Heterogeneous terminal structure of Ty1 and Ty3 reverse transcripts

Marcelle Wilhelm, Thérèse Heyman1, Sylvie Friant and François-Xavier Wilhelm*

Unité Propre de Recherche 9002 du Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, 67084 Strasbourg Cedex, France, 1Unité Mixte de Recherche 216 du Centre National de la Recherche Scientifique, Institut Curie Biologie, Centre Universitaire, 91405 Orsay Cedex, France

Received January 20, 1997; Revised and Accepted April 9, 1997

ABSTRACT

A specific terminal structure of preintegartive DNA is required for transposition of retroviruses and LTR-retrotransposons. We have used an anchored PCR technique to map the 3' ends of DNA intermediates synthesized inside yeast Ty1 and Ty3 retrotransposon virus-like particles. We find that, unlike retroviruses, Ty1 replicated DNA does not have two extra base pairs at its 3' ends. In contrast some Ty3 preintegartive DNA molecules have two extra nucleotides at the 3' end of upstream and downstream long terminal repeats. Moreover we find that some molecules of replicated Ty3 DNA have more than two extra nucleotides at the 3' end of the upstream LTR. This observation could be accounted for by imprecise RNase H cutting of the PPT sequence. The site of Ty1 and Ty3 plus-strand strong-stop DNA termination was also examined. Our results confirm that the prominent Ty1 and Ty3 plus-strand strong-stop molecules harbor 12 tRNA templated bases but also show that some Ty1 and Ty3 plus-strand strong-stop DNA molecules harbor less tRNA templated bases. We propose that these less than full length plus-strand molecules could be active intermediates in Ty retrotransposon replication.

INTRODUCTION

Retroviruses and retrotransposons such as yeast Tys and Drosophila copia elements replicate through RNA intermediates and alternate their genetic material between RNA and DNA (1–4). The genomic RNA of these mobile genetic elements is converted into double-stranded DNA by the process of reverse transcription. The replicated DNA is then integrated into the genomic DNA of the host cell where it can be transcribed to produce new molecules of genomic RNA. Synthesis of each strand of retrotransposon DNA begins with the synthesis of short DNA products called minus-strand and plus-strand strong-stop DNA. Minus-strand strong-stop DNA synthesis is initiated from the 3' hydroxyl group of a primer tRNA annealed at a primer binding site (PBS) located just downstream of the R-U5 sequence of the genomic RNA. Plus-strand strong-stop DNA synthesis commences from an RNase H resistant oligoribonucleotide spanning a purine rich sequence (the PPT, polypurine tract) located just upstream of the 3' U3-R sequence of the RNA. Minus-strand and plus-strand strong-stop DNA are elongated after they have moved from their sites of synthesis at one end of the template to an acceptor region at the other end of the template in a process called strand transfer.

As a consequence of the two strand transfers, the unique 3' U3 RNA sequence is duplicated at the 5' end of the preintegartive DNA to form the upstream LTR and the unique 5' U5 RNA sequence is duplicated at the 3' end of the DNA to form the downstream LTR. The final product of reverse transcription is a two LTRs linear double-stranded DNA molecule that is longer than the genomic RNA. For most retroviruses the replicated linear extrachromosomal DNA has 2 bp at each end which are not present at the end of integrated proviral DNA. The first step in retroviral integration is the cleavage of two nucleotides from the 3' end of preintegartive DNA. Thus, the 3' termini of retrovirus DNA intermediates represent a mixture of species with either two or no extra nucleotides. Recently Kirchner and Sandmeyer (5) have analyzed the ends of preintegartive Ty3 DNA and found that it has retrovirus-like two extra base pairs at each end. The terminal structure of Ty1 preintegartive DNA has not been directly examined. In the present study we have used an anchored PCR technique to map the 3' ends of DNA intermediates synthesized inside Ty1 and Ty3 VLPs. We find that Ty1 replicated DNA does not have two extra base pairs at its 3' ends whereas some Ty3 preintegartive DNA molecules have two extra nucleotides at the 3' ends of the upstream and downstream LTRs. Moreover we find that a few replicated extrachromosomal Ty3 molecules have more than two extra nucleotides at the 3' end of the upstream LTR. To explain this result we propose that the cleavage of the PPT primer by the RNase H activity of Ty reverse transcriptase is imprecise, leading to templated addition of extra nucleotides at the end of minus-strand after the second strand transfer.

The site of Ty1 and Ty3 plus-strand strong-stop DNA termination was also examined using the anchored PCR technique. In the case of Ty1, hybridization experiments (6,7) and our previous mapping results (8) indicated that most plus-strand strong-stop DNA molecules are extended into the tRNA primer. It has been recently reported that the majority of Ty1 and Ty3 plus-strand strong-stop DNA molecules harbor 12 tRNA templated

*To whom correspondence should be addressed. Tel: +33 3 8841 7006; Fax: +33 3 8860 2218; Email: wilhelm@ibmc.u-strasbg.fr
bases at the 3' end and may not be direct intermediates in retrotransposition (5,9). Our results confirm that the prominent Ty1 and Ty3 plus-strand strong-stop molecules harbor 12 tRNA templated bases but also show that some molecules harbor less tRNA templated bases. These less than full length plus-strand molecules could be active intermediates in Ty replication and could carry on a retrovirus-like second strand transfer involving complementarity between the tRNA inherited sequences and the PBS sequence at the end of the minus-strand DNA.

MATERIALS AND METHODS

Strains and plasmids

The yeast strain AGY9 (Mat α, his 4-539, lys 2-801, leu 2Δ1, trp 1Δ63, ura3-52, spt 3-202, GAL+) kindly provided by J.D. Boeke was used to minimize the amount of reverse transcript specified by endogenous Ty1 elements. Plasmid pJEF 1105 (10) kindly provided by J.D. Boeke is a high copy number (2 µm)
Figure 2. (A) Experimental approach used to analyze the 3′ ends of Ty DNA. (a) The DNA was extracted from purified VLPs. For simplicity only the DNA intermediate of step E in Figure 1 is represented. (b) and (c) After heat denaturation of the DNA intermediates, the 3′ ends were poly dA tailed by terminal nucleotidyltransferase and amplified by PCR using Ty specific primers and an oligo dT primer complementary to the added poly dA tail. d) The resulting PCR amplification products were digested by restriction enzyme R1 and R2 and subcloned in pSL1180 phagemid. e) DNA preparations of individual clones were used for double-stranded DNA sequencing. (B) Sequence analysis of PCR amplification products. Autoradiograms of four representative sequences of the 3′ ends of Ty3 plus-strand strong-stop DNA harboring 7, 10, 12 and 13 tRNA templated bases and three representative sequences of the 3′ ends of Ty3 minus-strand with 2, 4 and 5 extra nucleotides. The summary of 3′ end mapping of Ty1 and Ty3 DNA is given in Figures 3 and 4.

Analysis of Ty1-VLP DNA

AGY9 transformed with plasmid expressing Ty1 or Ty3 was used for VLPs production. Upon galactose induction, cells transformed by these plasmids produce large amounts of VLPs. Ty1-VLPs were isolated on a sucrose step gradient using a method described by Eichinger and Boeke (12) with minor modifications (7). Extraction of DNA from VLPs was as described (7).

Poly dA tailing, anchored PCR and subcloning of the amplified products

The poly dA tailing and the anchored PCR was done as described by Charneau et al. (13). Terminal nucleotidyltransferase was used to add homomer tails to DNA fragments extracted from purified VLPs. The DNA was heat-denatured at 95°C for 2 min and polydA tailing was achieved for 15 min at 37°C in a final volume of 20 µl in the presence of 50 U terminal nucleotidyltransferase EC2.7.7.31 (Boehringer Mannheim). 100 pmol of dATP, 0.75 mM CoCl2, 200 mM potassium cacodylate, 25 mM Tris–HCl pH 6.6, 0.25 mg/ml bovine serum albumin. The reaction was stopped by EDTA and the DNA was phenol/chloroform purified and precipitated by ethanol. The poly dA tailed DNA fragments were subjected to PCR amplification. The reaction conditions were 50 mM KCl, 10 mM Tris–HCl pH 8.4, 0.1 mg/ml gelatin, 1.5 mM MgCl2, 0.8 mM dNTPs, 16 ng/µl each of primers, 1 µl of DNA solution and Taq polymerase in a final volume of 100 µl. PCR amplification was carried out for 35 cycles at 92°C for 2 min, 50°C for 2 min and 72°C for 2 min. The resulting PCR amplification products were extracted with phenol/chloroform and digested by EcoRI and HindIII or KpnI and HindIII. The digested products were ligated to EcoRI and HindIII or KpnI and HindIII cut pSL1180 phagemid. Competent Escherichia coli cells were transformed. Small-scale phagemid DNA preparation of individual clones were used for double-stranded DNA sequencing.

Oligonucleotide primers

Oligonucleotide primers were synthesized using an Applied Biosystem 381 A DNA synthesizer. The following primers were used to amplify plus-strand 3′ ends, Ty1: oligo dT primer
Figure 3: Distribution of Ty1 and Ty3 plus-strand DNA 3′ ends. (A) Schematic structure of minus-strand DNA annealed to plus-strand initiated at PPT and double stranded full-length preintegrative DNA. The question marks indicate the 3′ ends which were characterized. (B) The graphs represent the number of cloned Ty1 and Ty3 fragments which terminate at the positions indicated by arrows above the sequence of Ty1 and Ty3 plus-strand DNA.

5′-ATCGAAGCTTTTTTTTTTTTTTT-3′ and primer 275 (positions 7254–7273 of pJEF1105) 5′-ATCGAATTCAAGAATTGTTGTA-GAATTGCAG-3′ or primer 276 (positions 7000–7020 of pJEF1105) 5′-ATCGAATTCAACACTGGCAGAGCATTAC-GC-3′. Ty3: oligo dT primer and primer 2321 (positions 5274–5295 of Ty3) 5′-ATCGAAATTCACTGGTTACTTCCCTAAGAC-3′.

The following primers were used to amplify the minus-strand 3′ ends, Ty1: oligo dT primer and primer 7680 (positions 1074–1093 of pJEF1105) 5′-TACGGTACCTTGGTTTTGGGTCA-TCA-TGC-3′. Ty3: oligo dT primer and primer 1845 (positions 817–843 of Ty3) 5′-TATGAAATTCCGCTGGCATTTCTGTGCATGC-3′.

RESULTS

The overall flow of the LTR retroelement replication process is presented in Figure 1. Question marks in step E and G indicate the 3′ ends of DNA intermediates which were analyzed using the anchored PCR technique described in Figure 2. In step E the 3′ end of plus-strand strong-stop DNA was mapped to determine how far in the tRNA molecule plus-strand synthesis is proceeding. In step G the question mark indicates that the anchored PCR technique was used to examine the terminal structure of Ty1 and Ty3 DNA after completion of full length DNA synthesis.

3′ ends of Ty1 and Ty3 plus-strand DNA

Specific Ty1 or Ty3 primers complementary to sequences in the U3 region and an oligo dT primer complementary to the 3′ poly dA tail were used to amplify the 3′ end of plus-strand DNA. This combination of primers (U3 + dT) allows amplification of the ends of plus-strand strong-stop DNA as well as full length plus-strand DNA and plus-strand DNA initiated at the central PPT2 sequence in the case of Ty1 (7,8). The 3′ ends of Ty1 plus-strand DNA were also amplified with a combination of primers (a primer located upstream of U3 and the poly dT primer) which allows separate analysis of the ends of full length plus-strand DNA.

3′ ends of Ty1 plus-strand. Figure 3 shows the distribution of 3′ ends of Ty1 plus-strand DNA analyzed with the U3 + dT primers. About 75% of the molecules terminate at the end of the LTR. The remaining 25% plus-strand molecules harbor tRNA-templated bases presumably generated by reverse transcription of the primer tRNA molecule attached to the minus-strand template. When the 3′ ends of Ty1 were analyzed with primers which allow separate analysis of the ends of full length plus-strand (6), we find that all full length plus-strand DNA fragments terminate at the end of the LTR (data not shown). This result is in agreement with sequence examination of Ty1 elements (14) predicting that the preintegrative DNA should not have two extra bases at the 3′
Figure 4. Distribution of Ty3 minus-strand DNA 3′ ends. (A) Schematic structure of double stranded full-length preintegrative DNA. The question mark indicates the 3′ ends which have been characterized. (B) The graph represents the number of cloned Ty3 fragments which terminate at the positions indicated by arrows above the sequence of Ty3 minus-strand DNA.

end of the downstream LTR since the site of priming of minus-strand which determines the 3′ end of downstream LTR coincides with the end of the LTR.

The 3′ ends of plus-strand strong-stop DNA were analyzed separately after recovering the 0.345 kb strong-stop fragment from an agarose gel. We find that the majority of plus-strand strong-stop DNA harbors 12 tRNA-templated bases, i.e. two bases more than the Ty1 PBS sequence. However, we also find that some molecules harbor less tRNA templated bases or terminate at the end of the LTR. One molecule has an extra base at the end of the 12 tRNA templated bases. Given that the PBS of Ty3 has only 8 nucleotides of complementarity to the tRNA primer, strong-stop fragments harboring more than 8 tRNA templated bases would clearly not be direct intermediates for a retrovirus-like second strand transfer. Only the few molecules harboring less tRNA templated bases (5 and 7) could hybridize without mismatches with the PBS after strand transfer to the 3′ end of the elongated minus-strand DNA.

3′ ends of Ty3 plus-strand. The 3′ ends of Ty3 plus-strand were analyzed with the combination of primers U3 + dT which allows amplification of all plus-strand intermediates including full-length plus-strand DNA. The result of the 3′ end mapping is shown in Figure 3. About half of the molecules harbor tRNA templated bases or terminate at the end of the LTR (Fig. 3B). One molecule has an extra base at the end of the 12 tRNA templated bases which could have been templated by base A64 of tRNA Met or could be a RT mediated terminal non-templated base as suggested by Gabriel et al. (17).

Figure 5. Plus-strand synthesis of Ty3 and mechanism of addition of extra nucleotides at the 3′ end of the upstream LTR. Plus-strand synthesis is initiated from the PPT fragment created by RNase H digestion of the genomic RNA (step E of Fig. 1). The end of the upstream LTR is determined by the site of initiation of plus-strand synthesis at the 3′ end of the PPT primer. (a) Synthesis of strong-stop plus-strand DNA. (Inset) To explain that the end of the upstream LTR is heterogeneous we propose that the cleavage of the PPT sequence by RNase H is imprecise resulting in PPT sequence of variable length. Three potential PPT sequences attached to the strong-stop plus-strand DNA are shown (uppercase letters: plus-strand DNA; lowercase letters: PPT primer). (b) The PPT primer is removed by RNase H cleavage. Plus-strand DNA is transferred at the end of minus-strand. (Inset) End sequence of transferred plus-strand DNA. (c) Reverse transcriptase completes DNA synthesis. (Inset) The extra nucleotides at the end of plus-strand DNA are copied resulting in a heterogeneous end structure of the upstream LTR.

3′ ends of Ty1 and Ty3 minus-strand DNA

3′ ends of Ty1 minus-strand. A specific Ty1 primer located downstream of the PBS sequence and the oligo dT primer were used to amplify the 3′ ends of Ty1 minus-strand DNA. The 3′ end of 20 Ty1 full length minus-strand DNA molecules were analyzed. We find that all molecules terminate at the end of the LTR (data not shown). No molecules terminate with two extra nucleotides at the end of the LTR. Thus the 3′ termini of full length minus-strand DNA coincides with the end of the LTR. This result is in agreement with the prediction from the sequence of Ty1 retrotransposon that preintegrative Ty1 DNA should have the same terminus as that of the integrated element.

3′ ends of Ty3 minus-strand. The distribution of Ty3 minus-strand 3′ ends is illustrated in Figure 4. The majority of molecules
terminate at the end of the LTR. A few molecules have two extra nucleotides CT at the end of the LTR. Surprisingly, molecules with 4, 5 and 14 extra nucleotides complementary to the PPT sequence have also been characterized. The 3′ end of minus-strand DNA is determined by the site of initiation of plus-strand synthesis at the 3′ end of the upstream LTR. We propose that incorrect cleavage of the 3′ end of Ty1 upstream LTR are likely dead-end products in the Ty3 life cycle.

**REFERENCES**