JOURNAL OF VIROLOGY, Apr. 1992, p. 2464–2472
0022-538X/92/042464-09$02.00/0

Vol. 66, No. 4

Interaction between Retroviral U5 RNA and the TψC Loop of the tRNA^{Trp} Primer Is Required for Efficient Initiation of Reverse Transcription

ASHOK AIYAR,[1] DAVID COBRINIK,[1]† ZHENG GE,[1] H.-J. KUNG,[2] AND JONATHAN LEIS[1]*

Departments of Biochemistry[1] and Molecular Biology,[2] Case Western Reserve University School of Medicine, Cleveland, Ohio 44106

Received 4 February 1991/Accepted 14 January 1992

The 5′ end of avian sarcoma and leukosis virus RNA near the primer binding site forms two RNA secondary structures, U5-inverted repeat (U5-IR) and U5-leader stems, which are required for efficient initiation of reverse transcription. Lying between these two secondary structures is a 7-base sequence that can anneal to the TψC loop of the tRNA^{Trp} primer. Base substitutions in U5 RNA which disrupt this potential interaction result in a defect in the initiation of reverse transcription both in vivo and in vitro. The defect can be complemented in vitro by base substitutions in the primer. The U5 RNA-TψC interaction is also dependent upon the presence of both the U5-IR and the U5-leader structures. These RNA secondary structures and primer interactions are conserved in other type C and D retroviruses, suggesting that there is a common mechanism for the initiation of reverse transcription in all of these retroviruses.

Reverse transcription is a complex multistep process that appears to be regulated at the level of initiation (5, 7). In avian sarcoma and leukosis viruses (ASLV), DNA synthesis is primed from the 3′OH end of a tRNA^{Trp} annealed to an 18-base viral RNA sequence near the 5′ terminus (8, 12, 33), called the minus-strand primer binding site (PBS). We have previously shown that two regions of RNA secondary structure, the U5-inverted repeat (U5-IR) and the U5-leader stem, are required for efficient initiation of reverse transcription as monitored by viral growth and the synthesis of early-replication cDNA products in melittin-permeabilized virions (5, 7). It has been previously reported that 7 bases of the TψC loop of the tRNA^{Trp} primer could anneal to sequences in U5 (15, 16), and this interaction may be required in an early step of reverse transcription (10). In this report, we demonstrate that the interaction between the TψC arm of tRNA^{Trp} and U5 RNA sequences is necessary for efficient initiation of reverse transcription in vitro as well as in permeabilized virions. In addition, an in vitro reconstituted system has been used to demonstrate that stimulation of reverse transcription by the U5 RNA-TψC loop interactions depends upon both the U5-leader and the U5-IR stem structures. Comparisons of the sequences and structures of the 5′ regions of other type C and D retrovirus RNAs indicate that all of these interactions are conserved, suggesting that these viruses utilize a common regulatory mechanism for the initiation of reverse transcription.

MATERIALS AND METHODS

Reagents. *Escherichia coli* DNA polymerase I, Klenow fragment (5 U/μl), various restriction enzymes, and RNase A were purchased from Boehringer Mannheim Biochemicals. T4 polynucleotide kinase (10 U/μl), T4 DNA ligase (400 U/μl), and restriction enzymes were purchased from New England Biolabs. T7 RNA polymerase (20 U/μl), RNasin, ribonuclease inhibitor (40 U/μl), and RQ1 DNase (1 U/μl) were purchased from Promega. Avian myeloblastosis virus (AMV) reverse transcriptase (RT) (30 U/μl) was obtained from Molecular Genetics Resources. Nucleocapsid (NC) protein was purified as previously described (21). The sequencing kit, Sequenase v2.0, was obtained from United States Biochemicals. All enzymes were used according to the manufacturers' specifications. Deoxynucleoside triphosphates and ribonucleoside triphosphates were obtained from Pharmacia. [α-^{32}P]dCTP (3,000 Ci/mmol) and [γ-^{32}P]ATP (6,000 Ci/mmol) were obtained from New England Nuclear. Oligodeoxynucleotides were synthesized on an Applied Biosystems 380B Nucleic Acid Synthesizer and purified as previously described (6). pGEM3z DNA was purchased from Promega. Other chemicals were of the highest grade available and were purchased from Sigma Biochemicals or Fisher Chemicals.

Bacterial strains and plasmid purifications. All vectors were transformed into the *E. coli* strain XL1Blue (obtained from Stratagene). pDC202B and pDC102B were also propagated as previously described (5). Plasmids were isolated by the alkaline lysis procedure and purified by equilibrium cesium chloride density gradient centrifugation (29).

Tissue culture. Quail QT-6 cells were grown in Dulbecco's modified Eagle's medium (Whittaker Bioproducts) supplemented with 4% fetal bovine serum (Irvine Scientific) and 0.4 mM L-glutamine in a 4% CO_2 atmosphere at 37°C. Cells were infected with virus in the presence of polybrene. RT assays using exogenous template-primers (5, 7) were used to detect virus in supernatants and to monitor viral growth. Transfections were performed as previously described (5, 7).

Analysis of reverse transcription in melittin-permeabilized virions. Examination of strong stop DNA (cDNA$_{101}$) or dAMP-tRNA^{Trp} synthesis in melittin-activated virions was performed as described earlier (4, 5).

Production of synthetic templates. An *Eco*RI (cuts at nucleotide 9238 in ASLV genome)-to-*Bgl*II (cuts at nucleotide 1630 in ASLV genome) (30) fragment from the vectors pDC101 (wild type), pDC103 (Δ3), pATVD6 (Δ6), pDC202 (S6), pDC203 (S4), and pDC204 (S4C) was each cloned into

* Corresponding author.
† Present address: Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142.

the *Eco*RI and *Bam*HI sites in the polylinker of the plasmid pGEM3z. Recombinants were identified by restriction mapping and sequencing and were named pGEMAA-WT, pGEMAA-Δ3, pGEMAA-Δ6, pGEMAA-S6, pGEMAA-S4, and pGEMAA-S4C, respectively. The various pGEMAA-xx plasmids were digested with *Sca*I (cuts at nucleotide 1072 in ASLV genome), and in vitro runoff transcripts were prepared by using T7 RNA polymerase as described previously (23, 32). Reactions were stopped by digesting the DNA template with RQ1DNase (4 U) for 30 min at 37°C. The reaction mixtures were then extracted with phenol-chloroform, and the remaining RNA was precipitated with 2 volumes of ethanol. The RNA pellets were stored in 80% ethanol or were suspended in diethyl pyrocarbonate-treated water. Prior to their use in reverse transcription reactions, the synthetic template RNAs were heated to 80°C for 5 min and slowly cooled to room temperature to allow the formation of stable secondary structures.

Production of synthetic RNA primers. The following four oligodeoxynucleotides were synthesized to prepare RNA primer transcripts: oligo A, 5′TAATACGACTCACTATA3′; oligo B, 5′TGGTGACCCCGACGTGATTCTATAGTGAGT CGTATTA3′; oligo C, 5′TGGTGACCCCGACGTGATTCGA ACACGCTATAGTGAGTCGTATTA3′; oligo D, 5′TGGTGA CCCCGACGTGATTCCTTGTGGCTATAGTGAGTCGTAT TA3′. The underlined sequence represents the complement to the S6 U5 RNA substitution. Oligo A (100 pmol) was annealed to 100 pmol each of oligo B, C, or D. The overhang templates in each case were filled in by a DNA polymerase I (Klenow)-catalyzed repair synthesis. The resultant duplexes were used directly to prepare runoff T7 transcripts as described above. The RNA products for the A/B, A/C, and A/D templates were a 20-base sequence complementary to the minus-strand PBS, a 28-base sequence complementary to the minus-strand PBS and extended by 8 nucleotides present in the TψC loop (PBS-TψC), and a 28-base sequence similar to that formed with A/C except that the 8-base extension is complementary to the S6 substitution (PBS-S6-TψC), respectively.

Reconstitution of reverse transcription in vitro. Reverse transcription reactions were reconstituted in vitro by using a modification of the procedure described by Prats et al. (27) and Barat et al. (2). Various synthetic template RNAs (1 μg) were each mixed with 150 pmol of unlabeled RNA primer (as indicated in Fig. 4C) and 25 pmol of the ^{32}P-end-labeled long terminal repeat (LTR) DNA primer (2×10^5 to 4×10^5 cpm/pmol) in a 10-μl annealing reaction that contained 40 mM Tris · HCl (pH 8.0), 60 U of AMV RT, 50 ng of Rous sarcoma virus (RSV) NC protein, and 60 mM NaCl. The reaction mixture was incubated at 40°C for 40 to 60 min, at which time it was brought to a final volume of 20 μl containing 60 mM NaCl, 25 mM Tris · HCl (pH 8.0), 50 μmol each of ATP, dATP, dGTP, and TTP, and 12.5 μmol of [α-^{32}P]dCTP (3,000 Ci/mmol, 60 μCi). An additional 30 U of AMV RT was added, and the reaction mixture was incubated for 40 min at 37°C. The DNA products from this reverse transcription reaction were treated with 2 μg of RNase A for 30 min at room temperature and then extracted with phenol-chloroform. One-third of each reaction mixture was analyzed by polyacrylamide gel electrophoresis under denaturing conditions (29). Newly synthesized cDNAs were detected by autoradiography using Kodak XR-OMAT film. Each experiment was repeated three times, with similar results. The amount of product DNA was estimated by densitometric tracing analysis of the autoradiograms.

Computer-assisted RNA secondary structure prediction. Sequences for several members of the ASLV family and other retroviruses were obtained from the GenBank data base by using the UWGCG FETCH program run on a VAX Server 3600 (3, 13). Secondary structure predictions and energy calculations were made by using the FOLD program run on either a VAX Server 3600 or a VAX Station 3200 and were displayed by using the UWGCG SQUIGGLES program (13, 34). Structures were refined as suggested by Frier et al. (14) and Puglisi et al. (28). Some predictions were made by using the program DNASIS run on an IBM 386SX compatible computer.

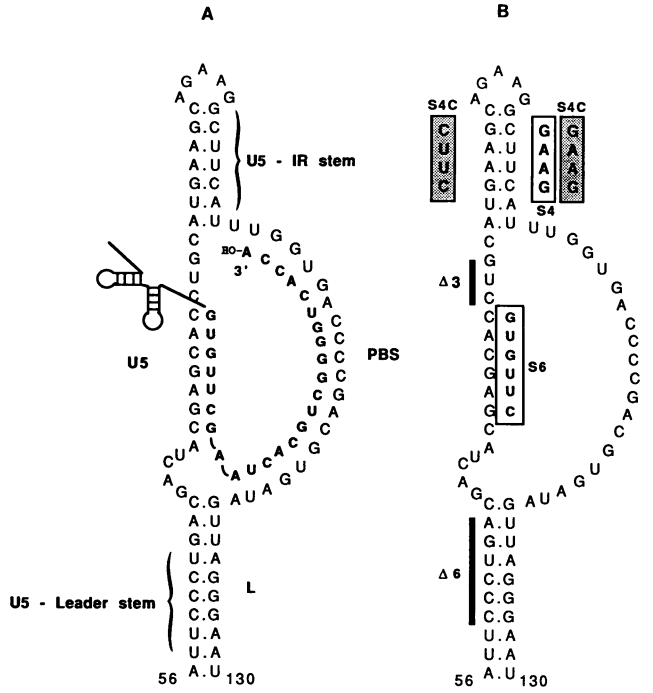

FIG. 1. Predicted ASLV RNA secondary structures for the 5′ noncoding region of RSV (SR-A) RNA near the PBS. (A) Numbers indicate nucleotides from the viral RNA 5′ end. U5, nucleotide sequences derived from the 5′ end of viral RNA; L, nucleotide sequences derived from the leader region of viral RNA; PBS, minus-strand primer binding site. The U5-leader stem and the U5-IR stem are as shown. The acceptor stem and TψC stem-loops of $tRNA^{Trp}$ have been unwound to permit interactions with the PBS and U5 RNA sequences. (B) U5-IR stem mutations Δ3, Δ6, S4, S4C, and S6. Sequence deletions are indicated by the thick bars, and sequence substitutions are boxed. The S4 substitution disrupts the U5-IR stem; the S4C substitution restores the U5-IR stem, but with altered sequence. The Δ6 mutation disrupts the U5-leader stem. The S6 substitution prevents the U5 RNA-TψC loop interaction.

RESULTS

An RSV (SR-A) RNA sequence extending from nucleotides 56 to 130 is shown in Fig. 1A. This sequence includes the minus-strand PBS and the U5-leader and U5-IR stem secondary structures (5, 7). Also shown is the potential interaction between U5 sequences, lying between the U5-IR and the U5-leader stem structures, and the TψC sequences of the tRNA primer. These U5 sequences appear to be strongly conserved among avian retroviruses (Table 1), suggesting that the potential interaction between the primer TψC sequences and U5 RNA may be important for viral replication. In order to test whether mutations that disrupt

TABLE 1. Avian retrovirus U5 RNA sequences complementary to the TψC loop of $tRNA^{Trp}$

Source of sequence[a]	U5 sequence complementary to TψC loop of $tRNA^{Trp}$	Base pairing potential with TψC loop
RSV Pr-C	CGAACAC	7/7
RSV Pr-A	CGAACAC	7/7
ASV UR2	CGAACAC	7/7
ev-2	CGAACAC	7/7
RAV-0	CGAACAC	7/7
ARV EV-1	CGAACAC	7/7
ACV MH2	CGAACAC	7/7
FSV	CGAACAC	7/7
AEV	CGAACAC	7/7
RSV SR-A	CGAGCAC	7/7
RSV Pr-B	CGAGCAC	7/7
RSV B77	CGAGCAC	7/7
AMCV HBI	CGAGCAC	7/7
ARV MH2-E21	CGAGCAC	7/7
AMCV MC29	CGCGAAC	5/7

[a] Virus and strain abbreviations are as follows: RSV Pr-C, RSV Prague C strain; Pr-A, Prague A strain; ASV UR2, avian sarcoma virus UR2; ev-2, endogenous virus 2; RAV-0, Rous associated virus-0; ARV EV-1, avian retrovirus EV-1; ACV MH2, avian carcioma virus MH2; FSV, Fujinami sarcoma virus; SR-A, Schmidt-Ruppin A strain; Pr-B, Prague B strain; AMCV HBI, avian myelocytomatosis virus HBI; B77, Bratislava 77 strain. Note that AMCV MC29 is replication defective (33).

the U5 RNA-TψC interaction result in a defect in reverse transcription, the S6 mutation (Fig. 1B) was constructed and its effect on viral replication was assessed. The S6 mutation is a 6-base substitution in U5 which prevents the above potential interaction.

Production of S6 mutant virus. The S6 mutation was cloned as an *Rsr*II-to-*Bst*EII fragment into the wild-type vector pDC101B (5). The resulting vector, pDC202B, was digested with *Hpa*I, ligated to *Hpa*I-digested pDC102B, and transfected into QT-6 quail fibroblasts, and stable transfectants were selected as described previously (5). These cell lines produced virus containing the S6 mutation in amounts comparable to those produced by wild-type-virus-producing cell lines prepared in the same way (data not shown).

Effect of the S6 mutation on viral growth. Equivalent amounts of virus, determined by RT activity by using exogenous template-primers, were collected from the S6- and the wild-type-virus-producing cell lines and used to infect QT-6 fibroblasts. The appearance and rate of production of progeny virus from these cells was monitored by RT activity, and the results are shown in Fig. 2. In four independent infection experiments, the S6 substitution mutation resulted in a delay in the appearance of progeny virus compared with what was seen with wild-type virus. The progeny virus collected from these cells also exhibited a delay in growth when infected into fresh QT-6 cells. This indicates that the late appearance of virus was not due to a reversion event. Since the S6 virus-producing cell lines released amounts of virus particles comparable to those released by the wild-type-producing cell lines, the slow growth observed for the S6 virus was not due to a defect in late events in viral replication, such as transcription, splicing, packaging, particle assembly, and budding. Furthermore, the RNA content of virions containing RNA with the S6 mutation, detected by hybridization to ^{32}P-labeled cDNA probes, was the same as in wild-type particles (data not shown). These results suggest that the S6 mutation causes an early defect to replication, similar to that observed by disruption of the U5-IR and U5-leader stem structures (5, 7).

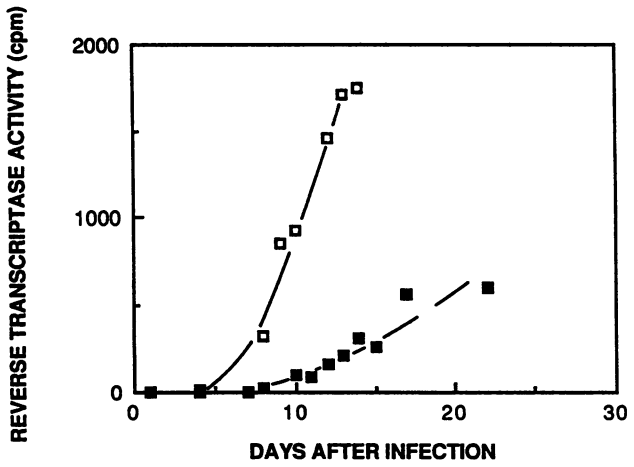

FIG. 2. Growth rates of wild-type and S6 mutant virus. Equal amounts of wild-type virus (□) or S6 virus (■) were infected into QT-6 fibroblasts, and the appearance of progeny virus was monitored by RT activity in cell supernatants as described in Materials and Methods.

Analyses of the amount of S6 viral DNA produced in acutely infected cells, as detected by Southern analysis, and of that integrated into the host cell DNA, as detected by polymerase chain reaction amplification followed by Southern analysis (5), indicated that each is reduced compared with amounts of wild-type DNA produced (data not shown). These results indicate that S6 virus manifests a replication defect in vivo.

Analysis of reverse transcription with endogenous primer-template complexes in permeabilized virions. Virus particles containing the S6 mutation were collected and treated with melittin, and the synthesis of both dAMP-$tRNA^{Trp}$ (Fig. 3A) and $cDNA_{101}$ (Fig. 3B) were examined. The amounts of these early replication intermediates were substantially smaller in virions containing the S6 mutation than in wild-type virions. This strongly suggests that the impairment in S6 reverse transcription in vivo is due to an initiation defect. Although these results could reflect a reduced amount of primer $tRNA^{Trp}$ bound at the PBS, this is considered unlikely since 35S RNA from ATVΔ21 (6), which has a 21-base U5 deletion that includes the TψC complementary sequence, bound as much $tRNA^{Trp}$ as did wild-type 35S RNA (data not shown).

Production of template and primer RNAs in vitro. Unlike other U5 substitution mutations (5, 7), the S6 mutation cannot be complemented by a base change in viral RNA. Instead, it requires sequence changes in the TψC arm of primer $tRNA^{Trp}$. To examine whether the S6 mutation could be complemented by base changes in the primer, we reconstituted a reverse transcription complex in vitro by utilizing as primers synthetic oligoribonucleotides that represent fragments of the $tRNA^{Trp}$ and a 1.1-kb segment of viral RNA containing part of U3, U5, PBS, the leader sequence, and part of *gag* as template (Fig. 4). The latter viral sequences were prepared by subcloning a DNA fragment extending from the *Eco*RI site in U3 to the *Bgl*II site in *gag* from each of the vectors pDC101B (wild type), pDC103B (Δ3), pDC202B (S6), pDC203B (S4), pDC204B (S4C), and pATVΔ6 (Δ6) into the plasmid pGEM3z as described in

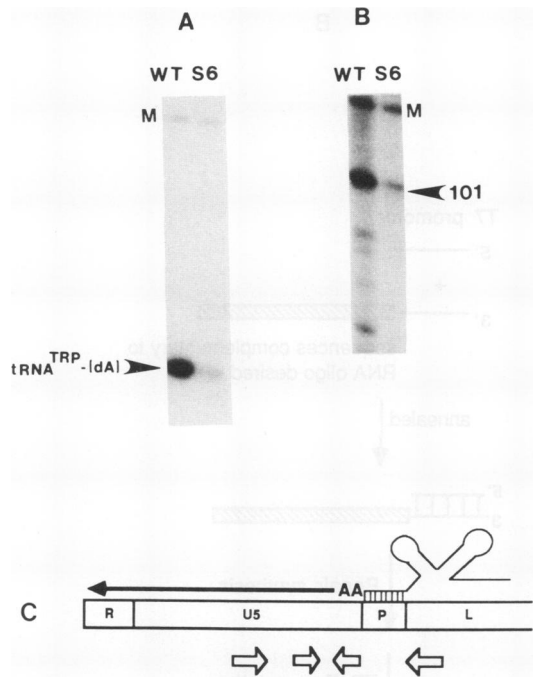

FIG. 3. Analysis of $tRNA^{Trp}$ primer extension products in melittin-activated virions. (A and B) The amount of $tRNA^{Trp}$ primer extension products formed after a 30-min incubation of virions with melittin and [α-^{32}P]dATP in the absence (A) or presence (B) of dCTP, dGTP, and TTP was as described in Materials and Methods. The wild type (WT) and the U5 mutant, S6, are indicated above the respective lanes. The migration distances of dA-tagged $tRNA^{Trp}$ and the $cDNA_{101}$ products are indicated. M denotes a ^{32}P-labeled 125-base DNA fragment added to each reaction upon termination. (C) Schematic representation of initiation of reverse transcription. AA, initial deoxynucleotides incorporated onto the 3'OH end of the $tRNA^{Trp}$ primer; closed arrow, $cDNA_{101}$ reverse transcription intermediate; open arrows, inverted repeat sequences; P, PBS; L, leader region.

Materials and Methods. Linearization of the resultant plasmids with *Sca*I, followed by in vitro RNA transcription catalyzed by T7 RNA polymerase, produced a 1.1-kb runoff transcript in each case (Fig. 4A).

The $tRNA^{Trp}$ primer fragments were prepared by the method of Milligan et al. (23). Oligodeoxynucleotide duplexes which contained a T7 promoter (both plus and minus strands) and the complement to the RNA sequences desired (in the minus strand) were filled in by repair DNA synthesis (Fig. 4B). The resultant duplexes were then utilized as templates to produce runoff RNA transcripts with T7 RNA polymerase (see Materials and Methods). Three RNA primers were synthesized. The first is a 20-mer (PBS) which can bind only the PBS. The latter two are 28-mers which, in addition to binding to the PBS, contain sequences that can base pair with wild-type U5 sequences (PBS-TψC) or S6 U5 sequences (PBS-S6-TψC).

Reconstitution of reverse transcription in vitro. Template RNA was annealed under physiological (low-salt) conditions to each of the above synthetic RNA primers in an RSV NC- and AMV RT-dependent reaction (Fig. 4C). Once annealed, the viral RNA-primer duplexes were analyzed for their abilities to support DNA synthesis catalyzed by RT in the presence of deoxynucleotides, one of which was labeled with ^{32}P. A runoff cDNA transcript primed from these RNAs is expected to be 164 bases in length after RNase treatment (Fig. 4C). A second oligodeoxynucleotide primer which was separately 5' ^{32}P end labeled was annealed to the RNA template (Fig. 4C) to serve as a control for the addition of equal amounts of template RNA to each reaction. A runoff cDNA transcript 105 bases in length is expected with this DNA primer.

When the PBS-TψC primer was used, the efficiency of reverse transcription from wild-type RNA was five times greater than when the same-sized PBS-S6-TψC primer was used (Fig. 5A, lanes b and c). The TψC sequences of the former but not the latter primer can anneal to the U5 sequences of wild-type RNA. When the experiment was repeated, substituting the S6 for wild-type viral RNA, a fivefold increase in reverse transcription was observed with the PBS-S6-TψC primer compared with what was observed with the PBS-TψC primer. Note that equal amounts of the $cDNA_{105}$ internal control product were obtained in all the reconstituted reactions (Fig. 5). These results indicate that there is a direct base pair interaction between U5 sequences and primer TψC sequences required for efficient initiation of reverse transcription. We note that when the PBS 20-mer primer, which lacks the TψC sequences, was used, the efficiency of priming appeared to be less than with the longer 28-mer primer, which could not anneal to the U5 RNA (Fig. 4 and 5). The reason for this behavior is not known, but it may be related to differences in the formation of the initiation complex caused by using primers of different sizes.

To examine whether the presence of the DNA primer decreases the efficiency of reverse transcription from the RNA primer, we performed reconstitution experiments with a wild-type RNA template containing a start site corresponding to that of authentic viral RNA and with the PBS-TψC RNA primer in the presence and absence of the DNA primer (Fig. 5B, lanes 1 and 2). The DNA product expected in this case should be 101 bases in length. No detectable change in the efficiency of reverse transcription from the RNA primer due to the presence of the DNA primer was observed. Similar results were obtained when the experiment was performed by using the plasmid pGEMAA-WT as template for RNA production (data not shown).

The U5-leader and U5-IR stem structures are required for the U5 RNA-TψC primer interaction to stimulate reverse transcription. Having established an in vitro system with purified components that closely mimics the initiation of reverse transcription observed in permeabilized virions, we sought to examine whether the U5-IR and U5-leader stem structures were required for the U5 RNA-TψC interaction to stimulate initiation. To answer this question, 1.1-kb viral RNA templates containing the Δ6 mutation, which disrupts the U5-leader stem, and the mutations S4 and S4C, which disrupt and reconstitute the U5-IR stem-loop, respectively (Fig. 1B), were prepared in vitro (5, 7, 23). In addition, an RNA containing the Δ3 mutation, which shortens the distance between the U5-IR and U5-leader stem structures, was also prepared (5) (Fig. 1B). Runoff transcripts utilizing the PBS and the PBS-TψC primers with these RNA templates are shown in Fig. 6. In each case, there was an increase in the efficiency of reverse transcription when the longer PBS-TψC primer was used. However, in contrast to what was observed with wild-type RNA, the degree of stimulation of reverse transcription was small for the Δ6, Δ3, and S4 RNAs, which disrupt either the U5-leader or U5-IR stem structures. In contrast, when RNA containing the S4C mutation was used as template, a three- to fourfold increase in reverse transcription was restored (Fig. 6). This mutation

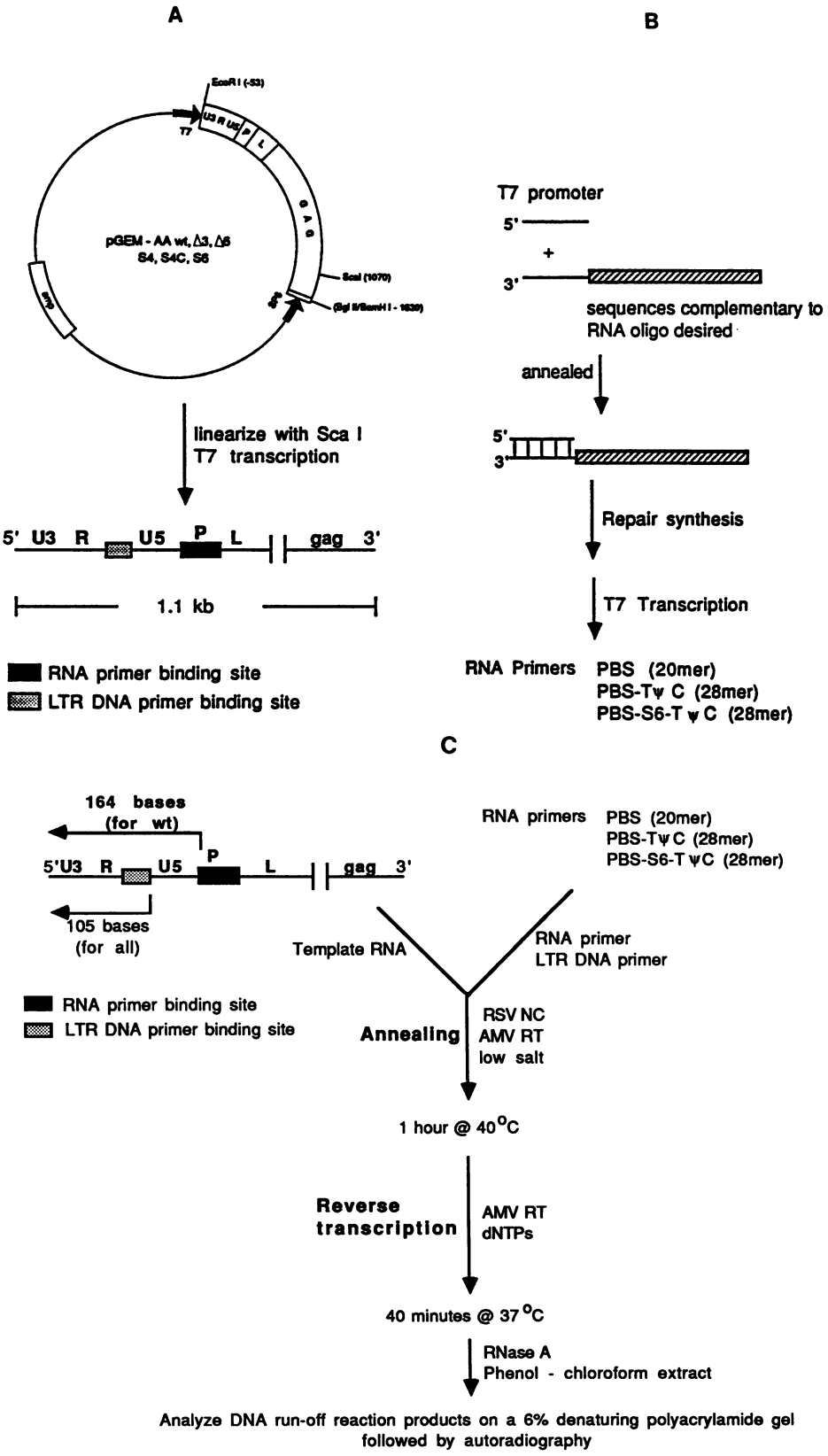

FIG. 4. Synthesis of templates and primers for reconstitution of reverse transcription in vitro. (A) The pGEM-AA series of vectors, containing the *Eco*RI-to-*Bgl*I fragment from wild-type (wt), Δ3, Δ6, S4, S4C, and S6 viral DNAs, was linearized with *Sca*I, and a runoff T7 RNA transcript was prepared for each vector as described in Materials and Methods. The 1.1-kb RNA transcripts in each case contains the terminal portion of U3, R, U5, PBS (indicated by black box marked with P), leader (L), and one-third of the *gag*. The stippled box shows

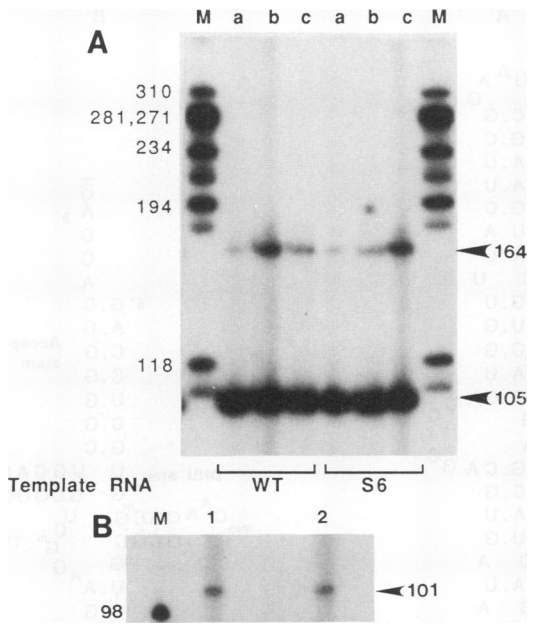

FIG. 5. (A) Efficient reverse transcription requires U5-TψC loop interactions. Reverse transcription was reconstituted with wild-type (WT) RNA or S6 RNA by utilizing the PBS (lanes a), PBS-TψC (lanes b), or PBS-S6-TψC (lanes c) primers as described in the legend to Fig. 4. A second LTR DNA primer was annealed to the viral RNA. The cDNA extension products utilizing the RNA primers and the LTR DNA primer are 164 and 105 bases in length, respectively (indicated by arrowheads). Lanes M contain 5′ ^{32}P-labeled, denatured, *Hae*III-digested φX174 DNA (sizes in bases indicated). (B) The presence of the DNA primer during the annealing and extension reactions does not affect the efficiency of extension from the RNA primer. Lane M, marker; lane 1, primer extension using an RNA template corresponding to authentic viral RNA that uses only the RNA primer, PBS-TψC; lane 2, primer extension performed on the same template using both the RNA and DNA primers. The RNA templates used in these reactions were produced by T7 transcription of a DNA template produced by polymerase chain reaction amplification of the R-U5-leader region of the plasmid pDC101.

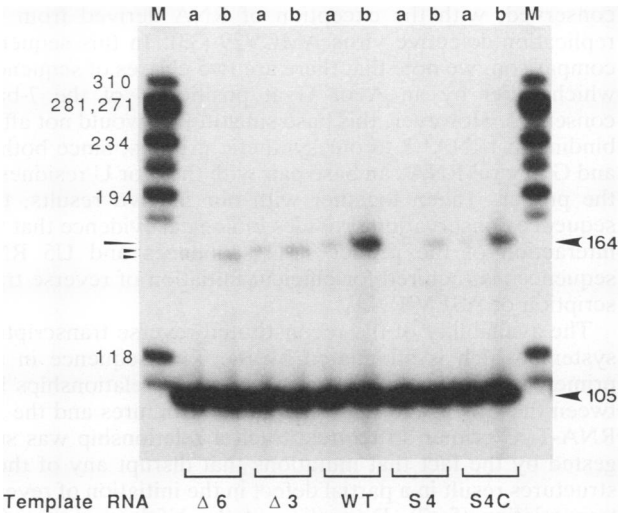

FIG. 6. Efficient reverse transcription requires intact U5-leader stem and U5-IR stem structures. Wild-type (WT) RNAs and RNAs containing the Δ6, Δ3, S4, and S4C mutations were used as templates (as indicated) in reverse transcription reactions that used either the PBS primer (lanes a) or the PBS-TψC primer (lanes b). The band at 164 bases represents the runoff product obtained with RNA primers annealed to wild-type, S4, and S4C RNAs (long pointed arrow at left); the bands at 161 and 158 bases (short pointed arrows at left) are runoff products from the Δ3 and Δ6 RNA templates, respectively. The band at 105 bases represents the runoff DNA formed with the LTR DNA primer. Lane M contains 5′ ^{32}P-labeled *Hae*III-digested φX174 DNA, as described in the legend to Fig. 5.

restores the U5-IR stem structure. These results indicate that both the U5-leader and U5-IR stem structures are necessary for the U5 RNA-TψC interaction to stimulate initiation of reverse transcription.

DISCUSSION

We have earlier shown that there are two regions of secondary structure in the 5′ nontranslated region of avian retrovirus RNA that contribute to efficient initiation of reverse transcription (5, 7). A schematic view of these structures is shown in Fig. 1A. In the present study, we have focused on the function of the sequences between these structured regions. Haseltine et al. (15, 16) had originally noted that these sequences could base pair with the TψC arm of the tRNA^{Trp} primer. Cordell and coworkers (10) later showed that tRNA^{Trp} fragments lacking the TψC loop did not reconstitute efficient reverse transcription of viral RNA compared with what was observed with longer fragments containing this region. We have now directly demonstrated that a 6-base substitution mutation in U5 RNA which disrupts the potential interaction with the TψC loop of tRNA^{Trp} results in a defect in the initiation of reverse transcription in permeabilized virions. In addition, the S6 defect can be complemented in vitro by changing the sequence in the TψC loop of the primer.

An analysis of U5 sequences from a variety of avian retroviruses (Table 1) has revealed that the 7-base region that interacts with the TψC loop of the primer is strongly

a region of U5 that can be bound by an oligodeoxynucleotide primer. (B) RNA primer fragments were prepared by using oligodeoxynucleotides containing a T7 promoter and sequences complementary to the tRNA primer of desired length (in the minus strand). The plus strand was filled in by DNA polymerase I (Klenow) repair synthesis and runoff T7 RNA transcripts prepared as described for panel A. RNA primers: PBS (20-mer), which can bind the PBS alone; PBS-TψC (28-mer), which can bind the PBS and wild-type U5 sequences as shown in Fig. 1; PBS-S6-TψC (28-mer), which can bind the PBS and mutant S6 U5 RNA. The sequences of the RNA primers produced are as follows: PBS, 5′GAAUCACGUCGGGGUCACCA3′; PBS-TψC, 5′GCGUGUUCGAAUCACGUCGGGGUCACCA3′; PBS-S6-TψC, 5′GCCACAAGGAA UCACGUCGGGGUCACCA3′. (C) Schematic representation of the reconstituted reverse transcription system. Template RNA was annealed to one of the three RNA primers (described in legend to panel B) and the LTR DNA primer in the presence of RSV NC protein and AMV RT, as described in Materials and Methods. Primer-dependent reverse transcription was reconstituted, the products were treated with RNase, and the newly synthesized DNAs were analyzed on a 6% denaturing polyacrylamide gel. The LTR DNA primer binds to nucleotides 21 to 40 of the RSV genome (sequence as in reference 30).

conserved, with the exception of RNA derived from the replication-defective virus AMCV29 (33). In this sequence comparison, we note that there are two classes of sequences which differ by an A or G at position 4 of the 7-base consensus. However, this base substitution would not affect binding of $tRNA^{Trp}$ to our synthetic primers, since both A and G in viral RNA can base pair with the ψ or U residues in the primer. Taken together with our in vivo results, this sequence conservation provides biological evidence that the interaction of the primer TψC sequences and U5 RNA sequences is required for efficient initiation of reverse transcription of ASLV RNA.

The availability of the reconstituted reverse transcription system, which is stimulated by the TψC sequence in the primer, has provided the means to study relationships between the U5-IR and U5-leader stem structures and the U5 RNA-TψC primer structures. Such a relationship was suggested by the fact that mutations that disrupt any of these structures result in a partial defect in the initiation of reverse transcription (5, 7). Disruption of the U5-leader stem (Δ6 mutation) or the U5-IR stem (S4 mutation) decreased initiation of reverse transcription in vitro. In contrast, a mutation that restored the U5-IR stem structure (S4C) restored the initiation of reverse transcription on the primer TψC loop sequences. Together, these experiments indicate that efficient reverse transcription requires all three interactions, i.e., the U5-leader and U5-IR stems and the U5 RNA-TψC structure. It is not clear how these interactions are interdependent. However, it is possible that the presence of all three causes the U5 region to be held in a rigid manner needed for recognition by RT or some other component of the reaction. Disruption of any one of these three interactions would disrupt this overall structure and result in a concomitant defect in reverse transcription. The Δ3 mutation, which decreases the distance between the two stem structures, may be defective in vivo and in vitro because it alters the orientation of each of the secondary structure regions to one another.

In the absence of primer tRNA, the U5 and leader regions of the viral RNA can form extensive hairpin structures (Fig. 7A). In order to form the reverse transcription complex depicted in Fig. 1A, both the viral RNA (Fig. 7A) and the tRNA (Fig. 7B) have to be denatured. The sum of the hydrogen bond energy lost in disrupting the viral RNA and the stem and TψC loops of the primer tRNA is calculated to be −56.9 kcal (−238 kJ) (see Materials and Methods). This is less than the hydrogen bond energy gained in forming the U5 RNA-tRNA structure shown in Fig. 1, which is calculated to be −70.2 kcal (−294 kJ).

Assembly of the primer complex (Fig. 1) in vitro requires both the viral NC and RT proteins. Barat et al. (2) have reported that NC functions by stabilizing the primer on the template RNA. The NC may also help to denature the extensive secondary structures in the 5′ noncoding region and in the primer, thereby promoting annealing (Fig. 8). Independent evidence for a role for NC in reverse transcription comes from the studies of Meric and Goff (22), who have reported that mutations in the NC abolish reverse transcription of viral RNA.

A role for RT in forming an initiation complex is also possible. RT is capable of promoting the unwinding of RNA-DNA and DNA-DNA duplexes but not RNA-RNA duplexes (9). RT also specifically interacts with the primer $tRNA^{Trp}$ (26). A large proportion of this interaction is maintained in primer fragments lacking the anticodon loop but retaining the TψC loop and adjacent sequences (17).

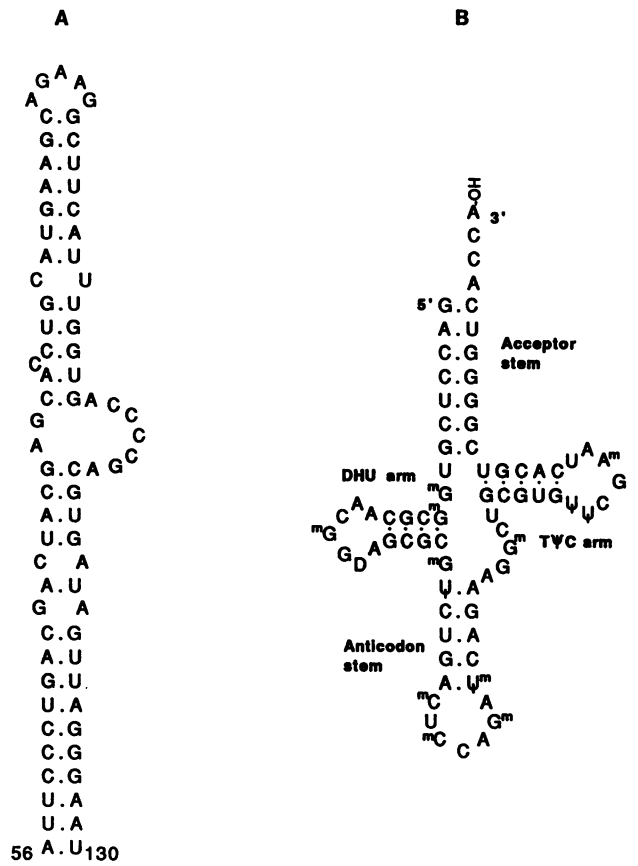

FIG. 7. Predicted secondary structure of ASLV viral RNA and primer $tRNA^{Trp}$. (A) ASLV viral RNA can form an extended hairpin loop in the region of the PBS in the absence of primer which includes sequences within the PBS. The numbers indicate nucleotides from the 5′ end of viral RNA. (B) $tRNA^{Trp}$ is presented in the cloverleaf structure. Modified bases are as indicated, without the actual position of methylation being devoted.

Thus, RT may interact with these sequences in the synthetic primers, thereby promoting the formation of an initiation complex.

It should be noted that the reconstituted system described here employed segments of tRNA prepared without modified bases; the sequence ψψC in $tRNA^{Trp}$ is UUC in our synthetic primer. The lack of modified bases, however, doesn't preclude detecting the functional consequences of the interaction of these sequences with U5 RNA. There are other known instances when synthetic tRNA molecules are biologically active. For example, $tRNA^{Phe}$ made as a tDNA as well as an unmodified T7 transcript can be charged with phenylalanine by the appropriate aminoacyl tRNA synthetase (20, 25). Also, the aminoacyl $tRNA^{Ala}$ synthetase charges a short ribonucleotide, prepared as an in vitro T7 transcript, corresponding to the acceptor stem and the TψC loop (31).

Can other retroviruses form similar U5 RNA structures and tRNA interactions? Having demonstrated requirements for the U5-IR stem (5), U5 RNA-TψC interaction, and U5-leader stem (7) in the ASLV, we examined whether other retroviruses could form similar interactions. The U5 sequences near the PBSs of several type C and D retroviruses, including Moloney leukemia virus (MuLV), spleen necrosis

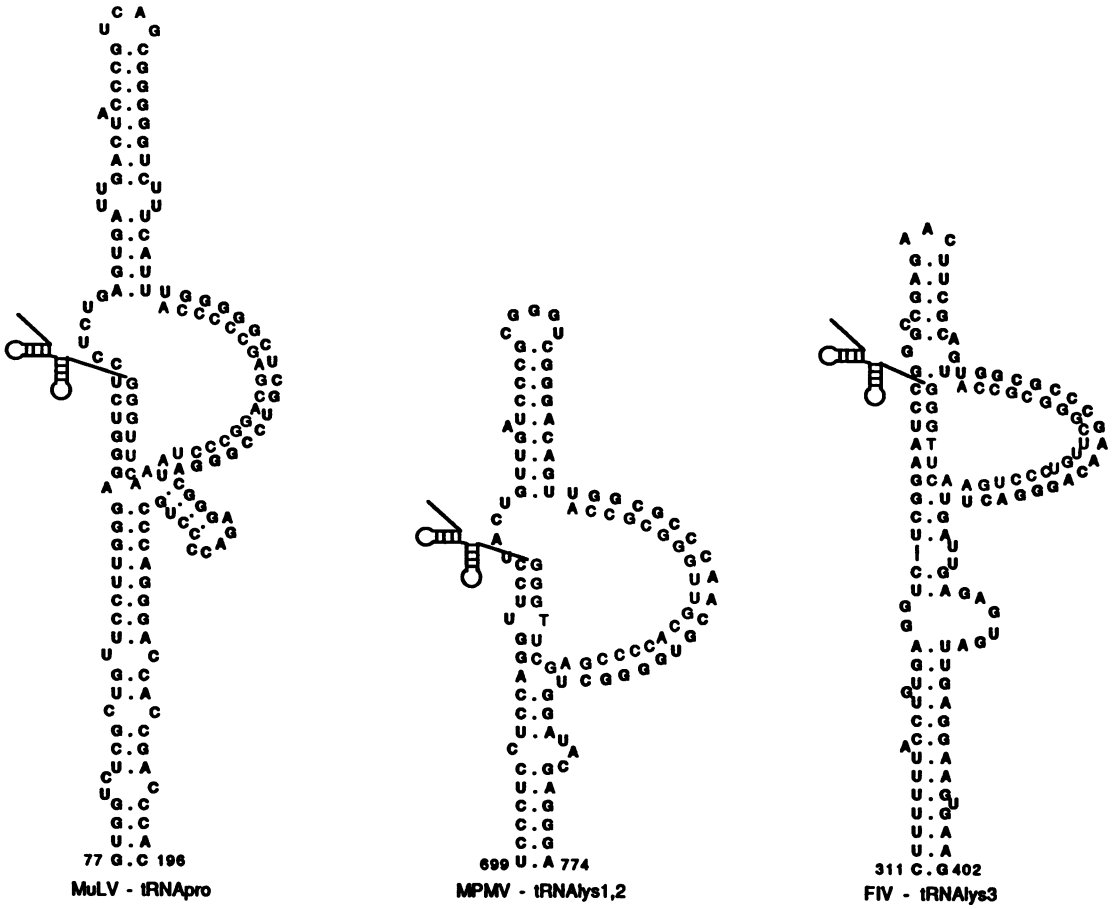

FIG. 8. Other type C and D retroviruses can potentially form U5 secondary structures and TψC interactions with the primer. The secondary structures (U5-leader and U5-IR stems) predicted to form near the PBS of MuLV, MPMV, and FIV RNAs are shown. The potential interactions between the PBS and U5 sequences with $tRNA^{Pro}$, $tRNA_{1,2}^{Lys}$, and $tRNA_3^{Lys}$, respectively, are indicated.

virus, simian sarcoma virus, bovine leukemia virus, mouse mammary tumor virus, human immunodeficiency virus type 1 (HIV-1) and HIV-2, gibbon ape leukemia virus, baboon endogenous virus, simian retrovirus type 1, equine infectious anemia virus, feline leukemia virus, Mason-Pfizer monkey virus (MPMV), and feline immunodeficiency virus (FIV), were examined. In all cases, U5-IR and U5-leader structures and U5 RNA-TψC loop interactions appeared to be conserved. The structures for MuLV, MPMV, and FIV, each of which utilizes a different tRNA primer, are shown in Fig. 8. While they differ from the ASLV structure shown in Fig. 1, their general features are the same. Each contains the U5-IR and U5-leader stem structures and the potential to form U5-TψC primer interactions. In the case of MuLV, mutations that would disrupt the U5-IR stem have been shown to result in a reverse transcription defect, similar to our results with ASLV U5-IR stem mutations (5, 24).

The finding that similar structures and interactions exist in several retroviral RNAs suggests that these viruses probably utilize a common mechanism to initiate reverse transcription. Virus particles contain all the enzymatic and structural components required for reverse transcription, yet initiation of reverse transcription does not proceed until a virus particle fuses with the cell in vivo or is treated with detergents or melittin in vitro. Thus, it is possible that the above structures and interactions cannot form in the constrained environment of the virus particle but do so upon entrance into a cell, thereby initiating reverse transcription.

There are now several known instances in retrovirus systems in which the secondary structure of the viral RNA plays an important role in replication. These include the hairpin structures required for (i) frameshifting during translation expression of the *pol* gene (18, 19); (ii) selective activation of 5′ LTR transcription, possibly through a cell-derived RNA enhancer protein (3a); (iii) transcription activation of the 5′ LTR by HIV *tat* interacting with the *tar* element (11); (iv) selective HIV-1 mRNA transport facilitated by *rev* interaction with the *rev*-responsive element (1); and (v) initiation of reverse transcription as reported here.

ACKNOWLEDGMENTS

We thank Han Huang for assistance in constructing viral mutants and Charles Huang for preparation of the wild-type T7 transcription vector.

This work was supported in part by Public Health Service grant CA 38046 from the National Cancer Institute and by Cancer Research Center grant P30 CA 43703. D.C. was supported in part by Public Health Service training grant T32 GM 07250 from the National Institutes of General Medical Sciences.

REFERENCES

1. **Ahmed, Y. F., S. M. Hanly, M. H. Malim, B. R. Cullen, and W. C. Greene.** 1990. Structure-function analyses of the HTLV-I

Rex and HIV-1 Rev RNA response elements: insights into the mechanism of Rex and Rev action. Genes Dev. **4**:1014–1022.
2. **Barat, C., V. Lullien, O. Schatz, G. Keith, M. T. Nugeyre, F. Gruninger-Leitch, F. Barre-Sinoussi, S. F. J. LeGrice, and J. L. Darlix.** 1989. HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA. EMBO J. **8**:3279–3285.
3. **Bilofsky, H. S., C. Burks, J. W. Fickett, W. B. Goad, F. I. Lewitter, W. P. Rindone, C. D. Swindell, and C. S. Tung.** 1986. The GenBank Genetic Sequence Data Bank. Nucleic Acids Res. **14**:1–4.
3a. **Boerkoel, C., and H.-J. Kung.** Unpublished observations.
4. **Boone, L. R., and A. M. Skalka.** 1981. Viral DNA synthesis in vitro by avian retrovirus particles permeabilized with melittin. I. Kinetics of synthesis and size of minus- and plus-strand transcripts. J. Virol. **37**:109–116.
5. **Cobrinik, D., A. Aiyar, Z. Ge, M. Katzman, H. Huang, and J. Leis.** 1991. Overlapping U5 sequence elements are required for efficient integration and initiation of reverse transcription. J. Virol. **65**:3864–3872.
6. **Cobrinik, D., R. Katz, R. Terry, A. M. Skalka, and J. Leis.** 1987. Avian sarcoma and leukosis virus *pol*-endonuclease recognition of the tandem long terminal repeat junction: minimum site required for cleavage is also required for viral growth. J. Virol. **61**:1999–2008.
7. **Cobrinik, D., L. Soskey, and J. Leis.** 1988. A retroviral secondary structure required for the efficient initiation for reverse transcription. J. Virol. **62**:3622–3630.
8. **Coffin, J. M.** 1990. Retroviridae and their replication, p. 1437–1500. *In* B. N. Fields, D. M. Knipe, R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman (ed.), Virology, 2nd ed. Raven Press, Ltd., New York.
9. **Collett, M. C., J. P. Leis, M. S. Smith, and A. J. Faras.** 1978. Unwinding-like activity associated with avian retrovirus RNA-directed DNA polymerase. J. Virol. **26**:498–509.
10. **Cordell, B., R. Swanstrom, H. M. Goodman, and J. M. Bishop.** 1979. $tRNA^{trp}$ as primer for RNA-directed DNA polymerase: structural determinants of function. J. Biol. Chem. **254**:1866–1874.
11. **Cullen, B. R.** 1990. The HIV-1 Tat protein: an RNA sequence-specific processivity factor. Cell **63**:655–657.
12. **Dahlberg, J. E., R. C. Sawyer, J. M. Taylor, A. J. Faras, W. E. Levinson, H. M. Goodman, and J. M. Bishop.** 1974. Transcription of DNA from the 70S RNA of Rous sarcoma virus. I. Identification of a specific 4S RNA which serves as primer. J. Virol. **13**:1126–1133.
13. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
14. **Frier, S. M., R. Kierzek, J. A. Jaeger, M. Sugimoto, M. H. Caruthers, T. Neilson, and D. H. Turner.** 1986. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA **83**:9373–9377.
15. **Haseltine, W. A., D. G. Kleid, A. Panet, E. Rothenberg, and D. Baltimore.** 1976. Ordered transcription of RNA virus genomes. J. Mol. Biol. **106**:109–131.
16. **Haseltine, W. A., A. M. Maxam, and W. Gilbert.** 1977. Rous sarcoma virus genome is terminally redundant: the 5′ sequence. Proc. Natl. Acad. Sci. USA **74**:989–993.
17. **Hu, J. C., and J. E. Dahlberg.** 1983. Structural features required for the binding of $tRNA^{trp}$ to avian myeloblastosis virus reverse transcriptase. Nucleic Acids Res. **11**:4823–4833.
18. **Jacks, T., M. D. Power, F. R. Masiarz, P. A. Luciw, P. J. Barr, and H. E. Varmus.** 1988. Characterization of ribosomal frameshifting in HIV-1 *gag-pol* expression. Nature (London) **331**:280–283.
19. **Jacks, T., K. Towsley, H. E. Varmus, and H. Majors.** 1987. Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus *gag*-related polyproteins. Proc. Natl. Acad. Sci. USA **84**:4298–4302.
20. **Khan, A. S., and B. A. Roe.** 1988. Aminoacylation of synthetic DNAs corresponding to *Escherichia coli* phenylalanine and lysine tRNAs. Science **241**:74–79.
21. **Leis, J., S. Johnson, L. S. Collins, and J. A. Traugh.** 1984. Effects of phosphorylation of avian retrovirus nucleocapsid protein pp12 on binding of viral RNA. J. Biol. Chem. **259**:7726–7732.
22. **Meric, C., and S. Goff.** 1989. Characterization of Moloney murine leukemia virus mutants with single amino acid substitutions in the Cys-His box of the nucleocapsid protein. J. Virol. **63**:1558–1568.
23. **Milligan, J., D. Groebe, G. Witherell, and O. Uhlenbeck.** 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res. **15**:8783–8798.
24. **Murphy, J. E., and S. Goff.** 1989. Construction and analysis of deletion mutations in the U5 region of Moloney murine leukemia virus: effects on RNA packaging and reverse transcription. J. Virol. **63**:319–327.
25. **Normanly, J., and J. Abelson.** 1989. tRNA identity. Annu. Rev. Biochem. **58**:1029–1049.
26. **Panet, A., W. A. Haseltine, D. Baltimore, G. Peters, F. Harada, and J. E. Dahlberg.** 1975. Specific binding of $tRNA^{trp}$ to avian myeloblastosis virus RNA-dependent DNA polymerase. Proc. Natl. Acad. Sci. USA **72**:2535–2539.
27. **Prats, A.-C., L. Sarih, C. Gabus, S. Litvak, G. Keith, and J. L. Darlix.** 1988. Small finger protein of avian and murine retroviruses has nucleic acid annealing activity and positions the replication primer tRNA onto genomic RNA. EMBO J. **7**:1777–1783.
28. **Puglisi, J. D., J. R. Wyatt, and I. Tinoco.** 1990. Solution conformation of an RNA hairpin loop. Biochemistry **29**:4215–4226.
29. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Schwartz, D. E., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell **22**:853–869.
31. **Shi, J.-P., C. Francklyn, K. Hill, and P. Schimmel.** 1990. A nucleotide that enhances the charging of RNA minihelix sequence variants with alanine. Biochemistry **29**:3621–3626.
32. **Sutton, R., and J. Boothroyd.** 1986. Evidence for *trans* splicing in trypanosomes. Cell **47**:527–533.
33. **Weiss, R., N. Teich, H. Varmus, and J. Coffin (ed.).** RNA tumor viruses, parts 1 and 2, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **Zuker, M., and P. Steigler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9**:133–148.