

Expression and characterization of a novel reverse transcriptase of the LTR retrotransposon Tf1

Noa Kirshenboim ^{a,1}, Zvi Hayouka ^b, Assaf Friedler ^b, Amnon Hizi ^{a,*,2}

^a Department of Cell and Developmental Biology, Sackler School of Medicine, Tel-Aviv University, Tel-Aviv, 69978, Israel

^b Department of Organic Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

Received 14 February 2007; returned to author for revision 15 March 2007; accepted 3 April 2007

Available online 23 May 2007

Abstract

The LTR retrotransposon of *Schizosacharomyces pombe*, Tf1, has several distinctive properties that can be related to the unique properties of its reverse transcriptase (RT). Consequently, we expressed, purified and studied the recombinant Tf1 RT. This monomeric protein possesses all activities typical to RTs: DNA and RNA-dependent DNA polymerase as well as an inherent ribonuclease H. The DNA polymerase activity shows preference to Mn⁺² or Mg⁺², depending on the substrate used, whereas the ribonuclease H strongly prefers Mn⁺². The most outstanding feature of Tf1 RT is its capacity to add non-templated nucleotides to the 3'-ends of the nascent DNA. This is mainly apparent in the presence of Mn⁺², as is the noticeable low fidelity of DNA synthesis. In all, Tf1 RT has a marked infidelity in synthesizing DNA at template ends, a phenomenon that can explain, as discussed herein, some of the features of Tf1 replication in the host cells.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Tf1 LTR retrotransposon; Retroelements; Retroviruses; Reverse transcriptase; RNA-dependent DNA polymerase; DNA-dependent DNA polymerase; Ribonuclease H; Non-templated addition of dNTPs

Introduction

Long terminal repeat (LTR) retrotransposons are a diverse group of retrovirus-like elements found in a broad variety of eukaryotic cells (Boeke and Devine, 1998; Coffin et al., 1997; Craig et al., 2002; Havecker et al., 2004). Their structure and mechanism of propagation are related to retroviruses and, similar to them, they encode the Gag, protease, reverse transcriptase (RT) and integrase (IN) proteins. As in all retroviruses, the RT of LTR retrotransposons plays a vital role in converting the (+) sense and single-stranded RNA genome into a double-stranded and integration-competent viral DNA. This reverse transcription process is catalyzed by the two activities of RT, the DNA polymerase (that can copy both RNA and DNA) and the ribonuclease H (RNase H), which

concomitantly cleaves the RNA strand in the DNA•RNA heteroduplex formed. The double-stranded DNA product is subsequently integrated into the host cell genomic DNA by IN. Similar to all retroviruses, in LTR retrotransposons, the synthesis of the (–) sense DNA strand is initiated from an RNA primer that is complementary to a specific genomic RNA sequence, which is located near the 5'-end of the RNA genome and is designated primer binding site (PBS) (Coffin et al., 1997; Le Grice, 2003; Skalka and Goff, 1993).

In retroviruses, the (–) strand primer initiates reverse transcription in a position that leaves two or three nucleotides beyond the 3'-end of the downstream LTR U5 sequence that are removed prior to the integration by a 3'-end processing activity of the viral IN. Conversely, in many LTR retrotransposons, such as Ty1, reverse transcription starts with DNA synthesis directly from the (–) strand primers without any addition of extra nucleotides. Hence, the integration is thought to proceed directly without any preceding processing, as was also suggested for Ty1 IN (Boeke and Devine, 1998; Eichinger and Boeke, 1990; Moore et al., 1995). This assumption contradicts the results of a very recent study of LTR

* Corresponding author. Fax: +972 36407432.

E-mail address: ahizy@post.tau.ac.il (A. Hizi).

¹ The results presented are in partial fulfillment of a Ph.D. thesis at Tel-Aviv University.

² A. Hizi is an incumbent of The Gregorio and Dora Shapira Chair for the Research of Malignancies at Tel-Aviv University.

retrotransposon Tf1 of the fission yeast *Schizosacharomyces pombe* (*S. pombe*). The cDNA species, produced *in vivo* in virus like particles (VLPs), which were isolated from *S. pombe* cells, exhibited a surprising heterogeneity of lengths (Atwood-Moore and Levin, 2005; Atwood-Moore et al., 2006). The dominant species contain the entire LTR and two extra nucleotides at the U3 end (as “T”s) in the upstream LTR that are probably templated by the last two nucleotides of the polypurine track (PPT). These observations suggest that the RNase H activity of Tf1 RT might be impaired or regulated. Another surprising feature of these Tf1 cDNA molecules was that a large portion of them terminated with one to eight non-templated nucleotides that are largely random in their sequence. While retroviruses and LTR retrotransposons are known to have cDNA species with non-templated additions, 85% for Tf1 cDNA molecules is the largest percentage ever observed among the various retroelements studied (Atwood-Moore and Levin, 2005; Atwood-Moore et al., 2006). The presence of such sequences can be explained by an unusual activity of Tf1 RT to add non-templated nucleotides. It should be noted that RTs are known to have a deoxynucleotidyl terminal transferase activity, which adds non-templated nucleotides to the 3'-end of the nascent DNA strand (Boutabout et al., 2001; Golinelli and Hughes, 2002; Patel and Preston, 1994; Peliska and Benkovic, 1992), though at a very low efficiency and certainly not as high as could be expected for Tf1 RT.

The prevailing notion for retroviruses and LTR retrotransposons (such as Ty1 and Ty3 of the yeast *Saccharomyces cerevisiae*) was that the synthesis of the (–) strand DNA is initiated from the 3'-end of a specific cellular tRNA, which serves as the primer (Coffin et al., 1997; Craig et al., 2002; Le Grice, 2003). A different priming mechanism, which does not involve the participation of any host tRNA, was proposed for initiating the synthesis of (–) strand DNA in Tf1 (Craig et al., 2002; Levin, 1997; Lin and Levin, 1997, 1998). This model suggests that a self-complementarity within the RNA genome induces intramolecular base-pairing with the PBS near the 5'-end of the RNA, causing this RNA end to fold back; thus, the scission of the duplex RNA provides a short 11 nucleotides self primer for the synthesis of the (–) strand DNA. *In vivo* studies of Tf1 have suggested that such a cleavage occurs between the 11th and 12th bases of the RNA transcript, generating the primer. Subsequent mutagenesis experiments implied that the RNase H domain of Tf1 RT may be required for this specific cleavage (Levin, 1996, 1997). A similar self priming mechanism for (–) strand DNA synthesis has been later proposed for a variety of LTR retroelements, such as the highly homologous Tf2 and Maggy, Skippy, Cft-1, Boty, copia of maize and Tf1/shushi of vertebrates (Butler et al., 2001). They all belong to a single lineage of the Ty3/gypsy group of LTR retrotransposons (Malik and Eickbush, 1999; Craig et al., 2002). These elements have probably diverged early in the evolution of LTR retrotransposons before retroviruses; thus, it is likely that the mechanism of self priming represents an early form of initiating reverse transcription in retroelements.

For all these reasons outlined, it was of interest to express the recombinant Tf1 RT and to study its catalytic features in

comparison with other retroviral RTs. We report here on the expression, purification and the study of the catalytic properties of the novel RT of Tf1 that belongs to the unique group of LTR retroelements. We show that this RT has many of the basic properties typical to RTs; however, it has also an unusual terminal transferase activity of adding non-templated nucleotides to the 3'-ends of the nascent DNA synthesized by the enzyme.

Results

Plasmids construction for the recombinant expression of Tf1 RT

As far as we know, the only RTs from LTR retrotransposons that were studied extensively *in vitro* were those of Ty1 and Ty3 (Pandey et al., 2004; Rausch et al., 2000; Wilhelm et al., 2005; Mules et al., 1998) that belong to groups different than that of Tf1, the Pseudoviridae and the Metaviridae, respectively. Given the great diversity of RT sequences among retroelements, it is not surprising that the amino acids sequence homologies between the pol genome of Tf1 and those of other LTR retrotransposons, let alone retroviruses, are low. Thus, the pairwise alignment scores calculated for the homology between Tf1 RT and the RTs of Ty3, Ty2, Ty1, human immunodeficiency virus type-1 (HIV-1) and HIV-2 were 22, 6, 9, 14 and 13%, respectively. Therefore, the design of recombinant Tf1 RT was based on several parameters: 1, previous information on Tf1 proteins; 2, the partial homology with the closest RT studied, that of Ty3; and 3, preliminary experiments conducted by us (to find which of the recombinant Tf1 RT versions generated has the highest DNA polymerase activity). We have already expressed in bacteria and studied the properties of Tf1 IN (Hizi and Levin, 2005). This protein starts at its amino terminus with the sequence H₂N-TDDFNQK-, coded by nucleotide 3100 and downstream of the Tf1 genome. Furthermore, based on the assumption, which is true for most LTR retroelements and retroviruses, that the RT protein is coded by the sequences located immediately upstream to the IN-coding region, the carboxyl terminus of the RT was chosen to be -NFVNQISI-COOH. To decide what the preferred amino terminus of the RT is, we have engineered two versions of Tf1 RT. The longer Tf1 RT version of 545 residues started at its amino terminus with H₂N-TTKLIINLN- (encoded by nucleotide 1465 and downstream). The shorter RT version is altogether 505 residues long, starting with the sequence encoded by nucleotide 1585; thus its amino terminus is H₂N-ISSSKHTL-. Expression plasmids were generated and the relevant bacterial clones were analyzed for Tf1 RT expression, using two methods. First, we have assessed the protein expression levels, employing Western blot analyses with anti six-histidine antibodies (as all RT versions were engineered to have a six-histidine tag on their amino termini). Then, the positive clones were tested for the recombinant RT-associated RNA-dependent DNA polymerase (RDDP) activity in bacterial lysates, as done earlier by us (Hizi and Hughes, 1988; Perach and Hizi, 1999; Taube et al., 1998). The expression level of the

longer 545-residues version was low with a very low RDDP activity. On the other hand, the 505 residues-long RT was well expressed with a high RDDP activity (data not shown). Consequently, we have proceeded by expressing, purifying and analyzing the 505 residues long RT. This protein is fused at its amino terminus to six-histidine plus three residues, derived from the coding sequence of the expressing plasmid; therefore, it is altogether 513 residues long. The plasmid designed to express this protein, pT5m-6His-Tf1 RT, induced the synthesis of a polypeptide with the expected molecular weight of approximately 56 kDa (see Fig. 1).

Purification of Tf1 RT

We have found that the levels of expression of recombinant Tf1 RT in bacteria strongly depend on the temperature under which the bacteria are grown and the duration of the induction employed (data not shown). This may result from the fact that the novel Tf1 RT is probably more sensitive to proteolysis by bacterial enzymes compared to other recombinant RTs previously expressed by us, those of HIV-1 and HIV-2, equine infectious anemia virus (EIAV), murine leukemia virus (MuLV), mouse mammary tumor virus (MMTV), bovine leukemia virus (BLV), porcine endogenous retrovirus (PERV) and bovine immunodeficiency virus (BIV) (Avidan et al., 2003, 2006; Hizi et al., 1988, 1991; Perach and Hizi, 1999; Taube et al., 1998). The optimal conditions determined for the induction of recombinant Tf1 RT are detailed under Materials and methods. Tf1 RT was purified to homogeneity, employing an affinity chromatography on Ni-NTA agarose, designed to selectively bind the six-histidine tag-containing proteins, followed by diethylaminoethyl (DEAE)-Sephacel ion-exchange chromatography and a second Ni-NTA agarose column (Fig. 1 and Table 1). A quantitative analysis of the purification steps is shown in Table 1, indicating an overall increase in the RT specific activity by almost 6000 fold. In all purification steps, there is an increase in the specific DNA polymerase activity of the RT, which is due to the combination of removing bacterial proteins and, not less importantly, to an increase in the total polymerase activity. The increase in the

Table 1

The quantitative pattern of Tf1 RT purification from bacterial extract

Purification step	Total (mg)	Total DNA polymerase activity (units)	Specific DNA polymerase activity (units/mg protein)
Crude bacterial extract	82	1.15×10^4	1.4×10^2 (1)
Ni-NTA agarose column	12.2	2.7×10^4	2.2×10^3 (16)
DEAE-Sephacel column	7.6	1.8×10^5	2.3×10^5 (170)
Ni-NTA agarose column	3.5	2.9×10^6	8.3×10^6 (5982)

Note. DNA polymerase activity is expressed in units. One RT unit is defined as the amount leading to the incorporation of 1 pmol of [3 H]dTTP into the trichloroacetic acid-insoluble material for 30 min at 37 °C in poly(rA)_n•oligo(dT)₁₂₋₁₈-dependent reaction. The numbers in parenthesis are the purification fold, namely the factor by which the specific RDDP activity of the purified protein increased relative to that of the crude bacterial extract.

total RT activity precludes the calculation of the relative recovery of the RT during the purification process. We have observed previously a similar phenomenon of an overall increase in the total DNA polymerase activity throughout the purification of the recombinant RTs of MMTV (Taube et al., 1998), BLV (Perach and Hizi, 1999) PERV (Avidan et al., 2003) and BIV (Avidan et al., 2006). This phenomenon can be explained by a removal of bacterial components that strongly inhibit RT activity.

Subunit composition of the Tf1 RT

Several retroviral RTs display a heterodimeric structure, where the smaller subunits share the same amino terminus with the larger ones, as they are derived from the larger subunits due to a cleavage at the carboxyl termini by the viral protease (Coffin et al., 1997; Skalka and Goff, 1993). This heterodimeric structure is typical to the RTs of lentiviruses and of the alpha retroviruses (such as avian sarcoma/leukosis virus, ASLV). Other RTs, for instance those of the beta retrovirus MMTV (Taube et al., 1998), the gamma retroviruses MuLV and PERV (Avidan et al., 2003) and the delta retrovirus BLV (Perach and

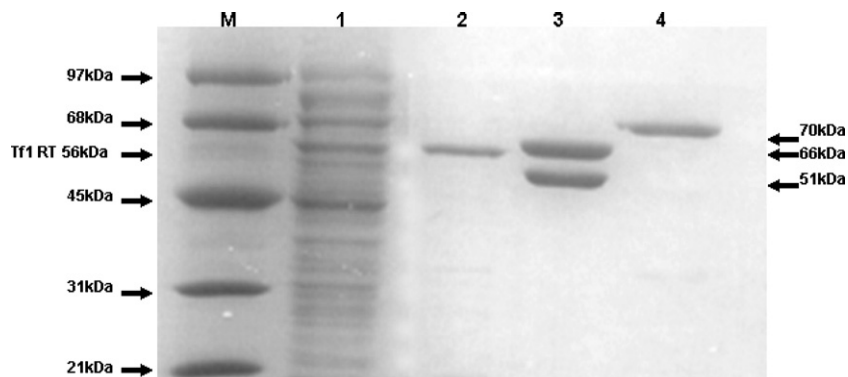


Fig. 1. SDS/PAGE analysis of the purified recombinant Tf1 RT. The purification of Tf1 RT was performed as describe in Materials and methods. Tf1 RT was analyzed by 10% SDS/PAGE and was stained with Coomassie Brilliant Blue. The molecular size of Tf1 RT is approximately 56 kDa, and those of HIV-1 RT and the MuLV RT, which were used as references, are 51/66 kDa and 70 kDa, respectively. M, molecular size markers. 1. Crude bacterial extract of the Tf1 RT-expressing bacteria. 2. Purified Tf1 RT. 3. Purified HIV-1 RT. 4. Purified MuLV RT.

Hizi, 1999), are all composed of a single subunit, hence show a monomeric organization. Likewise, the LTR retrotransposon Ty3 RT was shown to be monomeric (Rausch et al., 2000). The results of the size exclusion chromatography, shown in Fig. 2, indicate that Tfl RT behaves in solution as a protein with an apparent molecular weight of approximately 57 kDa. This figure is in an excellent agreement with the predicted size of a monomeric 513-long polypeptide (see above) as well as with the apparent subunit mass of approximately 56 kDa shown by a PAGE analysis of the purified protein (Fig. 1).

The relative DNA polymerase activities of Tfl RT

The enzymatic activities of the purified Tfl RT were analyzed in the presence of Mn^{+2} or Mg^{+2} , allowing a comparison of the activities and divalent cation preferences of the RT, to define the optimal assay conditions. In a preliminary study, we have found that the highest activities are obtained in the presence of either 5 mM Mg^{+2} or 0.5 mM Mn^{+2} (data not shown). Therefore, these were the divalent cation concentra-

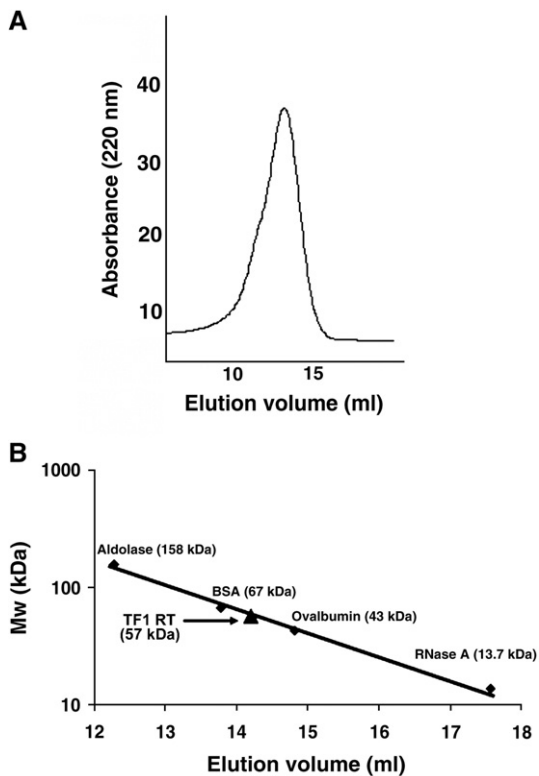


Fig. 2. Sizing purified Tfl RT by analytical gel filtration. (A) Analytical gel filtration of Tfl RT (10 μ M) was performed on an AKTA Explorer (Pharmacia Corp) using a Superose 12 analytical column 30 \times 1 cm (GE Healthcare-Pharmacia Corp) equilibrated with 10 mM Tris-HCl, pH 7.5, 25 mM NaCl, 0.1% Triton, 2 mM DTT, 1 mM EDTA and 20% glycerol. Protein was eluted with a flow rate of 1 ml/min at 4 $^{\circ}$ C and the elution profile was monitored by UV absorbance at 220 nm. The column was calibrated with molecular weight protein standards (GE Healthcare-Pharmacia Corp). (B) Determination of Tfl RT subunit stoichiometry by the size exclusion chromatography. Molecular weights of standard proteins have been indicated on the calibration curve.

Table 2

Specific catalytic activities, substrate and divalent cation preferences of Tfl RT

RT activity	Substrate	MgCl ₂	MnCl ₂
DDDP	dA•dT	184.6	1248.8
	dC•dG	3626.3	1155.3
	Activated DNA	1072.5	360
	Φ x174 phage DNA	35.4	41.4
RDDP	rA•dT	839.3	19453.2
	rC•dG	833	342.4
	16S Ribosomal RNA	70.2	52.1
RNase H	[³ H]rA•dT	7.8	175.4

Note. The specific activities of the DNA polymerase activities were expressed in pmol of dNTP incorporation into DNA in 30 min at 37 $^{\circ}$ C per μ g purified RT. For activated DNA, the dNTP incorporation was calculated for all dNTPs assuming equimolar incorporation. The specific RNase H activity was expressed in pmol [³H]AMP released from [³H]poly(rA)_n•poly(dT)_n in 30 min at 37 $^{\circ}$ C per μ g purified protein. The different catalytic activities were determined in the presence of either 5 mM Mg^{+2} or 0.5 mM Mn^{+2} . The values presented are averages calculated from four independent experiments (after subtracting background levels with no RT present in the reactions) with standard deviations of around 10%. The specific activities values marked in bold letters are the maximal ones obtained for each RT for every given substrate with either Mg^{+2} or Mn^{+2} .

tions used for testing the activities with different substrates. The relative specific activities of Tfl RT with a variety of substrates that have been used before to characterize RT activities (Avidan et al., 2003, 2006; Perach and Hizi, 1999) are presented in Table 2. Like all RTs studied so far, Tfl RT has RDDP and DNA-dependent DNA polymerase (DDDP) activities as well as an inherent RNase H activity. Previous data show that retroviral RTs differ in their divalent cation preferences that depend on the substrate used for assaying RT activities (Avidan et al., 2003; Perach and Hizi, 1999; Skalka and Goff, 1993; Taube et al., 1998). Similarly, we show here that Tfl RT has a strong preference for Mn^{+2} with poly(rA)_n•oligo(dT)_{12–18} and poly(dA)_n•oligo(dT)_{12–18}, whereas with all other substrates there is a substantial or a mild preference towards Mg^{+2} (and with Φ X174am3 template DNA there is no significant difference between the two divalent cations) (Table 2). Like the RDDP function of almost all studied RTs, Tfl RT exhibits the highest activity with poly(rA)_n•oligo(dT)_{12–18} and the substrate second in order is poly(rC)_n•oligo(dG)_{12–18}. For DDDP, the comparable order of preferences is as follows: poly(dC)_n•oligo(dG)_{12–18} > poly(dA)_n•oligo(dT)_{12–18} \approx activated DNA > Φ X174am3 DNA. As to the RNase H activity of Tfl RT, the apparent [³H]poly(rA)_n•poly(dT)_n-directed activity shows a high preference, of more than 20-fold, for Mn^{+2} over Mg^{+2} , which is even more pronounced in the second RNase H assay system used by us (see below).

Steady-state kinetic parameters of the DNA polymerase activity of Tfl RT

The kinetic parameters for the incorporation of the relevant dNTPs into the nascent DNA strand were determined from the double-reciprocal linear plots of the various initial enzymatic velocities as a function of varying substrate concentrations (data

not shown). For RDDP activity, we have tested variable concentrations of dTTP with poly(rA)_n•oligo(dT)_{12–18} and variable dGTP concentrations with poly(rC)_n•oligo(dG)_{12–18}. For the DDDP function, dTTP or dGTP were the variable substrates with activated gapped double-stranded DNA employed as template primer (with dATP and dCTP at a constant concentration). To examine the steady-state kinetic parameters under optimal conditions, the divalent cation used with the template primers was the preferable one; Mn²⁺ for poly(rA)_n•oligo(dT)_{12–18} and Mg²⁺ for poly(rC)_n•oligo(dG)_{12–18} and activated DNA (see Table 2). Mn²⁺ was also used in determining the steady-state kinetic constants of the RNase H activity. The results presented in Table 3 show that, with the synthetic DNA polymerase substrates, dTTP has a *K_m* value lower than that calculated for dGTP, whereas the opposite is true for activated DNA.

DNA synthesis under processive and nonprocessive conditions

During enzymatic polymerization reactions, the lengths of the nascent polymeric products, which are formed before the enzyme molecules dissociate from them, define the processivity of the polymerase (Avidan and Hizi, 1998; Von Hippel et al., 1994). The extent of product elongation in a single round of synthesis depends on parameters such as binding, single nucleotide addition, translocation and pausing. Retroviral RTs are inefficient in performing fully processive polymerization events (Avidan and Hizi, 1998; Avidan et al., 2003). Therefore, we have tested the processivity of the novel Tfl RT by examining its capacity to synthesize DNA (Fig. 3). This was done with a template of single-stranded ΦX174am3 DNA, primed with a short 5'-end-labeled DNA oligonucleotide, in the presence of Mg²⁺ or Mn²⁺ as the divalent cation. The reactions were performed in comparison with the well-studied RTs of HIV-1 and MuLV, each in the presence of its preferred divalent cation (Avidan and Hizi, 1998; Avidan et al., 2003). The extension of the 5'-end [³²P] labeled primer was performed in the absence or presence of a DNA trap. After preincubating the

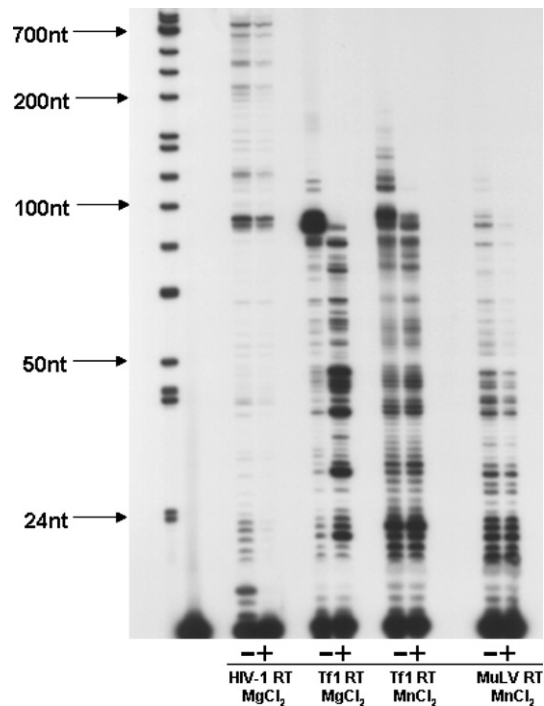


Fig. 3. DNA synthesis by Tfl RT under processive and nonprocessive conditions. Reactions were performed using 15nt synthetic 5'-end-labeled primer and an excess of the template single-stranded circular ΦX174am3 phage DNA. The sequence of the primer is shown in Materials and methods. (–), DNA extension performed without DNA trap; (+), DNA extension performed with 0.6 mg/ml of unlabeled trap of activated DNA. Reactions were done using either 5 mM MgCl₂ or 0.5 mM MnCl₂. Molecular mass marker was *Hinf*I-cleaved dephosphorylated double-stranded ΦX174am3 DNA fragments labeled with γ [³²P]ATP at their 5'-end by polynucleotide kinase.

Table 3
Steady-state kinetic parameters for DNA polymerase activities of Tfl RT

	rA · dT (dTTP)	rC · dG (dGTP)	Activated DNA (dTTP)	Activated DNA (dGTP)	RNase H
<i>K_m</i> (μM)	17.9	44.6	12.5	4.06	0.48
<i>K_{cat}</i> (s ⁻¹)	0.087	0.018	0.014	6.6 × 10 ⁻³	0.03
<i>K_{cat}/K_m</i> (s ⁻¹ M ⁻¹)	4.8 × 10 ³	4.0 × 10 ²	1.1 × 10 ³	1.7 × 10 ³	6.25 × 10 ⁴

Note. The RDDP and DDDP activities of Tfl RT were assayed with the specific substrates and variable concentrations of either dTTP or dGTP, as described in "Material and methods". All reactions were carried out for 10 min at 37 °C. The values, calculated for the dNTPs, were determined from the double reciprocal plots of the various substrate concentrations against the initial velocities of DNA synthesis. The plots were generated using linear regression analysis, yielding regression coefficients around 0.98 (not shown). The *K_{cat}* values are the steady-state rate coefficients calculated from the *V_{max}* values divided by the moles of the enzyme in the reaction.

enzymes with the labeled template primer (and prior to initiating polymerization by adding the four dNTPs), an excess of unlabeled activated DNA was added to trap the enzyme after its initial dissociation from the template primer (that takes place after finishing the first cycle of primer extension). As this DNA trap prevents further extensions of the primer due to re-association of the RT, each end-labeled nascent DNA strand is almost certainly restricted to a single round of DNA elongation. To prove this point, all tested RTs (those of Tfl, HIV-1 and MuLV) make longer DNA products when multiple rounds of synthesis are permitted without a DNA trap. The RTs were calibrated to have equal DDDP activities with ΦX174am3 DNA and the reactions were performed in the presence of either Mg²⁺ or Mn²⁺ (Fig. 3). As shown in Table 2, in the Tfl RT-directed DDDP reaction with ΦX174am3 DNA there is a slight preference for Mn²⁺ over Mg²⁺ indeed the results obtained in the processivity experiment support this conclusion (Fig. 3).

The extension products were also quantified and the elongation was calculated as a percentage of the total amount of both unextended and extended primers (Table 4). It is apparent that Tfl RT extends the majority of the 5'-end-labeled primer molecules, when multiple rounds of DNA synthesis are permitted (i.e., without a DNA trap). This is true for reactions conducted in the presence of Mg²⁺ or Mn²⁺ (of about 75% or

Table 4
Quantitative analysis of DNA primer extension by Tfl RT under processive and nonprocessive conditions

Enzyme	Product length (nt)	MgCl ₂			MnCl ₂		
		No trap	With trap	Ratio	No trap	With trap	Ratio
Tfl RT	16–50	23.8	45.4	1.9	50.9	47.9	0.9
	51–100	47.2	17.2	0.4	38.5	22.4	0.6
	101–200	4.5	0	0	8.7	0	0
	201–700	0	0	0	0	0	0
	>700	0	0	0	0	0	0
	Overall extension	75.5	62.6	0.83	98.1	70.3	0.72
HIV-1 RT	16–50	33.2	12.6	0.4			
	51–100	12.7	15.5	1.2			
	101–200	6.4	4.2	0.7			
	201–700	13.7	10.9	0.8			
	>700	1.22	0	0			
	Overall extension	66.9	43.2	0.65			
MuLV RT	16–50				62.2	64.5	1.03
	51–100				9.4	2.4	0.25
	101–200				1.6	0	0
	201–700				0	0	0
	>700				0	0	0
	Overall extension				73.2	66.9	0.91

Note. The radioactivity in the DNA bands in all polynucleotide length ranges was summed up and the values were divided by the sums of all extended and unextended primers (detecting in the scanning of the gels autoradiogram as shown in Fig. 3). The values are expressed in percentages of the extended primers of the total amount of the DNA product of each range. The calculations were conducted separately for gel lanes of reactions carried out in the absence or presence of an excess of the unlabeled DNA trap. The extensions in the presence of the DNA trap were divided by the comparable figures (obtained with no trap present), yielding the ratio values. The ratios calculated for the overall extensions are the relative processivity values. The values are the means calculated from two independent experiments (one of which is shown in Fig. 3) and the variations were less than 10%.

98%, respectively). In comparison, both HIV-1 RT (in the presence of Mg²⁺) and MuLV RT (in the presence of Mn²⁺) extended a lower percentage of the primer (67% and 73%, respectively). However, similar to MuLV RT, almost all products of Tfl RT are shorter than 100 nt, whereas more than 20% of the HIV-1 RT-synthesized products are longer than 100 nt. As can be seen in Fig. 3 and similar to other RTs (Avidan and Hizi, 1998; Avidan et al., 2002, 2003; Taube et al., 1998), there are specific pausings during DNA synthesis. These stop sites and the length of the products do not appear to depend on the divalent cation present. Taken together, the data suggest that, when multiple rounds of synthesis are allowed, the translocation (or what we define as “persistence of synthesis”) is lower for Tfl RT relative to HIV-1 RT and is slightly better for Tfl RT than for MuLV RT. The ratio between the products generated with the DNA trap relative to those synthesized without a trap indicates the relative overall processivity of DNA synthesis (see Table 4). It seems that the relative processivity of Tfl RT in the presence of Mn²⁺ is about the same as that observed with Mg²⁺ (of about 0.7 and 0.8, respectively). On the other hand, the processivity of Tfl RT is higher than that observed with HIV-1 RT (0.65) and lower than that of MuLV RT (0.9).

Non-templated nucleotide addition while copying DNA template ends by Tfl RT

A surprising feature of the Tfl cDNA molecules (isolated from Tfl VLPs in *S. pombe* cells) was that a large portion terminated at their 3'-ends with up to eight non-templated nucleotides with largely random sequences. Retroviruses and LTR retrotransposons are known to have cDNA species with non-templated additions; however, an apparent value of 85% for Tfl cDNA molecules are the largest percentage ever observed (Atwood-Moore and Levin, 2005; Atwood-Moore et al., 2006). The presence of such sequences can be explained by an unusual activity of Tfl RT to add non-templated nucleotides during DNA synthesis. RTs are known to have a deoxynucleotidyl terminal transferase activity, though at a very low efficiency that can hardly explain the high incidence of non-templated additional nucleotides observed *in vivo* in Tfl-transfected cells. Therefore, it was of interest to test whether the novel Tfl RT can be responsible for this phenomenon by exhibiting such an activity and whether the Tfl RT-associated activity is higher than that of HIV-1 RT.

The experiment described in Fig. 4 shows the elongation patterns of a 5'-end-labeled primer that forms a blunt 3'-end with the complementary DNA strand. The Tfl RT-directed reactions were performed with either Mn²⁺ or Mg²⁺ and with different combinations of dNTPs, all in comparison with an equal DNA polymerase activity of HIV-1 RT. The results indicate that Tfl RT has a substantial terminal transferase activity that is more pronounced in the presence of Mn²⁺. This activity is higher than the comparable activity of HIV-1 RT that shows little differences between the two divalent cations tested. The primer extension data were also quantified, as shown in Table 5. As to the additions of a single dNTP by Tfl RT with Mn²⁺, it is apparent that the products generated with dCTP or dTTP (as well as with all four dNTPs) are longer by a single nucleotide from the original 21 nt-long substrate (and slightly more than one half of the substrate was elongated) (Fig. 4A, Table 5). In contrast, with dATP, up to five consecutive nucleotides were added to the primer, while with dGTP there is an elongation of all primer molecules by up to two nucleotides (the majority of these molecules were extended by a single nucleotide). In all, the pattern of the non-templated addition in the presence of Mn²⁺, with a single dNTP present in the reaction mixture, is dATP > dGTP > dCTP ≈ dTTP ≈ all four dNTPs. A similar preference for adding purines over pyrimidines, with either Mn²⁺ or Mg²⁺, was evident also for the RTs of HIV-1 (Patel and Preston, 1994) and Tyl (Boutabout et al., 2001), albeit with a significantly lower efficiency compared to the Mn²⁺-dependent behavior of Tfl RT. Interestingly, with Tfl RT, the level of extension with all four dNTPs is significantly lower than that observed with either dATP or dGTP (Fig. 4A, Table 5), although the four dNTPs mixture contained the same concentrations of dATP and dGTP as in the single dNTP reactions. Since the concentrations of the four dNTPs were equimolar, it might be that the less efficiently incorporated pyrimidines (dCTP and dTTP) inhibit the addition of the more readily incorporated purines. In comparison, HIV-1 RT does not

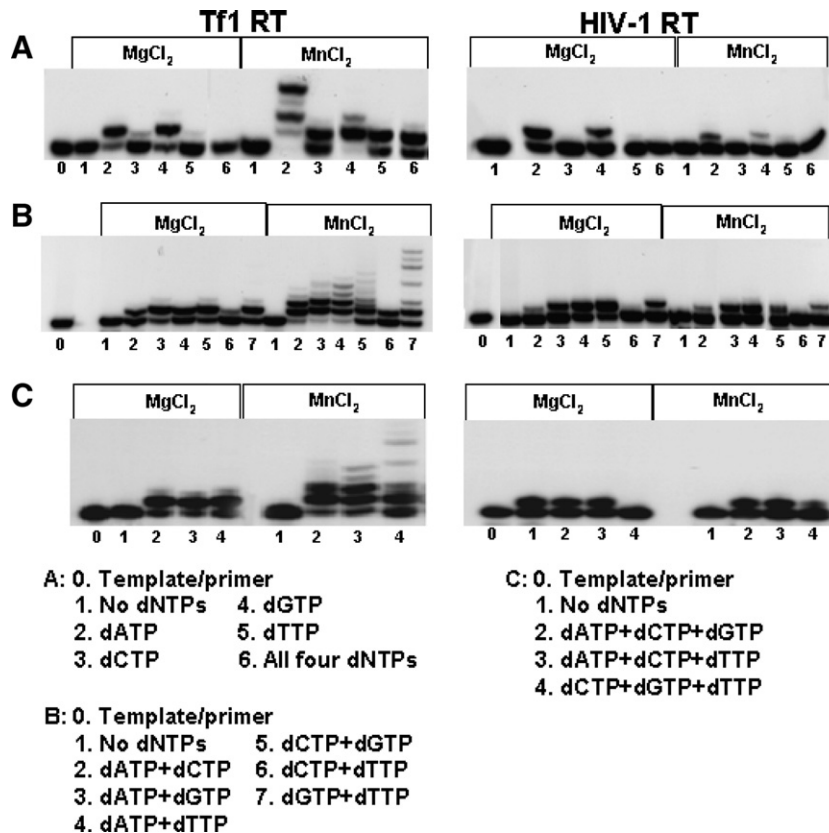


Fig. 4. Non-templated DNA synthesis by Tfl RT. We have followed the extension of a 21 bp double-stranded and blunt-ended DNA (with the U5-end sequence of HIV-1 DNA). The 5'-³²P end-labeled DNA was incubated with 50 ng of either Tfl RT or HIV-1 RT (with either 5 mM MgCl₂ or 0.5 mM MnCl₂) and the indicated combinations of dNTPs (each at a final concentration of 500 μM). All reactions components except for dNTPs were preincubating for 5 min at 37 °C. The reaction started by adding dNTPs followed by incubation for 30 min at 37 °C. The reaction products were analyzed by urea-PAGE and the extensions of the 5'-end-labeled primer are seen from an autoradiography of the gels.

show any significant elongation with dCTP, dTTP and all four dNTPs (Fig. 4A).

The pattern of incorporation becomes more complex when two dNTPs are present in each reaction (Fig. 4B and Table 5). Again, it is obvious that the non-templated addition by Tfl RT is more pronounced with Mn²⁺ compared to Mg²⁺ and that this activity is significantly higher than that of HIV-1 RT. The most efficient activity is always obtained when dATP is present (with an elongation of up to 5 nucleotides with a combination of dATP and dTTP) (Fig. 4B). Interestingly, the longest extension (by up to 8 nucleotides), although not necessarily the most efficient one, is seen with dGTP+dTTP (Table 5). This highlights an interesting discrimination between the most efficient non-templated additions and the longest extensions observed. There are differences between the efficiency of adding non-templated nucleotides altogether and the “persistence” of this activity (in terms of how longer can the primer be extended). Finally, when three dNTPs are present, the longest extensions (up to 7 nucleotides) are evident with dCTP+dGTP+dTTP (Fig. 4C). In summary, it is clear that Tfl RT is more efficient than the well-studied HIV-1 RT in incorporating non-templated nucleotides into the nascent DNA strand. The reasons for the unique preferences for a single nucleotide over a mixture of double, triple or all four nucleotides, by this exceptional activity, will be further studied by us.

The fidelity of DNA synthesis by Tfl RT

The fidelity of DNA synthesis was studied *in vitro* with many retroviral RTs, suggesting that RTs have relatively low fidelity. Similar to all RTs studied so far, Tfl RT lacks a 3' → 5' exonuclease proofreading activity, thereby permitting a direct kinetic analysis of primer extension. It was shown that the primary parameters for fidelity depend largely on the sequences of the nucleic acids copied, rather than on whether DNA or RNA templates are used (Bakhanashvili and Hizi, 1993a, 1993b; Rubinek et al., 1997; Taube et al., 1998). Consequently, we have analyzed DNA templates, as representing both DNA and RNA substrates.

The fidelity of 3'-end misinsertion was studied by following the incorporation of incorrect dNTPs opposite to the terminal template “A” nucleotide, in comparison with the incorporation of the correct dTTP (see Fig. 5 and Materials and methods). The reactions were performed with Tfl RT in the presence of Mn²⁺ or Mg²⁺ as the divalent cation, in comparison with the well studied HIV-1 RT (with Mg²⁺ as its preferred divalent cation). In the presence of the correct dNTP, dTTP, Tfl RT elongates the 5'-end-labeled primer by only a single dNTP with no further extensions, indicating no apparent misinsertions. However, when the wrong dNTPs are present, there are substantial differences in the extent of misincorporation by Tfl RT with the

Table 5
Quantitative analysis of the extension of the blunt-ended DNA by non-templated dNTP addition by Tfl RT

Nucleotide	MgCl ₂		MnCl ₂	
	Nucleotides	%	Nucleotides	%
dATP	1	57	5	97
dCTP	1	10	1	55
dGTP	1	50	2	98
dTTP	0	2	1	69
Mix dNTPs	0	0	1	67
dATP+dCTP	1	38	3	78
dATP+dGTP	2	62	4	92
dATP+dTTP	1	45	5	89
dCTP+dGTP	2	55	4	53
dCTP+dTTP	1	16	1	41
dGTP+dTTP	1	52	8	64
dATP+dCTP+dGTP	2	58	4	67
dATP+dCTP+dTTP	2	59	4	64
dCTP+dTTP+dGTP	2	51	7	55

Note. Quantification of the elongation by Tfl RT of the 21 nt blunt-ended template primer by the non-templated addition (see Materials and methods and Fig. 4). The nucleotide combinations that were used are described in the left column, the number of bases that were added to the template primer was derived from Fig. 4. The efficiency of the reaction is presented as the percentage of extension of the primer used for elongation by Tfl RT (100% represents the complete elongation of the primer) Quantification was done by ImageJ program.

two studied divalent cations. It is apparent that in the presence of Mn⁺² the levels of mispair formation (along with further elongation of the generated mispairs) are considerably higher than those observed with Mg⁺². The most noticeable elongation is with dATP, forming A•A mispair that is further elongated to form the correct A•T pair. This is followed by the formation of the incorrect A•G pair (18 nt) and by a three correct A•T pairs (19–21 nt length) that are further extended by the A•C and A•A mispairs (22 and 23 nt in lengths). With dCTP, the extent of elongation is lower than that with dATP, first forming a C•A mispair, which is further elongated to form C•T (17 nt) and

finally creating the correct C•G pair (18 nt). With the dGTP, there is only one apparent misinsertion forming the mispair G•A with no further extensions.

Unlike HIV-1 RT, Tfl RT shows a relatively low misincorporation in the presence of Mg⁺², when the highest extent of misinsertion is obtained with dCTP, forming a C•A mispair.

The RNase H activity

Sequence and structural analyses have shown that the RNase H domain is located in the carboxyl terminus of RTs, whereas the DNA polymerase domain is located at the amino terminal portion of the molecule. To confirm that Tfl RT has an RNase H activity and to characterize it, two assay methods were employed (Gao et al., 1998; Hizi et al., 1991; Sevilya et al., 2001). In the first one, the release of [³H] AMP was followed as a result of digestion of the [³H]poly(rA)_n•poly(dT)_n substrate. In the second, the pattern of cleavage of labeled RNA with a defined length in an RNA•DNA heteroduplex has been followed (see Materials and methods). The data show that Tfl RT possesses, like all RTs studied so far, an RNase H activity with a substantial preference for Mn⁺² over Mg⁺², but at lower efficiencies compared to that of HIV-1 RT (Fig. 6B, Tables 2 and 6).

The pattern of the cleavage of the 5'-end-labeled 267 nt RNA was used to characterize the RNase H activity of the RTs of HIV-1 and HIV-2 (Gao et al., 1998; Sevilya et al., 2001) as well as those of PERV, MuLV (Avidan et al., 2003), BIV (Avidan et al., 2006) and MMTV (Entin-Meer et al., 2002). Most RTs perform two cleavages. First, a primary one takes place at a position 17 nt away from the position corresponding to the 3'-end of the DNA heteroduplex (thus, producing in the specific assay employed, a 47 nt-long RNA product). Since it is assumed that the DNA polymerase domain of the RT binds the 3'-end of the DNA primer, this primary cleavage suggests that the

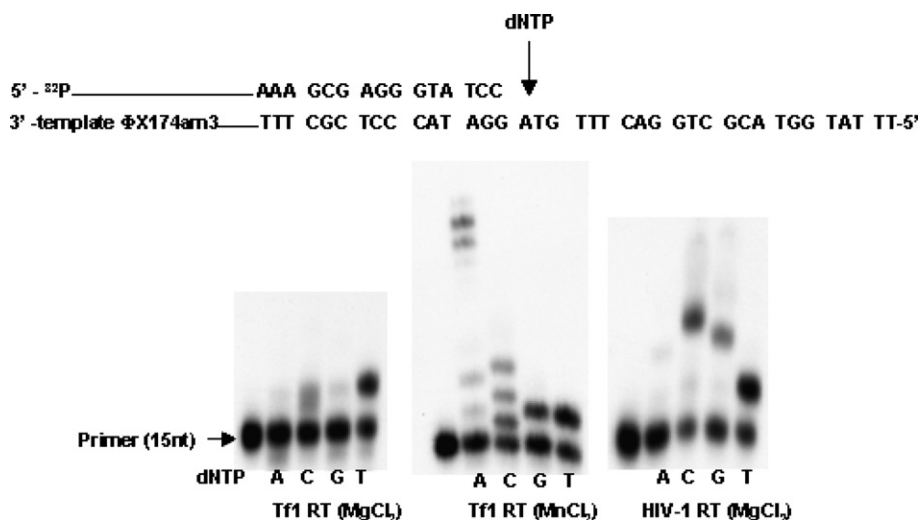


Fig. 5. The fidelity of misincorporation extension by Tfl RT. A single-strand circular ΦX174am3 DNA that was used as a template was annealed to the [³²P]-5'-end-labeled 15 nt DNA-primer. The primer was extended with an equal DNA polymerase activity of either Tfl RT or HIV-1 RT, in the presence of 1 mM of the single incorrect dNTP (A, C or G), or with 1 μM of the correct dNTP, dTTP. The sequence of the primer is shown on the top.

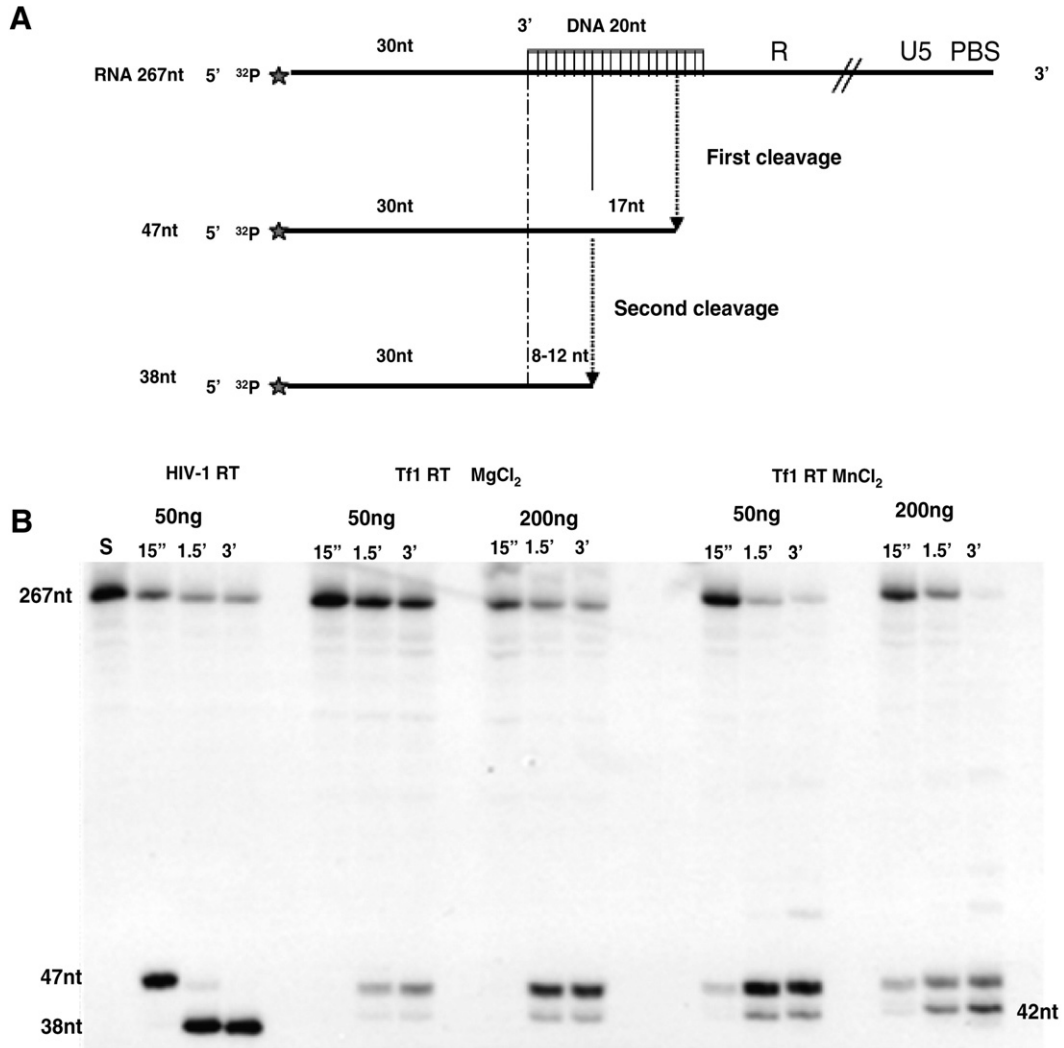


Fig. 6. The RNase H activity of Tf1 RT. (A) A scheme describing the RNase H reaction employed. (B) The RNase H activity of Tf1 RT was compared with that of HIV-1 RT. 50 ng and 200 ng of Tf1 RT, in a final volume of 10 μ l, were incubated with the RNase H substrate (the 5'-end-labeled 267 nt RNA and oligomeric DNA) at 37 $^{\circ}$ C for the indicated times (15 s, 1.5 min and 3 min), as described in Materials and methods. The reactions were performed with either 5 mM $MgCl_2$ or 0.5 mM $MnCl_2$. The products were resolved by 8% urea-PAGE.

molecular distance between the DNA polymerase active site and the RNase H active site equals to 17–18 nt. In many RTs, there is also a secondary RNase H cleavage at a position that corresponds to 8 nt away from the 3'-end of the DNA (producing in this assay 38 nt-long RNA molecules). This probably results from a secondary repositioning of the RT molecules on the RNA•DNA heteroduplexes after performing the primary cleavage (Gao et al., 1998; Palaniappan et al., 1996; Sevilya et al., 2001; Wisniewski et al., 2000). The primary cleavage performed by HIV-1, HIV-2, MMTV, MuLV and PERV RTs produces a similar 47 nt-long product (Avidan et al., 2003; Gao et al., 1999; Sevilya et al., 2001; Wisniewski et al., 2000). As seen in Fig. 6B, Tf1 RT exhibits a pattern of RNase H cleavage that is somewhat different than the one obtained with the well-studied HIV-1 RT. The first cleavage is indeed 47 nt-long RNA, suggesting that Tf1 RT behaves similar to the above mentioned RTs. Therefore, in Tf1 RT, similar to most RTs studied, the putative spatial distance between the RNase H and

the polymerase active sites equals to 17–18 nt. The only retroviral RT that exhibited a different molecular distance between the two active sites is the one from the lentivirus, BIV, recently studied by us (Avidan et al., 2006). With this enzyme the apparent distance is 20 nt. This indicates that, despite the much lower sequence homology to HIV RTs, Tf1 RT behaves in this manner closer to these RTs than BIV RT does (although BIV RT has a much higher sequence homology to the other lentiviral RTs studied). Interestingly, the secondary RNase H cleavage performed by Tf1 RT produces 42 nt-long RNA molecules that correspond to a 12 nt cleavage rather than the 8 nt cleavage produced by HIV-1 RT as well as by other RTs (Fig. 6B). This suggests that the repositioning of Tf1 RT, after performing the primary RNase H cleavage, is different than that of the RTs from HIV-1, HIV-2, MMTV as well as from BIV (Entin-Meer et al., 2002; Sevilya et al., 2001).

Finally, we have performed a quantitative analysis of the data obtained in three independent experiments (one of which is

Table 6
Quantitative analysis of RNase H cleavage products of Tf1 and HIV-1 RTs

	HIV-1 RT (50 ng)			Tf1 RT (50 ng)			Tf1 RT (200 ng)			Tf1 RT (50 ng)			Tf1 RT (200 ng)		
	MgCl ₂	1.5'	3'	MgCl ₂	1.5'	3'	MgCl ₂	1.5'	3'	MgCl ₂	1.5'	3'	MnCl ₂	1.5'	3'
1st cleavage	66.5±0.6	9.7±0.3	0	0	26.6±0.8	33.3±0.4	0	52.1±1.8	52.9±1.2	22.0±1.1	58.2±0.6	62.6±2.1	27.1±0.7	36.3±0.2	40.1±1.2
2nd cleavage	0	62±1.2	72.1±1.7	0	0	8.2±0.6	0	17.3±0.9	18.9±0.3	0	21.9±0.8	27.1±1.3	7.7±0.6	29.5±0.4	47.3±0.8
Total cleavage	66.5±0.6	71.7±1.5	72.1±1.7	0	26.6±0.8	41.5±1	0	69.4±2.7	71.8±1.5	22.0±1.1	80.1±1.4	89.7±3.4	34.7±1.3	65.8±0.6	87.4±2.0

Note. The radioactivity in the RNA bands of both first and second cleavage products of the 267 nt substrate were divided by the total amounts, which are the sums of all cleaved and uncleaved labeled RNA bands. The values of the cleavage products are expressed as percentages of the total amounts. The values are the average of three independent experiments.

presented in Fig. 6B, see Table 6). The data presented in Table 6 suggest that the specific enzymatic activity of Tf1 RT is significantly lower than that of HIV-1 RT. It is apparent that with 50 ng of Tf1 RT (in the presence of Mn²⁺) after 15 s of incubation about 22% of the substrate was cleaved. In comparison, about 66% of the substrate was cleaved by 50 ng HIV-1 RT. After longer incubation periods, the amounts of products generated by Tf1 RT were closer to those formed by HIV-1 RT as the reaction probably reaches saturation (HIV-1 RT converts all the 47 nt primary cleavage products into the secondary 38 nt products, even when there is still a significant amount of the uncleaved 267 nt substrate). This sequence of events was also observed with other RTs studied (Entin-Meer et al., 2002; Sevilya et al., 2001, 2003a, 2003b). On the other hand, the hydrolysis of the 267 nt substrate by Tf1 RT proceeds all through the incubation with very little accumulation of the secondary 42 nt cleavage products at the expense of the 47 nt products. Such a disparity in the kinetics can be explained by differences in the processivity of the cleavage. Thus, HIV-1 RT may have a higher processivity than that of Tf1 RT. According to this proposition, almost every molecule that undergoes a primary cleavage by HIV-1 RT is eventually cleaved again to create the shorter 38 nt products. While with Tf1 RT, after forming the first cleavage, most of the RT molecules fall off the 47 nt products and are recycled by interacting with additional uncleaved 267 nt molecules (rather than proceeding to create secondary cleavages). It will be of interest to study in the future these differences in the RNase H activity.

Discussion

The study presented here is devoted to the initial characterization of the enzymatic properties of the novel recombinant Tf1 RT. Tf1 belongs to a single lineage Ty3/gypsy group of LTR retrotransposons that probably diverged early in the evolution of LTR retrotransposons, well before retroviruses (Craig et al., 2002; Malik and Eickbush, 1999). However, even within this group of retrotransposons, the sequence homology between the RTs is relatively low, as evident from the homology score of 22% calculated for the RTs of Tf1 and Ty3. So far, the only LTR retrotransposon RTs studied in depth were those of Ty3 and Ty1 that have even a lower homology score of 9% (Pandey et al., 2004; Rausch et al., 2000; Wilhelm et al., 2005). Therefore, given the unique predicted properties of Tf1 RT (Atwood-Moore and Levin, 2005; Atwood-Moore et al., 2006; Levin, 1997; Lin and Levin, 1997), it was of interest to express and study this RT and to compare it to the other well studied RTs.

We have found here that the novel Tf1 RT has all activities expected from an RT, namely RNA and DNA-dependent DNA polymerase as well as an RNase H function. The most unique activity of Tf1 RT is its relatively high terminal transferase activity, as seen from its ability to add non-templated nucleotides to the blunt 3'-end of the nascent DNA strand (Fig. 4). This result is compatible with the *in vivo* study indicating that the ends of the cDNA molecules in VLPs have a surprising heterogeneity at their 3'-ends, and a large portion is

terminated with largely random sequences (Atwood-Moore and Levin, 2005; Atwood-Moore et al., 2006). This phenomenon could be explained by an exceptional terminal transferase activity of Tfl RT, a prediction supported by the study presented here.

Two major questions may arise from the high terminal transferase activity of Tfl RT. The first is why does Tfl RT need such an activity that is higher than that of other studied RTs of LTR retroelements? A possible answer is that these non-templated extra nucleotides are required to protect the 3'-ends of the cDNA from degradation by non-specific cellular 3' exonucleases. Alternatively, they may prevent aberrant annealing of the 3'-end due to micro-homology to sequences in the opposite LTR and the subsequent extensions. The obvious question arising from such an answer is how other retroelements manage without such a high RT activity? Since, as far as we know, the only RTs studied thoroughly for their terminal transferase activity are those of HIV-1 (Golinelli and Hughes, 2002; Patel and Preston, 1994; Peliska and Benkovic, 1992) and Ty1 (Boutabout et al., 2001), it might be that other RTs still possess a capacity to add non-templated nucleotides that is comparable to that reported herein for Tfl RT. Another explanation might be that RTs from other LTR retroelements do not really have such a pronounced terminal transferase activity since in cells transfected by these elements the levels of the non-specific nucleases are lower than those found in *S. pombe*. It might be as well that other cellular factors contribute to the protection of the cDNA ends of these other LTR retroelements and, hence, there is no need for a protection by excessive additions of non-templated nucleotides. These issues and related ones should be studied in the future.

A second question arises as well from the biochemical data presented here and from the *in vivo* data on Tfl (Atwood-Moore and Levin, 2005; Atwood-Moore et al., 2006). It is well established that in order for the cDNA, produced by the Tfl RT to be integration-competent, the extra 3'-end cDNA sequences must be removed prior to the integration by IN into the genomic host cell DNA. So, what mechanisms allow the removal of these sequences? One alternative is that cellular factors perform this sequence deletion. However, a highly likely possibility is that this activity is supplied by IN itself; hence, it is possible that unique features of Tfl IN compensate for the terminal transferase activity of RT by possessing a high 3'-end processing activity. It should be noted that, after the completion of cDNA synthesis in infected cells, most INs of retrotransposons studied can remove two or three nucleotides from the 3'-end of the viral DNA. Interestingly, we have found recently that, unlike these other studied INs, Tfl IN is capable of removing relatively long nucleotide extensions beyond the highly conserved "CA" sequence that is critical to the 3'-ends (Hizi and Levin, 2005). This supports the suggestion that Tfl IN can indeed compensate for the ability of RT to incorporate non-templated nucleotides to the 3'-ends of the nascent DNA.

The experiment conducted to assess the fidelity of DNA synthesis by Tfl RT indicates that Tfl RT has marked capacity of 3'-end misinsertion, mainly in the presence of Mn^{+2} (Fig. 5).

This infidelity is even higher than that of HIV-1 RT, which is considered to be less accurate than most other RTs studied (Bakhanashvili and Hizi, 1992, 1993b). Such an apparent low fidelity may suggest a potentially important mechanism for the *in vivo* generation of mutations in LTR retrotransposons.

The RTs of LTR retrotransposons studied most extensively are those of Ty1 and Ty3 that show pairwise alignment scores with Tfl RT of 9% and 22%, respectively (see above). The data previously reported for Ty1 RT (in the presence of 20 mM Mg^{+2}) show that it is less error than lentiviral RTs (of HIV-1 or HIV-2), whereas the ability to add non-templated nucleotides is roughly similar to that of HIV-1 RT (Boutabout et al., 2001). This is evidently different from the same two activities of Tfl RT. However, we have noticed the exceptional high terminal transferase activity and the low fidelity of Tfl RT mainly in the presence of Mn^{+2} rather than Mg^{+2} . Therefore, as no data were presented for Ty1 RT for the Mn^{+2} -dependent reactions, it is difficult to assess how different these two RTs are. The Ty1 RT-directed RNase H products indicate that unlike Tfl RT and most other RTs studied (where the cleavage takes place 17–18 nucleotides downstream from the nucleotide complementary to the 3'-end), the primary RNA cleavage of Ty1 RT is only 14–15 nucleotides (Wilhelm et al., 2005). This suggests that, like most other RTs, the molecular distance between the DNA polymerase and the RNase H active sites of Tfl RT equals to 17–18 nt, whereas the comparable distance in Ty1 RT is shorter (14–15 nucleotides). Interestingly, Ty3 RT shows a different cleavage pattern by its intrinsic RNase H, suggesting that the spatial distance between the active sites of the DNA polymerase and RNase H is higher (Rausch et al., 2000).

The data presented here for Tfl RT show that this protein is monomeric in solution (Fig. 2), as was also found for the RTs of Ty3, MuLV, PERV, BLV and MMTV, while other RTs from lentiviruses and ASLV RT are heterodimeric. It was suggested that the difference in the architectural organizations of MuLV RT and HIV-1 RT lies in a 27 residue-long peptide sequence that is located in the connection domain of MuLV RT and is absent in HIV-1 RT (Pandey et al., 2001). As we could not detect such a sequence in the putative connection subdomain of Tfl RT, there must be other reasons for the different molecular organizations of RTs. Still, it is likely that the three-dimensional structure of all RTs, including the novel Tfl RT, is similar to those of HIV and MuLV RTs that were extensively studied. In these RTs the folded DNA polymerase domain (located in the amino terminal portion) resembles a right hand, with "fingers", "palm" and "thumb" subdomains. This domain is connected through the "connection" subdomain to the RNase H domain, located in the carboxyl terminal part of the protein (Steitz, 1993; Ding et al., 1998; Geordiadis et al., 1995).

In summary, the new recombinant RT of Tfl shows *in vitro* some distinctive features that may explain various *in vivo* findings with this LTR retrotransposon, most notably the marked infidelity in synthesizing DNA at template ends. The features presented herein should be further studied to improve our understanding of the molecular mechanisms involved in the unique catalytic properties of Tfl RT. To this end, structure function relationship will be conducted to define the relevant

RT segments associated with these unique features. Other potential properties, suggested earlier for Tfl RT (Atwood-Moore and Levin, 2005; Atwood-Moore et al., 2006; Levin, 1997; Lin and Levin, 1997), will be also studied. Indeed, preliminary experiments conducted by us with Tfl RT show that this enzyme has an unusual capacity to cleave an 11 nt self primer from the 5'-end of Tfl genomic RNA and to use this RNA as a primer for DNA synthesis. Another study will look into the possibility that the RNase H activity associated with Tfl RT is impaired in removing the primers used for the synthesis of both the (–) sense and (+) sense DNA strands. Studies of this kind on Tfl RT can allow a better understanding of the reverse transcriptases in general and may explain how these enzymes have evolved from retrotransposons to retroviruses.

Materials and methods

Construction of the Tfl RT-expressing vector

The plasmid containing the Tfl proviral sequences was a generous gift from Dr. H. Levin from NIH. The Tfl wild-type RT used in this work is 505 amino acids long (plus a six-histidine tag at its amino terminus) that was expressed in *E. coli* and purified from bacterial extract. The Tfl RT gene was derived from the pET15b provirus-containing plasmid that was generated from the Tfl–107 NCYC 132 strain from *S. pombe* (Levin et al., 1990). The RT-encoding gene was synthesized by PCR, using the high fidelity Dynazyme polymerase (Finnzymes). The two primers used to synthesize the gene encoding the shorter 505-residue Tfl RT version studied here are as follows: the 5'-end primer: 5'-GCGGCCGCGCCATATGATATCTAGCAGTAAACACACGCTCTCTC-3', and the 3'-end primer: 5'-CGCGGAAGCTTCTATATCGAGATTTGATTAACAAAGATAC-3'.

The Tfl RT-encoding segments (nucleotides 1585–3099 of the proviral Tfl genome) were subcloned into pT5m(6-His-NdeI) plasmid that encodes a six-histidine tag at the amino terminus, using the *NdeI*-*HindIII* restriction sites. The positive clones were sequenced to verify the authenticity of the expressing genes. The Tfl RT-expressing plasmid was designated as pT5m-6His-Tfl RT.

The expression and purification of Tfl RT

The expression of Tfl RT was induced by 1 mM IPTG in the *E. coli* BL21(DE3)pLysS strain for 1 h at 30 °C (Sevilya et al., 2001). The induced bacteria were centrifuged and the pellets kept frozen at –80 °C. The RT, containing a six-histidine tag at its amino terminus, was purified to homogeneity, employing the outlined combination of columns. In preliminary experiments we have noticed that the Tfl RT is very sensitive to presumably proteases. Therefore, we added to the lysis buffer of the bacteria (20 mM HEPES, 100 mM NaCl, 300 µg/ml lysozyme) a cocktail of potent protease inhibitors (one tablet of Complete, Mini from Roche, per 10 ml) and 1 mM of phenylmethylsulfonyl fluoride (PMSF). The frozen bacteria pellet from 1 l broth

(approximately 3 g) were lysed in 30 ml of lysis buffer and sonicated on ice. All purification steps were performed at 4 °C. In order to precipitate nucleic acids, 0.3% of polyethylenimine was added and the lysate was centrifuged for 30 min at 27,000×g at 4 °C.

The first column was Ni-NTA agarose (Qiagen), which specifically binds the six-histidine-containing protein. The column was pre-equilibrated with 20 mM HEPES pH 7.8 and 500 mM NaCl. After loading, the column was extensively washed with 50 mM ammonium acetate, 500 mM NaCl and 10% glycerol (final pH 6.5). Elution was performed with a gradient of 25–500 mM imidazole in the washing buffer. The peak fractions were dialyzed against 10 mM NaCl, 20 mM Tris–HCl pH 8.0, 0.1% Triton X-100, 2 mM DTT, 1 mM EDTA pH 8.0 and 10% glycerol and then loaded on DEAE–Sephacrose, which was pre-equilibrated with 10 mM NaCl, 20 mM Tris–HCl pH 8.0, 0.1% Triton X-100, 2 mM DTT and 10% glycerol. Under these conditions, the majority of the RT did not bind the column and the flow-through fractions were dialyzed against 20 mM HEPES pH 7.8, 500 mM NaCl, 1 mM β-mercaptoethanol, 0.5 mM EDTA pH 8.0 and 10% glycerol. Then, another affinity chromatography was performed on Ni-NTA agarose, under conditions similar to those described above for the first column. The RT containing fractions were finally dialyzed against 20 mM Tris–HCl pH 7.5, 25 mM NaCl, 50% glycerol, 0.1% Triton X-100, 2 mM DTT and 1 mM EDTA and stored in aliquots at –80 °C (Hizi and Hughes, 1988).

Recombinant HIV-1 and MuLV RTs

For comparative studies, we have used these enzymes that were expressed in our laboratory and purified as described in detail previously (Avidan et al., 2002; Sevilya et al., 2003a, 2003b).

The relative DNA polymerase activities

The DNA polymerase activities of Tfl RT were analyzed and compared in the presence of either 0.5 mM Mn²⁺ or 5 mM Mg²⁺ as the divalent cation. The DDDP reactions were performed with the synthetic template primers, poly(dA)_n•oligo(dT)_{12–18} and poly(dC)_n•oligo(dG)_{12–18}, or with herring sperm activated gapped DNA, which was prepared as described previously (Hizi et al., 1991), or with Φx174am3 phage DNA (annealed to a 15 nt synthetic primer 5'-AAAGCGAGGGTATCC-3'). The RDDP reactions were performed with the synthetic template primers, poly(rA)_n•oligo(dT)_{12–18} and poly(rC)_n•oligo(dG)_{12–18} and the appropriate dNTPs as well as with *E. coli* ribosomal RNA annealed to a 16 nt synthetic DNA-primer (5'-ATT TCA CAT CTG ACT T-3'), which was hybridized to the 16S RNA. All reactions were carried out in 25 mM Tris–HCl, 40 mM KCl, 3 mM DTT and either 5 mM MgCl₂ or 0.5 mM MnCl₂, at a final pH 7.5 and 25 µM of each dNTP. The concentrations of the template primers were 5 µg/ml of the synthetic ones, 30 µg/ml of the activated DNA or 10 µg/ml of the Φx174am3 DNA and 50 µg/ml of 16S RNA.

DNA synthesis under processive and nonprocessive conditions

All reactions were performed with the 15-nucleotide synthetic 5'-end-labeled primer and an excess of the template single-stranded circular ϕ X174am3 phage DNA, as shown previously (Avidan and Hizi, 1998). The sequence of the primer is 5'-AAA GCG AGG GTA TCC-3' and it hybridizes to nt 588–602 of the ϕ X174am3 DNA. The final concentration of the template primer was 30 μ g/ml. The reactions were performed in 6.6 mM Tris–HCl (pH 8.0), 4 mM DTT, 24 μ g/ml BSA, as previously described (Avidan and Hizi, 1998; Avidan et al., 2002, 2003). The enzymes used have equal DDDP activities with ϕ X174am3 as DNA template. The reaction products were analyzed by electrophoresis through 8% polyacrylamide/urea gels, as described (Bakhanashvili and Hizi, 1992, 1993a, 1993b). The relative amount of the extension products was determined from the autoradiograms and the percentage of the total amounts (unextended and extended) of the labeled primers which were extended was calculated.

RT-dependent non-templated nucleotide addition during DNA synthesis

In this assay we have tested the capacity of RT to extend a 5'-end-labeled 21 nt-long primer (with the sequence 5'-GTGTGG-AAAATCTCTAGCAGT-3' that is annealed to the fully complementary 21-long oligonucleotide (with the sequence 5'-ACTGCTAGAGATTTTCCACAC-3'), thus forming a blunt-ended DNA duplex. These sequences are derived from the U5 end of HIV-1 LTR. The reactions were performed in 0.1 mg/ml BSA, 2 mM DTT, 20 mM Tris–HCl, pH 7.5, and either 5 mM Mg^{+2} or 0.5 mM Mn^{+2} . Before adding the dNTPs a preincubation of 5 min at 37 °C was done in order to allow Tfl RT to bind the template primer. Then, the dNTPs were added (each at a final concentration of 500 μ M), followed by incubation for 30 min at 37 °C. The polymerization reactions were performed in the presence of different dNTPs combinations (from a single one through all four dNTPs), as specified in Fig. 4. The reaction products were resolved by electrophoresis through 12% polyacrylamide/urea gels.

Fidelity of DNA synthesis

The reactions were performed by monitoring the standing-start misinsertion of Tfl RT, using the same method as previously described (Bakhanashvili and Hizi, 1993a, 1993b; Taube et al., 1998). Site-specific nucleotide misinsertion was assayed by following the incorporation of dNTPs opposite to the A residue. A single-strand circular ϕ X174am3 DNA was used as a template that was annealed to the [32 P] 5'-end-labeled 15 residues DNA primer described above in the DDDP reactions. This primer was extended in the presence of either dTTP or of each incorrect dNTPs (dATP, dCTP or dGTP). All reactions were incubated at 37 °C for either 2 min (for the correct incorporation) or for 5 min (for misincorporation). The

reaction products were analyzed by electrophoresis through 12% polyacrylamide/urea sequencing gels.

The RNase H activity

This activity was assayed by two methods. In the first, the release of [3 H]AMP from [3 H]poly(rA)_n annealed to poly(dT)_n (into the trichloroacetic acid-soluble fraction) was measured, as previously described (Hizi et al., 1991). In the second method, we have followed the pattern of cleavage of the labeled 5'-end 267 nt RNA, synthesized from the pBLRA30 plasmid as described (Gao et al., 1998) and annealed to a 20 nucleotide-long oligomeric DNA with a complementary sequence: 5'-AGTTAGCCAGAGAGCTCCCA-3' (see scheme in Fig. 6A). The reaction mixtures contained approximately 0.1 pmol of the [32 P] 5'-end-labeled RNA transcript, annealed to 0.4 pmol of the synthetic oligonucleotide DNA. The 3'-end of the DNA oligonucleotide is complementary to the region starting 30 nt away from the 5'-end of the RNA transcript.

Acknowledgments

We are very grateful to Dr. Henry Levin from NIH for introducing us to the field of Tfl and for his gift of the plasmid containing the Tfl sequences. We would also like to thank Drs. S. Loya and I. Oz Gleenberg from our laboratory for critically reading the manuscript and for helpful suggestions.

References

- Atwood-Moore, A., Levin, H.L., 2005. Specific recognition and cleavage of the plus-strand primer by reverse transcriptase. *Virology* 79, 14863–14875.
- Atwood-Moore, A., Yan, K., Judson, R.L., Levin, H.L., 2006. The self primer of the long terminal repeat retrotransposon Tfl is not removed during reverse transcription. *Virology* 80, 8267–8270.
- Avidan, O., Hizi, A., 1998. The processivity of DNA synthesis exhibited by drug-resistant variants of human immunodeficiency virus type-1 reverse transcriptase. *Nucleic Acids Res.* 26, 1713–1717.
- Avidan, O., Meer, M.E., Oz, I., Hizi, A., 2002. The processivity and fidelity of DNA synthesis exhibited by the reverse transcriptase of bovine leukemia virus. *Eur. J. Biochem.* 269, 859–867.
- Avidan, O., Loya, S., Tonjes, R.R., Sevilya, Z., Hizi, A., 2003. Expression and characterization of a recombinant novel reverse transcriptase of a porcine endogenous retrovirus. *Virology* 307, 341–357.
- Avidan, O., Bochner, R., Hizi, A., 2006. The catalytic properties of the recombinant reverse transcriptase of bovine immunodeficiency virus. *Virology* 351, 42–57.
- Bakhanashvili, M., Hizi, A., 1992. Fidelity of the RNA-dependent DNA synthesis exhibited by the reverse transcriptases of human immunodeficiency virus types 1 and 2 and of murine leukemia virus: mispair extension frequencies. *Biochemistry* 31, 9393–9398.
- Bakhanashvili, M., Hizi, A., 1993a. Fidelity of DNA synthesis exhibited in vitro by the reverse transcriptase of the lentivirus equine infectious anemia virus. *Biochemistry* 32, 7559–7567.
- Bakhanashvili, M., Hizi, A., 1993b. The fidelity of the reverse transcriptases of human immunodeficiency viruses and murine leukemia virus, exhibited by the mispair extension frequencies, is sequence dependent and enzyme related. *FEBS Lett.* 319, 201–205.
- Boeke, J.D., Devine, S.E., 1998. Yeast retrotransposons finding a nice quiet neighborhood. *Cell* 93, 1087–1089.
- Boutabout, M., Wilhelm, M., Wilhelm, F., 2001. DNA synthesis fidelity by the reverse transcriptase of the yeast retrotransposon Ty1. *Nucleic Acids Res.* 29, 2217–2222.

- Butler, M., Goodwin, T., Simpson, M., Singh, M., Poulter, R., 2001. Vertebrate LTR retrotransposons of the Tf1/sushi group. *J. Mol. Evol.* 52, 260–274.
- Coffin, J.M., Hughes, S.H., Varmus, H.E., 1997. *Retroviruses*. Cold Spring Harbor Laboratory Press.
- Craig, N.L., Craigie, R., Gellert, M., Lambowitz, A.M., 2002. *Mobile DNA II*. ASM Press.
- Ding, J., Das, K., Hsiou, Y., Sarafianos, S.G., Clark Jr., A.D., Jacobo-Molina, A., Tantillo, C., Highes, S.H., Arnold, E., 1998. Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8 Å resolution. *J. Mol. Biol.* 284, 1095–1111.
- Eichinger, D.J., Boeke, J.D., 1990. A specific terminal structure is required for Ty1 transposition. *Genes Dev.* 4, 324–330.
- Entin-Meer, M., Sevilya, Z., Hizi, A., 2002. The role of phenylalanine-119 of the reverse transcriptase of mouse mammary tumor virus in DNA synthesis, ribose selection and drug resistance. *Biochem. J.* 367, 381–391.
- Gao, H.Q., Boyer, P.L., Arnold, E., Hughes, S.H., 1998. Effects of mutations in the polymerase domain on the polymerase, RNase H and strand transfer activities of human immunodeficiency virus type 1 reverse transcriptase. *J. Mol. Biol.* 277, 559–572.
- Geordiadis, M.M., Jessen, S.M., Ogata, C.M., Telesmitsky, A., Goff, S.P., Hendrickson, W.A., 1995. Mechanistic implications from the structure of catalytic fragment of Moloney murine leukemia virus reverse transcriptase. *Structure* 3, 879–892.
- Golinelli, M., Hughes, S., 2002. Non-templated nucleotide addition by HIV-1 reverse transcriptase. *Biochemistry* 41, 5894–5906.
- Havecker, E.R., Gao, X., Voytas, D., 2004. The diversity of LTR retrotransposons. *Genome Biol.* 5, 225–231.
- Hizi, A., Hughes, S., 1988. Expression in *Escherichia coli* of Moloney murine leukemia virus reverse transcriptase whose structure closely resembles the viral enzyme. *Gene* 66, 319–323.
- Hizi, A., Levin, H.L., 2005. The integrase of the long terminal repeat-retrotransposon Tf1 has a chromodomain that modulates integrase activities. *J. Biol. Chem.* 280, 39086–39094.
- Hizi, A., McGill, C., Hughes, S.H., 1988. Expression of soluble, active human immunodeficiency virus reverse transcriptase in *Escherichia coli* and analysis of mutant. *Proc. Natl. Acad. Sci. U.S.A.* 85, 1218–1222.
- Hizi, A., Tal, R., Shaharabany, M., Loya, S., 1991. Catalytic properties of the reverse transcriptases of human immunodeficiency viruses type 1 and type 2. *J. Biol. Chem.* 266, 6230–6239.
- Le Grice, F.J.S., 2003. “In the beginning”: initiation of minus strand DNA synthesis in retroviruses and LTR-containing retrotransposon. *Biochemistry* 42, 14349–14355.
- Levin, H.L., 1996. An unusual mechanism of self-primed reverse transcription requires the RNase H domain of reverse transcriptase to cleave an RNA duplex. *Mol. Cell. Biol.* 16, 5645–5654.
- Levin, H.L., 1997. It's prime time for reverse transcriptase. *Cell* 88, 5–8.
- Levin, H.L., Weaver, C.D., Boeke, D.J., 1990. Two related families of retrotransposon from *Schizosaccharomyces pombe*. *Mol. Cell. Biol.* 10, 6791–6798.
- Lin, J.H., Levin, H.L., 1997. A complex structure in the mRNA of Tf1 is recognized and cleaved to generate the primer of reverse transcription. *Genes Dev.* 11, 270–285.
- Lin, J.H., Levin, H.L., 1998. Reverse transcription of a self-primed retrotransposon requires an RNA structure similar to the U5-IR stem-loop of retroviruses. *J. Virol.* 18, 6859–6869.
- Malik, H.S., Eickbush, T.H., 1999. Modular evolution of the integrase domain in the Ty3/Gypsy class of LTR retrotransposons. *J. Virol.* 73, 5186–5190.
- Moore, S.P., Powers, M., Garfinkel, D.J., 1995. Substrate specificity of Ty1 integrase. *J. Virol.* 69, 4683–4692.
- Mules, E.H., Uzun, O., Gabriel, A., 1998. In vivo Ty1 reverse transcription can generate replication intermediates with untidy ends. *J. Virol.* 72, 6490–6503.
- Palaniappan, C., Fuentes, G.M., Rodriguez-Rodriguez, L., Fay, P.J., Bambara, R.A., 1996. Helix structure and ends of RNA/DNA hybrids direct the cleavage specificity of HIV-1 reverse transcriptase RNase H. *J. Biol. Chem.* 271, 2063–2070.
- Pandey, P.K., Kaushik, N., Talele, T.T., Yadav, P.N., Pandey, V.N., 2001. Insertion of a peptide from MuLV RT into the connection subdomain of HIV-1 RT results in a functionally active chimeric enzyme in monomeric conformation. *Mol. Cell. Biochem.* 225, 135–144.
- Pandey, M., Patel, S., Gabriel, A.I., 2004. Insights into the role of an active site aspartate in Ty1 reverse transcriptase polymerization. *J. Biol. Chem.* 279, 47840–47848.
- Patel, P.H., Preston, B.D., 1994. Marked infidelity of human immunodeficiency virus type-1 reverse transcriptase at RNA and DNA template ends. *Proc. Natl. Acad. Sci. U.S.A.* 91, 549–553.
- Peliska, J.A., Benkovic, S.J., 1992. Mechanism of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase. *Science* 258, 1112–1118.
- Perach, M., Hizi, A., 1999. Catalytic features of the recombinant reverse transcriptase of bovine leukemia virus expressed in bacteria. *Virology* 259, 176–189.
- Rausch, J.W., Grice, M.K., Henrietta, M., Nymark-McMahon, Miller, J.T., Le Grice, S.F.J., 2000. Interaction of p55 reverse transcriptase from the *Saccharomyces cerevisiae* retrotransposon Ty3 with conformationally-distinct nucleic acid duplexes. *J. Biol. Chem.* 275, 13879–13887.
- Rubinek, T., Bakhanashvili, M., Taube, R., Avidan, O., Hizi, A., 1997. The fidelity of 3' misinsertion and mispair extension during DNA synthesis exhibited by two drug-resistant mutants of the reverse transcriptase of human immunodeficiency virus type 1 with Leu74→Val and Glu89→Gly. *Eur. J. Biochem.* 247, 238–247.
- Sevilya, Z., Loya, S., Hughes, H.S., Hizi, A., 2001. The ribonuclease H activity of the reverse transcriptases of human immunodeficiency viruses type 1 and type 2 is affected by the thumb subdomain of the small protein subunits. *J. Mol. Biol.* 311, 957–971.
- Sevilya, Z., Loya, S., Duvshani, A., Adir, N., Hizi, A., 2003a. Mutagenesis of Cysteine 280 of the reverse transcriptase of human immunodeficiency virus type-1: the effects on the ribonuclease H activity. *J. Mol. Biol.* 327, 19–30.
- Sevilya, Z., Loya, S., Adir, N., Hizi, A., 2003b. The ribonuclease H activity of the reverse transcriptases of human immunodeficiency viruses type-1 and type-2 is modulated by residue 294 of the small subunit. *Nucleic Acids Res.* 31, 1481–1487.
- Skalka, A.M., Goff, S.P., 1993. *Reverse Transcriptase*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Steitz, T., 1993. DNA- and RNA-dependent DNA polymerases. *Curr. Opin. Struct. Biol.* 3, 31–38.
- Taube, R., Loya, S., Avidan, O., Perach, M., Hizi, A., 1998. Reverse transcriptase of mouse mammary tumour virus: expression in bacteria, purification and biochemical characterization. *Biochem. J.* 329, 579–587.
- Von Hippel, P.H., Fairfield, F.R., Dolejsi, M.K., 1994. On the processivity of polymerases. *Ann. N. Y. Acad. Sci.* 726, 118–131.
- Wilhelm, F.-X., Wilhelm, M., Gabriel, A., 2005. Reverse transcriptase and integrase of the *Saccharomyces cerevisiae* Ty1 element. *Cytogenet. Genome Res.* 110, 269–287.
- Wisniewski, M., Balakrishnan, M., Palaniappan, C., Fay, P.J., Bambara, R.A., 2000. Unique progressive cleavage mechanism of HIV reverse transcriptase RNase H. *Proc. Natl. Acad. Sci. USA* 97, 11978–11983.