Published online 4 February 2017

Nucleic Acids Research, 2017, Vol. 45, No. 7 ***3967–3984***
doi: 10.1093/nar/gkx073

Arm-specific cleavage and mutation during reverse transcription of 2′,5′-branched RNA by Moloney murine leukemia virus reverse transcriptase

Jessica Döring and Thomas Hurek*

Department of Microbe-Plant Interactions, CBIB (Center for Biomolecular Interactions Bremen), University of Bremen, PO Box 330440, D-28334 Bremen, Germany

Received April 06, 2016; Revised January 13, 2017; Editorial Decision January 16, 2017; Accepted January 30, 2017

ABSTRACT

Branchpoint nucleotides of intron lariats induce pausing of DNA synthesis by reverse transcriptases (RTs), but it is not known yet how they direct RT RNase H activity on branched RNA (bRNA). Here, we report the effects of the two arms of bRNA on branchpoint-directed RNA cleavage and mutation produced by Moloney murine leukemia virus (M-MLV) RT during DNA polymerization. We constructed a long-chained bRNA template by splinted-ligation. The bRNA oligonucleotide is chimeric and contains DNA to identify RNA cleavage products by probe hybridization. Unique sequences surrounding the branchpoint facilitate monitoring of bRNA purification by terminal-restriction fragment length polymorphism analysis. We evaluate the M-MLV RT-generated cleavage and mutational patterns. We find that cleavage of bRNA and misprocessing of the branched nucleotide proceed arm-specifically. Bypass of the branchpoint from the 2′-arm causes single-mismatch errors, whereas bypass from the 3′-arm leads to deletion mutations. The non-template arm is cleaved when reverse transcription is primed from the 3′-arm but not from the 2′-arm. This suggests that RTs flip ∼180° at branchpoints and RNases H cleave the non-template arm depending on its accessibility. Our observed interplay between M-MLV RT and bRNA would be compatible with a bRNA-mediated control of retroviral and related retrotransposon replication.

INTRODUCTION

Reverse transcriptase (RT) is the virus-encoded enzyme responsible for converting the single-stranded (ss) RNA genome of retroviruses and related retrotransposons into integration-competent double-stranded (ds) DNA. The enzyme possesses both an RNA- and DNA-dependent DNA polymerase and a ribonuclease H (RNase H) activity (1,2). RNase H hydrolyzes RNA in RNA/DNA hybrids via a two-magnesium-ion catalytic mechanism to generate 5′-phosphate and 3′-hydroxyl ends (3,4). The enzymatic activities of RT are localized in two separate protein domains. The DNA polymerase domain occupies the N-terminal region and the RNase H domain comprises the C-terminal region of the protein (5). Crystal structures of RT complexed with ds nucleic acids have revealed that the distance in nucleotides between the polymerase and RNase H active sites is 17–18 base pairs (6,7). Further, the two active sites are arranged so that they are positioned on opposite strands of the nucleic acid (8). In the polymerization-competent binding mode of RT, the DNA polymerase active site is positioned over the 3′-end of the primer strand to initiate DNA synthesis, while the RNase H active site is located over the template strand (9). RT first uses the positive-sense viral genomic RNA as template to synthesize minus-strand DNA. DNA synthesis is primed from a transfer RNA (tRNA) molecule annealed adjacent to the 5′-end of genomic RNA. Synthesis continues through the 5′ unique sequence (U5) and the 5′repeat (R) region until the 5′-end of the RNA template, generating the so-called minus-strand strong-stop DNA [(−)ssDNA]. After removal of the template RNA by RT's RNase H activity, the (−)ssDNA is transferred from the 5′- to the 3′-end of the same or the second viral RNA molecule (minus-strand transfer) using the complementarity between the 5′ R region of the DNA strand and the 3′ R region at the 3′-end of genomic RNA. After the (−)ssDNA hybridizes to the 3′-end R region, minus-strand DNA synthesis on the viral RNA genome resumes (1,2). During and following minus-strand DNA synthesis RNase H degrades the RNA template using the three known modes of cleavage: internal, DNA 3′-end-directed, and RNA 5′-end-directed (5,10). Specific purine-rich RNA fragments, designated polypurine tracts (PPTs), are left over and serve as primers to initiate the plus-strand DNA synthesis on the minus-strand DNA template (1,2). RT binds the RNA primer in the polymerase-competent orientation and positions thereby the RNase H active site over the DNA tem-

*To whom correspondence should be addressed. Tel: +49 421 218 62861; Fax: +49 421 218 62873; Email: thurek@uni-bremen.de

plate (9). Cleavage of the non-template RNA requires a reorientation of the replicating enzyme on the hybrid (11,12). The enzyme flips around 180°, RNase H is positioned over non-template strand RNA and is active, while DNA polymerase is positioned over template strand DNA and is inactive (9,11,12).

The coordination of DNA polymerase and RNase H activity of RT is still under debate. Modeling studies (13), crystallographic (14) and biochemical data (15) suggested that either the DNA polymerase or the RNase H catalytic site can engage the RNA/DNA hybrid at a time. From these results, it was concluded that DNA polymerization and RNase H cleavage are mutually exclusive and occur sequentially (13–15). Recently, it was demonstrated that RNase H can still be active when both catalytic sites engage the hybrid simultaneously (16,17). This result was obtained by trapping the DNA 3′-end in the polymerase active site. From this finding, it was proposed that both active sites can engage the hybrid at the same time (16,17) and that both activities can work simultaneously (17). Although, there is no general agreement on whether both activities of RT occur sequentially or simultaneously, it is widely accepted that RT's RNase H activity is inefficient during continuous DNA polymerization (15,17) and efficient during pausing of DNA synthesis (18–21).

Secondary structures such as hairpins in the template cause pausing of DNA synthesis and RT RNase H carries out one to two pause-related cleavages in the RNA/DNA hybrid downstream from the pause site in the RNA template (18–21). Genomic template degradation at pause sites promotes template switching and increases the retroviral recombination rate during minus-strand DNA synthesis (22,23). Apart from secondary structures, RNA can also form branched structures containing a 2′,5′-phosphodiester bond. Branched RNA (bRNA) molecules contain two RNA strands which branch out from a nucleotide (branchpoint nucleotide) into a 2′-arm and a 3′-arm via vicinal 2′,5′- and 3′,5′-phosphodiester bonds, respectively (24,25). When RTs bypass branchpoints, they misread simultaneously the branched nucleotide through the 2′,5′- (26–28) and 3′,5′-phosphodiester bond (29). Branched lariat RNAs are naturally formed by the spliceosome (30,31) and by group II introns (32–34). Additional branched nucleic acid structures such as Y-like *trans*-splicing intermediates were found in eukaryotes (35–37) and multicopy single-stranded DNAs were discovered in prokaryotes (38,39). Furthermore, it has been proposed that the retrotransposon Ty1 genomic RNA has a 2′,5′-branched lariat structure (40). Cheng and Menees (40) suggested that joining of the 5′ R region (representing the 2′-arm) to a branched nucleotide near the 3′ R region (representing the 3′-arm) of the same RNA molecule might facilitate the intramolecular minus-strand transfer. Based on this intriguing model, it has been proposed that the 2′,5′-branched form of genomic RNA is common among retroviruses and related retrotransposons (41,42). However, the 2′,5′-branched lariat form of viral genomic RNA is controversially discussed (43–45) because direct evidence for this structure is lacking.

It is well established that branchpoint nucleotides in template RNA cause pausing of DNA synthesis by RTs (44,46–50). It is likely, therefore, that RT RNase H generates pause-related cleavages in the RNA template while the enzyme pauses during synthesis at the branchpoint. However, the pause-related cleavage pattern induced by a branchpoint has not been described previously. Moreover, DNA synthesis of bRNA can be primed from two different template strands (template arms). The questions arise (i) whether the non-template arm is cleaved by RT RNase H when DNA synthesis occurs on the other arm and (ii) whether during RT pausing at the branchpoint, RT's RNase H generates different cleavage pattern when DNA synthesis is primed from the 2′- or 3′-arm. This might impact the minus-strand transfer from the 2′- to the 3′-arm in the lariat model proposed by Cheng and Menees (40) and affect replication of retroviruses and related retrotransposons. Furthermore, the question is raised (iii) whether RT's DNA polymerase misreads differently the branched nucleotide through the 2′,5′- than through the 3′,5′-phosphodiester bond. To address these questions, we synthesized a long-chained bRNA oligonucleotide by splinted-ligation (Figure 1) (51,52). This bRNA construct allowed us to examine the RNA cleavage and mutational patterns generated by Moloney murine leukemia virus (M-MLV) RT during DNA synthesis from the 2′- and 3′-arm. For such assays, it is essential to prepare long-chained bRNA templates because appropriate cleavage and primer binding sites have to be provided. Currently, two methods are described for construction of bRNA molecules: the solid-phase and the deoxyribozyme-mediated synthesis. The solid-phase method permits a sequence-independent synthesis of bRNA (sequence can be freely chosen) (53)

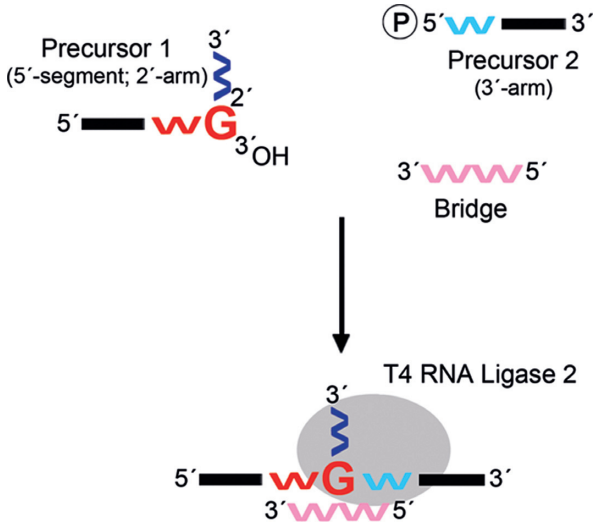

Figure 1. Scheme of the splinted-ligation method in bRNA construction. In this method, a 2′-5′ linked ribo-guanosine (G)-nucleoside in an RNA strand containing the 5′-segment and 2′-arm (precursor 1) is transformed into a branchpoint nucleotide by ligation to an RNA strand representing the 3′-arm (precursor 2). To do so, the two precursors are hybridized partially to a complementary RNA bridge. In this way, the 5′-phosphate of precursor 2 is brought close to the free 3′-hydroxyl of the 2′-5′ linked nucleoside of precursor 1. The two oligonucleotides are then joined by T4 RNA Ligase 2. Red, blue, and pink symbols 'w' represent RNA; the black line represents DNA. The 2′-5′ linked ribo-G-nucleoside in precursor 1 at nucleotide (nt) position 37 is highlighted. Nucleic acids downstream of a 2′-5′ linkage are plotted vertically in linear and branched oligonucleotides.

but synthesized bRNAs (53–55) are not suitable for cleavage and mutational analysis because of their size limitation. The deoxyribozyme-mediated technique permits synthesis of long-chained bRNA molecules but this method does not allow for a sequence-independent construction (56–58). The ligase-mediated construction (52) of bRNA is thus currently the only sequence-independent method to prepare long-chained bRNA molecules because long solid-phase synthesized oligonucleotides are used.

We found that pause-related cleavage and misprocessing of the branched nucleotide by M-MLV RT proceed arm-specifically. Our sequence analysis of arm-specific complementary DNAs (cDNAs) revealed that the bypass of the branchpoint from the 3′-arm leads to significantly more deletion mutations than from the 2′-arm. Conversely, bypass from the 2′-arm causes significantly more single mismatch errors at the position of the branched nucleotide than from the 3′-arm. Our cleavage analysis showed that M-MLV RT RNase H cleaves the non-template 2′-arm when DNA synthesis occurs on the 3′-arm template. In contrast, RT RNase H does not cleave the non-template 3′-arm when DNA synthesis occurs on the 2′-arm template. Our finding suggests that when DNA synthesis is primed from the 3′-arm of our bRNA, M-MLV RT flips around 180° at the branchpoint and RNase H cleaves in this binding mode the non-template RNA. Moreover, we found that the branchpoint in our bRNA induces multiple pause-related cleavages in both the 2′- and 3′-arm RNA templates. Our observations of the interplay between RT and the branched nucleotide led us to hypothesize that the proposed branchpoint in genomic RNA of Ty1 (40), of retroviruses and related retrotransposons (41,42) might have some functional importance to the virus life cycle. On one hand, the branched nucleotide in genomic RNA would promote template degradation to facilitate the minus-strand transfer. On the other hand, the branchpoint would reorient RT to delay template degradation after minus-strand transfer.

MATERIALS AND METHODS

General

The sequences of oligonucleotides used in this study are listed in Table 1. The sequences and hybridization temperatures of DNA probes can be found in Supplementary Table S1. Ethanol precipitation of nucleic acids was carried out with Pellet Paint Co-Precipitant according to the manufacturer's protocol (Merck Millipore) and resuspended in Tris-ethylenediaminetetraacetic acid (Tris–EDTA, TE) buffer (10 mM Tris–HCl, 1 mM disodium EDTA, pH 8.0) unless stated otherwise. Negative controls were carried out simultaneously without adding the enzyme. Polymerase chain reaction (PCR) was done in a final volume of 50 μl containing 1× Go Taq reaction buffer (Promega), 0.05 mM deoxynucleoside triphosphates (dNTPs) (Life Technologies), 25 pmol each of the forward and reverse primer (either 2–5 or 3–5), and 1.25 unit of GoTaq Polymerase (Promega). For reverse transcription (RT)-PCR with bRNA or precursor 1 as templates, 0.1 or 0.004 volume of the RT reactions, respectively, were subjected to PCR unless indicated otherwise. The PCR consisted of an initial denaturation step at 94°C for 2 min, followed by 35 cycles of 94°C for 30 s, 62.5°C for 30 s and 72°C for 30 s, and a final extension step at 72°C for 5 min. Amplicons were column-purified using the QIAquick Nucleotide Removal Kit according to the manufacturer's protocol (Qiagen). UV spectrophotometric quantification of DNA was done by using a Nanodrop 2000 spectrophotometer (Thermo Scientific). Restriction digestions were performed in a final volume of 10 μl containing 1× CutSmart buffer (New England BioLabs, NEB) and 3 units of restriction endonuclease (NEB). The reactions were incubated at 37°C for 60 min. Heat inactivation of enzymes was carried out at 70°C for 15 min. For purification of nucleic acids from polyacrylamide gels, the gel slice nutated in three volumes of TE buffer at 37°C overnight. After ethanol precipitation, gel-purified double-stranded or single-stranded nucleic acids were dissolved in TE buffer containing 15 mM sodium chloride (NaCl) or TE buffer, respectively. Double-stranded oligonucleotides were prepared by mixing 200 fmol of complementary single strands in a final volume of 4 μl TE containing 60 mM NaCl. The mixture was heated to 94°C for 10 s, cooled down fast to 78°C, and then cooled down slowly to 39°C. For gel electrophoretic analysis, 10 and 20 fmol of ds oligonucleotides, between 25 and 100 fmol of ss oligonucleotides, or 0.1 volume of the RT-PCR reaction were used. Unlabeled or fluorescently labeled single-stranded oligonucleotides purchased from Eurofins MWG Operon, or a Low Molecular Weight DNA Ladder (NEB) were used as size markers for gel electrophoresis. Size markers in base pairs (bp) or in nucleotides (nt) are given on the left of the gel. All experiments were repeated at least twice if not indicated otherwise.

Gel electrophoresis and Imaging

Denaturing polyacrylamide gels consisted of 15% acrylamide, 0.17% *N*,*N*′-methylenebisacrylamide, 8 M urea in 1× Tris–borate–EDTA (TBE) buffer (89 mM each Tris-base and boric acid, 2 mM EDTA, pH 8.3). Samples were mixed with two volumes of urea loading buffer (8 M urea, 50 mM Tris–HCl pH 7.5, and 20 mM EDTA pH 8.0) and heated at 70°C for 10 min before loading. Native polyacrylamide gels consisted of 12% acrylamide, 0.17% *N*,*N*′-methylenebisacrylamide in 1× TBE. One-sixth volume of Gel loading dye, Blue (6×) (NEB) was added to the samples. One-tenth volume of loading-dye, (10×) containing 25% Ficoll-400 and 0.4% xylene cyanol was added to the samples for gel electrophoresis on a 4% agarose gel in 1× Tris-acetate–EDTA (TAE) buffer (40 mM Tris-acetate and 1 mM EDTA, pH 8.3). Gel electrophoreses were performed according to standard protocols. Polyacrylamide gels were stained with SYBR Gold following the manufacturer's instructions (Life Technologies), and agarose gels were stained with ethidium bromide solution (0.5 μg/ml). Stained or fluorescently labeled nucleic acids are presented with a dark or clear background, respectively, and were visualized using a Mode Imager Typhoon FLA 9500 (GE Healthcare). Fluorescently labeled nucleic acids were quantified using ImageQuant software v5.2 (Molecular Dynamics). The percentage of a particular fluorescently labeled nucleic acid was calculated from the sum of all fluorescently labeled nucleic acids per lane.

Table 1. List of oligonucleotides used in this study

Oligonucleotide[a]	Sequence[b] (5' to 3')
Chimeric RNA/DNA oligonucleotide with a 2'-5' linkage (60), precursor 1[c,d]	CTGTAAGAACACTAAGACGC**UCGAACGGCGCGCCG GG2',5'GAUCCUCUAGAGUCGACCUGCAC**
5'-phosphorylated chimeric RNA/DNA oligonucleotide (59), precursor 2[e]	**CGAGAUGCCAGCGCUACCGUCGU**CTGAGGTCTTA GTGTTCTTGAACACGGTCGCGAGAG
Bridge RNA oligonucleotide (22)	**GCUGGCAUCUCGCCCGGCGCGC**
Forward primer F (20)[f]	CTGTAAGAACACTAAGACGC
Reverse primer 2-5 (16)[g]	GTGCAGGTCGACTCTA
Reverse primer 3-5 (21)	ACCGTGTTCAAGAACACTAAG
Reverse primer O (50)	CAGACGACGGTAGCGCTGGCATCTCGCCCGGCGCG CCGTTCGAGCGTCTT
Oligonucleotide representing cDNA until the 2'-5' linkage and 2',5'-branch (23)	GTGCAGGTCGACTCTAGAGGATC
Oligonucleotide representing cDNA until the 3',5'-branch (51)	ACCGTGTTCAAGAACACTAAGACCTCAGACGACGGT AGCGCTGGCATCTCG
Oligonucleotide representing full-length cDNA through the 2'-5' linkage and 2',5'-branch (60)	GTGCAGGTCGACTCTAGAGGATCCCCGGCGCGCCG TTCGAGCGTCTTAGTGTTCTTACAG
Oligonucleotide representing full-length cDNA through the 3',5'-branch (88)	ACCGTGTTCAAGAACACTAAGACCTCAGACGACGGT AGCGCTGGCATCTCGCCCGGCGCGCCGTTCGAGC GTCTTAGTGTTCTTACAG
Oligonucleotide representing the 2'-arm (23)	GATCCTCTAGAGTCGACCTGCAC
Chimeric RNA/DNA oligonucleotide containing exclusively 3',5' linkages (60)	CTGTAAGAACACTAAGACGC**UCGAACGGCGCGCCG GGGAUCCUCUAGAGUCGACCUGCAC**
Oligonucleotide with the sequence of the bRNA construct without 2'-arm (96)	CTGTAAGAACACTAAGACGCTCGAACGGCGCGCCG GGCGAGATGCCAGCGCTACCGTCGTCTGAGGTCTTA GTGTTCTTGAACACGGTCGCGAGAG

[a]Oligonucleotides were purchased from Eurofins MWG Operon unless stated otherwise. The nucleotide length of each oligonucleotide is indicated in parentheses.

[b]RNA sequences are shown in bold and DNA sequences in normal capitals.

[c]The red ribonucleotide at the nucleotide position 37 is connected via a 2',5'-phosphodieser bond to the succeeding ribonucleotide.

[d]Oligonucleotide was synthesized by ChemGenes Corporation.

[e]Oligonucleotide was purchased from Eurogentec.

[f]5'-cy3-labeled primer was used for amplification of cDNA for T-RFLP analysis.

[g]5'-cy5-labeled primer was used to prime DNA synthesis on linear oligonucleotides.

Hybridization analysis

Nucleic acids were blotted to a positively charged nylon membrane (GE Healthcare Life Sciences) overnight by capillary transfer. Unless otherwise indicated, membranes were hybridized with the digoxigenin-labeled probe in hybridization buffer [5× saline sodium citrate (SSC) buffer (20× SSC buffer is 3 M NaCl and 0.3 M sodium citrate, pH 7.0), 5× Denhardt's solution and 0.5% sodium dodecyl sulfate (SDS)] overnight. Membranes were sequentially washed in 5× SSC buffer containing 0.1% SDS, 2× SSC buffer containing 0.1% SDS and 0.1× SSC buffer containing 0.1% SDS at the hybridization temperature for 15 min. Blot development using alkaline phosphatase-conjugated anti-digoxigenin antibody and the chemiluminescent substrate CDP-Star, as well as stripping the membrane was done as per manufacturer's instructions (Roche Life Science). Efficiency of stripping was monitored by re-development of the blot. After detecting no chemiluminescent signal, stripped blots were reprobed. For hybridization analysis of DNA targets, reprobed blots were used, and for analysis of RNA targets, blots were used only once. Chemiluminescent signals were detected with a Luminescent Image Analyzer LAS-3000 mini (Fujifilm Life Science).

Ligase-mediated construction of branched RNA

For bRNA construction, 250 pmol each of precursor 1, precursor 2 and RNA bridge were mixed in a final volume of 10 μl TE buffer containing 100 mM NaCl. The mixture was heated to 79°C for 10 s and allowed to cool slowly to 42°C. Fifty pmol of ds oligonucleotide from the hybridization reaction was used for ligation of the 5′-phosphate with the 3′-hydroxyl. The ligation reaction was performed in 20 μl containing 1× T4 Rnl2 reaction buffer (NEB), 12.5% (w/v) polyethylene glycol (PEG) 8000 or PEG 4000 and 5 units of T4 RNA Ligase 2 (NEB). The reaction and negative control were incubated at 37°C overnight and subsequently ethanol

precipitated. Ligation of the two precursors was confirmed by hybridization analysis (Supplementary Figure S1). The bRNA oligonucleotide was gel-purified twice.

Preparation of the 3′-fluorescein-labeled precursor 1

To label the 3′-end of precursor 1 using Terminal Deoxynucleotidyl Transferase (TdT), DNA has to be attached to its 3′-end because 3′ terminal ribonucleotides are poorly accepted by the enzyme. Addition of a poly-deoxyadenosine (poly-dA) tail to the 3′-end of precursor 1 was done in a final volume of 25 μl containing 60 pmol of precursor 1, 1× Poly(A) Polymerase reaction buffer (Affymetrix), 0.5 mM deoxyadenosine triphosphate (dATP) (Life Technologies) and 600 units of Yeast Poly(A) Polymerase (Affymetrix). The reaction was incubated at 37°C for 8 h and was column-purified using the QIAquick Nucleotide Removal Kit according to the manufacturer's instructions (Qiagen). The eluted oligonucleotide was ethanol precipitated and desalted by drop dialysis. The 3′-end labeling of precursor 1 was done in a final volume of 30 μl containing 1× TdT buffer (Life Technologies), 100 μM fluorescein-12-uridine triphosphate (Fluorescein-12-UTP) (Roche Life Science) and 30 units of TdT (Life Technologies). The reaction was incubated at 37°C for 60 min and the 3′-fluorescein-labeled precursor 1 was phenol/chloroform extracted, ethanol precipitated and gel-purified.

Primer extension reaction

For primer extension reactions, 250 fmol of template and 25 pmol of unlabeled reverse primer (either 2–5 or 3–5) or 500 fmol of 5′-cy5-labeled reverse primer 2–5 were mixed in a final volume of 6.2 μl TE buffer containing 15 mM NaCl. For primer extensions from the 2′- and 3′-arms, reverse primers 2–5 and 3–5, respectively, were used. In reactions with the fluorescein-labeled precursor 1, 500 fmol of 3′-fluorescein-labeled precursor 1, 25 pmol of reverse primer 3–5 (and 250 fmol of branched RNA) were used. All mixtures were heated to 94°C for 10 s, cooled down fast to 78°C, and then cooled down slowly to 27°C. Primer extension reactions were carried out in 25 μl containing 1× M-MLV RT reaction buffer (Promega), 0.5 mM dNTPs (Life Technologies) and 200 units of either M-MLV RT (H−) or M-MLV RT (H+) (Promega). The reaction and negative control were incubated at 42°C for 1 to 18 h, and RTs were heat-inactivated. Samples were ethanol precipitated and separated by polyacrylamide gel electrophoresis.

The primer extension reaction from the 2′-arm of our bRNA using M-MLV RT (H+) was phenol/chloroform extracted before ethanol precipitation. One-third volume of this sample was analyzed by gel electrophoresis and 0.3 volume was used for primer extension analysis on cleavage products using 100 fmol of reverse primer O and reverse primer 3–5, respectively.

Purification of bRNA and terminal-restriction fragment length polymorphism (T-RFLP) analysis

Purification of bRNA. After primer extension from the 2′-arm of our bRNA using M-MLV RT (H−), a proteinase K treatment was carried out in a final volume of 60 μl containing 100 mM Tris–HCl pH 7.5, 150 mM NaCl, 12.5 mM EDTA pH 8.0, 1% SDS and 12 μg proteinase K (Roche Life Science). The reaction was incubated at 55°C for 60 min. Nucleic acids were phenol/chloroform extracted, ethanol precipitated, and separated on a native polyacrylamide gel. The truncated primer extension product of bRNA was gel-purified and used for a second round of reverse transcription (RT). For RT reaction, one-fifth volume of the gel-purified product was mixed with 25 pmol of reverse primer 2–5 in a final volume of 17.5 μl RNase/DNase free water (MP Biomedicals). The mixture was heated to 94°C for 30 s, kept on ice, and the reverse transcription was started by adding 5 μl 5× M-MLV RT reaction buffer (Promega), 1.25 μl 10 mM dNTPs (Life Technologies), and 1 μl M-MLV RT (H−) (200 units) (Promega). The ethanol precipitated sample was resuspended in 10 mM Tris–HCl, pH 8.5 containing 7.5 mM NaCl. One half of the reverse transcribed fragments were used for restriction digestion with DdeI. The restriction endonuclease was heat-inactivated, the fragments were ethanol precipitated and used for a third round of reverse transcription, which was performed exactly as the second RT reaction. After ethanol precipitation, the sample was resuspended in TE buffer.

Monitoring of bRNA purification by T-RFLP. To monitor depletion of full-length cDNA from linear RNA and enrichment of full-length cDNA from bRNA, we used T-RFLP analysis, a common DNA fingerprinting technique. This method is based on restriction endonuclease digestion, electrophoretic separation, and quantification of fluorescently end-labeled PCR products. Length and quantity of the end-labeled restriction fragments reflect the polymorphism of the restriction site. We attempted to use equal amounts of input cDNA for amplification with a 5′-cy3-labeled forward primer. To achieve this, one-twenty fifth volume of the first RT reaction was diluted 1:8 in TE buffer and 1 μl from the dilution was subjected to PCR (unpurified). One-tenth volume of the second RT reaction was used for PCR (gel-purified). One-half volume of the precipitated reverse transcribed fragments from the third reaction were subjected to PCR (gel-purified + DdeI treated). RT-PCR amplicons were column-purified and quantified. Sixty ng of DNA was used for restriction digestion with BamHI-HF (High Fidelity) or NaeI. No heat inactivation of the enzymes was performed. The restriction digestions and negative control were analyzed by gel electrophoresis, and the end-labeled restriction fragments were quantified.

Sequencing of RT-PCR amplicons and bioinformatic analysis

RT-PCR products obtained from full-length cDNAs through the 3′,5′-branch were ethanol precipitated and gel-purified. One-tenth volume of the gel-purified amplicons was used in the cloning reaction. RT-PCR products obtained from full-length cDNAs through the 2′,5′-branch were treated with DdeI before ethanol precipitation. In this case one-fifth volume of the precipitated amplicons was used for cloning. RT-PCR products obtained from full-length cDNAs through the 2′-5′ linkage of precursor 1 were column-purified and quantified. Three ng of DNA was used

in the cloning reaction. All amplicons were cloned using CloneJET PCR Cloning Kit following the manufacturer's protocol (Thermo Scientific). Colony PCR was carried out using pJET1.2 forward and reverse primers (components of the kit) to analyze the colonies for presence of the insert. Inserts of positive clones were analyzed by Sanger sequencing using the microtitre plate (MTP) sequencing service provided by LGC Genomics. CLC Genomics Workbench v7.5 (Qiagen) was used for aligning the sequences to the corresponding reference sequence and evaluating the sequencing data.

Statistical analysis

Statistical analysis was done with GraphPad Prism v6.07 (Graphpad software). Data were analyzed using one-way ANOVA (analysis of variance) followed by Holm-Sidak's multiple comparison test. $P > 0.05$ was considered not significant.

RESULTS

Constructing the branched RNA oligonucleotide

Ligase-mediated construction of branched DNA oligonucleotides by the splinted-ligation technology (51) was reported previously by Mendel-Hartvig *et al*. (52). As illustrated in Figure 1, we used this method to construct our bRNA oligonucleotide. We chose RNA/DNA chimeric precursors for the bRNA construction because the DNA regions in bRNA (Figure 2A) cannot be cleaved by M-MLV RT RNase H. The non-cleavable DNA sequences of the 5′-segment and 3′-arm served as hybridization sites for probes, and simplified subsequent analysis of RNA cleavage products. The RNA sequences in our bRNA construct (Figure 2A) served also as hybridization sites for probes and facilitated subsequent examination of RNase H-mediated RNA hydrolysis. To test the suitability of the site-specific probes for analyzing RNA cleavage products, we performed hybridization analyses using oligonucleotides as positive and negative hybridization controls (Figure 2B).

RNase H-deficient M-MLV RT generates two cDNAs and the wild-type M-MLV RT only one cDNA from the branched RNA

Previous work on intron RNA lariats had shown that branched nucleotides stall DNA synthesis by RTs during primer extension from the 3′-arm of the branchpoint (47–50). Because of this, RTs generate truncated cDNA until the 3′,5′-branch from bRNA (47–50). However, Tuschl *et al*. (29) found that avian myeloblastosis virus (AMV) RT (H+) and Superscript II M-MLV RT (H−) also pass the branched nucleotide from this arm and generate two cDNAs from bRNA: (i) truncated cDNA until the 3′,5′-branch and (ii) full-length cDNA through the 3′,5′-branch. We wanted to examine whether wild-type M-MLV RT with (H+) and M-MLV RT without (H−) RNase H activity generate these two cDNAs from our bRNA. For this purpose, we performed primer extension analysis using the 3–5 reverse primer binding to the 3′-arm of our bRNA. When M-MLV RT extends this primer until the 3′,5′-branch, the enzyme produces truncated cDNA of 51 nucleotides (nt) in length (Figure 3A). When M-MLV RT extends this primer through the 3′,5′-branch to the end of the template, the enzyme generates full-length cDNA of 88 nt in length (Figure 3A). Primer extension analysis revealed that both RTs synthesized truncated cDNA until the 3′,5′-branch of our bRNA construct (Figure 3B). To our surprise, hybridization analysis of primer extension products clearly showed that only M-MLV RT (H−) could generate full-length cDNA through the 3′,5′-branch (Figure 3B, lane 3, right panel). Instead, multiple bands were visible on the gel when the primer extension on bRNA was catalyzed by M-MLV RT (H+) (Figure 3B, lane 4, left panel). Subsequent hybridization analyses revealed that these bands represented RNA cleavage products generated by RT RNase H during DNA synthesis from the 3′-arm (see below).

To increase the detection limit of full-length cDNA through the 3′,5′-branch, we subjected the cDNAs generated by both RTs to PCR. RT-PCR analysis revealed that the full-length PCR amplicon (88 bp) was only generated when reverse transcription was catalyzed by the RNase H-deficient RT (Figure 3C, lane 3). An amplicon somewhat smaller than the full-length PCR product was visible in the gel when reverse transcription was catalyzed by the wild-type RT (Figure 3C, lane 4). To determine whether the smaller amplicon corresponds as well to full-length cDNA or is an unspecific PCR product, we subjected the amplicons to hybridization using a probe specific for full-length cDNA through the 3′,5′-branch. Hybridization analysis clearly illustrated that the wild-type of M-MLV RT is not able to pass the branchpoint of our bRNA (Figure 3D, lane 2, right panel). In addition, RT-PCR analysis revealed that a 111 bp product was amplified by PCR from cDNA generated by both RTs (Figure 3C and D). Sequence analysis showed that this amplicon consisted of precursor 1 (5′-segment and 2′-arm) ligated at its 2′-arm to precursor 2 (3′-arm) (data not shown). When DNA synthesis is primed from the 3′-arm, both RTs are able to generate full-length cDNA from this unbranched (linear) ligation side-product. The unwanted ligation side-product was produced during the ligase-mediated construction of our bRNA. Apparently, this side-product migrated at the same position in the gel as the bRNA oligonucleotide and was co-purified with it.

Purifying the branched RNA oligonucleotide for sequence analysis of full-length cDNA through the 2′,5′-branch

Previous work on intron RNA lariats had shown that RNase H-deficient RTs incorporate deoxyguanosine monophosphate (dGMP) opposite to the branch guanosine when the enzymes read through the 2′,5′-branch (27,28,59). Because of this misincorporation, sequenced RT-PCR amplicons contain a cytidine at the position of the branch guanosine (G→C transversion) in the sense-strand (27,28,59). We wanted to confirm the insertion of a dGMP opposite to the branchpoint when M-MLV RT (H−) synthesizes DNA starting from the 2′-arm of our bRNA. For this purpose, full-length cDNAs through the 2′,5′-branch should be amplified with the forward and 2–5 reverse primers followed by sequencing of the resulting amplicons.

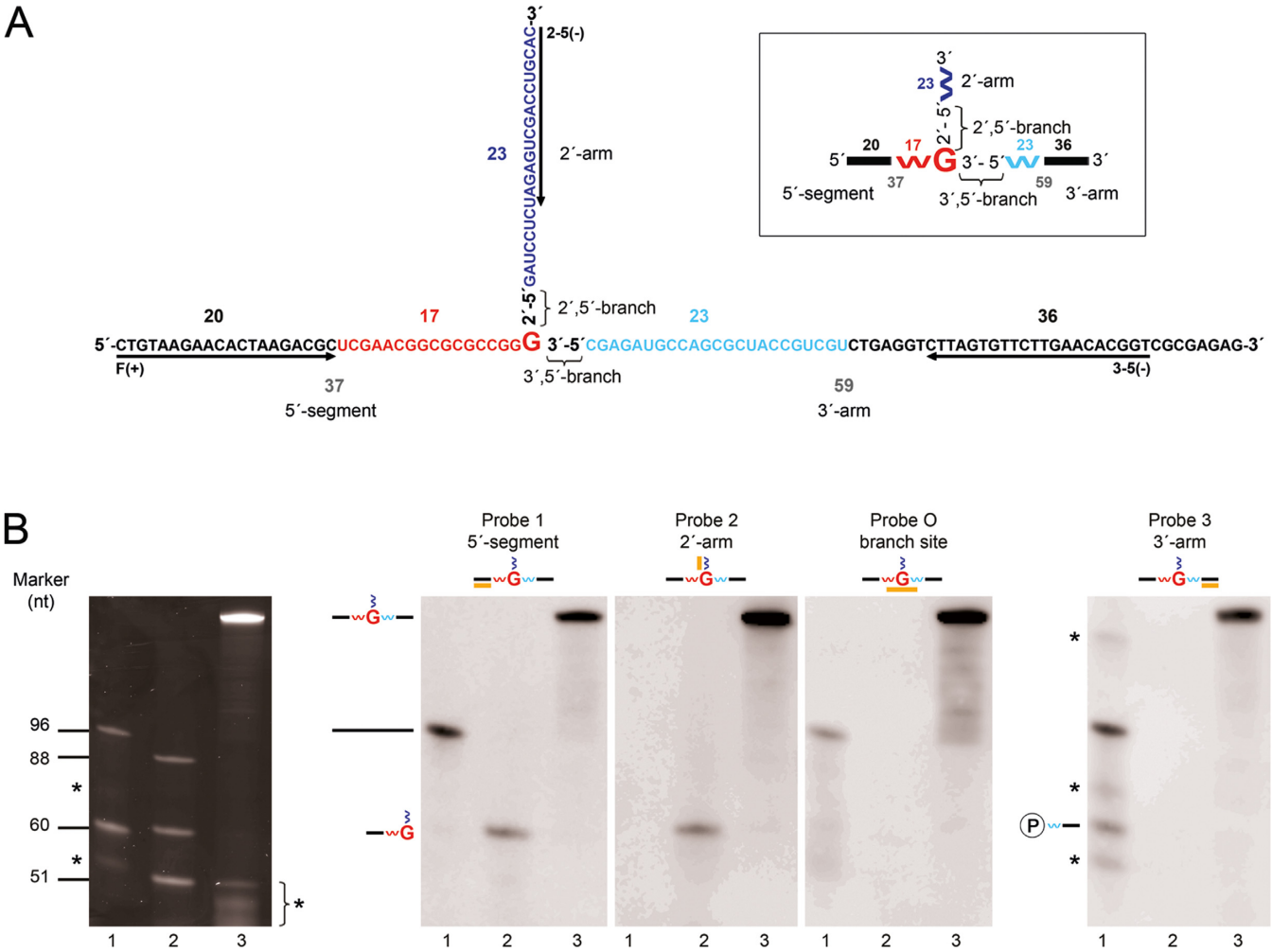

Figure 2. Branched RNA oligonucleotide and site-specific detection by probe hybridization. (**A**) Sequence and length of the bRNA construct. The bRNA is composed of a 37-mer 5′-segment, a 23-mer 2′-arm, and a 59-mer 3′-arm. The branchpoint nucleotide guanosine (branch guanosine) at nt position 37 is highlighted. The 2′,5′-phosphodiester bond of the branch guanosine is designated as 2′,5′-branch and its 3′,5′-phosphodiester bond as 3′,5′-branch. Red and blue nucleotides represent RNA; black nucleotides represent DNA. Numbers refer to the length in nucleotides of the respective nucleic acid. The orientation of the shown bRNA construct was defined as 'sense'. Primer binding sites are indicated by arrows. The primer with the suffix (+) is the forward primer and those with the (−) are the reverse primers. The primers 2–5 and 3–5 were used to prime DNA synthesis from the 2′- and 3′-arm, respectively. A schematic presentation of the constructed bRNA oligonucleotide is boxed in black. (**B**) Site-specific probes used in hybridization analyses. Lane 1: 96 nucleotides (nt) long oligonucleotide with the sequence of the bRNA construct without 2′-arm and precursor 2 (59 nt). Lane 2: oligonucleotide representing full-length cDNA through the 3′,5-branch (88 nt), precursor 1 (60 nt), and oligonucleotide representing cDNA until the 3′,5′-branch (51 nt). Lane 3: negative control for primer extension reaction using bRNA and reverse primer 3–5. Due to its unusual shape, the bRNA oligonucleotide exhibits an anomalous electrophoretic mobility in polyacrylamide gels and migrates more slowly than the corresponding linear oligonucleotide (oligonucleotide containing the 5′-segment and 3′-arm, 96 nt) (85). Samples were loaded on a 15% denaturing polyacrylamide gel (left), blotted and corresponding blots (right) were hybridized with site-specific probes (probes 1–3, probe O) as indicated. Probe 1 is specific for the 5′-segment of bRNA, probe 2 for the 2′-arm of bRNA, probe O for the overlapping branch site of bRNA, and probe 3 for the 3′-arm of bRNA. Probe-target regions are plotted in pictograms for each blot, with probes shown in orange. Hybridizing oligonucleotides from lanes 1–3 are schematically presented on the left of the blots. By-products in solid-phase synthesis of oligonucleotides loaded in lane 1 and of reverse primer 3–5 loaded in lane 3 are labeled with asterisks.

During our study, we noted that our bRNA construct was contaminated with a ligation side-product (Figure 3C and D) which consisted of precursor 2 ligated to precursor 1 at its 2′-arm. Similarly, precursor 1 might also be present. The presence of precursor 1 and the ligation side-product during DNA synthesis will complicate sequence analysis. Because a branched nucleotide is an obstacle to DNA synthesis by RT (44,46–50), M-MLV RT (H−) will frequently terminate at the branchpoint resulting in an incomplete cDNA/RNA hybrid (truncated bRNA hybrid). Linear RNA which lacks this obstacle will be favored over bRNA as template in the synthesis of full length cDNA. Consequently, the majority of full-length cDNAs will be produced from the linear RNA and only a small fraction will be the desired product derived from the bRNA.

We established two methods to purify our bRNA construct from linear RNA. The principle of the two purification methods is to deplete full-length hybrids and to recycle truncated bRNA hybrids for further rounds of primer extension. First, our (contaminated) bRNA was reverse transcribed with M-MLV RT (H−) using the 2–5 reverse primer binding to the 2′-arm. The resulting full-length hy-

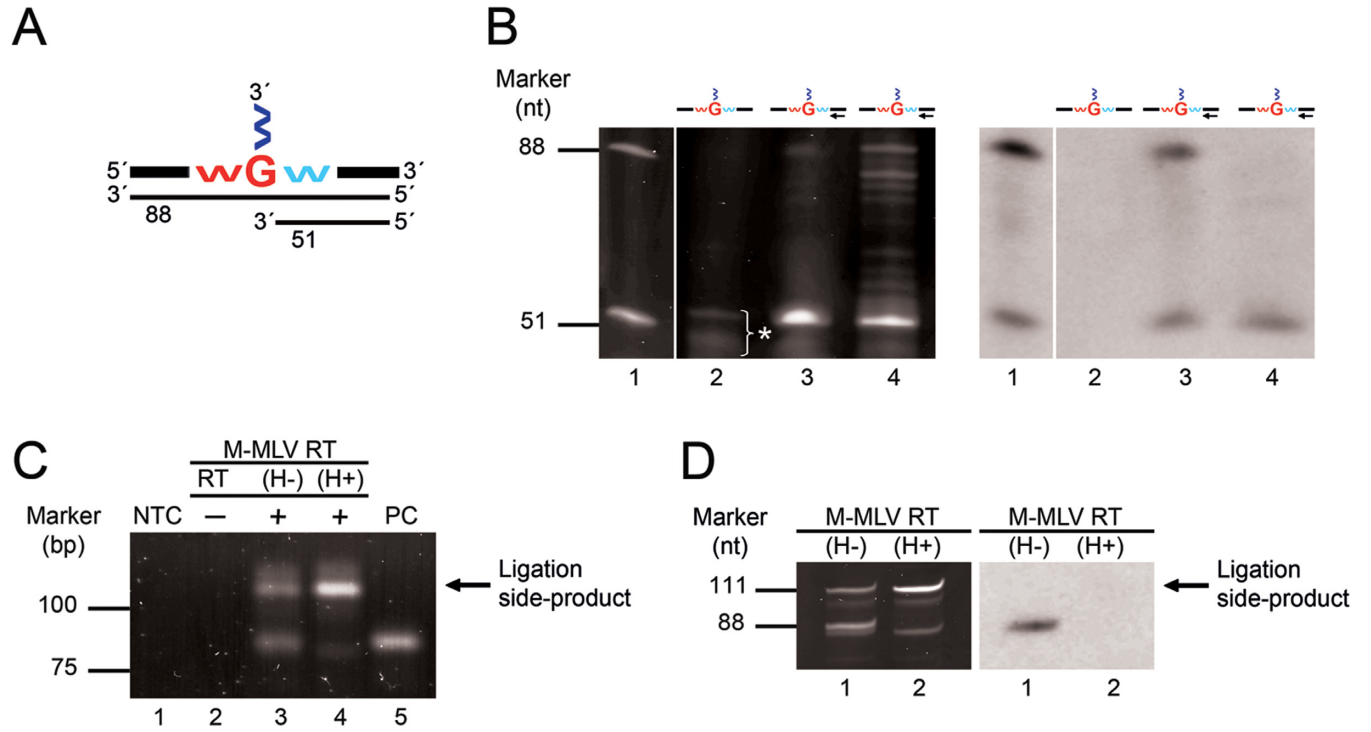

Figure 3. Detection of full-length cDNA through the 3′,5′-branch and truncated cDNA until the 3′,5′-branch. (**A**) Schematic presentation of full-length and truncated cDNA generated by primer extension from the 3′-arm of our bRNA construct. Numbers below lines refer to the lengths in nucleotides of the two cDNAs. (**B**) Analysis of primer extension products synthesized by M-MLV RT (H−) and (H+) from the 3′-arm by probe hybridization. Lane 1: oligonucleotide representing full-length cDNA through the 3′,5′-branch (88 nt) and oligonucleotide representing truncated cDNA until the 3′,5′-branch (51 nt). Lane 2: negative control for primer extension reaction. Asterisk indicates by-products in solid-phase synthesis of reverse primer 3–5. Lanes 3 and 4: primer extension reactions from the 3′-arm of bRNA using M-MLV RT (H−) and (H+), respectively. The black arrow in the pictogram shows the primer-target region and direction of primer extension. Oligonucleotides from lane 1 were used as size markers for gel electrophoresis and as positive hybridization controls. Samples were separated on a 15% denaturing polyacrylamide gel (left), blotted and hybridized with probe tc3-5 (right). This probe is specific for full-length cDNA through the 3′,5′-branch and truncated cDNA until the 3′,5′-branch. (**C**) Detection of full-length cDNA through the 3′,5′-branch of our bRNA construct by RT-PCR analysis. One fmol of oligonucleotide corresponding to full-length cDNA through the 3′,5′-branch was used as a positive control for PCR. Lane 1: no template control (NTC) for PCR. Lane 2: negative control for RT-PCR (no RT was added). Lanes 3 and 4: RT-PCR amplicons. Lane 5: positive control (PC). The RT-PCR product corresponding to the unwanted ligation side-product is noted. Samples were loaded on a 4% agarose gel. (**D**) Analysis of RT-PCR amplicons by probe hybridization. Sixty ng of column-purified RT-PCR amplicons (lanes 1 and 2) shown in panel (C) were separated on a 15% denaturing polyacrylamide gel (left), blotted and hybridized with probe flc3-5 (right). This probe is specific for full-length cDNA through the 3′,5′-branch. The ligation side-product is noted as in panel (C). Colors as in Figure 1.

brids between full-length cDNA and template, and truncated bRNA hybrids were separated by gel electrophoresis on a native polyacrylamide gel (Figure 4A, lane 4). The single product band visible in the gel was determined as the truncated bRNA hybrid using constructed hybrids as size markers (Figure 4A). In order to detect full-length hybrids, we performed hybridization analysis using a probe specific for full-length cDNA. Because hybridization analysis revealed that truncated bRNA hybrids were clearly separated from full-length extension products (Figure 4A, lane 4, right panel), we have gel-purified them. We then removed the bound truncated cDNA from bRNA by heat-denaturation and used the free bRNA as template for another round of primer extension from the 2′-arm. Amplicons of the subsequent PCR were analyzed by restriction digestion.

We exploited the ability of M-MLV RT (H−) to incorporate the correct deoxycytosine monophosphate (dCMP) opposite to the 2′-5′ linked ribo-G-nucleoside in linear RNA (Supplementary Table S2) and the incorrect dGMP opposite to the branch guanosine in bRNA (27,28,59). The resulting cDNA polymorphism can be analyzed by restriction digestion of RT-PCR products. For this purpose, we designed the sequence of precursor 1 so that the correct incorporation of dCMP generates a BamHI and the misincorporation of dGMP generates a NaeI restriction site in RT-PCR amplicons (Figure 5). To facilitate analysis of BamHI and NaeI restriction digestions of RT-PCR products, we amplified the full-length cDNA with a 5′-cy3-labeled forward primer and quantified the BamHI and NaeI restricted end-labeled PCR products by terminal-restriction fragment length polymorphism (T-RFLP) analysis.

The restriction digestion with BamHI showed that the gel-purified truncated bRNA hybrids were still contaminated with linear RNA (Supplementary Figure S2, lane 2), suggesting that linear RNA templates co-migrated with truncated bRNA hybrids in the gel and were co-purified. Next, to make the generated full-length hybrids of the primer extension reaction on gel-purified truncated bRNA hybrids inaccessible to RT-PCR, we truncated them by restriction digestion. We designed the sequence of precursor 1 to contain a DdeI restriction sequence in the 5′-segment. When M-MLV RT (H−) synthesizes full-length

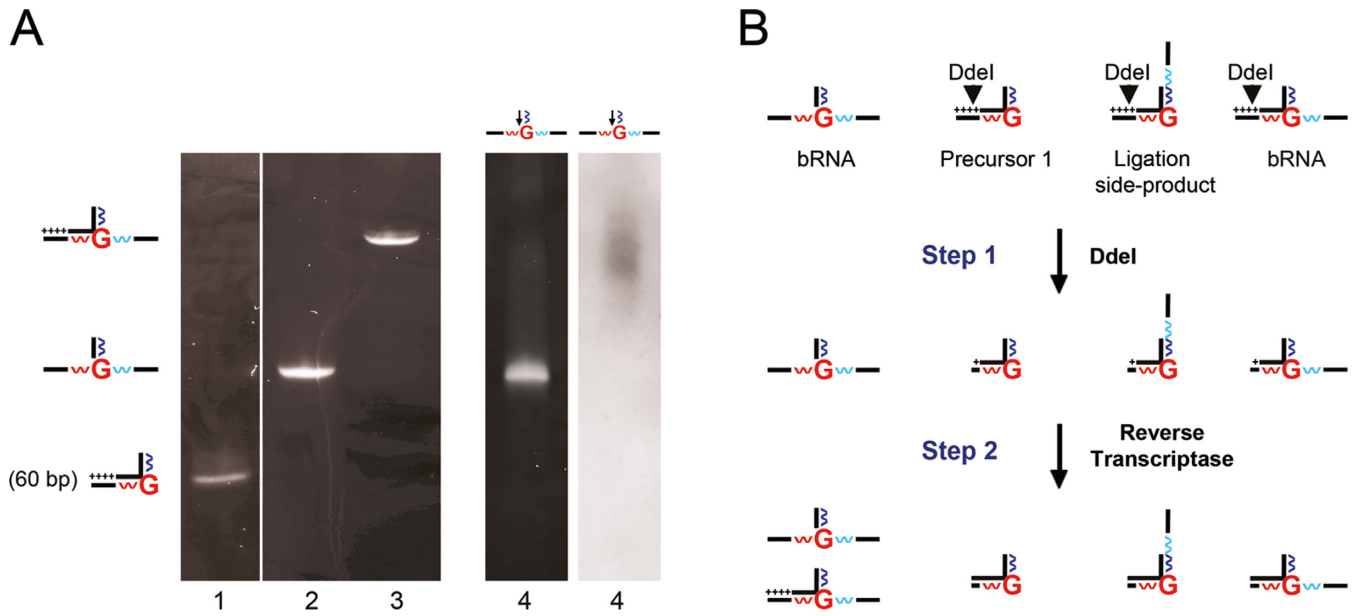

Figure 4. Depletion of full-length hybrids from bRNA. (**A**) Gel purification of truncated bRNA hybrids. Lane 1: hybrid between precursor 1 and full-length cDNA. Lane 2: hybrid between bRNA and truncated cDNA until the 2′,5′-branch. Lane 3: hybrid between bRNA and full-length cDNA. Lane 4: products of primer extension reaction from the 2′-arm of bRNA using M-MLV RT (H−). Hybrids from lanes 1–3 are schematically presented on the left of the gel. Pluses in pictograms indicate the forward primer binding site (PBS) of full-length cDNA. Hybrids (lanes 1–3) were used as size markers for electrophoresis on a 12% native polyacrylamide gel. The primer extension reaction from lane 4 was blotted and hybridized with probe flc2-5 (right). This probe is specific for full-length cDNA through the 2′-5′-linkage of linear RNA and bRNA. Arrows in pictograms as in Figure 3B. (**B**) Scheme showing truncation of the DNA duplex of full-length hybrids by restriction digestion. Truncated and full-length bRNA hybrids, as well as full-length linear RNA hybrids are presented schematically at the top. Pluses in pictograms as in panel (A). The DdeI restriction site in the DNA duplex of full-length hybrids is shown. DdeI restriction digestion removes the forward PBS of full-length cDNAs and truncates templates (step 1). In the next round of primer extension from the 2′-arm, M-MLV RT (H-) generates truncated and full-length cDNA from bRNA, and truncated cDNAs from truncated templates (step 2). Colors as in Figure 1.

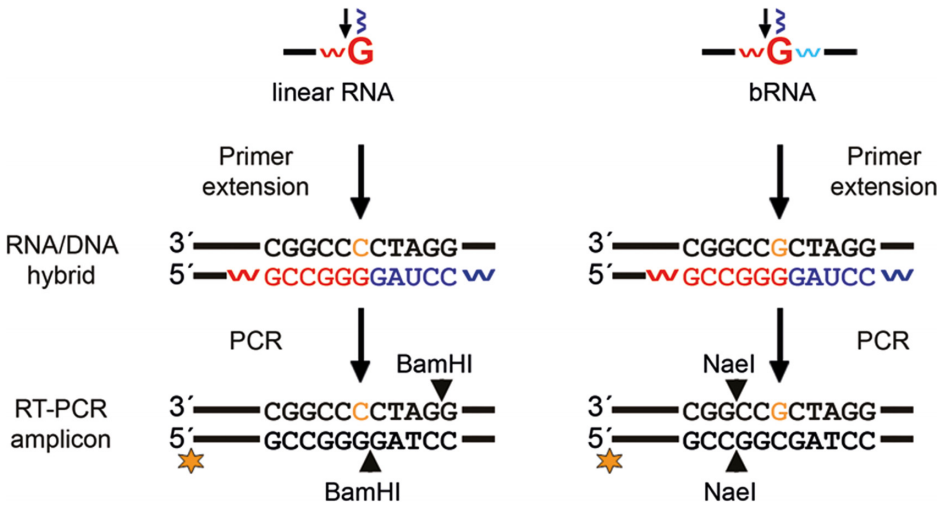

Figure 5. RT-PCR amplicons from linear RNA and bRNA can be distinguished by restriction digestion. Linear and bRNA templates are presented schematically at the top. Arrows in pictograms as in Figure 3B. Orange nucleotides indicate the incorporated nucleotide by M-MLV RT (H−) opposite to the 2′-5′ linked ribo-G-nucleoside in linear RNA and in bRNA. They can be identified in RT-PCR amplicons using terminal-restriction fragment length polymorphism (T-RFLP) analysis. For this purpose, the forward primer was labeled at the 5′-end with the fluorophore cy3 (indicated by a star). Colors as in Figure 1.

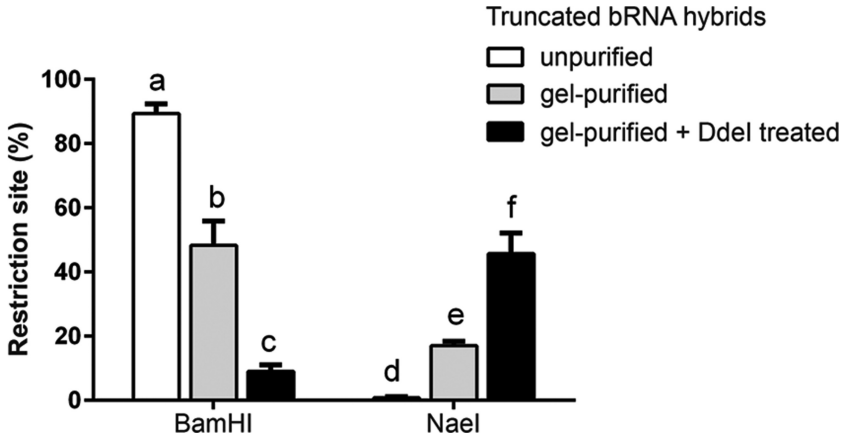

Figure 6. Frequency distribution of BamHI and NaeI restriction sites in RT-PCR amplicons. T-RFLP patterns of BamHI and NaeI digested RT-PCR amplicons obtained from unpurified, gel-purified or gel-purified and DdeI treated truncated bRNA hybrids shown in Supplementary Figure S2 were evaluated quantitatively. Error bars above columns indicate standard deviations from three independent experiments. Columns headed by different letters are significantly different from each other at $P < 0.05$ according to analysis of variance (ANOVA) followed by Holm-Sidak's multiple comparison test.

cDNA from either linear or bRNA, a DdeI restriction site is generated within the 5′-segment of the DNA duplex (Figure 4B, Supplementary Figure S3). DdeI restriction digestion of full-length hybrids achieves two goals: (i) DdeI removes the forward primer binding site of full-length cDNAs and makes them inaccessible to amplification by PCR, and (ii) DdeI truncates the 5′-segment of templates. In consequence, in the next round of primer extension from the 2′-arm full-length cDNAs cannot be generated from these RNA templates (Figure 4B). In truncated bRNA hybrids, cDNA terminates at the 2′,5′-branch and because of this, the 5′-segments of these bRNAs remain single-stranded and escape restriction by DdeI. Branched RNAs of truncated bRNA hybrids can then be used for another round of primer extension from the 2′-arm to generate full-length cDNA through the 2′,5-branch (Figure 4B). After DdeI restriction digestion, heat-denaturation, and primer extension, we amplified full-length cDNA and digested the RT-PCR products with BamHI and NaeI to monitor the efficiency of our restriction digestion method.

Evaluation of the T-RFLP patterns of BamHI and NaeI digested RT-PCR products (Supplementary Figure S2) revealed that depletion of full-length hybrids and recycling of truncated bRNA hybrids for further rounds of primer extension drastically reduced contamination of bRNA by linear RNA (Figure 6). As expected, without purification of the bRNA construct, the vast majority of RT-PCR products were obtained from linear and not from bRNA templates since 89 ± 3% (mean ± standard deviation) of the amplicons were restricted by BamHI compared to only 0.8 ± 0.3% by NaeI. After application of the two purification methods, only 9 ± 2% of the amplicons were restricted by BamHI, while the percentage of amplicons restricted by NaeI increased to 46 ± 6% (Figure 6). This indicates that after purification, bRNA was predominantly the template for RT-PCR, and therefore RT-PCR amplicons from gel-purified and DdeI treated bRNA were used for subsequent sequence analysis.

Analysis of mutations in full-length cDNAs traversing the branchpoint from the 2′- and 3′-arms of the branched RNA

To explore whether M-MLV RT (H−) misreads differently the branch guanosine in our bRNA construct during DNA synthesis from the 2′- versus the 3′-arm, we performed sequence analysis of full-length RT-PCR products. Full-length cDNAs generated by priming on the 2′-arm of gel-purified and DdeI treated bRNA (2′-arm-specific cDNA) and on the 3′-arm of bRNA (3′-arm-specific cDNA) were amplified, cloned, and sequenced. The sequencing reads surrounding the branch guanosine obtained from 2′- and 3′-arm-specific cDNAs are shown in Supplementary Tables S3 and S4, respectively. The most common mutations around the branchpoint found in arm-specific cDNAs are presented in Table 2. Sequencing of these cDNAs revealed that the arms induce characteristic mutation profiles of bRNA (Table 2).

DNA synthesis by the wild-type M-MLV RT terminates to a higher extent at the 2′-5′ linkage in an RNA template than DNA synthesis by the RNase H-deficient M-MLV RT

Lorsch *et al*. (60) showed by primer extension analysis that AMV RT (H+) and Superscript M-MLV RT (H−) can read through a 2′-5′ linkage in a linear RNA template. However, AMV RT (H+) terminated DNA synthesis at the 2′,5′-phosphodiester bond to a higher extent than Superscript RT (H−). Lorsch *et al*. (60) suggested that this result is due to the fact that the two RTs from different viruses have different fidelities and specificities. An alternative explanation for this observation would be that termination of DNA synthesis at the 2′-5′ linkage is related to RT's RNase H activity. To test this hypothesis, we performed primer extension analysis on precursor 1 containing the 2′-5′ linkage at nt position 37 (Table 1) using M-MLV RT with (H+) or without (H−) RNase H activity. The 5′-cy5-labeled reverse primer 2–5 was used to prime DNA synthesis, facilitating detection and quantification of truncated cDNA until the 2′-5′ linkage and full-length cDNA through the 2′-5′ linkage by scanning the gel. Quantitative evaluation of cy5-labeled cD-

Table 2. Frequencies of transversion and deletion mutations in arm-specific, full-length cDNAs after cloning.

	Mutation frequency[a]	
	Mean ± SD (%)	
Mutation type	2′,5′-branch	3′,5′-branch
Transversion G→C[b]	67.2 ± 3.7 (A)[c]	19.4 ± 10.1 (B)
Deletion[d]	10.4 ± 2.3 (B)	73.2 ± 6.4 (A)

[a]The mutation frequency was determined by sequence analysis of cloned RT-PCR products (32–46 clones each) obtained from three independent experiments listed in Supplementary Tables S3 and S4. SD, standard deviation.
[b]Only unambiguous transversions at the branchpoint position in the DNA sense-strand were included.
[c]Values followed by different letters in parentheses are significantly different from each other at $P < 0.0001$ according to ANOVA followed by Holm-Sidak's multiple comparison test. Values follwed by identical letters are not significantly different from each other at $P > 0.05$.
[d]All sequences with one or more deletions around the branchpoint (sense-strand) were included independent of a transversion at the branchpoint position.

NAs showed that DNA synthesis by M-MLV RT (H+) terminated thrice as strong at the 2′,5′ phosphodiester bond as DNA synthesis by M-MLV RT (H−) (Figure 7). Consistent with this finding, the yield of full-length cDNA generated by the wild-type RT was reduced to 40% in comparison to the RNase H-deficient RT. When DNA synthesis was primed on the control template with only 3′-5′ linkages, polymerization was not affected and both RTs generated nearly equal amounts of full-length cDNAs (Figure 7). These data indicate (i) that RT's RNase H activity promotes termination of DNA synthesis at the 2′-5′ linkage and (ii) that the absence of RNase H activity enables RT to read efficiently through 2′-5′ linkages.

M-MLV RT RNase H generates more cleavage products when RT pauses at a branchpoint than at a 2′-5′ linkage

Previous studies have shown that structure-induced pausing of RT during DNA polymerization leads to pause-related cleavages in the RNA template (18–21). We wanted to investigate whether pausing of RT at a weakly structured pause site (2′-5′ linkage) induces a different RNA cleavage pattern than pausing of RT at a highly structured pause site (branchpoint). To assess the relationship of the structural property of a pause site and RNA degradation by RT's RNase H, it is important to use the same template sequence because RNase H cleaves RNA in a sequence-specific manner (61,62). We, therefore, carried out primer extension and RNA degradation analysis on precursor 1 and our bRNA construct using the wild-type M-MLV RT. Both templates had the same sequence (5′-segment and 2′-arm), but the former contained the 2′-5′ linked ribo-G-nucleoside while the latter had the branch guanosine at nt position 37 (Figures 1 and 2A). DNA synthesis was primed from the 2′-arm of both templates and M-MLV RT (H+) stalled at both pause sites, as indicated by the accumulation of cDNA at the 2′-5′ linkage in precursor 1 (Figure 7A, lane 3) and at the 2′,5′-branch in bRNA (Supplementary Figure S4). To analyze RNA hydrolysis by RNase H without pausing of RT, we used the same template sequence containing exclusively 3′-5′ linkages (control template). M-MLV RT RNase H generated two cleavage products of ∼58 and ∼42 nt in length from the control template and precursor 1 (Figure 8A, lanes 4 and 5). These products were probably generated by internal cleavages. In the internal cleavage mode, RT RNase H behaves as a typical endonuclease and cleaves RNA along the length of an RNA/DNA hybrid (5,10). When the 3′-5′ linked ribo-G-nucleoside at position 37 was replaced by a 2′-5′ linked ribo-G-nucleoside, a pause-related cleavage product of ∼45 nt in length was generated by RT RNase H (Figure 8A, lane 5). The pause-related cleavage pattern induced by the 2′-5′ linkage in precursor 1 resembled the cleavage pattern caused by the base of a stable hairpin in an RNA template (20,21). Similar to the RNA hairpin base (20,21), the pausing M-MLV RT probably dissociated from the DNA 3′ terminus at the 2′-5′ linkage, slid forward, and RNase H carried out a pause-related cleavage 8 nt away from the 2′-5′ linkage. To analyse RNA cleavage products from the control template and precursor 1, we hybridized them with probe 1 complementary to the 5′ DNA sequence of both templates. Hybridization analysis revealed that RNase H cleavages occurred at template positions 42, 45 and 58 (Figure 8B, lanes 4 and 5), as already indicated by the lengths of the cleavage products.

When the 2′-5′ linked ribo-G-nucleoside was substituted by a branch guanosine, M-MLV RT RNase H generated at least six pause-related cleavage products (Figure 8A, lane 6). Only a small amount of bRNA (upper band) was not cleaved at all (Figure 8A, lane 6). To relate these cleavage products to cleavages in the RNA template, we performed hybridization analysis using probe 2 complementary to the 2′-arm. This analysis showed that most of the cleavage products did not hybridize with this probe (Figure 8B, lane 6), suggesting that the RNA template was shortened by RNase H and could no longer serve as a target for probe 2. An alternative possibility would be that cleavages in the 5′-segment or in the non-template arm (3′-arm) also contribute to this cleavage pattern.

Only the template arm is cleaved when DNA synthesis is primed from the 2′-arm of the branched RNA

To examine whether M-MLV RT RNase H cleaved, in addition to the template, the 5′-segment or the non-template arm (3′-arm) of our bRNA during DNA synthesis from the 2′-arm, we performed primer extension analyses on the generated cleavage products using RNase H-deficient RT. For this purpose, we chose two primers, reverse primer O and reverse primer 3–5. To detect cleavages in the 5′-segment, we used primer O binding to the overlapping branch site of bRNA. When M-MLV RT (H−) extends this primer to the end of the 5′-segment, the enzyme generates cDNA of 63

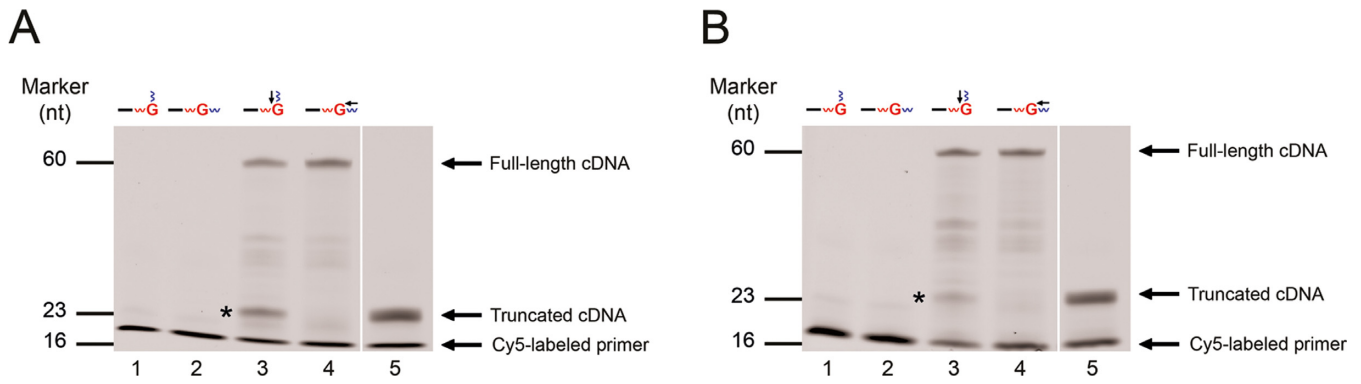

Figure 7. Effect of the 2′-5′ linked ribo-G-nucleoside in precursor 1 on DNA synthesis by M-MLV RT (H+) and (H−). (**A**) DNA synthesis catalyzed by the wild-type RT. Primer extension reactions were primed with the 5′-cy5-labeled reverse primer 2_5. Lanes 1 and 2: negative controls for primer extension reaction on precursor 1 and control template with only 3′-5′ linkages, respectively. Lanes 3 and 4: primer extension reactions on precursor 1 and control template, respectively. Lane 5: primer extension reaction on oligonucleotide representing the 2′-arm using M-MLV RT (H−). Colors as in Figure 1 and arrows in pictograms as in Figure 3B. Samples were separated on a 15% denaturing polyacrylamide gel. Full-length cDNA and truncated cDNA until the 2′-5′ linkage, as well as the cy5-labeled primer are noted. (**B**) DNA synthesis catalyzed by the RNase H-deficient RT. Legend as in panel (A). The cDNA that stops at the 2′,5′-linkage is labeled with an asterisk.

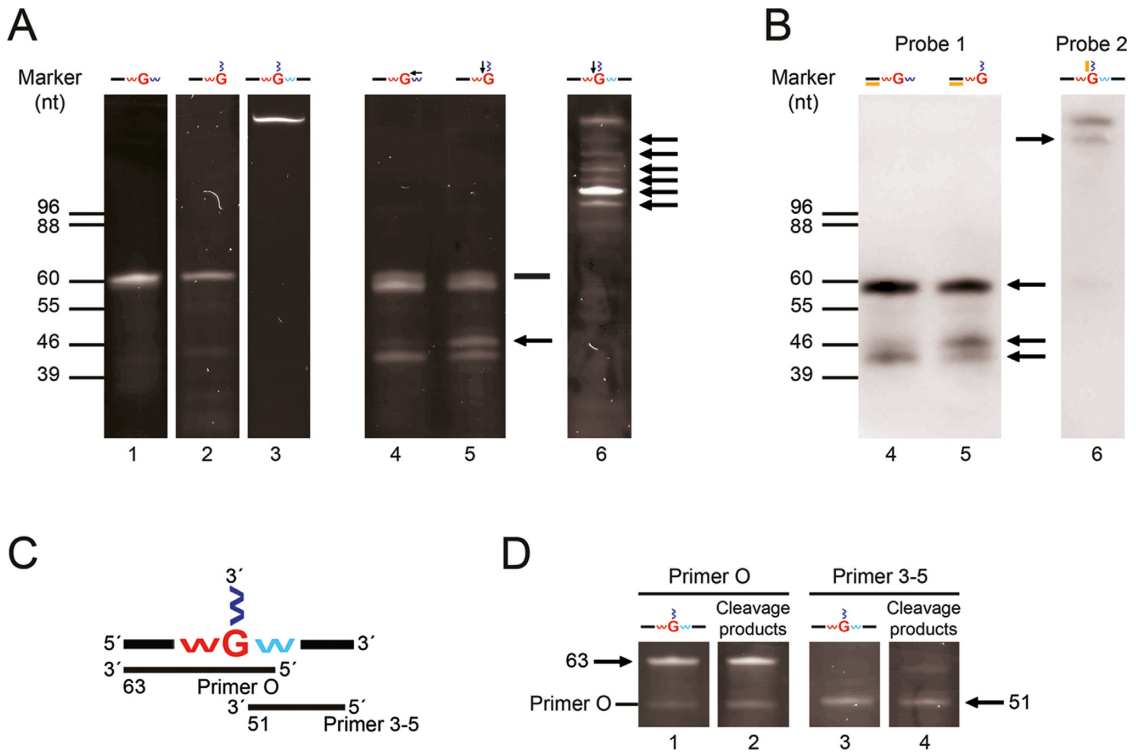

Figure 8. Characterizing RNA cleavage products generated by M-MLV RT RNase H during DNA synthesis from the 2′-arm. (**A**) Typical RNA cleavage patterns of oligonucleotides containing either a 3′-5′ linkage (control template), a 2′-5′ linkage (precursor 1), or a branched nucleotide (bRNA) at nt position 37. Lanes 1–3: negative controls for RNase H cleavage of the control template, precursor 1, and bRNA, respectively. Lanes 4–6: RNase H cleavages of the control template, precursor 1 and bRNA, respectively, during DNA synthesis from the 2′-arm. Arrows in pictograms as in Figure 3B. Samples were electrophoresed on a 15% denaturing polyacrylamide gel. Pause-related cleavage products are labeled with an arrow, and full-length cDNA from linear templates is indicated by a black bar (**B**) Analysis of RNA cleavage products by probe hybridization. RNA cleavage patterns (lanes 4–6) shown in panel (A), were blotted and hybridized with either probe 1 (lanes 4 and 5) or probe 2 (lane 6) as indicated. Probe specificities and probe-target regions as in Figure 2B. Hybridizing cleavage products are labeled with an arrow. (**C**) Scheme showing the primer extension method. Branched RNA and cDNAs produced with the reverse primers O and 3–5 are presented. Numbers below lines refer to the lengths in nucleotides of the generated cDNAs. (**D**) Primer extension analysis of bRNA and RNA cleavage products using M-MLV RT (H−). Lanes 1 and 2: primer extension reactions with primer O on bRNA and RNA cleavage products, respectively. The primer band of primer O is noted. Lanes 3 and 4: primer extension reactions with primer 3–5 on bRNA and RNA cleavage products, respectively. Samples were analyzed on a 15% denaturing polyacrylamide gel. Complementary DNAs generated from bRNA and RNA cleavage products are labeled with arrows, where the numbers indicate their lengths in nucleotides. Colors as in Figure 1.

nt in length (Figure 8C). To detect cleavages in the 3′-arm, we used primer 3–5 binding to the 3′-arm of bRNA. When the RNase H-deficient RT extends this primer until the 3′,5-branch, the enzyme generates cDNA of 51 nt in length (Figure 8C). We used equal amounts of cleavage products and bRNA as control for the primer extension reactions. Primer extension analyses revealed that M-MLV RT (H−) generated equal amounts of the two cDNAs from the cleavage products and the control (Figure 8D), indicating that the cleavage products contained the branchpoint, the complete 5′-segment and 3′-arm. This result shows that all cleavages occur in the template and that the non-template strand is not cleaved by M-MLV RT RNase H when DNA synthesis is primed from the 2′-arm of our bRNA.

Both arms of the branched RNA are cleaved when DNA synthesis is primed from the 3′-arm

We next investigated the RNA cleavage pattern induced by the branchpoint during DNA synthesis from the 3′-arm of our bRNA construct. We have shown in Figure 3B, lane 4 that the wild-type M-MLV RT paused at the branchpoint during DNA synthesis, as indicated by the accumulation of cDNA at the 3′,5′-branch. M-MLV RT RNase H generated several predominant and minor pause-related cleavage products (Figure 9A). To analyze the RNA cleavage products, we performed hybridization analyses using site-specific probes (Figure 2B). To detect cleavages in the RNA template (3′-arm), we hybridized the cleavage products with probe 3 (3′-arm) and probe O (overlapping branch site). To detect cleavages in the 5′-segment and non-template arm (2′-arm) of our bRNA, we hybridized the cleavage products with probe 1 and probe 2, respectively. Hybridization analyses showed that the cleavage products labeled with I and II resulted from cleavages in the RNA template because probe 3 did not hybridize (Figure 9A). To our surprise, the predominant products migrating between 60 and 51 nt (cleavage products from cluster III) resulted from cleavages in the RNA template and in the non-template 2′-arm because neither probe 2 nor probe 3 showed hybridization signals (Figure 9A). Additionally, the cleavage products of cluster III, except the top one, did not hybridize with probe O (Figure 9A), indicating that the RNA template was further shortened by RT RNase H and could not longer serve as a target for this probe. In our hybridization analysis, all cleavage products hybridized with probe 1 (Figure 9A), showing that no cleavages occurred in the 5′-segment of our bRNA construct. As already illustrated in Figure 3B, the predominant band with 51 nt in length (product labeled with IV) could be identified as truncated cDNA until the 3′,5′-branch.

To exclude the possibility that the observed cleavage of the 2′-arm was caused by accidental RNA/DNA hybid formation, we performed primer extension analyses on a 3′-fluorescein-labeled precursor 1 using the wild-type M-MLV RT. The fluorescein-labeled precursor 1 (fl-precursor) allowed us to monitor cleavages in its 2′-arm by scanning the gel. We first tested whether mispriming of reverse primer 3–5 to the 2′-arm led to cleavage of the non-template strand. For this purpose, we carried out a primer extension reaction with primer 3–5 and the fl-precursor. Primer extension analysis revealed that the 2′-arm of the fl-precursor was not cleaved by RT RNase H (Figure 9B, lane 2), indicating that cleavage of the non-template arm of bRNA was not caused by mispriming of primer 3–5. We next tested whether mispriming of the generated cDNA from the 3′-arm or our bRNA construct to the 2′-arm led to degradation of the non-template strand. For this purpose, we performed a primer extension reaction with primer 3–5, the fl-precursor and bRNA. Primer extension analysis showed that the 2′-arm of the fl-precursor was not cleaved by RT RNase H as well (Figure 9B, lane 4). This result indicated that cleavage of the non-template arm of bRNA was not caused by mispriming of either cDNA or bRNA. We suggest, therefore, that the template and the non-template strands are cleaved by M-MLV RT RNase H when DNA synthesis is primed from the 3′-arm of our bRNA.

DISCUSSION

In this study, we succeeded in the construction of long-chained branched RNA molecules. Our bRNA construct allowed us to investigate for the first time the arm-directed processing of the branched nucleotide within a single bRNA molecule. We found that misprocessing of the branch guanosine by M-MLV RT (H−) proceeds arm-specifically. When RNase H-deficient RT encounters the branchpoint from the 2′-arm, the enzyme incorporates predominantly dGMP opposite to the branch guanosine. The 2′-arm-induced mutation leads to a single mismatch error at the branchpoint position in the DNA sense-strand which is commonly used to pinpoint branched nucleotides in intron lariats by RT-PCR (26–28) and more recently by high-throughput methods (59,63). To detect branchpoints in intron lariats, RT reactions are widely carried out with RNase H-deficient RTs. Inactivation of RNase H lowers the mismatch error rate of RT's polymerase at other template positions (64), allowing an unbiased detection of branchpoints. We found that when M-MLV RT (H−) encounters the branchpoint from the 3′-arm, the enzyme misincorporates dGMP and skips predominantly one to several template nucleotides located upstream of the branch guanosine. The 3′-arm-induced mutation leads to a deletion mutation in the DNA sense-strand. This mutational profile was also reported by Tuschl *et al*. (29). The reason for this arm-specific processing of bRNA may be related to the structural difference at the branchpoint (here designated the branched nucleotide, and the 2′- and 3′-nucleotides). Conformational analysis showed that the branched nucleoside base stacks with the 2′-linked nucleobase, while the 3′-linked base is unstacked and located in a coplanar manner with respect to the branched nucleoside base (65,66). This base unstacking in the template strand may explain why deletion mutations are drastically more abundant in full-length cDNAs when synthesized from the 3′- than from the 2′-arm. The reason for this is that strand slippage can occur in response to an unstacked template base, which results in deletion mutations during DNA synthesis (67).

We also found that, in contrast to the RNase H-deficient M-MLV RT, wild-type M-MLV RT cannot read through the branchpoint of our bRNA construct and the ability of wild-type RT to read efficiently through the 2′-5′ linkage of the linear precursor 1 is decreased. We showed that both

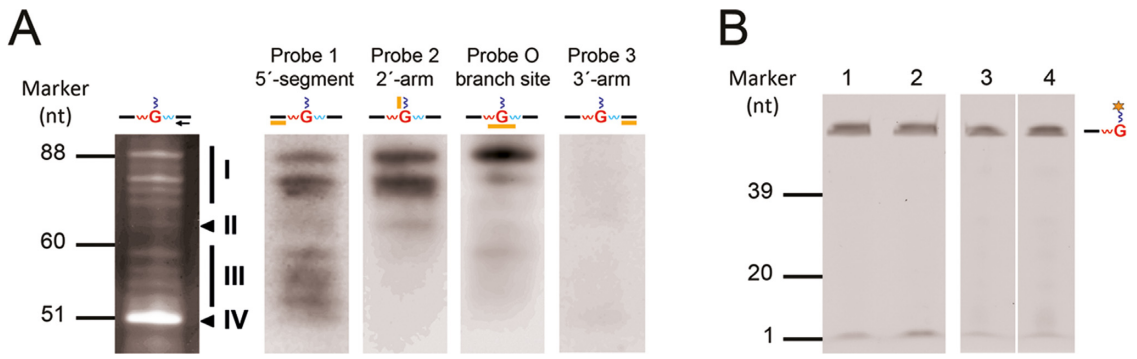

Figure 9. Characterizing RNA cleavage products generated by M-MLV RT RNase H during DNA synthesis from the 3′-arm. (**A**) Hybridization analyses of cleavage products. Typical RNA cleavage pattern of the bRNA construct separated on a 15% denaturing polyacrylamide gel is shown on the left (same gel as in Figure 3B, lane 4). RNA cleavage patterns were blotted and corresponding blots (right) were hybridized with specific probes (probes 1–3, probe O) as indicated. Probe specificities and probe-target regions as in Figure 2B. Products were grouped into four clusters (I–IV) according to their hybridization with probes as indicated on the right of the gel. Cleavage products of cluster I hybridized with probes 1, 2 and O, cleavage product of cluster II with probes 1 and 2, cleavage products of cluster III with probe 1, and product of cluster IV was identified as truncated cDNA until the 3′,5′-branch (see Figure 3B, lane 4). (**B**) Primer extension analyses on the 3′-fluorescein-labeled precursor 1 using M-MLV RT (H+). Lanes 1 and 3: negative controls for primer extension reaction, respectively. Lane 2: primer extension reaction using reverse primer 3–5. Lane 4: primer extension reaction using reverse primer 3–5 and bRNA. The 3′-fluorescein-labeled precursor 1 is schematically presented on the right, where the star indicates the position of the fluorophore. Samples were separated on a 15% denaturing polyacrylamide gel. The bands migrating at the bottom of the gel represented the loading dye. Colors as in Figure 1.

the branchpoint and the 2′,5′-phosphodiester bond represent pause sites for either RT. Contrary to the RNase H-deficient RT, wild-type RT can generate pause-related cleavages in the RNA template. These cleavages cause the template to dissociate from the DNA primer 3′ terminus and DNA synthesis by wild-type RT is terminated (21). On the other hand, the RNA template stays intact and the stalled DNA synthesis is resumed by RNase H-deficient RT with time to yield full-length cDNA. The observation that some wild-type RTs can read through branchpoints (29,46) may be related to different stalling times at branchpoints and to different RNase H activities of these enzymes.

Our bRNA oligonucleotide provided for the first time insights into the branchpoint-mediated RNA hydrolysis by RT RNase H. Using the same template sequence, we found that M-MLV RT RNase H cleaves the RNA template more frequently when pausing of the enzyme is caused by the branch guanosine than by the 2′-5′ linked ribo-G-nucleoside. Furthermore, we observed that RNase H also cleaved the RNA template several times when DNA synthesis was started from the 3′-arm of our bRNA, which differs in sequence from the 2′-arm. Based on these observations, we concluded that multiple pause-related cleavages in the template are probably always induced when branchpoints stall RT-catalyzed DNA synthesis (44,46–50). Because pause-related cleavage caused by a stable hairpin is dependent on forward sliding of non-polymerizing RT on the RNA/DNA hybrid (20,21), it is conceivable that forward sliding is involved in pause-related cleavages caused by a branchpoint as well. To induce multiple cleavages in the RNA template, M-MLV RT (H+), presumably slides in a stop-and-go manner on the hybrid of the 2′- and 3′-arm while passing the branch guanosine in our bRNA. Whenever the RT pauses, RNase H generates a pause-related cleavage in the RNA/DNA hybrid.

DNA synthesis of bRNA can be initiated from two different templates (arms). Here, we investigated whether the non-template arm of our bRNA construct is cleaved by M-MLV RT (H+) when DNA polymerization takes place on the other arm. We found that RT RNase H cleaves the non-template arm of the bRNA when DNA synthesis occurs on the 3′-arm but not on the 2′-arm. Cleavage of the non-template 2′-arm is an unexpected observation as RT's RNase H usually cleaves the template strand in the known pause-related cleavage mode. However, the observed non-template cleavage can be explained by a reorientation of M-MLV RT during pausing at the branchpoint and repositioning of the RNase H domain to the single-stranded 2′-arm. In particular, RT can bind to ds nucleic acids in a polymerization-competent and in a flipped, polymerization-incompetent orientation (9). In the flipped RT binding mode, RT is flipped ∼180 degrees so that the polymerase active site is located over the template strand, separated from the primer 3′ terminus, and the RNase H active site is located over the primer strand (non-template strand) (9). RT can flip between the two binding modes by a hopping mechanism (68). In this mechanism, RT dissociates from one bound state into a pseudo-bound state (RT and nucleic acid remain sufficiently close), RT rotates ∼180 degrees, and associates to the other bound state (68). Flipping of RT can be spontaneous (9,69) or introduced by a pause site (11,12,70). Liu *et al.* (70) observed that a stem-loop structure within the RNA template reorients the replicating RT into the flipped binding orientation during initiation of minus-strand DNA synthesis. They speculated that RT's reorientation prevents further DNA polymerization (70). Apparently, during initiation of plus-strand DNA synthesis from the polypurine tract (PPT), a pause site on the DNA template also triggers the reversal in binding orientations of the polymerizing enzyme (11,12). This reorientation allows RT RNase H to cleave the non-template RNA (11,12). Because branched nucleotides represent pause sites for RTs (Figure 3B, Supplementary Figure S4) (44,46–50), we suggest that the branched nucleotide in our bRNA imposes a reorientation of RT as well.

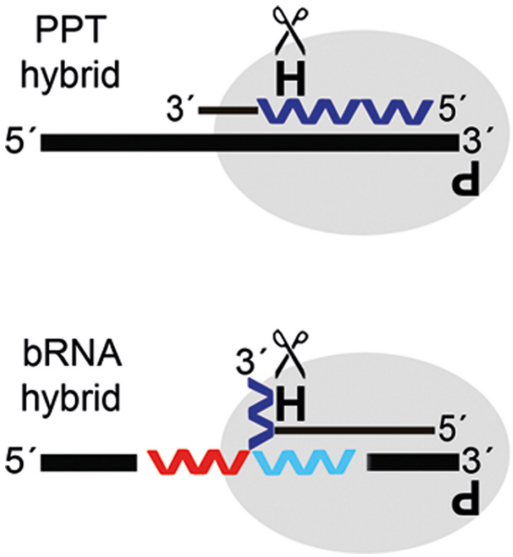

Figure 10. Flipped bound state of RT on polypurine tract (PPT) and bRNA hybrids. The relative positions of the RNase H and DNA polymerase catalytic sites of RT bound to the PPT and bRNA hybrid are presented. H with the scissor and the rotated P on the enzyme represent the active RNase H and inactive DNA polymerase catalytic site, respectively. The non-template strand is coloured in dark blue, otherwise colors as in Figure 1.

Although RT's RNase H activity is specific to RNA/DNA hybrids, in some cases, RNase H is capable of cleaving single-stranded RNA adjacent to an RNA/DNA duplex. Cleavage of ssRNA by human immunodeficiency virus 1 (HIV-1) and M-MLV RT RNases H (71), and *Escherichia coli* RNase HI (72) can occur when a DNA oligonucleotide is annealed to a longer RNA template. *Escherichia coli* RNase HI may interact with the RNA/DNA hybrid and position the active site onto the 3′-protruding ssRNA for cleavage close to the hybrid (72). We assume that M-MLV RT (H+) uses the same mechanism to cleave the ss 2′-arm. Based on these considerations we propose the following scenario at the branchpoint: DNA synthesis is primed from the 3′-arm of our bRNA construct and the replicating M-MLV RT pauses at the branchpoint. Pausing of RT occurs initially in the polymerization-competent orientation. The branch guanosine triggers the flip in binding orientations. RT rotates ∼180° and rebinds the RNA/DNA hybrid of the 3′-arm in the flipped state. In the bound state of RT, the RNase H active site is positioned in vicinity to the ss 2′-arm for cleavage. Cleavage of the ss non-template 2′-arm in our bRNA resembles cleavage of the non-template RNA in the ds polypurine tract of retroviruses and related retrotransposons. The alignment of either non-template strand to the RNase H active site is the same as when RT binds to the PPT and bRNA hybrid in the flipped state (Figure 10). When DNA synthesis was primed from the 2′-arm of our bRNA, we found that M-MLV RT RNase H does not cleave the non-template strand despite the anticipated flipped bound state of RT during pausing at the branchpoint. The reason for this may be that the RNase H active site cannot access the non-template 3′-arm for cleavage.

Long terminal repeats (LTR) retrotransposons and retroviruses have similar genetic structures and replication mechanisms (73). Here, we describe a possible flipping mechanism of the retroviral M-MLV RT at a branched nucleotide. We propose that protein flipping at branchpoints is common among RTs of LTR retrotransposons and retroviruses although the RTs differ in their structure. The dynamic properties of the heterodimeric retroviral HIV-1 RT, i.e. sliding and flipping on nucleic acids (9,68–70), regulate various phases of reverse transcription of HIV-1 genomic RNA, including plus-strand DNA synthesis from PPT, RNA degradation, and strand-displacement synthesis (69). The monomeric retroviral M-MLV RT (74,75) and homodimeric LTR retrotransposon Ty3 RT (76) have to be dynamic as well to catalyze these processes. It is likely, therefore, that in spite of differences in sequence and structure (77), retroviral and LTR retrotransposon RTs, such as M-MLV, HIV-1, and Ty3 RTs, share the same dynamic flexibility at branchpoints and position their RNase H active site close to the non-template ssRNA arm. However, it is also possible that the RTs of retroviruses and related retrotransposons may differ with respect to cleavage of non-template strands at branchpoints, and that branchpoints may trigger the reversal in binding orientation of these RTs for a different period of time.

Retroviruses and LTR retrotransposons are thought to have evolved from group II self-splicing introns (78,79). Based on the work of Cheng and Menees (40) it has been speculated that 2′,5′-branched genomic RNA is common among them (41,42). However, although the 2′,5′-branched form of genomic RNA is controversial (43,45), it is well established that the host-encoded RNA lariat debranching enzyme, Dbr1 plays a critical role in the life cycle of LTR retrotransposons and retroviruses (42,45,80–84). Dbr1 is a 2′,5′-phosphodiesterase that cleaves the 2′,5′ linkage at the branchpoint of branched nucleic acids (25,85). Cheng and Menees (40) suggested that yeast Dbr1 cleaves a 2′,5′-branched lariat structure present in Ty1 genomic RNA. The 2′,5′ linkage joins the 5′ repeat (R) region with the 3′ nt of the 3′ unique sequence (U3) upstream of the 3′ R region in the same RNA molecule (40). They proposed that synthesis of the minus-strand strong-stop DNA [(−)ssDNA] corresponding to the 5′ R-U5 region continues until the branchpoint (40). Complementarity between the R regions facilitates (−)ssDNA to hybridize to the 3′-end of genomic RNA (minus-strand transfer) (1,2). Subsequent removal of the branched nucleotide by yeast Dbr1 allows RT to complete minus-strand DNA synthesis after strand transfer (40). The observations that yeast *dbr1* mutant or deletion strains were defective in formation of full-length Ty1 and Ty3 minus-strand DNA (81,83,84) are compatible with debranching of genomic bRNAs by Dbr1 for the completion of minus-strand synthesis. Moreover, short hairpin RNA-mediated knockdown of human Dbr1 expression caused as well a reduction in the formation of full-length HIV-1 minus-strand DNA, but notably, inhibition of Dbr1 activity had no effect on the formation of (−)ssDNA (42,45). This result confirmed that debranching of genomic bRNA is required after

minus-strand transfer (40). Although the lariat form of genomic RNA remains controversial (43,45), sequence analysis of reverse transcribed Ty1 (40) and M-MLV (86) genomic RNAs revealed a single mismatch error at the position of the presumed branched nucleotide. This mutational profile (26–28,59,63) supported the hypothesis that the 3′ nt of the U3 region forms a 2′,5′-phosphodiester bond with the 5′-end of genomic RNA of Ty1 (40) and M-MLV (41).

Apart from Dbr1, our observed interplay between RT and the branched nucleotide may also play an important role in the life cycle of retroviruses and LTR retrotransposons. When tRNA-primed (-)ssDNA synthesis stalls at the proposed branched nucleotide at the 5′-end of genomic RNA (40) and RT pauses at the branchpoint, the branched nucleotide induces multiple pause-related cleavages in the RNA template (5′ R-U5 region) and would thereby promote the clearance of (-)ssDNA to allow the strand transfer. After strand transfer, the (-)ssDNA would hybridize in close vicinity to the proposed branched nucleotide located between the U3 and 3′ R region. Again, when RT pauses at the branchpoint, RT RNase H generates multiple pause-related cleavages in the RNA template (3′ R region) and this would cause the template to dissociate from the DNA 3′ terminus and DNA synthesis to terminate. When debranching by the nuclear Dbr1 does not occur in time, template degradation by RT RNase H would compromise full-length minus-strand DNA synthesis. This would explain why a too early start of reverse transcription is deleterious for retroviral replication (87–90).

After minus-strand transfer, the chance for completion of DNA synthesis downstream of the branchpoint would be dependent on two branchpoint-mediated processes: flipping of RT at the branchpoint (Figure 10) would delay template degradation, whereas debranching by Dbr1 would clear genomic RNA for DNA synthesis (40). Flipping of RT at the branchpoint would relocate the RNase H active site from the template (3′ R region) to the non-template strand (full-length or truncated 5′ R-U5 region). This would slow down template degradation during the transport of the reverse transcription complex from the cytoplasm to the nucleus of the cell (45). The non-template strand would be full-length when an intermolecular minus-strand transfer occurs, whereas it would be truncated when an intramolecular transfer takes place. It would be interesting to study whether the delay of template degradation is dependent on RT structure and RNase H activity in the flipped RT binding orientation and is influenced by the length of the non-template strand. This may provide clues to the development of new therapeutics, in case the retroviral genomic RNA is 2′,5′-branched.

In conclusion, we successfully constructed a long-chained bRNA oligonucleotide by splinted-ligation. We showed that RNase H-deficient M-MLV RT generates two cDNAs, the truncated and full-length cDNA, whereas the wild-type M-MLV RT generates only truncated cDNA from our bRNA. When we investigated the mutational and cleavage patterns produced by M-MLV RT, we found that bypass from the 2′-arm causes significantly more single mismatch errors at the branched nucleotide position than from the 3′-arm. Whereas bypass from the 3′-arm induces significantly more deletions downstream from the branchpoint than from the 2′-arm. Arm-specific mutation may be related to the structural difference at the branchpoint. Furthermore, we found that a template-associated branched nucleotide induces multiple pause-related cleavages in the RNA template. We showed that both arms of our bRNA construct are cleaved by M-MLV RT RNase H when DNA synthesis is primed from the 3′-arm but not when synthesis is primed from the 2′-arm. Arm-specific cleavage is probably defined by the accessibility of RNase H active site to the non-template strand while RT pauses in the flipped bound state at the branchpoint. Our observed interplay between RT and the branchpoint would presumably play a crucial role in a bRNA-mediated control of retrovirus and LTR retrotransposon replication.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Sina Duschek for expert technical assistance and preparation of the clone libraries. We thank Prof Reimer Stick for critical reading of the manuscript and for his helpful comments and suggestions, Dr Claudia Sofía Burbano, Dr Abhijit Sarkar and M.Sc. Wiebke Bünger for their careful reading of the manuscript. We are indebted to Prof Barbara Reinhold-Hurek for provision of laboratory facilities and support.

FUNDING

Funding for open access charge: Oxford University Press (full waiver). Funding for excess page charges: University of Bremen.
Conflict of interest statement. None declared.

REFERENCES

1. Herschhorn,A. and Hizi,A. (2010) Retroviral reverse transcriptases. *Cell. Mol. Life Sci.*, **67**, 2717–2747.
2. Le Grice,S.F.J. and Nowotny,M. (2014) *Nucleic Acid Polymerases*. Springer-Verlag, Berlin Heidelberg, NY, Vol. **30**, pp. 189–214.
3. Krug,M.S. and Berger,S.L. (1989) Ribonuclease H activities associated with viral reverse transcriptases are endonucleases. *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 3539–3543.
4. DeStefano,J.J., Buiser,R.G., Mallaber,L.M., Bambara,R.A. and Fay,P.J. (1991) Human immunodeficiency virus reverse transcriptase displays a partially processive 3′ to 5′ endonuclease activity. *J. Biol. Chem.*, **266**, 24295–24301.
5. Schultz,S.J. and Champoux,J.J. (2008) RNase H activity: structure, specificity, and function in reverse transcription. *Virus Res.*, **134**, 86–103.
6. Jacobo-Molina,A., Ding,J., Nanni,R.G., Clark,A.D. Jr, Lu,X., Tantillo,C., Williams,R.L., Kamer,G., Ferris,A.L., Clark,P. *et al.* (1993) Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 6320–6324.
7. Sarafianos,S.G., Das,K., Tantillo,C., Clark,A.D. Jr, Ding,J., Whitcomb,J.M., Boyer,P.L., Hughes,S.H. and Arnold,E. (2001) Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J.*, **15**, 1449–1461.
8. Arnold,E., Jacobo-Molina,A., Nanni,R.G., Williams,R.L., Lu,X., Ding,J., Clark,A.D Jr, Zhang,A., Ferris,A.L., Clark,P. *et al.* (1992) Structure of HIV-1 reverse transcriptase/DNA complex at 7 Å resolution showing active site locations. *Nature*, **357**, 85–89.

9. Abbondanzieri,E.A., Bokinsky,G., Rausch,J.W., Zhang,J.X., Le Grice,S.F. and Zhuang,X. (2008) Dynamic binding orientations direct activity of HIV reverse transcriptase. *Nature*, **453**, 184–189.
10. Champoux,J.J. and Schultz,S.J. (2009) Ribonuclease H: properties, substrate specificity and roles in retroviral reverse transcription. *FEBS J.*, **276**, 1506–1516.
11. Götte,M., Maier,G., Onori,A.M., Cellai,L., Wainberg,M.A. and Heumann,H. (1999) Temporal coordination between initiation of HIV (+)-strand DNA synthesis and primer removal. *J. Biol. Chem.*, **16**, 11159–11169.
12. Götte,M. (2011) Initiation of HIV reverse transcription: is enzyme flipping required? *Viruses*, **4**, 331–335.
13. Nowotny,M., Gaidamakov,S.A., Ghirlando,R., Cerritelli,S.M., Crouch,R.J. and Yang,W. (2007) Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription. *Mol. Cell*, **28**, 264–276.
14. Lapkouski,M., Tian,L., Miller,J.T., Le Grice,S.F. and Yang,W. (2013) Complexes of HIV-1 RT, NNRTI and RNA/DNA hybrid reveal a structure compatible with RNA degradation. *Nat. Struct. Mol. Biol.*, **20**, 230–236.
15. DeStefano,J.J., Buiser,R.G., Mallaber,L.M., Myers,T.W., Bambara,R.A. and Fay,P.J. (1991) Polymerization and RNase H activities of the reverse transcriptases from avian myeloblastosis, human immunodeficiency, and Moloney murine leukemia viruses are functionally uncoupled. *J. Biol. Chem.*, **266**, 7423–7431.
16. Beilhartz,G.L., Wendeler,M., Baichoo,N., Rausch,J., Le Grice,S. and Götte,M. (2009) HIV-1 reverse transcriptase can simultaneously engage its DNA/RNA substrate at both DNA polymerase and RNase H active sites: implications for RNase H inhibition. *J. Mol. Biol.*, **388**, 462–474.
17. Li,A., Li,J. and Johnson,K.A. (2016) HIV-1 reverse transcriptase polymerase and RNase H active sites work simultaneously and independently.*J. Biol. Chem.*,**291**, 26566–26585
18. Suo,Z. and Johnson,K.A. (1997a) Effect of RNA secondary structure on the kinetics of DNA synthesis catalyzed by HIV-1 reverse transcriptase. *Biochemistry*, **36**, 12459–12467.
19. Suo,Z. and Johnson,K.A. (1997b) Effect of RNA secondary structure on RNA cleavage catalyzed by HIV-1 reverse transcriptase. *Biochemistry*, **36**,12468–12476.
20. Purohit,V., Balakrishnan,M., Kim,B. and Bambara,R.A. (2005) Evidence that HIV-1 reverse transcriptase employs the DNA 3′ end-directed primary/secondary RNase H cleavage mechanism during synthesis and strand transfer. *J. Biol. Chem.*, **280**, 40534–40543.
21. Purohit,V., Roques,B.P., Kim,B. and Bambara,R.A. (2007) Mechanisms that prevent template inactivation by HIV-1 reverse transcriptase RNase H cleavages. *J. Biol. Chem.*, **282**, 12598–125609.
22. Lanciault,C. and Champoux,J.J. (2006) Pausing during reverse transcription increases the rate of retroviral recombination. *J. Virol.*, **80**, 2483–2494.
23. Basu,V.P., Song,M., Gao,L., Rigby,S.T., Hanson,M.N. and Bambara,R.A. (2008) Strand transfer events during HIV-1 reverse transcription. *Virus Res.*, **134**, 19–38.
24. Wallace,J.C. and Edmonds,M. (1983) Polyadenylylated nuclear RNA contains branches. *Proc. Natl. Acad. Sci. U.S.A.*, **80**, 950–954.
25. Nam,K., Hudson,R.H., Chapman,K.B., Ganeshan,K., Damha,M.J. and Boeke,J.D. (1994) Yeast lariat debranching enzyme. Substrate and sequence specificity. *J. Biol. Chem.*, **269**, 20613–20621.
26. Vogel,J., Hess,W.R. and Börner,T. (1997) Precise branch point mapping and quantification of splicing intermediates. *Nucleic Acids Res.*, **15**, 2030–2031.
27. Vogel,J. and Börner,T. (2002) Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. *EMBO J.*, **21**, 3794–3803.
28. Gao,K., Masuda,A., Matsuura,T. and Ohno,K. (2008) Human branch point consensus sequence is yUnAy. *Nucleic Acids Res.*, **36**, 2257–2267.
29. Tuschl,T., Sharp,P.A. and Bartel,D.P. (1998) Selection in vitro of novel ribozymes from a partially randomized U2 and U6 snRNA library. *EMBO J.*, **17**, 2637–2650.
30. Padgett,R.A., Konarska,M.M., Grabowski,P.J., Hardy,S.F. and Sharp,P.A. (1984) Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. *Science*, **225**, 898–903.
31. Konarska,M.M., Grabowski,P.J., Padgett,R.A. and Sharp,P.A. (1985) Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors. *Nature*, **313**, 552–557.
32. Kruger,K., Grabowski,P.J., Zaug,A.J., Sands,J., Gottschling,D.E. and Cech,T.R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*, **1**, 147–157.
33. Peebles,C.L., Perlman,P.S., Mecklenburg,K.L., Petrillo,M.L., Tabor,J.H., Jarrell,K.A. and Cheng,H.L. (1986) A self-splicing RNA excises an intron lariat. *Cell*, **44**, 213–223.
34. van der Veen,R, Arnberg,A.C., van der Horst,G., Bonen,L., Tabak,H.F. and Grivell,L.A. (1986) Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. *Cell*, **44**, 225–234.
35. Murphy,W.J., Watkins,K.P. and Agabian,N. (1986) Identification of a novel Y branch structure as an intermediate in trypanosome mRNA processing: evidence for trans splicing. *Cell*, **47**, 517–525.
36. Sutton,R.E. and Boothroyd,J.C. (1986) Evidence for trans splicing in trypanosomes. *Cell*, **47**, 527–535.
37. Krause,M. and Hirsh,D. (1987) A trans-spliced leader sequence on actin mRNA in C. elegans. *Cell*, **49**, 753–761.
38. Yee,T., Furuichi,T., Inouye,S. and Inouye,M. (1984) Multicopy single-stranded DNA isolated from a gram-negative bacterium, Myxococcus xanthus. *Cell*, **38**, 203–209.
39. Furuichi,T., Dhundale,A., Inouye,M. and Inouye,S. (1987) Branched RNA covalently linked to the 5′ end of a single-stranded DNA in Stigmatella aurantiaca: structure of msDNA. *Cell*, **48**, 47–53.
40. Cheng,Z. and Menees,T.M. (2004) RNA branching and debranching in the yeast retrovirus-like element Ty1. *Science*, **303**, 240–243.
41. Perlman,P.S. and Boeke,J.D. (2004) Ring around the retroelement. *Science*, **9**, 182–184.
42. Ye,Y., De Leon,J., Yokoyama,N., Naidu,Y. and Camerini,D. (2005) DBR1 siRNA inhibition of HIV-1 replication. *Retrovirology*, **2**, 63.
43. Coombes,C.E. and Boeke,J.D. (2005) An evaluation of detection methods for large lariat RNAs. *RNA*, **11**, 323–331.
44. Pratico,E.D. and Silverman,S.K. (2007) Ty1 reverse transcriptase does not read through the proposed 2′,5′-branched retrotransposition intermediate in vitro. *RNA*, **13**, 1528–1536.
45. Galvis,A.E., Fisher,H.E., Nitta,T., Fan,H. and Camerini,D. (2014) Impairment of HIV-1 cDNA synthesis by DBR1 knockdown. *J. Virol.*, **88**, 7054–7069.
46. Conklin,J.F., Goldman,A. and Lopez,A.J. (2005) Stabilization and analysis of intron lariats in vivo. *Methods*, **37**, 368–375.
47. Krainer,A.R., Maniatis,T., Ruskin,B. and Green,M.R. (1984) Normal and mutant human beta-globin pre-mRNAs are faithfully and efficiently spliced in vitro. *Cell*, **36**, 993–1005.
48. Rodriguez,J.R., Pikielny,C.W. and Rosbash,M. (1984) In vivo characterization of yeast mRNA processing intermediates. *Cell*, **39**, 603–610.
49. Domdey,H., Apostol,B., Lin,R.J., Newman,A., Brody,E. and Abelson,J. (1984) Lariat structures are in vivo intermediates in yeast Pre-mRNA splicing. *Cell*, **39**, 611–621.
50. Padgett,R.A., Konarska,M.M., Aebi,M., Hornig,H., Weissmann,C. and Sharp,P.A. (1985) Nonconsensus branch-site sequences in the in vitro splicing of transcripts of mutant rabbit beta-globin genes. *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 8349–8353.
51. Moore,M.J. and Query,C.C. (2000) Joining of RNAs by splinted ligation. *Methods Enzymol.*, **317**, 109–123.
52. Mendel-Hartvig,M., Kumar,A. and Landegren,U. (2004) Ligase-mediated construction of branched DNA strands: a novel DNA joining activity catalyzed by T4 DNA ligase. *Nucleic Acids Res.*, **32**, e2.
53. Katolik,A., Johnsson,R., Montemayor,E., Lackey,J.G., Hart,P.J. and Damha,M.J. (2014) Regiospecific solid-phase synthesis of branched oligoribonucleotides that mimic intronic lariat RNA intermediates. *J. Org. Chem.*, **79**, 963–975.
54. Montemayer,E.J., Katolik,A., Taylor,A.B., Scheurmann,J., Combs,D.J., Johnsson,R., Holloway,S.P., Stevens,S.W., Damha,M.J. and Hart,P.J. (2014) Structural basis of lariat RNA recognition by the intron debranching enzyme Dbr1. *Nucleic Acids Res.*, **42**, 10845–10855.
55. Schwer,B., Khalid,F. and Shuman,S. (2016) Mechanistic insights into the manganese-dependent phosphodiesterase activity of yeast Dbr1

with bis-p-nitrophenylphosphate and branched RNA substrates. *RNA*, **22**, 1819–1827.
56. Wang,Y. and Silverman,S.K. (2003) Deoxyribozymes that synthesize branched and lariat RNA. *J. Am. Chem. Soc.*, **125**, 6880–6881.
57. Wang,Y. and Silverman,S.K. (2005) Efficient one-step synthesis of biologically related lariat RNAs by a deoxyribozyme. *Angew. Chem., Int. Ed. Engl.*, **44**, 5863–5866.
58. Pratico,E.D., Wang,Y. and Silverman,S.K. (2005) A deoxyribozyme that synthesizes 2′,5′-branched RNA with any branch-site nucleotide. *Nucleic Acids Res.*, **33**, 3503–3512.
59. Mercer,T.R., Clark,M.B., Andersen,S.B., Brunck,M.E., Haerty,W., Crawford,J., Taft,R.J., Nielsen,L.K., Dinger,M.E. and Mattick,J.S. (2015) Genome-wide discovery of human splicing branchpoints. *Genome Res.*, **25**, 290–303.
60. Lorsch,J.R., Bartel,D.P. and Szostak,J.W. (1995) Reverse transcriptase reads through a 2′-5′ linkage and a 2′-thiophosphate in a template. *Nucleic Acids Res.*, **23**, 2811–2814.
61. Schultz,S.J., Zhang,M. and Champoux,J.J. (2010) Multiple nucleotide preferences determine cleavage-site recognition by the HIV-1 and M-MuLV RNases H. *J. Mol. Biol.*, **397**, 161–178.
62. Schultz,S.J., Zhang,M. and Champoux,J.J. (2009) Preferred sequences within a defined cleavage window specify DNA 3′ end-directed cleavages by retroviral RNases H. *J. Biol. Chem.*, **284**, 32225–32238.
63. Bitton,D.A., Rallis,C., Jeffares,D.C., Smith,G.C., Chen,Y.Y., Codlin,S., Marguerat,S. and Bähler,J. (2014) LaSSO, a strategy for genome-wide mapping of intronic lariats and branch points using RNA-seq. *Genome Res.*, **24**, 1169–1179.
64. Álvarez,M., Barrioluengo,V., Afonso-Lehmann,R.N. and Menéndez-Arias,L. (2013) Altered error specificity of RNase H-deficient HIV-1 reverse transcriptases during DNA-dependent DNA synthesis. *Nucleic Acids Res.*, **41**, 4601–4612.
65. Damha,M.J. and Ogilvie,K.K. (1988) Conformational properties of branched RNA fragments in aqueous solution. *Biochemistry*, **27**, 6403–6416.
66. Koole,L.H., Buck,H.M., Kuijpers,W.H.A., Balgobin,N., Nyilas,A., Remaud,G., Vial,J.-M. and Chattopadhyaya,J. (1988) Lariat formation in splicing of pre-messenger RNA. Conformation and base stacking at the lariat branch point studied using 500-MHz 1H NMR and CD spectroscopy. *Recl. Trav. Chim. Pays-Bas*, **107**, 663–667.
67. Manjari,S.R., Pata,J.D. and Banavali,N.K. (2014) Cytosine unstacking and strand slippage at an insertion-deletion mutation sequence in an overhang-containing DNA duplex. *Biochemistry*, **53**, 3807–3816.
68. Ganji,M., Docter,M., Le Grice,S.F. and Abbondanzieri,E.A. (2016) DNA binding proteins explore multiple local configurations during docking via rapid rebinding. *Nucleic Acids Res.*, **44**, 8376–8384.
69. Liu,S., Abbondanzieri,E.A., Rausch,J.W., Le Grice,S.F. and Zhuang,X. (2008) Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. *Science*, **322**, 1092–1097.
70. Liu,S., Harada,B.T., Miller,J.T., Le Grice,S.F. and Zhuang,X. (2010) Initiation complex dynamics direct the transitions between distinct phases of early HIV reverse transcription. *Nat. Struct. Mol. Biol.*, **17**, 1453–1460.
71. Gao,H.Q., Sarafianos,S.G., Arnold,E. and Hughes,S.H. (1999) Similarities and differences in the RNase H activities of human immunodeficiency virus type 1 reverse transcriptase and Moloney murine leukemia virus reverse transcriptase. *J. Mol. Biol.*, **17**, 1097–1113.
72. Lima,W.F. and Crooke,S.T. (1997) Cleavage of single strand RNA adjacent to RNA-DNA duplex regions by Escherichia coli RNase H1. *J. Biol. Chem.*, **272**, 27513–27516.
73. Deininger,P.L. and Batzer,M.A. (2002) Mammalian retroelements. *Genome Res.*, **12**, 1455–1465.
74. Verma,I.M. (1975) Studies on reverse transcriptase of RNA tumor viruses III. Properties of purified Moloney murine leukemia virus DNA polymerase and associated RNase H. *J. Virol.*, **15**, 843–854.
75. Gerard,G. F. and Grandgenett,D. P. (1975) Purification and characterization of the DNA polymerase and RNase H activities in Moloney murine sarcoma-leukemia virus. *J. Virol.*, **15**, 785–797.
76. Nowak,E., Miller,J.T., Bona,M.K., Studnicka,J., Szczepanowski,R.H., Jurkowski,J., Le Grice,S.F. and Nowotny,M. (2014) Ty3 reverse transcriptase complexed with an RNA-DNA hybrid shows structural and functional asymmetry. *Nat. Struct. Mol. Biol.*, **21**,389–396.
77. Skalka,A.M. (2014) Retroviral DNA transposition: themes and variations. *Microbiol. Spectrum*, **2**, MDNA3-0005-2014.
78. Malik,H.S. and Eickbush,T.H. (2001) Phylogenetic analysis of ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses. *Genome Res.*, **11**, 1187–1197.
79. Boeke,J.D. (2003) The unusual phylogenetic distribution of retrotransposons: a hypothesis. *Genome Res.*, **13**, 1975–1983.
80. Chapman,K.B. and Boeke,J.D. (1991) Isolation and characterization of the gene encoding yeast debranching enzyme. *Cell*, **65**, 483–492.
81. Karst,S.M., Rutz,M.L. and Menees,T.M. (2000) The yeast retrotransposons Ty1 and Ty3 require the RNA Lariat debranching enzyme, Dbr1p, for efficient accumulation of reverse transcripts. *Biochem. Biophys. Res. Commun.*, **268**, 112–117.
82. Salem,L.A., Boucher,C.L. and Menees,T.M. (2003) Relationship between RNA lariat debranching and Ty1 element retrotransposition. *J. Virol.*, **77**, 12795–12806.
83. Griffith,J.L., Coleman,L.E., Raymond,A.S., Goodson,S.G., Pittard,W.S., Tsui,C. and Devine,S.E. (2003) Functional genomics reveals relationships between the retrovirus-like Ty1 element and its host Saccharomyces cerevisiae. *Genetics*, **164**, 867–879.
84. Mou,Z., Kenny,A.E. and Curcio,M.J. (2006) Hos2 and Set3 promote integration of Ty1 retrotransposons at tRNA genes in Saccharomyces cerevisiae. *Genetics*, **172**, 2157–2167.
85. Ruskin,B. and Green,M.R. (1985) An RNA processing activity that debranches RNA lariats. *Science*, **229**, 135–140.
86. Kulpa,D., Topping,R. and Telesnitsky,A. (1997) Determination of the site of first strand transfer during Moloney murine leukemia virus reverse transcription and identification of strand transfer-associated reverse transcriptase errors. *EMBO J.*, **16**, 856–865.
87. Thomas,J.A. and Gorelick,R.J. (2008) Nucleocapsid protein function in early infection processes. *Virus Res.*, **134**, 39–63.
88. Lori,F., di Marzo Veronese,F., de Vico,A.L., Lusso,P., Reitz,M.S. Jr and Gallo,R.C. (1992) Viral DNA carried by human immunodeficiency virus type 1 virions. *J. Virol.*, **66**, 5067–5074.
89. Trono,D. (1992) Partial reverse transcripts in virions from human immunodeficiency and murine leukemia viruses. *J. Virol.*, **66**, 4893–4900.
90. Zhang,H., Dornadula,G., Orenstein,J. and Pomerantz,R.J. (2000) Morphologic changes in human immunodeficiency virus type 1 virions secondary tointravirion reverse transcription: evidence indicating that reverse transcription may not take place within the intact viral core. *J. Hum. Virol.*, **3**, 165–172.