Nucleic Acids Research, 1997, Vol. 25, No. 8 **1641–1648**

BC1 RNA, the transcript from a master gene for ID element amplification, is able to prime its own reverse transcription

M. Richard Shen[1,*], Jürgen Brosius[2] and Prescott L. Deininger[1,3]

[1]Department of Biochemistry and Molecular Biology, Neuroscience Center of Excellence, Stanley S.Scott Cancer Center, Louisiana State University Medical Center, 1901 Perdido Street, New Orleans, LA 70112, USA, [2]Institute for Experimental Pathology, ZMBE, University of Münster, Von-Esmarch-Strasse 56, D-48149 Münster, Germany and [3]Laboratory of Molecular Genetics, Alton Ochsner Medical Foundation, New Orleans, LA 70121, USA

Received September 10, 1996; Revised and Accepted February 21, 1997

ABSTRACT

ID elements are short interspersed elements (SINEs) found in high copy number in many rodent genomes. BC1 RNA, an ID-related transcript, is derived from the single copy BC1 RNA gene. The BC1 RNA gene has been shown to be a master gene for ID element amplification in rodent genomes. ID elements are dispersed through a process termed retroposition. The retroposition process involves a number of potential regulatory steps. These regulatory steps may include transcription in the appropriate tissue, transcript stability, priming of the RNA transcript for reverse transcription and integration. This study focuses on priming of the RNA transcript for reverse transcription. BC1 RNA gene transcripts are shown to be able to prime their own reverse transcription in an efficient intramolecular and site-specific fashion. This self-priming ability is a consequence of the secondary structure of the 3′-unique region. The observation that a gene actively amplified throughout rodent evolution makes a RNA capable of efficient self-primed reverse transcription strongly suggests that self-priming is at least one feature establishing the BC1 RNA gene as a master gene for amplification of ID elements.

INTRODUCTION

ID elements are short interspersed elements (SINEs) found in rodent genomes and are believed to be ancestrally derived from an alanine tRNA gene (1–3). They vary in copy number from a few hundred copies per haploid genome in guinea pigs to 130 000 copies per haploid genome in rats (4). BC1 RNA is a homogeneous ID-related transcript found in high abundance in the rodent brain and in low abundance in other tissues (5–7). cDNA clones of BC1 RNA have revealed that the RNA can be divided into three distinct domains. The 5′-end of the transcript contains the ID body (75 nt), followed by an A-rich internal region (50 nt), which is followed by a unique region at the 3′-end (Fig. 1; 6). The transcription terminator signal for RNA polymerase III (pol III) is a stretch of T residues. BC1 RNA contains two to four U residues at the 3′-end, an initial indication that it is transcribed by RNA pol III (6,8,9). *In vitro* transcription assays have confirmed that the BC1 RNA is transcribed by RNA pol III and that 5′ as well as internal sequence elements of the single copy BC1 RNA gene are required for efficient transcription (5,6).

Recent studies have demonstrated that some SINE-containing loci are much more effective than others at the amplification process (4,10; reviewed in 11). In fact, the vast majority of SINEs appear to be incapable of active amplification and the process is dominated by a relatively few master, or source, genes (4; reviewed in 11). This is based largely on the findings that there are distinct subfamilies in most SINEs and that only very specific subfamilies appear to have been active at any given time. The BC1 RNA gene is believed to be the original master gene for ID element amplification (4). This conclusion is based on the observation that the ID repeats in various rodent species show the same sequence changes as are found at the BC1 RNA locus. Additionally, it is clear that the BC1 locus is one of the very oldest ID-containing loci. Thus, the BC1 locus has shaped the evolution of ID elements in the rodent genome by making a significant portion of the copies throughout rodent evolution (12,13).

Expression of the BC1 locus in germline tissue is one potential explanation for its efficiency at retroposition (H.Tiedge, Z.Zakeri and J.Brosius, unpublished data; 15). However, other transcripts for IDs and other SINEs are also produced which are not apparently involved in the amplification process (14–17). For that matter, there must also be some feature(s) that chooses the SINE transcripts for highly efficient amplification relative to the vast bulk of other RNAs in the cell (4).

There are many potential rate limiting steps that may be involved in determining whether an individual SINE element will amplify effectively and therefore serve as a master, or source, gene. These may include transcription in germline cells at the appropriate time, transcript stability, priming and elongation of the RNA for reverse transcription and integration (reviewed in 4,18–20). The genomic location of the inserted SINE sequence

*To whom correspondence should addressed at present address: Biology and Biotechnology Research Program, L-452, Lawrence Livermore National Laboratory, PO Box 808, Livermore, CA 94551, USA. Tel: +1 510 422 5657; Fax: +1 510 422 2282; Email: shen3@llnl.gov

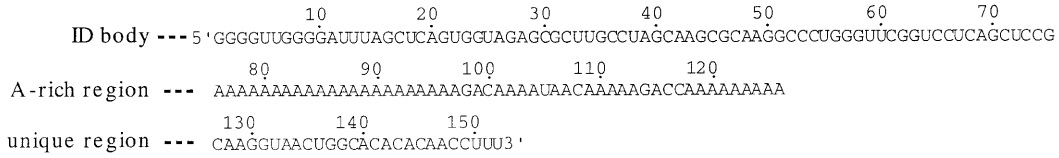

Figure 1. Nucleotide sequence of the BC1 RNA (adapted from DeChiara and Brosius, accession no. m16113; 6). BC1 RNA contains three domains, the ID body, the middle A-rich region and the unique region.

Table 1. Oligonucleotides used to generate templates for variant BC1 RNAs

Name	5'→3' sequence	Comments
T7	TAATACGACTCACTATAGGG	5' primer containing a T7 polymerase promoter
Tr	TAATACGACTCACTATAGGGAGACTG-GAGTGCAGTGGCGGGTTCGGTCCT	5' primer to create a 5'-truncated BC1 RNA
M8	TAATACGACTCACTATAGGGGTTGGG-GATT	5' primer for M8-3UA
3U	AAAGGTTGTGTGTGCCAGTT	3' primer for wild-type BC1 3' end (T7-3U)
3A	TTTGGTTGTGTGTGCCAGTT	3' primer for Tr-3A
3UO	AAATTTTGTGTGTGCCAGTT	3' primer for T7-3UO
3UA	AAAGGTTGTGTGTGCCAGTT-ACCTTTTTTT	3' primer for M8-3UA
2U	AAGGTTGTGTGTGCCAGTT	3' primer for T7-2U
2ULs	AAGGTAATGTGTGCCAGTT	3' primer for T7-2ULs
3'G	CAAGGTTGTGTGTGCCAGTT	3' primer for T7-3'G
3'GLs	CAAGGTAATGTGTGCCAGTT	3' primer for T7-3'GLS
4U	AAAAGGTTGTGTGTGCCAGT	3' primer for T7-4U

is the principal factor that would make one SINE more efficient at the above processes than another. The context in which the SINE is located may provide the tissue-specific *cis*-regulatory elements for high level expression (5,6,16,21–23). As SINE sequences, like ID elements, do not encode a RNA pol III termination signal, the transcripts would all vary in sequence at their 3′-unique ends. In a few SINEs, this 3′-end variation may positively influence RNA stability and may include secondary structure that allows the RNA to self-prime. One of the earliest models of SINE amplification proposed that the short stretch of U residues typically found terminating RNA pol III-derived transcripts would be able to self-prime efficiently on the A-rich region at the 3′-end of the SINE (24). The kinetics of a self-primed mechanism are pseudo-first order and would be many times faster than a second order reaction in which another RNA molecule primes on the A-rich region. To the best of our knowledge there has been no SINE demonstrated to self-prime and the BC1 RNA gene is the only SINE master locus that has been identified thus far. Thus, we wished to determine whether the proposed self-priming mechanism (24) might be part of what makes the BC1 RNA gene an effective master gene.

MATERIALS AND METHODS

Construction, *in vitro* transcription and isolation of the BC1 RNA 3′ variants

Individual BC1 RNA 3′ variant templates were created by PCR amplification of a full-length cDNA clone of BC1 RNA (pBCX607; 13). The amplification primers were designed to hybridize to the T7 RNA polymerase promoter at the 5′ position and to the BC1 RNA gene unique region at the 3′-end. The sequences of the PCR primers used to create the templates for transcription are presented in Table 1. Conditions for PCR amplifications were as follows: 100 ng pBCX607, 100 pmol each primer, 0.4 mM dNTPs, 500 mM KCl, 100 mM Tris–HCl, pH 9.0, 1.0% Triton X-100 and 2.5 U AmpliTaq DNA polymerase (Perkin-Elmer Cetus) in a volume of 100 µl. $MgCl_2$ concentrations were individually optimized for each primer pair and all primer pairs amplified well at $MgCl_2$ concentrations of 2–3 mM. The cycle conditions in the Perkin-Elmer Cetus GeneAmp PCR System 9600 were 94°C, 2 min; 94°C, 15 s, 55°C, 15 s, 72°C, 30 s for 20 cycles; followed by a 72°C incubation for 10 min. PCR products were purified through a 1.8% agarose–1× TBE (90 mM Tris–borate, 2 mM EDTA) gel containing 0.1 µg/ml ethidium bromide. DNA quantities were estimated by the intensity of ethidium bromide staining. PCR products were isolated from the gel with DEAE paper and resuspended in 10 µl TE buffer (25).

Conditions for *in vitro* transcription of the BC1 RNA 3′ variants were as described by Gurevich *et al.* (26). Briefly, the reaction conditions were as follows: 80 mM HEPES, pH 7.5, 2 mM spermidine, 20 mM dithiothreitol (DTT), 12 mM $MgCl_2$, 0.5 mM NTPs, 1 µCi/µl [α-^{32}P]CTP (3000 Ci/mmol, 10 mCi/ml; Amersham), 8 ng/µl PCR product, 2 U/µl RNasin (Promega) and 1.4 U/µl T7 RNA polymerase (Pharmacia). After incubation at 37°C for 1 h, RNase-free DNase I (Ambion) was added to a concentration of 0.04 U/µl and the reaction incubated for a further 15 min at 37°C. The *in vitro* transcribed RNAs were ethanol precipitated and resuspended in 10 µl RNA loading buffer (80% v/v formamide, 1 mM EDTA, pH 8.0, 0.1% bromophenol blue, 0.1% xylene cyanol; 27). The *in vitro* transcription products were then separated on a 6% sequencing gel and the region corresponding to the radiolabeled RNA was excised. The RNA was eluted

from the polyacrylamide gel slice with 400 µl elution buffer (2 M NH_4OAc, 1% SDS; 27) at 37°C for 4–16 h with shaking. The RNA was ethanol precipitated from the supernatant, washed with 70% ethanol, air dried and resuspended in 20 µl diethylpyrocarbonate-treated H_2O. The amount of radiolabeled [α-^{32}P]CTP that was incorporated into the RNA was used to quantify the amount of RNA synthesized.

Reverse transcription and analysis of the self-primed RNAs

The reverse transcription reaction was carried out in the buffer supplied by the vendor (Gibco-BRL). The RNA was heated at 85°C for 5 min in 1.44× reverse transcription buffer and then ice/H_2O quenched. The remaining reagents were added to final concentrations as follows: 0.5 mM dNTPs, 1 µCi/µl [α-^{32}P]dATP (3000 Ci/mmol, 10 mCi/ml; Amersham), 10 mM DTT, 2 U/µl RNasin, 4 U/µl Superscript RNase H$^-$ reverse transcriptase (Gibco-BRL) (the final RNA concentration was 0.04 pmol/µl). In some cases, the 3′-unique region primers were added to a concentration of 0.3 pmol/µl The reactions were incubated at 37°C for the indicated times and stopped by heating at 85°C for 5 min. RNase A and RNase One (Promega) were added to a final concentration of 0.2–1 µg/µl and 0.1–0.5 U/µl respectively and further incubated at 37°C for 15 min prior to termination with 0.6 vol stop/loading dye (0.05% bromophenol blue, 0.05% xylene cyanol, 20 mM EDTA in deionized formamide; 27). RNase One degrades RNA to cyclic nucleotide monophophates and is able to cleave RNA at all four bases (28). The reverse transcribed samples were then denatured at 85°C for 5 min and loaded onto a 6% sequencing gel. The gel was dried and exposed to X-ray film or to a phosphorimager screen (Molecular Dynamics). Quantification of the amount of self-primed product was carried out with ImagQuant (Molecular Dynamics).

Southern blot of self-primed, reverse transcribed, *in vivo* isolated BC1 RNA

Rat brain, testes and liver total cytoplasmic RNAs were a gift from J.Kim and were isolated by the method of Chomczynski and Sacchi (29). The reverse transcription reaction was carried out as described above except that the final RNA concentration was 0.2 µg/µl, RNase digestion was for 30 min and radiolabeled dATP was omitted from the reactions. The reverse transcribed products (from 1.2 µg total RNA) were separated on a 6% sequencing gel. The separated products were electroblotted onto Hybond-N filters, at 1 A for 30 min in 0.5× TBE. The DNA was cross-linked to the membrane with 150 mJ UV light (BioRad GS Genelinker) and prehybridized in 6× SSC (20× SSC is 3 M NaCl, 0.3 M sodium citrate·2H_2O, pH 7.0; 27), 0.5% SDS, 100 µg/ml denatured salmon sperm DNA, 5× Denhardt's solution and 50% formamide for 1 h at 37°C. The 75 bp radiolabeled ID body probe (generated by PCR with primer pair GGGGTTGGGGATTT and CGGAGCTGAGGACC; 30) was added to the prehybridization solution at a final concentration of 2.5×10^5 c.p.m./ml and hybridized for 12 h at 37°C. The hybridized membranes were washed twice in 2× SSC, 0.1% SDS at room temperature for 20 min each. The washed membrane was exposed to film (78 h) at –80°C with an intensifying screen.

RESULTS

Reverse transcription of the *in vitro* transcribed BC1 RNAs

PCR primers T7 and 3U were used to construct a T7 RNA polymerase-driven transcript of BC1 RNA that terminates with the typical three U residues associated with most RNA pol III-directed transcripts (T7-3U RNA). The RNA formed a band, approximately eight bases in width, on a 6% sequencing gel. The heterogeneity in the *in vitro* transcribed RNA is likely the result of *Taq* DNA polymerase slippage along the middle A-rich region of the template during PCR amplification. It is also possible that T7 RNA polymerase may show heterogeneity at the site of transcript initiation and termination. The nature of the sample's heterogeneity was determined by dividing the RNA into two equal size fractions. Reverse transcription reactions were performed on both the high molecular weight (HMW) and low molecular weight (LMW) fractions independently. Only the HMW fraction yielded significant amounts of reverse transcribed products, whereas the LMW fraction showed a very weak signal of reverse transcribed products (data not shown). The priming site appeared to be at position ~128 ± 3 of the BC1 RNA sequence. Primer extension using a sense primer of the T7-3U cDNA confirmed the self-priming site to be at base position ~128 of the BC1 RNA (data not shown). Primer extension of the T7-3U RNA using the 3U primer gave a product size of ~153 ± 3 bases (Fig. 2). Secondary structure analysis of BC1 RNA (31) revealed that the 3′-end may form a hairpin structure that will enable the RNA to prime its own reverse transcription at base position 128 of the BC1 RNA (Fig. 3A). This 3′-end was not involved in any potential competing structures even at free energy values 20% lower than the optimal structure. The closest alternative structure for the 3′-terminus would involve pairing of the three terminal U residues with the runs of A. As demonstrated below, priming on the A stretches does not compete with the more specific self-priming reaction until the length of the U residues is increased. The loss of priming in the LMW fraction would be most consistent with the size heterogeneity being the result of 3′ truncations that eliminated sequences essential for self-priming. In subsequent experiments on *in vitro* transcribed RNA, the HMW and LMW fractions were not segregated. A time course was performed on the reverse transcription reaction for the T7-3U RNA. A primer extension reaction on the RNA using an excess of 3U oligonucleotide was also included in these reactions to approximate 100% priming efficiency. Time points were taken every 15 min, for 120 min (Fig. 2). The T7-3U RNA self-primed products showed a gradual increase (Fig. 4A).

C-tailed cDNA clones of *in vivo* BC1 RNA showed some heterogeneity in the number of U residues at the 3′-end (6). Two *in vitro* transcribed RNAs (T7-2U and T7-4U) were used to investigate the potential effects of the 3′ heterogeneity in the self-priming reaction. T7-2U and T7-4U contain two and four U residues at their 3′-ends respectively. T7-2U RNA reverse transcription showed a rapid increase in self-primed products in <15 min, followed by a slow secondary increase (Fig. 4A). This data demonstrates that removing the third U from the transcript increases the efficiency of the self-priming reaction. Appearance of the self-primed product for T7-4U RNA showed a time course similar to that seen for T7-3U (Fig. 4A). However, in addition to the ~128 base product that was found in all self-primed RNAs in this study, a secondary priming site (position ~106) and a smear between the two sites was observed (Fig. 5). This suggests that

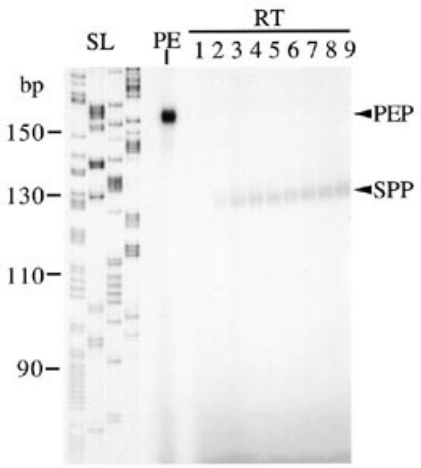

Figure 2. Reverse transcription of self-primed T7-3U RNA (BC1 RNA). *In vitro* transcribed T7-3U RNA was reverse transcribed as described in Materials and Methods and samples taken at 15 min intervals, starting at 0 min (lanes 1–9). A primer extension (lane PE) reaction using the 3U primer (Table 1) was also performed. The primer extension product (PEP) and self-primed product (SPP) are indicated by arrows. A sequencing ladder (SL) was used to determine the length of the reverse transcribed products in bases, as indicated by the numbers.

increasing the number of U residues at the 3′-end can lead to some direct priming on the A-rich region, as was originally proposed for priming of SINE retroposition (24). However, this also demonstrates that the number of U residues (two to four) found on BC1 would not effectively compete with the more specific priming reaction.

Disruption of the self-priming reaction

The predicted self-priming hairpin structure is stabilized by two G-C bp interrupting three A-U bp (Fig. 3A). We determined whether three A-U bp without the G-C bp were sufficient for self-priming. The two C residues (positions 148 and 149) near the 3′-end of the unique region were changed to A residues in T7-3UO RNA. The reverse transcription reaction for T7-3UO RNA showed that it did not self-prime at a detectable level (data not shown). Therefore, three U-A bp were not sufficient for efficient self-priming of the RNA at any site and the G-C bp are required for the self-priming hairpin to be stabilized in this reaction.

Enhancement of the self-priming reaction

The secondary structure analysis and the T7-2U RNA self-priming reactions suggested that if position 128 in the BC1 RNA is the priming site, then BC1 RNAs with three or four U residues at the 3′-end might not self-prime efficiently because the 3′-terminus would not base pair (Fig. 3A). However, we did observe self-primed products from T7-3U and T7-4U, probably due to 3′-end truncations in the *in vitro* transcribed RNAs, as described above. BC1 RNAs with three or four U residues at the 3′-end mismatch with a C residue at position 127 when the RNA 3′-end folds onto the self-priming region. This mismatch is at the last base of T7-3U RNA and the penultimate base in T7-4U RNA (Fig. 3A). Correction of the mismatch at position 127 will allow the 3′ base of the RNA to anneal and allow efficient reverse transcription from that position. M8-3UA RNA corrects the

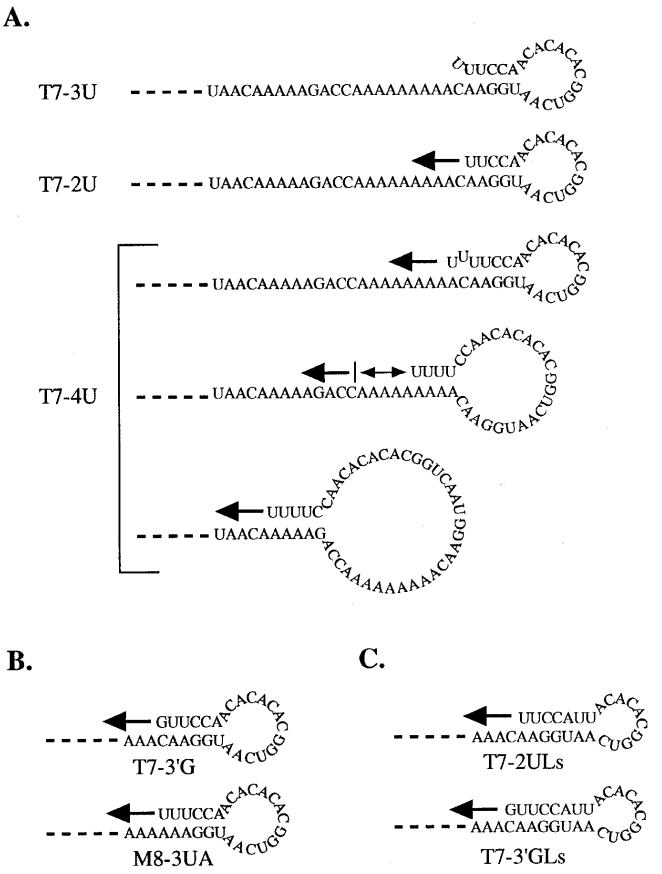

Figure 3. Hypothetical secondary structures for the 3′-ends of the various *in vitro* transcribed RNAs. (**A**) *In vitro* transcribed RNAs with 2 U (T7-2U), 3 U (T7-3U) or 4 U (T7-4U) residues at their 3′-ends. T7-4U RNA is potentially able to self-prime at several sites (Fig. 5). (**B**) Enhancement of the self-priming reaction. The 3′-ends of these RNAs match the priming site exactly (unlike T7-3U or T7-4U). (**C**) Increasing the number of base pairs involved in the stem of the hairpin structure compared with the RNAs T7-2U (T7-2ULs) and T7-3′G (T7-3′GLs).

mismatch at position 127 (Fig 3B). Reverse transcription of M8-3UA RNA shows that this RNA self-primes with a highly efficient time course, similar to T7-2U RNA (Fig. 4B). There is a large accumulation of self-primed products in <15 min, followed by a steady increase. In another modification of the 3′-end, T7-3′G RNA contains a G residue at the 3′-end and this G residue is able to base pair with the C residue at position 127 when the self-primed hairpin forms (Fig. 3B). Reverse transcription of T7-3′G RNA also shows the same time course as the efficient templates T7-2U and M8-3UA RNA (Fig. 4B).

The effect of increasing the number of bases pairs in the stem of the hairpin on self-priming efficiency was also investigated. T7-2ULs RNA contains two more base pairs (longer stem) relative to T7-2U RNA that help stabilize the stem in the hairpin (Fig. 3C). The rate of product accumulation was similar to T7-2U, M8-3UA and T7-3′G RNAs (Fig 4C). However, in a control where 2ULs oligonucleotide was included to serve as a primer in the reverse transcription reaction, the self-priming reaction competed with the 2ULs oligonucleotide primer extension reaction (data not shown). This indicated that the self-priming reaction was able to compete successfully against a >8-fold molar

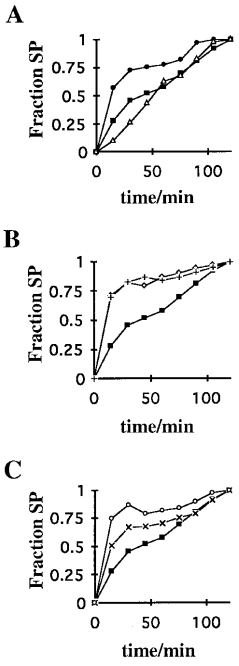

Figure 4. Time course of the self-primed reverse transcription reaction. *In vitro* synthesized RNAs were reverse transcribed and the products were quantified as described in Materials and Methods. The fraction SP is the ratio of reverse transcribed product at each time point relative to the 120 min time point. Each data point represents the mean of two independent experiments and each data set is compared with the T7-3U RNA time course. (**A**) Time course comparison of 2 U (●, T7-2U), 3 U (■, T7-3U) or 4 U (Δ, T7-4U) residues at the 3′-end of the RNA. These RNAs represent the different *in vivo* 3′-ends found in the BC1 RNA. (**B**) Enhancement of the self-priming reaction. The 3′-end base of these RNAs matches the priming site exactly with a G-C base pair (+, T7-3′G) or a U-A base pair (◊, M8-3UA), compared with 3 U (■, T7-3U) residues. (**C**) An increase of 2 U-A base pairs in the stem of the hairpin structure in T7-2U RNA (×, T7-2ULs) and T7-3′G RNA (○, T7-3′GLs) when compared with T7-3U RNA (■, T7-3U).

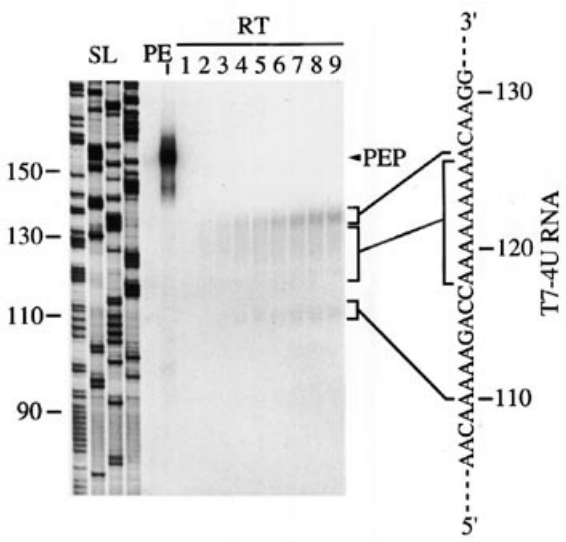

Figure 5. Time course of reverse transcription for T7-4U *in vitro* transcribed RNA. The reverse transcription reaction conditions are as described in Materials and Methods. Samples were taken at 15 min intervals, starting at 0 min (lanes 1–9). A primer extension reaction is shown in the PE lane. The primer extension product is denoted PEP. The self-primed products are bracketed and the priming sites that gave rise to these products are denoted along the sequence by lines and brackets. The numbers adjacent to a region of the T7-4U sequence represent the distance (bases) from the 5′-end of the transcript. The sequencing ladder (SL) defines the size of the reverse transcribed products in bases, as denoted by the numbers adjacent to the sequencing ladder.

excess of the 2ULs oligonucleotide primer. In a similar reaction T7-3′GLs RNA contains two more base pairs in the stem of the hairpin than T7-3′G RNA (Figs 3C and 4C). The T7-3′GLs RNA self-priming reaction also successfully competed against the 3′GLs oligonucleotide primer (data not shown).

Intramolecular versus intermolecular priming of BC1 RNA

Kinetically, the priming mechanism for BC1 RNA should favor an intramolecular event (self-priming) over an intermolecular event. To demonstrate self-priming more directly, we used a competitive assay in which we prepared Tr-3A RNA which was 5′ truncated with respect to the ID portion of BC1 RNA but contains all of the internal A-rich region and all of the unique region, except that the three U residues at the 3′-end have been replaced by three A residues. These modifications make Tr-3A RNA unable to prime its own reverse transcription and if primed intermolecularly by T7-3U RNA, Tr-3A RNA will produce a shorter product. A constant amount of Tr-3A RNA (1 pmol) was mixed with different amounts of T7-3U RNA (100× range, from 0.1× to 10×) to see whether T7-3U RNA would prime Tr-3A RNA in an intermolecular fashion for reverse transcription. No detectable signal was seen for the intermolecular reaction. The only signal that was observed was from the self-priming reaction by T7-3U RNA (Fig. 6). Primer extension reactions were performed in these reactions to show the relative amounts of RNA and that intermolecular priming may occur. The primers are 20 bases in length and out-compete the self-priming reaction due to their greater stability and 2-fold greater abundance.

A conformational control was also designed in order to show that the priming site in Tr-3A RNA was not blocked by modifications at the 5′-end. Tr-3U RNA is identical to Tr-3A RNA except that the three A residues at the 3′-end of Tr-3A RNA are replaced by three U residues in Tr-3U RNA. In the reverse transcription reaction, Tr-3U RNA was able to prime its own reverse transcription in the presence and absence of T7-3U RNA (Fig. 6). These experiments, taken together with the primer extension controls, demonstrate that the dominant priming mechanism for BC1 RNA is an intramolecular reaction and not an intermolecular reaction in our assay.

Reverse transcribed, self-primed BC1 RNA from rodent tissues

To detect self-priming of authentic BC1 RNA, a BC1 probe (double-stranded) was hybridized to a Southern blot of reverse transcribed total RNA from brain, liver and testes. The Southern blot revealed a high abundance of ID-hybridizing cDNA in the brain, a low amount of hybridizing signal in the testes and no detectable signal in the liver (Fig. 7). The signal was confirmed with strand-specific oligonucleotides and represents only the antisense (first) strand of ID cDNA (data not shown). The reverse transcription reaction was carried out without exogenously added primers, therefore, we conclude that the primer for the reverse

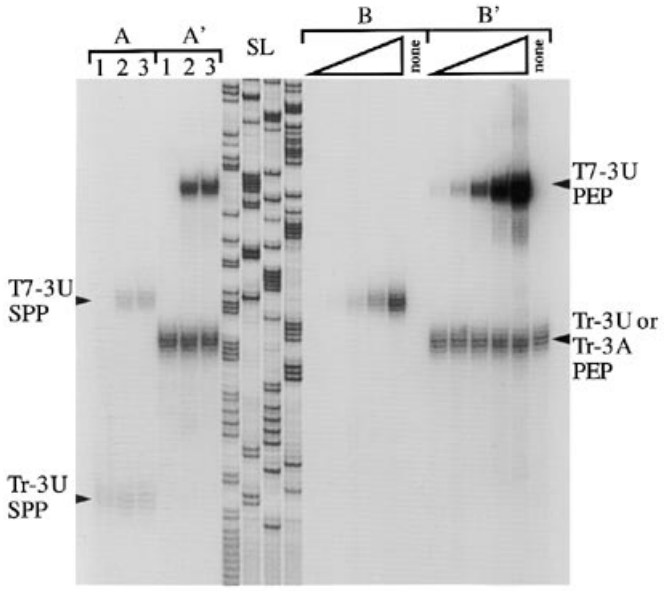

Figure 6. Intramolecular versus intermolecular priming. Two RNA species, T7-3U and Tr-3A RNAs are able to self-prime and unable to self-prime respectively. Tr-3A RNA also differs from T7-3U RNA by a 5′ truncation, but Tr-3A RNA contains the same priming site as T7-3U RNA. (**A**) A control RNA (Tr-3U) with the same 5′ truncation as Tr-3A RNA but able to self-prime was reverse transcribed without T7-3U RNA (lane 1) and with T7-3U RNA (duplicates, lanes 2 and 3) in equal molar ratios. The appearance of the Tr-3U self-primed product (Tr-3U SPP) in the absence and presence of T7-3U RNA suggests that the priming site in Tr-3A RNA is not 'hidden' due to the 5′ truncation. (**A′**) A primer extension reaction was performed for the RNA samples in (A) to show the relative amounts of T7-3U (T7-3U PEP) and Tr-3U (Tr-3U PEP) RNAs. (**B**) A constant amount of Tr-3A RNA was titrated with 0.1×, 0.3×, 1×, 3×, 10× and 0 molar ratios of T7-3U RNA in a reverse transcription reaction. T7-3U RNA reverse transcribed products were readily detected (T7-3U SPP), however, no reverse transcribed products were observed for Tr-3A RNA. This suggests that the priming mechanism is an intramolecular event and not an intermolecular event. (**B′**) A primer extension reaction was also performed on the RNA samples in (B) to show primer–template intermolecular reactions and the relative amounts of Tr-3A (Tr-3A PEP) and T7-3U (T7-3U PEP) RNAs.

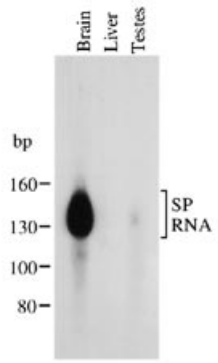

Figure 7. Southern blot of self-primed reverse transcribed total RNA from rat brain, liver and testes. Total RNA from the indicated tissues were reverse transcribed and transferred to Hybond-N filters. The filter was probed with the ID body of the BC1 RNA gene. The numbers indicate the approximate size in bases and the brackets indicate the self-primed reverse transcribed ID-related cDNAs.

transcription reaction was present in the RNA preparation. The size distribution of the hybridizing signal in the brain ranged from ~120 to ~150 bases and correlates well with the size distribution of C-tailed cDNA clones of BC1 RNA from the brain (6). The tissue distribution of the hybridizing signal also correlates with expression levels for BC1 RNA in those tissues. (6,15,32–35).

DISCUSSION

ID elements in rodents, like other retroposons, amplify through a RNA intermediate. Reverse transcription of these RNAs necessitates a primer. One of the original proposals for the priming mechanism was a self-priming reaction involving the U residues at the 3′-end of the transcripts (RNA pol III derived) folding back onto the oligo(A)-rich region of the source RNA (24). We have shown that three U residues at the 3′-end of the BC1 RNA are not sufficient to efficiently prime reverse transcription *in vitro*, probably because the base pairing is not sufficiently stable; instead, a few GC base pairs are required to stabilize the hairpin structure. Inserted SINE elements do not carry the self-priming structure into the new genomic location and, as a consequence, the newly inserted SINE element (to become an efficient master gene) must acquire a self-priming secondary structure at the 3′-end by chance. The observation that the BC1 RNA, which has been demonstrated to be a dominant master gene in the amplification and evolution of rodent ID elements (4), is capable of efficiently directing self-primed reverse transcription strongly suggests that self-priming is one of the factors in SINE amplification efficiency. However, since the abundance of ID elements in rodents spans three orders of magnitude (4), with guinea pig and rats constituting the extremes, involvement of the 3′ secondary structure in priming reverse transcription is clearly not the sole parameter, since the 3′ secondary structures do not differ at all in, for example, rat (130 000 ID copies) and syrian hamster (2000 ID copies). Despite the fact that in rat additional master genes contribute to ID dispersal (4), other factors such as appropriate temporal and spatial transcription in the germline, transcript stability, availability and activity of reverse transcriptase or integration of the cDNA copy into the genome may be equally important.

BC1 RNA is the only example of a SINE master gene that has been demonstrated to be able to prime its own reverse transcription. However, there are examples of functional transcripts that have inserted abundant truncated copies of themselves by a self-primed mechanism. Human U3 small nuclear (sn)RNA retropseudogenes have 3′ truncations and are flanked by direct repeats. In an *in vitro* assay using AMV reverse transcriptase and isolated *in vivo* transcribed U3 snRNA, it was shown that U3 snRNA was able to prime its own reverse transcription. The size and sequence of the *in vitro* produced U3 cDNA correlated well with the size and sequence of the genomic U3 retropseudogenes (36). Other examples include the human 7SK RNA, of which there are ~3000 copies in the human genome (of these <0.5% may represent true 7SK genes). The vast majority represent 7SK pseudogenes and include some that are formed through DNA-mediated duplication and some due to reverse transcription of a RNA intermediate. Sequences of some of the 7SK pseudogenes have revealed that most have 3′ truncations and are flanked by direct repeats, suggesting formation through a RNA intermediate (37). 7SK RNA in an *in vitro* assay is also able to prime its own reverse transcription and was shown to contain a possible hairpin structure at its 3′-end (38,39). However, the self-primed 7SK

cDNA was larger than the genomic copies of the 7SK retropseudogenes. This may indicate differences in the process of integration between U3 snRNA and 7SK RNA retropseudogenes. Human and rat U4 snRNA pseudogenes that have 3′ truncations and are flanked by direct repeats are also believed to have arisen by integration of self-primed reverse transcribed U4 snRNAs (40,41). The abundant nature of these retropseudogenes attests to the effectiveness of a self-primed, reverse transcribed mechanism for retroposition.

The expression of BC1 RNA in spermatogonia (H.Tiedge, Z.Zakeri and J.Brosius, unpublished data) suggests that BC1 RNA is available for germline retroposition. Branciforte and Martin have shown that prepubertal mouse testes contain full-length L1 transcripts and ORF1 protein (42). They also showed that ORF1 protein is localized to the spermatocytes, the cell type undergoing meiosis. These full-length L1 transcripts could conceivably provide a source of reverse transcriptase (RT) encoded by ORF2 (43). In humans, a functional RT is encoded in ORF2 by L1.2A, a putatively functional human L1 element (44,45). This L1 RT uses several primer–template combinations (45). Other sources of RT may be provided by endogenous retroviruses or through a retroviral infection (1). The source of RT and whether the RT may prefer certain primer–template combinations in the retroposition process is unknown. However, L1.2A RT showed a <2-fold preference in primer–template combinations when DNA primers were used on RNA templates (45). This evidence suggests that BC1 RNA accessible to RT *in vivo* would be reverse transcribed.

The BC1 3′ structure may play a role other than or in addition to self-priming, such as interacting with RNA pol III termination (46). Another possible role would be that it has been selected for its ability to interact with parts of the retroposition mechanism, similar to 3′ recognition of R2 retrotransposon RNA by its RT (47). Alternatively, it is likely that a 3′ hairpin structure stabilizes the RNA against exonuclease degradation, which will also contribute to the availability of the RNA for the retroposition process.

The finding that BC1 RNA is able to self-prime extends our understanding of the evolution and formation of a SINE master gene. There are many factors that are required for the formation of an efficient retroposon master gene. Among these could be upstream sequences for efficient transcription of the retroposon (in germline tissue), a stable RNA and a unique region that allows for hairpin formation (for self-primed reverse transcription); these last two elements may be interrelated (5,22–24). These specific requirements would significantly limit the number of potential SINE master genes. Insertion of a retroposon in a genomic location does not guarantee that all of the above conditions are satisfied. Only a relatively small subset of SINE genes seem capable of active transcription (16,21). It is also clear that some SINE loci which are not actively involved in retroposition are expressed (14,16,17). Even with loci which express RNA, the RNA pol III terminator may be located in a region that does not allow stable hairpin formation. This may affect RNA termination and stability, as well as the self-priming ability of the transcript (48).

In this model, the probability of forming a new and effective master gene is the product of the probabilities of a locus being transcriptionally active, forming an effective self-priming hairpin and perhaps of other features which influence the stability, reverse transcription or integration of the SINE transcript or cDNA. A requirement for even a small hairpin at the 3′-end could represent a significant level of selection for SINE loci which are capable of retroposition versus those that cannot. Multiplied by factors for a limited number of genomic loci supporting active pol III transcription (which are also likely to be one or two orders of magnitude at least), unknown factors for the probability of the RNA pol III terminator being close enough to the oligo(dA)-rich region of the SINE or sequences in the RNA transcript resulting in an unstable transcript or inability to bind RT, it is easy to see why even with an excess of >100 000 copies, many SINE families seem to have only a few highly active master genes at any one time (49; reviewed in 11,18,19).

ACKNOWLEDGEMENTS

We thank Susan Hoffman, Harvey Mohrenweiser and Santosh Arcot for comments on the manuscript, Joomyeong Kim for helpful discussions, Mark Batzer for the oligonucleotides and Bart Beeman and Rick Wooten for help with the graphics. This work was supported by Public Health Service grant GM45668 to P.L.D. and, in part, MH38819 to J.B. Work was performed under the auspices of the US DOE by LLNL under contract no. W-7405-ENG-48.

REFERENCES

1 Deininger,P.L. (1989) In Berg,D.E. and Howe,M.M. (eds), *Mobile DNA*. American Society for Microbiology, Washington, DC, pp. 619–636.
2 Daniels,G.R. and Deininger,P.L. (1985) *Nature (Lond.)*, **317**, 819–822.
3 Sakamoto,K. and Okada,N. (1985) *Nucleic Acids Res.*, **13**, 7195–7206
4 Kim,J., Martignetti,J.A., Shen,M.R., Brosius,J. and Deininger,P. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 3607–3611.
5 Martignetti,J.A. and Brosius,J. (1995) *Mol. Cell. Biol.*, **15**, 1642–1650.
6 DeChiara,T.M. and Brosius,J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 2624–2628.
7 Martignetti,J.A. and Brosius,J. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 9698–9702.
8 Geiduschek,E.P. and Tocchini-Valentini,G.P. (1988) *Annu. Rev. Biochem.*, **57**, 873–914.
9 Gabrielsen,O.S. and Sentenac,A. (1991) *Trends Biochem. Sci.*, **16**, 412–416.
10 Matera,G., Hellmann,U. and Schmid,C. (1990) *Mol. Cell. Biol.*, **10**, 5424–5432.
11 Deininger,P.L., Batzer,M.A., Hutchison,C.A. and Edgell,M.H. (1992) *Trends Genet.*, **8**, 307–311.
12 Brosius,J. and Tiedge,H. (1996) *Virus Genes* **11**, 163–179.
13 Cheng,J.-G., Tiedge,H. and Brosius,J. (1996) *DNA Cell Biol.*, **15**, 549–559.
14 Liu,W.-M., Maraia,R., Rubin,C. and Schmid,C. (1994) *Nucleic Acids Res.*, **22**, 1087–1095.
15 Kim,J., Kass,D.H. and Deininger,P.L. (1995) *Nucleic Acids Res.*, **23**, 2245–2251.
16 Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D. (1992) *J. Mol. Biol.*, **226**, 689–706.
17 Maraia,R.J., Driscoll,C.T., Bilyeu,T., Hsu,K. and Darlington,G.J. (1993) *Mol. Cell. Biol.*, **13**, 4233–4241.
18 Deininger,P.L. and Batzer,M.A. (1993) *Evol. Biol.*, **27**, 157–196.
19 Schmid,C. and Maraia,R. (1992) *Curr. Opin. Genet. Dev.*, **2**, 874–882.
20 Weiner,A.M., Deininger,P.L. and Efstratiadis,A. (1986) *Annu. Rev. Biochem.*, **55**, 631–661.
21 Slagel,V.K. and Deininger,P.L. (1989) *Nucleic Acids. Res.*, **17**, 8669–8682.
22 Ullu,E. and Weiner,A.M. (1985) *Nature (Lond.)*, **318**, 371–374.
23 Young,L.S., Takahashi,N. and Sprague,K.U. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 374–378.
24 Jagadeeswaran,P., Forget,B.G. and Weissman,S.M. (1981) *Cell*, **26**, 141–142.
25 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Plainview, NY.

26 Gurevich,V.V., Pokrovskaya,I.D., Obukhova,T.A. and Zozulya,S.A. (1991) *Anal. Biochem.*, **195**, 207–213.
27 Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (eds), (1992) *Current Protocols In Molecular Biology.* Green Publishing Associates and Wiley-Interscience, New York, NY.
28 Meador,J. Cannon,B., Cannistraro,V.J. and Kennell,D. (1990) *Eur. J. Biochem.*, **187**, 549–553.
29 Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.
30 Mertz,L.M. and Rashtchian,A. (1994) *Anal. Biochem.*, **221**, 160–165.
31 Jaeger,J.A., Turner,D.H. and Zuker,M. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 7706–7710.
32 McKinnon,R.D., Danielson,P., Brow,M.A., Bloom,F.E. and Sutcliffe,J.G. (1987) *Mol. Cell. Biol.*, **7**, 2148–2154.
33 Sapienza,C. and St-Jacques,B. (1986) *Nature (Lond.)*, **319**, 418–420.
34 Sutcliffe,J.G., Milner,R.J., Gottesfeld,J.M. and Lerner,R.A. (1984) *Nature (Lond.)*, **308**, 237–241.
35 Sutcliffe,J.G., Milner,R.J., Bloom,F.E. and Lerner,R.A. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4942–4946.
36 Bernstein,L.B., Mount,S.M. and Weiner,A.M. (1983) *Cell*, **32**, 461–472.
37 Murphy,S., Altruda,F., Ullu,E., Tripodi,M., Silengo,L. and Melli,M. (1984) *J. Mol. Biol.*, **177**, 575–590.
38 Suh,D., Yuan,Y., Henning,D. and Reddy,R. (1989) *Eur. J. Biochem.*, **186**, 221–226.
39 Wassarman,D.A. and Steitz,J.A. (1991) *Mol. Cell. Biol.*, **11**, 3432–3445.
40 Hammarström,K., Westin,G. and Pettersson,U. (1982) *EMBO J.*, **1**, 737–739.
41 Saba,J.A., Busch,H. and Reddy,R. (1985) *Biochem. Biophys. Res. Commun.*, **130**, 828–834.
42 Branciforte,D. and Martin,S.L. (1994) *Mol. Cell. Biol.*, **14**, 2584–2592.
43 Sakaki,Y., Hattorri,M., Fujita,A., Yoshioka,K. and Georgiev,G.P. (1986) *Cold Spring Harbor Symp. Quant. Biol.*, **51**, 465–469.
44 Dombroski,B.A., Mathias,S.L., Nanthankumar,E., Scott,A.F. and Kazazian,H.H.,Jr (1991) *Science*, **254**, 1805–1808.
45 Mathias,S.L., Scott,A.F., Kazazian,H.H.,Jr, Boeke,J.D. and Gabriel,A. (1991) *Science*, **254**, 1808–1810.
46 Maraia,R.J., Kenan,D.J. and Keene,J.D. (1994) *Mol. Cell. Biol.*, **14**, 2147–2158.
47 Luan,D.D. and Eickbush,T.H. (1995) *Mol. Cell. Biol.*, **15**, 3882–3891.
48 Maraia,R.J., Chang,D.Y., Wolffe,A.P., Vorce,R.I. and Hsu,K. (1992) *Mol. Cell. Biol.*, **12**, 1500–1506.
49 Shen,M.R., Batzer,M.A. and Deininger,P.L. (1991) *J. Mol. Evol.*, **33**, 311–320.