DNA polymerization by the reverse transcriptase of the human L1 retrotransposon on its own template in vitro

Olga Piskareva, Vadim Schmatchenko*
Institute of Biochemistry and Physiology of Microorganisms RAS Pushchino, Prospekt Nauki 5, 142290 Pushchino, Moscow region, Russia

Received 27 September 2005; revised 2 December 2005; accepted 27 December 2005
Available online 3 January 2006
Edited by Horst Feldmann

Abstract L1 elements (LINE-1s) account for 17% of the human genome and have achieved this abundance by transpositions via an RNA intermediate, or retrotransposition. Reverse transcription is a crucial event in the retrotransposition of the active human L1 element and is carried out by the L1-encoded ORF2 protein. Previously, we performed biochemical characterization of the human L1 ORF2 protein with reverse transcriptase (RT) activity (referred to as L1 RT), expressed in baculovirus-infected insect cells. In the present study, we describe the properties of DNA- and RNA-dependent DNA synthesis catalyzed by the L1 RT on the L1 templates in vitro. We found that L1 RT synthesized at least 620 of nucleotides per template binding event utilizing L1 RNA in vitro. Under processive conditions the L1 RT synthesized cDNA over 5 times longer than that Moloney murine leukemia virus RT on the heteropolymeric RNA template used in these studies. These data are the first to demonstrate that RT from the human L1 element is a highly processive polymerase among RT enzymes. This report also presents a strong evidence of lack of RNase H activity for the L1 ORF2 protein in vitro, distinguishing L1 RT from retroviral RTs. Finally, we found strong pausing for of the L1 RT during DNA polymerization within the 3′ untranslated region of L1 mRNA, that is result of contribution both rGs runs of the polypurine stretch and immediately adjacent stem-loop structure. A mechanism facilitating minus-strand DNA synthesis during reverse transcription of L1 element in vivo is discussed.

© 2005 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Non-LTR retrotransposons; LINE-1; Retroviruses; Reverse transcriptase; Processivity; Pausing

1. Introduction

The long interspersed elements (LINEs, L1s) are non-long terminal repeat (LTR) retrotransposons, which inhabit mammalian genomes. L1 elements account for 17% of the human genome and have achieved this abundance by transpositions via an RNA intermediate, or retrotransposition. Active L1 elements have the capacity to cause mutation, disease, genetic variation and polymorphism and their inactive copies appear to be involved in recombination and rearrangement [1]. Full-length human L1 elements are about 6 kb long and contain a 5′ untranslated region (UTR), two non-overlapping open reading frames (ORFs) and a 3′ UTR ending in a poly(A) tail [1]. The product of ORF1 is an RNA-binding protein, which associates with L1 RNA [2]. ORF2 encodes a protein of about 149 kDa, which has endonuclease [3] and reverse transcriptase [4] activities. The ORF2 polypeptide chain consists of N-terminal domain of AP-like endonuclease, reverse transcriptase domain and C-terminal domain of unknown function containing a putative zinc-binding motif [1]. L1-encoded proteins possess a cis preference, i.e. act preferentially on the L1 mRNA that encodes them [5]. The great majority insertions of L1s are 5′ truncated and transpositionally inactive [1,6], but the truncation mechanism is unknown up to now.

Reverse transcription is a crucial event in the retrotransposition of the human L1 element and is carried out by the L1 ORF2 protein. The cell culture-based retrotransposition assay studies have demonstrated that deletions and mutations in ORF2 inactivate retrotransposition in cultured human cells [7]. It has been shown that the L1 ORF2 protein uses the nicked DNA as a primer to initiate cDNA synthesis on the RNA template in target-primed reverse transcription (TPRT) reaction in vitro [8]. This mechanism of reverse transcription of non-LTR retrotransposons was originally demonstrated in experiments with reverse transcriptase encoded by the R2 element from Bombyx mori [9]. The experiments showed that R2 RT uses the free 3′ end at the target-site DNA nick to prime minus-strand DNA synthesis on the R2 RNA template in vitro. Lately, high processivity and template switching of DNA polymerization have been demonstrated for this reverse transcriptase [10,11]. Features of DNA synthesis catalyzed by the ORF2 protein of the human L1 element remain unclear so far.

Recently, we have expressed the human L1 ORF2 protein in baculovirus-infected insect cells and purified it using its reverse transcriptase activity (hereafter referred to as L1 RT). Additional biochemical studies have been performed on L1 RT [12].

In the present report, we describe the properties of DNA- and RNA-dependent DNA polymerization catalyzed by the L1 RT on the L1 templates in vitro. It has been observed that L1 RT synthesizes DNA on the minus ssDNA of element L1 effectively in vitro. We found that L1 RT polymerized at least 620 of nucleotides per template binding event utilizing its own RNA. Under processive conditions the L1 RT are able to synthesize cDNA much longer than that MMLV RT on the heteropolymeric RNA template used in these studies. These data demonstrated for the first time that RT of the human L1 element has higher processivity than that of most retroviral RTs. We also report that the L1 ORF2 protein displays a lack of RNase H activity in vitro, a facility that is required for all retroviral RTs. Moreover, we found strong pausing for of the L1 RT during DNA
polymerization within the 3' UTR of L1 mRNA, which is a result of contribution of both rG runs of the polyurine stretch and the immediately adjacent stem–loop structure. This mechanism, which facilitates minus-strand DNA synthesis during reverse transcription of L1 element in vivo is proposed here.

2. Materials and methods

2.1. Enzymes
The ORF2 protein with reverse transcriptase activity (L1 RT) of the active L1.2 retrotransposon used in this report is recombinant polymerase expressed in insect cells (Sf21) and purified as described in detail previously [12]. The specific RT activity of the purified L1 RT was about 400 U/μg protein determined as described earlier [12].

MMLV reverse transcriptase was obtained from Promega.

2.2. Templates and primers
The L1 sequence was derived from pSM42 (a gift of H.H. Kazazian) carried ORF2 and 3' UTR of the human L1 element L1.2A [4,13]. The numbering of the L1.2 sequence is that of Genbank Accession No. M80343.

The DNA primer for the L1 DNA template was 26 nt oligonucleotide s+3′L1 (5′-TCTACAACTGTAGTACTGATGAGAG) corresponding to nucleotide positions 5629–5654 of L1 element. DNA primer for the L1 RNA templates was oligonucleotide a-L1-3′ UTR (5′-CGATTTCGACCTCGACGCTT) complementary to pSM42-derived 20 nt at 3′ end of the RNA templates.

RNAs spanning nucleotides 5674–6026 and 4838–6026 of L1 element were generated in vitro using the T7 transcription kit (Fermentas) from Clal-linearized plasmids pT7-455 and pT7-1300 following extraction. Briefly, an 373 bp BcoI–ClaI fragment or 1209 bp BamHI–ClaI fragment from pSM42 were cloned into the expression vector pT7-7 (USB) resulted in pT7-455 and pT7-1300, respectively, and RNA was transcribed from the T7 promoter. Synthesized 455 nt and 1294 nt RNAs carried a short vector-derived sequence at the 5′ end and pSM42-derived 20 nt at 3′ end. Both L1 RNA templates having the same 3′ end were hybridized with the DNA primer L1-3′ UTR. The pT7-7 vector-derived 5′ terminus extremitas of the RNA transcripts contained 21 nt self-complementary sequence directly at 5′ end, which was able to form a strong hairpin structure.

Single-strand L1 DNA from 3′ end of L1 element was generated by polymerase chain reaction (PCR) and enzyme digestion. At the first the 418 nt DNA fragment comprising of L1 3′ sequence (nt positions 5629–6026) was amplified by PCR from plasmid pSM42 using non-phosphorylated primer a-L1-3′ UTR (antisense) and primer s+3′L1 (sense), which was previously phosphorylated by T4 kinase. Following PCR amplification, the phosphorylated plus-strand of the PCR product was removed by digestion with lambda exonuclease (Fermentas). The minus-strand, which served as a template, was extracted with phenol/chloroform and precipitated with ethanol. The DNA primer s+3′L1 was annealed to the 3′ end of the generated 418 nt single-strand DNA template. The primer–template complexes were prepared as follows: the primer and template were mixed at a 1:1 molar ratio, heated to 90 °C for 5 min, and then slowly cooled to room temperature for annealing.

2.3. Primer extension and processivity assays

Primer extension reactions were performed in 50 mM Tris–HCl (pH 8.0), 50 mM KCl, 5 mM MgCl2, 10 mM dithiothreitol, 1 U RNAsin, 0.1 μmol RNA/DNA hybrid and the L1 RT (concentrations described in Fig. 4) in a total volume of 10 μl. The samples were incubated for 1 h at 37 °C. The reaction was stopped by addition of an equal volume of formamide gel loading buffer. The RNA cleavage products were analyzed by electrophoresis in 8% denaturing polyacrylamide–7 M urea gel, and detected by autoradiography. MMLV RT with RNase H activity (Promega) was used as a positive control.

2.4. Assay for the RNase H activity of the L1 RT

The RNase H activity assay was performed by following the cleavage of the RNA portion of an RNA/DNA hybrid. The internally [α-32P]-labeled 455 nt RNA, synthesized in vitro using the T7 transcription kit (Fermentas) from the pT7-455 plasmid linearized by Clal, was annealed to the complementary 39 nt long synthetic DNA oligonucleotide at a molar ratio of 1:1. The reaction mixture contained 50 mM Tris–HCl (pH 8.0), 50 mM KCl, 5 mM MgCl2, 10 mM dithiothreitol, 1 U RNAsin, a 0.1 pmol RNA/DNA hybrid and the L1 RT (concentrations described in Fig. 4) in a total volume of 10 μl. The samples were incubated for 1 h at 37 °C. The reaction was stopped by addition of an equal volume of formamide gel loading buffer. The RNA cleavage products were analyzed by electrophoresis in 8% denaturing polyacrylamide–7 M urea gel, and detected by autoradiography.

3. Results

3.1. Primer extension by the L1 RT on the L1 DNA and RNA templates

Purified recombinant L1 RT was examined for its capacity to synthesize DNA on DNA and RNA templates derived from the 3′ end portion of the human L1 element in primer extension assay in vitro. Fig. 1 shows the size of DNA products synthesized by DNA- and RNA-dependent DNA polymerase activity of L1 RT on the templates. The DNA template for this reaction was a 418 nt minus-strand DNA molecule of L1 3′ end with 26 nt DNA primer annealed. We found that L1 RT generated a single extended DNA product, whose length corresponded to that of L1 DNA template (Fig. 1B, lane 1). For the RNA template, a 455 nt RNA molecule, containing the L1 3′ end of L1 element with 21 nt DNA primer annealed was used. The primer for the initiation of the reverse transcription reaction on the L1 RNA transcripts was designed to initiate DNA elongation of the template starting from an rA (Fig. 3). Fig. 1B, lane 2, shows that L1 RT generated a range of extended cDNA products, the largest of which (at 455 nt), matched that of the RNA template used. These results demonstrate that, the L1 RT is able to extend the primers efficiently and to synthesize high molecular weight products on DNA as well as RNA templates of the L1 element.

3.2. Highly processive DNA synthesis by L1 reverse transcriptase

The processivity of a DNA polymerase can be defined as the number of nucleotides incorporated in polymeric form before the enzyme dissociates from the template–primer complex. Primer extension on the L1 RNA transcripts was performed using an excess competitive substrate to determine the processivity of RNA-dependent DNA synthesis catalyzed by L1 RT. Processivity assays containing commercial MMLV RT were included for comparison studies. The cDNA polymerization was carried out both in the absence and presence of an excess poly(rC)-oligo(dG) as a competitive substrate. This trap for RT molecules was added after incubating RTs with the RNA/primer complex and prior to the addition of the labeled dNTPs mix. Hence, the result was a reduction in the overall yield of product, which is directly proportional to the processivity of the enzyme.

All polymerization reactions were stopped by addition of an equal volume of gel loading buffer. Reaction products were analyzed by electrophoresis in a 6% sequencing gels containing 7 M urea, and detected by autoradiography. The relative amounts of the primer extension products were determined by densitometric scanning of the gel autoradiographs. To determine the size of the RT products, sequencing reactions were carried out with the same primers and homologous DNA.

polymerase molecules are trapped with the competitive substrate after their initial dissociation from template–primer, preventing further primer extensions by rebound of the enzyme molecules. Thus, the extent of primer elongation in the presence of trap is directly proportional to relative processivity of the polymerases. The efficiency of poly(rC)–oligo(dG) as a trap was estimated in a primer extension reaction. Addition of the competitive substrate together with RT, and the RNA/primer complex resulted in no visible polymerization products (data not shown).

Fig. 2 illustrates the length of cDNA products polymerized by L1 RT under non-processive and processive conditions. The template for this reaction was a 455 nt RNA molecule, containing the L1 3’ end (nt positions 5674–6029) as described in Section 2. The L1 RT (12 ng) was preincubated with either the DNA · DNA or the RNA · DNA template–primer for 5 min at 37 °C and the reactions were initiated by the addition of all four dNTPs with [α-32P]dATP. The reaction products were resolved on a 6% sequencing gel and visualized by autoradiography. Products of primer extension by L1 RT on the 418 nt DNA template (lane 1) and on the 455 nt RNA template (lane 2) for 20 min. The positions of the full-length products are indicated by arrows on the left. Nucleotide positions are indicated by numbers on the right.

The primer extension assay with the 1294 nt RNA template revealed accumulated similar pause DNA intermediates as those of the 455 nt RNA template and also a termination product about 600 nt long both in the absence and presence of trap (Fig. 2A, lanes 3 and 4). The results demonstrate that the L1 RT is highly processive and able to polymerize on heteropolymeric RNA more than 600 nt per template binding event. Regarding the primer extension patterns observed for MMTV RT, one could see that the longest cDNA product synthesized by the polymerase under processive conditions was 120 nt (Fig. 2B, lane 2) in comparison with 620 nt cDNA generated by L1 RT under the same conditions. Therefore, it is clear that polymerase processivity of L1 RT is at least 5 times higher than that of MMLV RT.

It should be noted that full-length extension products have not been detected during L1 RT polymerization on long
1294 nt RNA templates under processive conditions (data not shown). Since the probability of termination decreases exponentially, the processivity of the L1 RT is in range from 620 to 1294 nt during DNA synthesis on the native L1 mRNA template starting from its 3′ end. The level of L1 RT processivity is greater than that of most retroviral RTs on heteropolymeric RNA [14,15]. For example, HIV RT synthesizes cDNA products on RNA templates of random base compositions are in the range of 50–100 nucleotides [14]. However, a similar high processivity was demonstrated for RT of non-LTR retrotransposon R2 from *Bombyx mori* in Bibillo and Eickbush studies [10].

3.3. L1 RT pausing during cDNA synthesis on L1 3′ RNA sequence

DNA synthetic activity of most RT studied so far is characterized by pause sites during polymerization on heteropolymeric templates. We found several strong pause sites for L1 RT during cDNA polymerization on L1 RNA templates, which correspond with termination product 141, 152, 156 and 164 nt long (Fig. 2, lanes 1–4). All pauses, the strongest of which was P2, occurred on the templates within the same region of L1 3′ UTR.

The pauses of polymerization may have resulted from either the dissociation of RT from the template or from a stalling of some RT molecules at a pause site, which remain bound to the 3′ end of elongated primer and continue DNA synthesis at a greatly reduced rate [16]. It is known that either nucleotide similarity or secondary structure of template results in pauses of polymerization catalyzed by RT [16–18]. Examination of both primary and secondary structure of the 3′ end portion of L1 RNA used demonstrated possible causes of pausing.

We found that the main L1 RT pause sites (P1, P2, P3 and P4) are located on L1 RNA template in L1 3′ UTR containing rG runs of polypurine stretch (Fig. 3). Furthermore, a potential stem–loop structure (nt positions 5896–5923) deduced from the dyad symmetry is also in the pausing region (Fig. 3A). The P1 stop of L1 RT occurs in the stem of the secondary structure with a calculated stability of 13.1 kcal/mol at 37°C using RNA folding program, Mfold [19] (Fig. 3B). A single weak pause site corresponding to 620 nt termination product, separate from the pause sites within 3′ UTR, has been observed during cDNA synthesis on 1294 nt RNA by L1 RT and appears to be induced by the rG run. The presented data demonstrate that main pauses of L1 RT happened at rGs runs of polypurine stretch within the stable hairpin structure of L1 3′ UTR during minus-strand synthesis by the polymerase on L1 RNA templates. These findings led to the conclusion that this region of L1 mRNA with unusual structure is evidently a major barrier for reverse transcription process catalyzed by L1 RT.

In the case of MMLV RT, both the polypurine stretch and the hairpin structure in 3′ UTR of L1 RNA template had a dramatic effect on polymerization by the enzyme and blocked the DNA synthesis even without a trap (Fig. 2A). Under non-processive conditions MMLV RT yielded a wide range of cDNA products, which were less than 164 nt in length (Fig. 2B, lane 1). In contrast with MMLV RT, L1 RT on L1 RNA template showed consistent diminished pausing during DNA polymerization, an evidence of its high template–primer binding affinity (Fig. 2A).
Interestingly, no pausing was found for L1 RT during DNA synthesis on ssDNA (Fig. 1B, lane 1), which was complementary to the L1 RNA template and contains a sequence, which is able to form a secondary structure. This could be explained by the much lower energy formation threshold for DNA hairpins than those for RNA, as the RT could melt it and pass along template without stops. Furthermore, the majority extended dTs and dCs runs of L1 ssDNA template did not affect synthesis by L1 RT, whereas, for example, HIV-1 RT was blocked by runs of dTs and dAs on DNA template[17]. These findings suggest that L1 RT is able to carry out effectively plus-strand synthesis on the L1 DNA sequence.

3.4. Lack of RNase H activity for the L1 ORF2 protein

The sequences of reverse transcriptases encoded by non-LTR retrotransposons are highly divergent from the LTR retrotransposon and retroviral polymerases. Most non-LTR elements do not contain an RNase H domain, which is present in retroviruses and in all LTR retrotransposons [20]. It is important to confirm the prediction of the polymerase’s domain organization in silico study with experimental data. We therefore examined the RNase H activity of the purified full-length ORF2 protein of the L1 element. The assay was performed in the presence of an internally labeled RNA transcript 455 nt long annealed with complementary DNA oligonucleotide (Fig. 4A). Fig. 4B shows the pattern of the RNA cleavage of the RNA–DNA heteroduplexes by L1 RT and commercially available MMLV RT, which possesses low RNase H activity. Incubation of this RNA–DNA hybrid with MMLV RT yielded two RNA fragments of 199 and 217 nt (Fig. 4B, lane 1). Contrary to the results observed with MMLV RT, no directed RNase H activity was detected for the L1 RT, even in the presence of 3-fold excess L1 RT in reaction (Fig. 4B, lanes 2 and 3).

4. Discussion

Reverse transcription is a crucial event in the retrotransposition of the human L1 element and is carried out by the RT activity of ORF2 protein. In a target-primed reverse transcription reaction [9] ORF2 protein nicks target DNA and use the generated 3'-OH to prime the cDNA synthesis on the L1 RNA [8]. To complete the process of L1 mobilization, the L1 RT must polymerase approximately 6000 nucleotides of both strands. Currently, the characteristics of DNA syntheses

---

Fig. 3. Pause sites for L1 RT on the 3' portion of L1 mRNA sequence. (A) Location of L1 RT pauses on the L1 RNA from nucleotides 5674 to 6026. The data reflect the results of Fig. 2 (lane 1). The start and the end of the L1 3' sequence of synthesized RNA template are indicated by upright arrows. The pause sites are marked by triangles. Location of the stem–loop structure sequence is outlined. Polypurine stretch is underlined. DNA primer is a-L1-3' UTR (see Section 2). (B) Secondary structure of the L1.2-derived hairpin with the locations of L1 RT pause sites within the stem and immediately contiguous polypurine stretch. The folding free energy of the hairpin is 13.1 kcal/mol at 37 °C by Mfold. The pause sites are marked by triangles, and polypurine stretch is underlined. Nucleotide positions refer to the L1.2 sequence [13].
catalyzed by the reverse transcriptase of human L1 element remained unknown.

In the present report, we examined the efficiency of DNA polymerization by the L1 RT on RNA and DNA sequences derived from the 3′ end portion of L1 element in vitro and demonstrated that the enzyme was able to polymerize high molecular weight DNA products on these templates (Fig. 1). We found that L1 RT synthesized at least 620 nucleotides per template binding event, utilizing heteropolymeric RNA (Fig. 2). In primer extension reactions under processive conditions, cDNA products generated by L1 RT on L1 RNA template were several times longer than those generated by commercial MMLV RT. These findings demonstrate for the first time that RT from the human L1 element is a highly processive enzyme.

Recently, similar processivity has been shown for RT of non-LTR retrotransposon R2 from Bombyx mori, which was greater than that of most retroviral RTs [10]. The highly processive DNA synthesis ability exhibited by RTs of non-LTR retrotransposons requires to copy thousands of nucleotides and complete TPRT taking place on chromosomal DNA in nuclei.

Unlike non-LTR retrotransposon RTs, retroviral RTs possess low or moderate processivity [14,15,10], but the process of converting the viral single-stranded RNA into double-stranded DNA is efficient because the cDNA synthesis occurs within viral particles [21,22]. Therefore, the higher processivity of L1 RT observed as compared with retroviral RTs seems to be explained by the unfavorable conditions of the reverse transcription process.

The DNA polymerase domain of retrovirus RTs resembles a right hand with the fingers, palm and thumb subdomains forming a template binding cleft [23,24]. The subdomains of non-LTR retrotransposon RTs reveals structure similarity, but larger than that of retroviruses [25,26]. It has been proposed that the presence of additional amino acid sequences observed in the fingers and thumb subdomains of R2 RT and its absence in retroviral RTs appear to increase the RT processivity [10]. Consequently, the same may be true for RT of the human L1 element also. The essential role of fingers subdomain in determining polymerase processivity is encouraged by experiments, which demonstrated that the recombinant mutant HIV-1 RT with extended fingers was more processive than wild type enzyme [27]. The insertions into the β3-β4 hairpin flexible loop of HIV-1 RT fingers subdomain also lead to increased polymerase processivity [27]. Another opportunity for polymerases to increase their processivity is utilization of accessory factors. It is known that most DNA polymerases interact with a processivity factor, which increases their binding to the primer–template and accordingly their processivity [28]. The L1 ORF2 polymerase is a single polypeptide composed of an endonuclease domain at the N-terminus, an RT domain in the central part a cysteine-rich region at the C-terminus [29]. But the function of the C-domain remains unclear. Whereas, in the retroviral RTs the C-terminal one-third of the polypeptide chain is the RNase H domain [30], the human L1 retrotransposon has no sequences corresponding to RNase H [20] and the L1 ORF2 protein lack of RNase H activity in vitro as shown in this study. However the cysteine-rich C-domain contains a CCHC zinc knuckle structure [1] and may be involved in RNA binding. Thus, C-terminal domain of ORF2 protein is an attractive candidate as an intramolecular processivity factor that provides a more stable interaction of the polymerase and L1 mRNA, decreasing its dissociation, while at the same time, allowing the enzyme to translocate on the template. This suggestion is supported by the recent data of Kazazian’s laboratory, which demonstrated that the C-terminal residues of L1 ORF2 protein substantially affected the genomic insertion length and retrotransposition frequency of the L1 element in a cultured cell assay and thus may be important in the binding of ORF2 protein to L1 RNA to facilitate reverse transcription [31]. However, in the absence of a crystal structure of the polymerase a definite model for increasing of processivity by the C-domain remains unproved. From another aspect, the cis sites presented on the C-domain of ORF2 protein may participate in recognizing L1 RNA. Previously, L1 RNA cis-acting sequences for binding of the L1 ORF1 protein were described by Hohjoh and Singer [2]. Such cis-acting elements of L1 RNA, which preferably interact with L1-encoded proteins, are likely to reflect a cis preference [5].

Fig. 4. Examination of RNase H activity for L1 RT. (A) Schematic representation of the RNase H assay. The internally labeled 455 nt RNA at the site of the annealed 39 nt DNA oligonucleotide is cleaved to create 199 and 217 nt fragments. (B) Analysis of the RNA cleavage products on a denaturing 8% polyacrylamide-7 M urea gel after RNA/DNA hybrid treated by enzymes. Reactions were performed as described in Section 2. Lane 1, reaction in the presence of MMLV RT with RNase H activity (Promega, 40 ng). Lanes 2 and 3, reactions in the presence of L1 RT (12 and 36 ng). Lane 4, reaction done without enzyme. The sizes and positions of the initial RNA and cleavage products are indicated on the left of the figure.
Therefore, it is possible, that the interaction of the C-domain with the L1 RNA cis-acting elements allows ORF2 protein molecule to bind L1 RNA template after its dissociation, and thereby reintiate reverse transcription. In that case, amino acid substitutions in the cis-binding sites of C-domain may not result in changes of the level of RT processivity, but may essentially influence the retrotransposition efficiency of L1 element.

In the present report, we also describe the pausing of L1 RT during minus-strand DNA synthesis on RNA sequences derived from the 3’ end portion of L1 element. We found the main pause sites for L1 RT on the L1 RNA templates are located within 3’ untranslated region (Figs. 2 and 3). These L1 RT pauses are likely to be induced by the rG runs of a polypurine stretch and adjoining the 27 nt stem–loop structure, which are in this template region (Fig. 3). The strongest pause site, P2, and another two strong sites, P3 and P4, are in rG-rich polypurine stretch. Another site P1 is within the stem of the hairpin (Fig. 3). This is the first data demonstrating an association between L1 RT strong pausing and 3’ RNA structure of the human L1 element.

It has been shown previously that the polypurine stretch in the L1 3’ UTR has potential secondary structure; because of rG runs of the sequence can form antiparallel intrastrand guanine tetrads [32]. Therefore, described L1 RT pauses at the rG runs may be caused by this secondary structure, but a biological role of the pauses is unclear. Since the polypurine stretch in the 3’ UTR of L1 retrotransposons is evolutionarily conserved [32], it might be proposed that the strong pausing of L1 RT in this region with unusual structure serve specific function in the transposition of L1 elements in vivo.

When discussing the processivity of L1 RT from the results presented here, it should be emphasized that L1 polymerase has higher processivity compared with retroviral RTs. However, the obstacles faced by TPRT in vivo already outlined are disadvantageous to successful synthesis of long L1 cDNA, in which fewer RT molecules associated with its mRNA appear to participate in the reaction. In this case, reverse transcription by L1 RT will generally happen under one template binding event, and the synthesis of first strand of DNA will be incomplete. Thus, transposition of full-length L1 copies seems to require a mechanism facilitating minus-strand DNA synthesis during reverse transcription of the L1 element. Such a mechanism may be anchoring the L1 5’ terminus at the 3’ UTR and forming an L1 RNA loop. The 5’ terminus of L1 mRNA is likely to be added upon in cis-acting by molecules of L1 RT associated with 3’ UTR of the element, either at the polymerase’s strong pause region (stalling site) or at an L1 3’ polyA tract, bringing it in proximity to the insertion site. Since polypurine stretches of the 5’ UTR and 3’ UTR are highly homologous, the 5’ end of L1 RNA has the potential to base pair with polypyrimidine stretches of growing cDNA in the L1 RT pause site and may bind via RNA · DNA triple helical formation. Previously, it has been shown that the polypurine-polypyrimidine sequence at 3’ UTR of the rat L1 element can assume DNA triplex structure [33]. Closed L1 RNA cis-acting elements transiently entrap molecules of L1 RT and prevent the last to leave the place of DNA synthesis. Under these conditions rebinding of L1 RT molecules with 3’-OH growing cDNA will happen with greater frequency after enzyme-template dissociation, which may suggest successful synthesis of the first DNA strand. The L1 ORF1-coded protein probably also participates in the process, as melting and renaturing activity of ORF1 protein from the mouse L1 retrotransposon have been described [34]. As L1 RT lacks RNase H activity, intact L1 mRNA in the heteroduplex apparently serves a stabilization factor during generation of first strand DNA, and later is either cleaved by cellular RNases or displaced by L1 RT during second strand DNA synthesis.

The DNA-dependent DNA polymerase activity of L1 RT described here directs efficient polymerization on L1 DNA template and we would expect converting L1 mRNA onto double-strand DNA followed by insertion of L1 element. Such a mechanism assumes formation of full-length L1 insertions and requires intact 5’ UTR and 3’ UTR of L1 element.

The proposed model of L1 RNA looping may explain a bimodal distribution of L1 insertion lengths in the human genome, when short 5’ truncated (less than 1 kb) and full-length (6 kb) insertions of L1 element are encountered more often [35–38]. Respectively, in the event of forming the L1 loop, L1 reverse transcriptase will reach the 5’ end of the L1 RNA successfully. In this case, full-length insertions may result from completion of TPRT. Nevertheless, when L1 5’ termini anchoring does not happen and a compact transcriptional structure does not form, rebinding L1 RT molecules with L1 template for elongation of cDNA is unlikely. Therefore, DNA polymerization by L1 RT is likely to be performed generally under one template binding event. 5’ truncated and variable L1 copies (mostly less than 1 kb) then accumulated, the length of which are distributed in accordance with the probability of premature termination for L1 RT polymerase.

In conclusion, our report illustrates features of DNA polymerization catalyzed by the ORF2 protein encoded by the L1.2A element. It has been shown recently that the L1.2A element has a mutation in the carboxyl terminal of the L1 ORF2 protein that decreases its ability to retrotranspose in cultured cells [31,39]. It remains unclear whether this mutation affects the processivity of the L1 ORF2 protein. Thus, it may be very interesting to investigate DNA synthesis carried out by ORF2 proteins encoded by the L1.2B and L1.3 elements that have higher retrotransposition efficiency than L1.2A in cultured human cells.

Acknowledgments: We thank Marina Zakharova and Sergey Chernyshov for help with the primer extension assays.

References


