg of aprotinin per ml, 0.5 μg of leupeptin per ml, 0.7 μg of pepstatin per ml, 3 mM dithiothreitol, and 2 mM phenylmethylsulfonyl fluoride. A 5-ml portion of supernatant recovered from a 3,000-rpm SS34 centrifugation (5 min) of the cell extract was loaded onto a 20 to 70% linear gradient of sucrose in buffer B/EDTA and centrifuged for 24 h at 25,000 rpm in a Beckman SW28 rotor. To isolate nucleic acid from particles, 1.2-ml fractions were collected and 0.8-ml portions each of fractions 26 to 30 were pooled and diluted in buffer B up to a volume of 11.5 ml. This diluted sample was spun at 35,000 rpm in a Beckman SW41 for 1 h to pellet the particles. Pellets were resuspended in 0.4 ml of 25 mM EDTA–0.1% sodium dodecyl sulfate (SDS)–50 μg of proteinase K per ml (Boehringer Mannheim) and incubated at room temperature for 2 h. This solution was then phenol extracted and ethanol precipitated. The material from one 500-ml culture was then resuspended in 0.4 ml of Tris-EDTA buffer (TE) and split into two samples, with one receiving the addition of 1 μl of 2-mg/ml RNase A. Both samples were incubated for 1 h at 37°C and then phenol extracted and ethanol precipitated. These pellets were resuspended in sequencing stop buffer and loaded onto a 6% polyacrylamide–46% urea sequencing gel.

Electrotransfer of sequencing gel and DNA blot procedure. The sequencing gels were electrotransferred in a Hoefer TE77 onto GeneScreen Plus (DuPont) for 1 h with 300 mA in 0.5× Tris-borate-EDTA buffer (TBE). The filter was UV

TABLE 2. Oligonucleotides used in the construction of plasmids and the analysis of Tf1 DNA

Oligonucleotide	Sequence	Use
HL38	GGAAGAGGAATCCTGGC	PCR oligonucleotide used as 5′ flank for all *Xho*I-*Avr*II fragments
HL39	CCAATTGCTTCCAGTCTTTG	PCR oligonucleotide used as 3′ flank for all *Xho*I-*Avr*II fragments
HL42	GCGTAGCTAACAATAAGTGAACTCTTG	Top-strand fusion oligonucleotide of overlapping pair with 5th-position mutant in PBS
HL43	TCACAAGAGTTCACTTATTGTTAG	Bottom-strand fusion oligonucleotide of overlapping pair with 5th-position mutant in PBS
HL44	GTAGCTAACAATAACTCAACTCTTGTG	Top-strand fusion oligonucleotide of overlapping pair with 7th-position mutant in PBS
HL45	GATCACAAGAGTTGAGTTATTGTTAG	Bottom-strand fusion oligonucleotide of overlapping pair with 7th-position mutant in PBS
HL46	CATATCCTCGAGTCCACTTATGAGCTATA	PCR oligonucleotide with *Xho*I site used to create the mutation in the 5th position of the 5′ end
HL47	GGCATATCCTCGAGTTGAGTTATGAGC	PCR oligonucleotide with *Xho*I site used to create the mutation in the 7th position of the 5′ end
HL77	AGTTCAGTTATGAGCT	Probe for DNA blot (see Fig. 4C)
HL78	TGTTAGCTACGCAGTTT	Used as sequencing primer for ladder on DNA blot
HL89	AAACTGCGTAGCTAACA	Probe for DNA blot (see Fig. 5)
HL90	GCTAACAATAACTGAACT	Probe for DNA blot (see Fig. 5)
HL92	GAGCTATATAAATGATAG	Probe for DNA blot (see Fig. 5)
JB341	ATACTCAGTTGTTACTC	Probe for DNA blot (see Fig. 5)

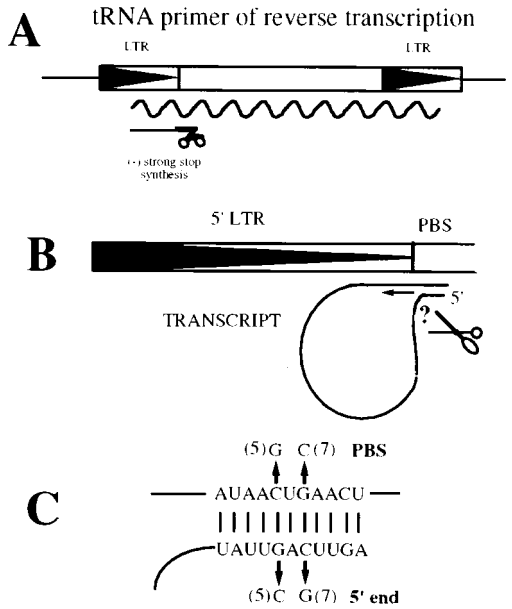

FIG. 1. Mechanisms of priming reverse transcription. (A) Accepted mechanism of tRNA-mediated priming of minus-strand strong-stop DNA for LTR-containing retroelements. The largest rectangle represents the retroelement, and the smaller, triangle-containing rectangles are the LTRs. The wavy line depicts the full-length transcript, and below this is a straight line that represents the minus-strand strong-stop DNA with a tRNA primer at its right end. (B) Proposed self-priming structure of Tf1 mRNA with an RNA duplex of 11 bp that forms between the first 11 bases of the transcript and the PBS. Formation of this structure may be followed by a cleavage event that exposes a 3′ OH for priming. The horizontal triangle is the 5′ LTR of Tf1. The looped structure is the mRNA, and the arrow shows the direction of DNA synthesis after priming. (C) RNA duplex sequence of the 11 bp that forms in the Tf1 mRNA. The arrows indicate the positions of the single point mutations made in the fifth and seventh base positions of the PBS and of the 5′ end. Compensatory pairs of changes made at the fifth and seventh base positions were also expressed. Figure 2 shows the effects on transposition of all these mutations.

cross-linked for 7 min in a Stratalinker (Stratagene) at the default energy setting. The filter was then prehybridized for 1 h in 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% SDS–0.1 mg of sheared herring sperm DNA per ml. Phosphorylated oligonucleotide probe (10^7 cpm) was added, and the hybridization was allowed to occur overnight. The filters were washed three times for 5 min each in 6× SSC–0.1% SDS.

RESULTS

Effect on transposition of mutations in the putative self-priming duplex. Initial support for a self-priming mechanism of reverse transcription was the observation that the first 11 bases of the Tf1 transcript has perfect complementarity to the Tf1 PBS (Fig. 1C). To evaluate a possible role of this putative structure in priming of reverse transcription, the effect on transposition of point mutations in the PBS as well as in the first 11 bases of the transcript was determined. The assay for transposition utilized *S. pombe* cells with a plasmid copy of Tf1 fused to a strong inducible promoter that overproduced a full-length Tf1 transcript (17). To detect transposition events, a bacterial *neo* gene was placed in a nonessential region of the transposon. The presence of *neo* allows *S. pombe* to grow in high concentrations of G418. After expression is induced, reverse transcription of the Tf1 mRNA and the subsequent insertion of the *neo*-marked cDNA into the host genome yields transposition events. Cells that gain newly transposed copies of Tf1 are identified by their ability to grow on G418 once the plasmid copy of *neo* is evicted (15).

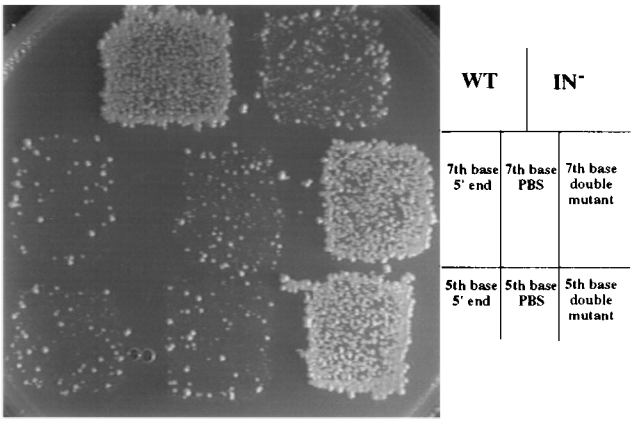

FIG. 2. Transposition activity of a plasmid version of *neo*-marked Tf1 with point mutations in the putative primer and PBS sequences. The level of growth on this G418 plate represents the frequency of transposition events within each of these strains. From left to right, the top row contains the wild type (WT) (YHL1282) and Tf1 with a frameshift in IN (YHL1554); the second row contains Tf1 with mutations in the seventh-position base pair located in the 5′ end of the transcript (YHL2738), in the PBS (YHL2722), and in both the 5′ end and the PBS (YHL2786); and the third row contains Tf1 with mutations in the fifth-position base pair located in the 5′ end of the transcript (YHL2754), in the PBS (YHL2706), and in both the 5′ end and the PBS (YHL2770).

The wild-type Tf1-*neo* plasmid in strain YHL1282 produced high levels of transposition that resulted in a patch of confluent growth on G418 medium, as shown in Fig. 2. YHL1554, a negative control strain that contained the Tf1 plasmid with a frameshift mutation in the IN domain, produced background levels of growth that are the result not of transposition but of low-level homologous recombination (23). Since single-base changes in the PBS of the *Saccharomyces cerevisiae* retrotransposon Ty1 had large effects on the transposition frequency, the first mutations made in Tf1 as part of this study were also single-base changes (10). Strain YHL2706 contained a Tf1 assay plasmid with a C-to-G base change at the fifth base position of the PBS (Fig. 1C). This single mutation greatly reduced the frequency of transposition, as seen by the inability to grow on the G418 plate in Fig. 2. Strain YHL2754 contained a similar G-to-C mutation made in the 5′ end of the Tf1 transcript at the base proposed to hybridize to the fifth base of the PBS. This mutation also greatly reduced the transposition frequency. To determine if these transposition defects were due to the lack of hydrogen bonding between these two bases, strain YHL2770, in which complementarity was reestablished by combining the PBS and the 5′-end mutations just described, was constructed. YHL2770 showed transposition frequencies similar to the wild type (Fig. 2), indicating that complementarity between the PBS and the 5′ end, rather than the exact sequence at these two sites, was important for transposition. To establish definitively that complementarity within the 11-base region of the putative RNA duplex is required for transposition, the seventh base of the PBS was mutagenized as well as its counterpart at the 5′ end of the transcript (Fig. 1C). YHL2722 has a G-to-C mutation at the seventh base of the PBS, while YHL2738 has a C-to-G mutation in the position of the 5′ end opposing the seventh base of the PBS. Both of these mutant strains showed greatly reduced transposition frequencies, confirming that the sequence of the PBS and the 5′ end of the transcript are each required for normal transposition. YHL2786, a strain that contained both mutations, had wild-type levels of transposition. Thus, the defects in transposition observed in the individual mutations were repaired by combin-

TABLE 3. Quantitative transposition frequencies of strains mutated in the PBS or the 5′ end of the transcript[a]

Transposon[b]	No. of colonies		% With transposition
	5-FOA^r	G418^r	
WT (YHL1282)	2,603	30	1.10
IN fs (YHL1554)	2,743	2	0.070
Base 5 position			
5′ end (YHL2754)	2,875	0	<0.030
PBS (YHL2706)	2,321	0	<0.040
Double (YHL2770)	2,753	22	0.80
Base 7 position			
5′ end (YHL2738)	2,600	0	<0.040
PBS (YHL2722)	2,900	0	<0.030
Double (YHL2786)	2,521	22	0.87

[a] The strains were tested for transposition by growing patches of cells on EMM-uracil agar for 2 days at 32°C. The cells were then diluted and grown in liquid EMM from an optical density at 600 nm of 0.1 to saturation. The cells were spread onto 5-FOA plates, and the resulting colonies were printed to YES plates containing 500 μg of G418 per ml.

[b] WT, wild type; fs, frameshift.

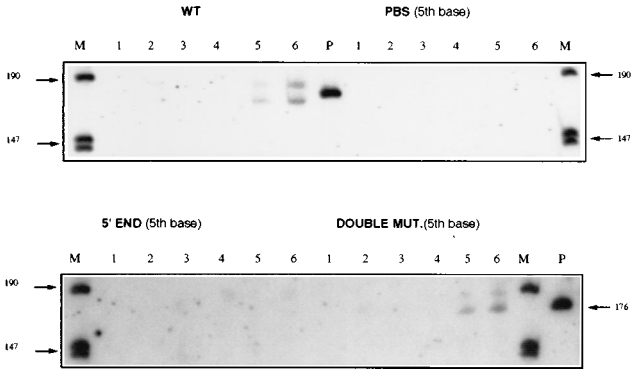

FIG. 3. DNA blot detection of minus-strand strong-stop species isolated from virus-like particles. Extracts from strains YHL1282 (WT), YHL2706 (PBS, 5th base), YHL2754 (5′ END, 5th base), and YHL2770 (DOUBLE MUT., 5th base) were loaded onto sucrose gradients, and the nucleic acids from all fractions were analyzed by DNA blot hybridization analysis. Pools 1 through 6 each contain material from five consecutive gradient fractions beginning from the top of the gradients and ending at the bottom. The minus-strand strong-stop species of DNA were detected with JB341, a radiolabeled oligonucleotide, as the specific probe of the minus strand. Lanes M contain molecular size markers (in bases), and lanes P have a synthetically produced Tf1 minus-strand strong-stop DNA that is 5 bases larger than the expected size and was made by restriction digestion of a Tf1 plasmid with *Bgl*II sites at the PBS and the 5′ end. To determine if the absence of the strong-stop signal in the single-point mutant strains could be due to degradation of the nucleic acid in the sample or reduced loading compared with the wild type and the double mutant, RNA hybridization blots were run on the extracted nucleic acid from all four strains. Equal amounts of Tf1 transcript in the particle fractions of the sucrose gradients were observed. The gel was a 0.8-mm-thick 6% sequence gel that was electrotransferred to GeneScreen Plus (Amersham). The oligonucleotides were hybridized at 36°C in 6× SSC–0.5% SDS.

ing two compensatory mutations that reestablish base pairing between the PBS and the 5′ end (Fig. 2).

Quantitative analysis of the transposition defects caused by the point mutations shown in Fig. 1C was conducted by determining what percentage of colonies contain cells derived from a progenitor that had suffered a transposition event (17). Compared with the wild-type version of Tf1, the IN frameshift mutant showed at least a 15-fold lower level of transposition activity (Table 3). For all of the mutations with single-base changes, the frequency dropped at least 27-fold compared with the wild type. These extremely low levels of transposition were rescued to 80 and 87% of normal in the double mutants with compensatory changes at the fifth and seventh positions, respectively.

Measurement of the minus-strand strong-stop DNA levels. Although the transposition data demonstrate a surprising requirement for an interaction between the 5′ end of the transcript and the PBS, it was important to determine if the point mutations actually disrupted the synthesis of minus-strand strong-stop DNA as predicted in the model for self-priming. Toward this end, a DNA blot was made of nucleic acid extracted from pools of sucrose gradient fractions that contained Tf1 virus-like particles (Fig. 3). This blot was designed to detect minus-strand strong-stop DNA. Therefore, the nucleic acids were fractionated on a 6% sequencing gel and, after transfer, the resulting membrane was probed with labeled oligonucleotides that hybridized specifically to the minus strand of the strong-stop region. A doublet of about the size of an artificially produced Tf1 strong-stop control fragment (P) was observed in pools 5 and 6 of a sucrose gradient that contained virus-like particles from the wild-type transposon strain, YHL1282 (Fig. 3). The doublet that indicated the presence of minus-strand strong-stop DNA in the wild type was absent in strains with a point mutation in the PBS or the 5′ end of the transcript. As expected, the strong-stop signal was restored in the strain with the double mutation. These results provide strong evidence that the drop in transposition frequencies caused by the single mutations was due to an inability to synthesize minus-strand strong-stop DNA as predicted for defects that block priming.

Because the 11-base sequence at the 5′ end of the Tf1 transcript is also present in the 3′ LTR, the identical fifth- and seventh-position changes were made in the 3′ LTR. The results indicated that these 3′ LTR bases mutations have no effect on transposition (data not shown).

To exclude the possibility that the changes in transposition observed in these experiments are due to altered levels of Tf1 expression, the amounts of Gag protein were measured by immunoblot analysis. A wild-type level of Gag was observed in each of the mutants described above. In addition, all these mutants produced wild-type levels of particles as determined by immunoblot analysis of sucrose gradient fractions that also contained normal levels of Tf1 mRNA (data not shown).

Characterization of the minus-strand strong-stop DNA species. The sizes of the minus-strand strong-stop doublet species, as well as the positions of their termini, were examined for information that might relate to the priming mechanism. To determine the precise sizes of the two major strong-stop DNA species seen in Fig. 3, particle nucleic acid was run on a thin sequence gel and electrotransferred to a filter that was probed with labeled JB341, an oligonucleotide that was complementary to Tf1-specific sequence of the minus strand in the center of the strong-stop species (the identities of all of the oligonucleotide probes are given in Fig. 4, and the blot is shown in Fig. 5A). The sizes of the two strong-stop species were determined with a sequence ladder that was produced with HL78, an oligonucleotide that had a 5′ end that corresponded to the 5′ end sequence of the strong-stop DNA after the material was treated with RNase A (17). The strong-stop signal in the RNase A lane of Fig. 5A was a doublet with a top band (170 to 172 bases) equivalent in size to that predicted for a full-sized strong-stop product that has a 3′ end that extends up to the first base of its template, the Tf1 transcript (plus sign in Fig. 4). The bottom band (159 to 161 bases) was the size predicted for a minus-strand strong-stop species that had a 3′ end shortened by about 11 bases (asterisk in Fig. 4). The untreated sample

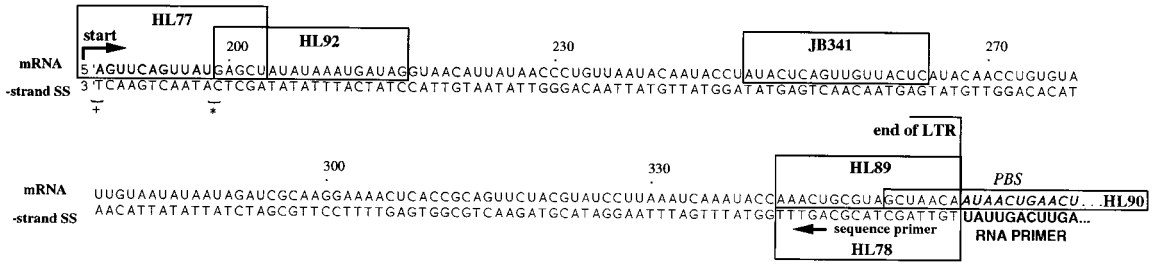

FIG. 4. Structure of the minus-strand strong-stop DNA. This figure shows the sequence of the 5′ end of the Tf1 transcript (mRNA) and the structure of the minus-strand strong-stop species (−strand SS). The numbers above the mRNA sequence refer to the base positions relative to the published sequence of Tf1 (16). The arrow marked "start" indicates the beginning of Tf1 transcription from its native promoter (15). The start site as expressed from the *nmt1* promoter is 1 base upstream (14). The boxes labeled with the names of oligonucleotides indicate their sequences relative to the sequences of the mRNA and the strong-stop species. The bracket labeled "end of LTR" shows the 3′ end of the 5′ LTR. The plus sign indicates the position of the 3′ end of the largest strong-stop species, while the asterisk marks the 3′ end of the smaller strong-stop product. The 11 bases in boldface type at the 5′ end of the mRNA indicate the position of the proposed primer of reverse transcription. The 11 bases shown in boldface italic type indicate the position of the RNA primer after reverse transcription has occurred. The sequence of the PBS in the mRNA is shown in italics.

also contained a doublet, but in this case the previously observed lower band was absent and a larger species appeared that was about 11 to 13 bases larger than the size predicted for the strong-stop DNA, presumably because of the presence of the RNA primer at the 5′ end. Figure 5B contains an autoradiogram of the same filter probed before exposure to JB341 with HL90, an oligonucleotide that was designed to hybridize to the RNA primer at the 5′ end of the strong-stop DNA predicted to have the same sequence as the first 11 bases of the Tf1 transcript (Fig. 4). The sample treated with RNase A showed no signal, while the lane that contained material not treated with RNase A showed the same two bands subsequently detected by JB341. This result indicated that the two bands in the untreated material contain RNA at their 5′ ends that had the same sequence as the beginning of the Tf1 transcript. Also, the lack of signal from both bands in the treated sample indicated that RNase treatment of the samples in Fig. 5A caused the banding pattern seen by shortening each of the major bands in the untreated lane by about 11 bases, the size predicted for the RNA primer by the self-priming model.

To determine if the differences in the structures at the 5′ ends of the two major strong-stop species treated with RNase versus those not treated with RNase were limited to the absence of the RNA primer, the filter used in Fig. 5A and B was stripped and reprobed with HL89, an oligonucleotide that is complementary to the sequence just 3′ of the RNA primer (Fig. 5C). Since both major species produced a signal, regardless of whether they were treated with RNase, and since HL90 detected the two major species only before RNase treatment, both strong-stop products had 5′ DNA ends that corresponded to the 3′ end of the LTR as shown in Fig. 4.

The 3′ ends of the strong-stop products were mapped with three identical DNA blot filters that were probed with either JB341, HL77 (an oligonucleotide complementary to the predicted 3′ end of the strong-stop sequence), or HL92 (an oligonucleotide complementary to sequence beginning 11 bases upstream from the 3′ end). Figure 4 shows the positions of these oligonucleotides. Although HL92 (Fig. 6B) detected both major strong-stop species before and after RNase treatment, HL77 (Fig. 6C) hybridized only to the larger of the two. The lack of signal from the shorter species, in both the RNase-treated and untreated lanes, indicated that the difference in the sizes of the strong-stop species was due to sequence absent from the 3′ end of the smaller product. Figure 6A contains the same material on an independent filter probed with JB341 and

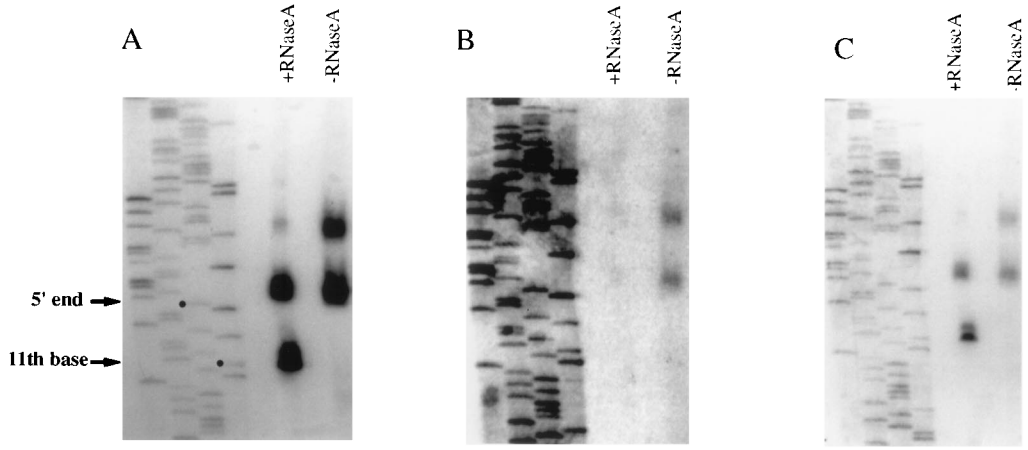

FIG. 5. Identification of the RNA primer at the 5′ end of the strong-stop species. Nucleic acid was extracted from Tf1 particles and loaded onto a sequence gel. After being electrotransferred to a filter, it was probed in succession with three labeled oligonucleotides. The sequence ladder was produced with HL78 as a primer. (A) DNA blot produced with JB341 as the probe, after the filter had been stripped of signal from HL90. The positions of the 5′ end of the transcript and the 11th base of the transcript are shown on the sequence ladder. (B) Result of the first probing of this filter. The HL90 probe was used to detect the presence of the RNA primer. (C) After the filter was probed with JB341, it was stripped and reprobed with HL89 to detect sequence adjacent to the RNA primer.

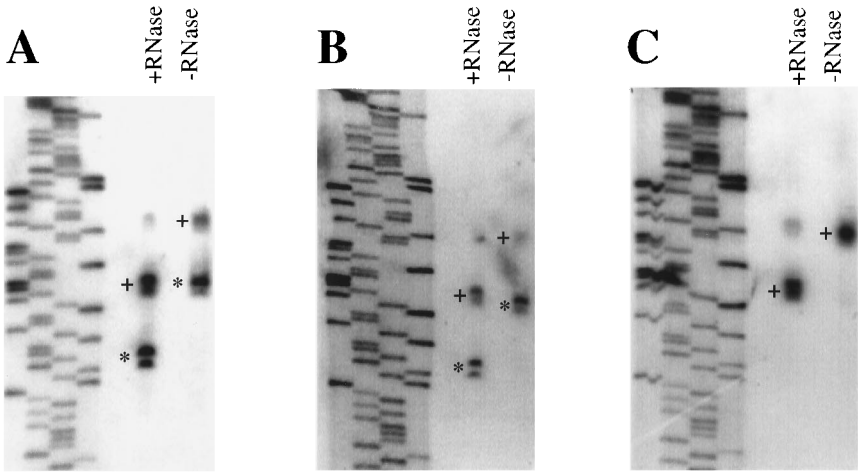

FIG. 6. Mapping the 3′ ends of the strong-stop species with DNA blots. The sizes of the minus-strand strong-stop doublet species, as well as the positions of their termini, were examined for information that might relate to the priming mechanism. Nucleic acid was extracted from Tf1 particles and loaded onto a sequence gel. Three sets of lanes from the gel were then electrotransferred onto three identical filters. The same sequence reaction mixture with HL78 was included on each filter for size comparisons. In addition, material from the particles was either pretreated with RNase A or loaded onto the gel after a mock incubation. Each blot was marked with a plus sign to indicate the position of the long strong-stop DNA and an asterisk to label the position of the short strong-stop DNA. (A) Filter probed with JB341. (B) Filter probed with HL92. (C) Filter probed with HL77.

shows that approximately equal amounts of the major strong-stop species were present in this set of blots.

DISCUSSION

Taken together, the mutations described above demonstrate the presence of sequences essential for transposition at the PBS as well as at the 5′ end of the transcript. While any model for reverse transcriptase priming would predict a detrimental effect of mutations in the PBS, there is no obvious aspect of cDNA synthesis that can account for essential sequences in the 5′ end of the transcript other than the self-priming model.

That transposition could be restored in four defective elements by simply recreating complementarity between the PBS and the 5′ end of the transcript supports the key feature of the self-priming model in Fig. 1B, namely, that the PBS must anneal to the 5′ end of the transcript for transposition to occur. The correlation between transposition activity and production of the minus-strand strong-stop DNA indicated that the mutations that blocked transposition did so at or before the priming step of reverse transcription. This is consistent with the prediction that the single mutations blocked transposition at the priming step and not during any later processes in the transposition pathway. The presence of wild-type amounts of particles in sucrose gradient fractions that also contained normal levels of Tf1 mRNA indicated that the point mutations were not blocking a step in transposition earlier than priming.

The size analysis of the strong-stop DNA and the identification by hybridization of the RNA primer provided strong support for the self-priming model. The size of the larger band in samples treated with RNase A was that predicted for a full-length strong-stop DNA that began at the end of the LTR and extended all the way to the 5′ end of the transcript. Although the sizes of the strong-stop species correlate well with predicted lengths, the broadness and, in the case of RNase A treatment, the doublet appearance of the bands make it difficult to rule out the possibility of heterogeneity in the strong-stop species. The blots probed with HL90 and HL89 suggested that the DNA sequences of the two major strong-stop species have similar 5′ ends that possess RNA primers that were identified as containing the sequence of the 5′ end of the transcript. Probes HL77 and HL92 produced signals indicating that several bases of DNA were absent from the 3′ end of the smaller strong-stop band. These results and those of the previously performed primer extension reactions (17) that showed uniform 5′ ends of the strong-stop sequence both strongly support the conclusion that the 11-base difference in sizes of the major strong-stop species was due to sequence absent from the 3′ end of the smaller species. The absence of 11 bases at the 3′ end of the strong-stop DNA should not interfere with subsequent steps in reverse transcription since the missing nucleotides are within the repeated R region of the LTR and would be regenerated from sequence in the 3′ LTR. The amount of the shorter strong-stop species suggested that about half of the Tf1 transcripts that served as template during reverse transcription were either modified or truncated 11 bases from their 5′ ends, destroying the ability of RT to copy the final 11 bases of mRNA into DNA. Although the looped structure shown in Fig. 1B cannot alone provide a primer for reverse transcription, one possibility is that an endonuclease cuts between the 11th and 12th bases of the transcript and exposes a free 3′ OH. This cleavage would have to occur in a small fraction of mRNA, such as the mRNA undergoing reverse transcription and RNase H degradation, since cleavage was not observed in particle-purified Tf1 transcripts as determined by primer extension (14). If cleavage of the transcript is the mechanism responsible for priming, the observation that half of the strong-stop material is full length implies that priming can occur in *trans.*

Another possible mechanism for the initiation of Tf1 cDNA synthesis is by 2′-OH priming of DNA with a nucleotide within the first 11 bases of the Tf1 transcript. 2′-OH self-priming of a transcript by reverse transcriptase has been demonstrated for bacterial retron elements from in vitro reactions with purified components (9, 12). If this mechanism is used for Tf1 priming, the presence of full-length strong-stop species suggests that either an enzyme can debranch the 2′-OH linkage or priming can occur in *trans.*

The analysis of sequences of other LTR-containing retroelements for putative looped structures that include the PBS sequences resulted in evidence that a family of retrotransposons, including CfT-I isolated from the fungal tomato patho-

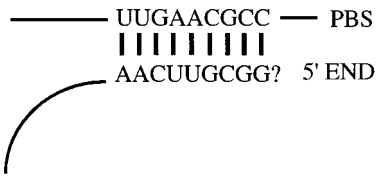

FIG. 7. Sequence of the putative self-priming region of CfT-I. This represents the putative self-priming structure of CfT-I containing 9 bases thought to be the PBS because of its location adjacent to the 3′ end of the 5′ LTR. The bottom 9 bases are within the LTR. ?, unknown position of the transcription start site.

gen *Cladosporium fulvum* (19), use a self-priming mechanism for reverse transcription. Alignments of reverse transcriptase sequences from all retroelements show that CfT-I is the closest known relative to Tf1 (22, 29). Interestingly, CfT-I has a poor tRNA candidate for a primer (19), and its transcript has the potential to loop back and form a 9-base duplex between the PBS (bases 544 to 552) and a region within the 5′ LTR (bases 343 to 351). This CfT-I duplex structure is shown in Fig. 7. Both of these regions are in positions within CfT-I that correspond to the two sites within Tf1. Although the site of transcription initiation is unknown for CfT-I, the 9 bases of perfect complementarity with the PBS region and the close phylogenetic relationship between Tf1 and CfT-I strongly suggest that these elements are members of a family that self-prime their reverse transcription. Tf2, another element from *S. pombe*, contains sequences highly related to Tf1 in the LTR, PBS, and Pol regions but has little similarity within the Gag section of the element (16). Tf2 appears to be a member of this family of self-priming transposons, since it contains the same self-complementary sequences as Tf1 (28). Because the self-priming mechanism is probably used in retroelements isolated from organisms as different as the fission yeast *S. pombe* and the filamentous phytopathogen *C. fulvum*, other examples of retroelements may exist that self-prime cDNA synthesis, and some of these could include retroviruses.

Other retroelements with novel mechanisms for priming reverse transcription include the hepadnaviruses, the mitochondrial plasmid Mauriceville, and the non-LTR retrotransposon R2Bm. The reverse transcriptases of hepadnaviruses use a covalent bond to themselves instead of a 3′-OH link to an RNA as a priming mechanism (26). In comparison, the reverse transcriptase encoded by the Mauriceville mitochondrial plasmid of *Neurospora* spp. can initiate cDNA synthesis de novo and appears to require no priming (27). R2Bm encodes a reverse transcriptase that cleaves the target site DNA and uses the newly created insertion site 3′ OH to prime cDNA synthesis without any complementarity between the primer and the template (18). These unusual mechanisms may represent evolutionary intermediates in the development of the Tf1 mechanism, which may in turn have evolved into the successful tRNA priming strategy used by other LTR-containing retroelements.

Despite the enormous diversity among retroviruses and LTR-containing retrotransposons, several properties are key to their lifestyle and are thus conserved among all known examples of these elements. Some of these properties include the LTRs, PBSs, and polypurine stretches, all of which are essential for cDNA synthesis. Because the retroviruses and LTR-containing retrotransposons that have been sequenced contain PBSs with significant complementarity to specific tRNAs, the possibility existed that some essential aspect of retroelement propagation placed special constraints on the structure of the primer that required a tRNA molecule. The self-priming of Tf1, a typical LTR-containing retrotransposon, demonstrates not only that tRNAs are not essential primers of minus-strand reverse transcription but also that a mobile element's own transcript can prime cDNA synthesis. The lack of dependence on the host for a specific tRNA primer raises the intriguing possibility that self-priming retroelements could transfer horizontally into other species, since they would bring their own primer with them.

ACKNOWLEDGMENTS

I thank Dan Voytas for pointing out that CfT-I possesses the same looped structure that Tf1 and Tf2 has. I am grateful to Michael Benedik, Jef Boeke, Tom Dever, Alan Hinnebusch, and Reed Wickner for their critical reading of the manuscript.

REFERENCES

1. **Boeke, J. D.** 1988. Retrotransposons, p. 59–103. *In* E. Domingo, J. J. Holland, and P. Ahlquist (ed.), RNA genetics, vol. II. Retroviruses, viroids, and RNA recombination. CRC Press, Inc., Boca Raton, Fla.
2. **Boeke, J. D., J. Trueheart, G. Natsoulis, and G. R. Fink.** 1987. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol. **154:**164–175.
3. **Chapman, K. B., A. S. Byström, and J. D. Boeke.** 1992. Initiator methionine tRNA is essential for Ty1 transposition in yeast. Proc. Natl. Acad. Sci. USA **89:**3236–3240.
4. **Curcio, M. J., and D. J. Garfinkel.** 1991. Single-step selection for Ty1 element retrotransposition. Proc. Natl. Acad. Sci. USA **88:**936–940.
5. **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.
6. **Eichinger, D. J., and J. D. Boeke.** 1990. A specific terminal structure is required for Ty1 transposition. Genes Dev. **4:**324–330.
7. **Gabriel, A., and J. D. Boeke (ed.).** 1993. Retrotransposon reverse transcription. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
8. **Garfinkel, D. J., J. D. Boeke, and G. R. Fink.** 1985. Ty element transposition: reverse transcriptase and virus-like particles. Cell **42:**507–517.
9. **Hsu, M., S. Eagle, S. Inouye, and M. Inouye.** 1992. Cell-free synthesis of the branched RNA-linked msDNA from retron-Ec67 of *Escherichia coli*. J. Biol. Chem. **267:**13823–13829.
10. **Keeney, J. B., K. B. Chapman, V. Lauermann, D. F. Voytas, S. V. Astrom, U. von Pawel-Rammingen, A. Bystrom, and J. D. Boeke.** 1995. Multiple molecular determinants for retrotransposition in a primer tRNA. Mol. Cell. Biol. **15:**217–226.
11. **Kikuchi, Y., Y. Ando, and T. Shiba.** 1986. Unusual priming mechanism of RNA-directed DNA synthesis in *copia* retrovirus-like particles of *Drosophila*. Nature (London) **323:**824–826.
12. **Lampson, B. C., M. Inouye, and S. Inouye.** 1989. Reverse transcriptase with concomitant ribonuclease H activity in the cell-free synthesis of branched RNA-linked msDNA of *Myxococcus xanthus*. Cell **56:**701–707.
13. **Leis, J., A. Aiyar, and D. Cobrinik.** 1993. Regulation of initiation of reverse transcription of retroviruses, p. 33–47. *In* A. Skalka and S. Goff (ed.), Reverse transcriptase. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
14. **Levin, H.** Unpublished data.
15. **Levin, H. L., and J. D. Boeke.** 1992. Demonstration of retrotransposition of the Tf1 element in fission yeast. EMBO J. **11:**1145–1153.
16. **Levin, H. L., D. C. Weaver, and J. D. Boeke.** 1990. Two related families of retrotransposons from *Schizosaccharomyces pombe*. Mol. Cell. Biol. **10:**6791–6798.
17. **Levin, H. L., D. C. Weaver, and J. D. Boeke.** 1993. Novel gene expression mechanism in a fission yeast retroelement: Tf1 proteins are derived from a single primary translation product. EMBO Journal **12:**4885–4895.
18. **Luan, D. D., M. H. Korman, J. L. Jakubczak, and T. H. Eickbush.** 1993. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell **72:**595–605.
19. **McHale, M. T., I. N. Roberts, S. M. Noble, C. Beaumont, M. P. Whitehead, D. Seth, and R. O. Oliver.** 1992. CfT-I: an LTR-retrotransposon in *Cladosporium fulvum*, a fungal pathogen of tomato. Mol. Gen. Genet. **233:**337–347.
20. **Moreno, S., A. Klar, and P. Nurse.** 1991. Molecular genetic analysis of fission yeast *Schizosaccharomyces pombe*. Methods Enzymol. **194:**795–823.
21. **Rose, M. D., F. Winston, and P. Hieter.** 1990. Methods in yeast genetics: a laboratory course manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Springer, M. S., and R. J. Britten.** 1993. Phylogenetic relationships of reverse transcriptase and RNase H sequences and aspects of genome structure in the gypsy group of retrotransposons. Mol. Biol. Evol. **10:**1370–1379.

23. **Sweeney, E., and H. Levin.** Unpublished data.
24. **Varmus, H., and P. Brown (ed.).** 1989. Retroviruses. American Society for Microbiology, Washington, D.C.
25. **Varmus, H. E., and R. Swanstrom.** 1984. Replication of retroviruses, p. 369–512. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses, vol. 1. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Wang, G. W., and C. Seeger.** 1992. The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis. Cell **71:**663–670.
27. **Wang, H., and A. Lambowitz.** 1993. The Mauriceville plasmid reverse transcriptase can initiate cDNA synthesis de novo and may be related to reverse transcriptase and DNA polymerase. Cell **75:**1071–1081.
28. **Weaver, D. C., G. V. Shpakovski, E. Caputo, H. L. Levin, and J. D. Boeke.** 1993. Sequence analysis of closely related retrotransposon families from fission yeast. Gene **131:**135–139.
29. **Xiong, Y., W. Burke, and T. Eickbush.** 1993. Pao, a highly divergent retrotransposable element form Bombyx mori containing long terminal repeats with tandem copies of the putative R region. Nucleic Acids Res. **21:**2117–2123.