JOURNAL OF VIROLOGY, Feb. 1994, p. 611–618
0022-538X/94/$04.00+0

Vol. 68, No. 2

A Specific Orientation of RNA Secondary Structures Is Required for Initiation of Reverse Transcription

ASHOK AIYAR, ZHENG GE, AND JONATHAN LEIS*

Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106

Received 2 September 1993/Accepted 19 October 1993

The 5′ end of avian retrovirus RNA near the primer-binding site (PBS) forms two secondary structures, the U5-inverted repeat (U5-IR) and the U5-leader stems, and contains a 7-nucleotide sequence that anneals to the TψC loop of the $tRNA^{Trp}$ primer. Mutations that disrupt any of these base pair interactions cause defects in initiation of reverse transcription both in vivo and in vitro (D. Cobrinik, A. Aiyar, Z. Ge, M. Katzman, H. Huang, and J. Leis, J. Virol. 65:3864–3872, 1991; A. Aiyar, D. Cobrinik, Z. Ge, H.-J. Kung, and J. Leis, J. Virol. 66:2464–2472, 1992). We have now examined the effect of perturbing the non-base-paired intervening "spacer" sequences between these secondary-structure elements. Small deletions or insertions in these intervening sequences decreased initiation of reverse transcription in vitro. In contrast, base substitutions, which maintain the spacing distances between the structures, had no detectable effect. Additionally, a small deletion at the 3′ end of the PBS caused a significant decrease in initiation of reverse transcription whereas substitution mutations again had no effect. Together, these results indicate that reverse transcriptase forms a complex in which the different structural elements are maintained in a specific orientation that is required for efficient initiation of reverse transcription. Specific sequence recognition of the duplex structures by reverse transcriptase is also required since mosaic RNAs that combine the human immunodeficiency virus type 1 PBS with avian sequences is not efficiently utilized for reverse transcription even though the primer used can anneal to the substituted PBS.

Retroviruses are competent to initiate DNA synthesis shortly after budding from the cell, yet for most viruses there is no evidence that significant reverse transcription occurs in intact, extracellular virions. This suggests the existence of a mechanism that delays initiation of DNA synthesis until infection of a host cell. Details of initiation of DNA synthesis are not fully understood. However, the process is influenced by a series of viral 5′ noncoding RNA secondary structures which are capable of modulating the efficiency of the reaction (1, 7, 8). These secondary structures are near the primer-binding site (PBS) and are called the U5-leader stem, the U5-IR stem, and the U5-TψC interaction region. Direct evidence for the formation of some of these structures has been provided by nuclease sensitivity studies of viral RNA transcripts (5, 6). The U5-leader stem is formed by sequences in the middle of U5 and in the leader (8). Mutations which disrupt the structure impair initiation of reverse transcription in infected cells, whereas mutations which alter the sequence but retain the structure have no effect. The U5-IR stem is composed of an inverted repeat in U5 directly adjacent to the PBS. Disruption of this structure results in the same defect in initiation of reverse transcription (7). Independent studies by Murphy and Goff (16) with murine leukemia virus support a role for U5-IR sequences in reverse transcription. In Rous sarcoma virus (RSV) DNA, the U5-IR stem falls within U5 sequences that have been shown to be recognized and utilized by IN for integration (11). Thus, mutations that disrupt the U5-IR stem structure and alter its nucleotide sequence cause defects in integration as well as reverse transcription (7).

In addition to the U5-IR and U5-leader stems, an interaction between seven nucleotides of the TψC arm of the $tRNA^{Trp}$ and the U5 sequences lying between the stem structures is necessary for efficient initiation of reverse transcription by avian sarcoma/leukosis virus reverse transcriptase (RT). An in vitro reconstituted system was used to demonstrate that stimulation of reverse transcription by the U5 RNA-TψC interactions depends on both the U5-leader and U5-IR stem structures (1). In the present study, we show that the non-base-paired sequences between the various RNA secondary structures are also important for efficient reverse transcription, most probably by maintaining the duplex regions at a specific distance or orientation with respect to each other.

* Corresponding author. Mailing address: Department of Biochemistry, Case Western Reserve University, 2119 Abington Rd., Cleveland, OH 44106. Phone: (216) 368-3360. Fax: (216) 368-4544. Electronic mail address: jxl8@biochemistry.cwru.edu

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 various restriction enzymes were purchased from New England Biolabs. T7 RNA polymerase (20 U/μl), RNasin (RNase inhibitor) (40 U/μl), and RQ1DNase (1 U/μl) were purchased from Promega. Avian myeloblastosis virus (AMV) reverse transcriptase (RT) (30 U/μl) was obtained from Molecular Genetics Resources. Human immunodeficiency virus type 1 (HIV-1) RT (15 U/μl) was a gift from S. LeGrice, Department of Medicine, Case Western Reserve University. Vent DNA polymerase (2 U/μl) was obtained from New England Biolabs. The sequencing kit (Sequenase v2.0) and shrimp alkaline phosphatase (20 U/μl) were obtained from United States Biochemicals. All enzymes were used as specified by the manufacturers. 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 (1). 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 introduced into *E. coli* DH5α (Bethesda Research Laboratories). pDC101B and pDC103B were propagated as previously described (8). Plasmids were isolated by the alkaline lysis procedure and purified by equilibrium cesium chloride density gradient centrifugation (22).

Mutagenesis. All mutations were made by using the PCR mutagenesis method described by Aiyar and Leis (2). The following mutagenesis oligonucleotides were used:

Δ3LS	5′CCCTGACGA...CGAGCACCTGCATGAAGC3′
S5LS	5′CCCTGACCTTCGCGAGCACCTGCATGAAGC3′
I2LS	5′CCCTGACGATGATTCGAGCACCTGCATGAAGC3′
I4LS	5′CCCTGACGACTGACTACGAGCACCTGCATGAAGC3′
I4US	5′GCTTCATGCTGGCAGGTGCTCG3′ (minus strand)
S4US	5′ACGAGCACGCATATGAAGCAGAAGGC3′
Lys3PBS	5′GGCTTCATTTGGCGCCTGAACAGGGACAGTTAGGGAATAGTGGTCGG3′
Δ3PBS	5′GGCTTCATTTGGTGACCCCGACGT...AGTTAGGGAATAGTGGTCGG3′
S3PBS	5′GGCTTCATTTGGTGACCCCGACGTACGAGTTAGGGAATAGTGGTCGG3′

The nomenclature used to described these mutations is as follows. The first letter indicates that there is a deletion (Δ), insertion (I), or substitution (S); Lys3 indicates substitution of the HIV-1 PBS for the RSV PBS; the number indicates the size of the deletion, insertion, or substitution; the remaining letters indicate the site, i.e., lower spacer (LS), upper spacer (US), or primer-binding site (PBS) (see Fig. 1A). Within the sequence the underlined letters represent either an insertion or a substitution. The periods indicate a deletion. The Lys3PBS mutation contains an unintended C-to-T mutation at position 71. This mutation is boxed in Fig. 2A. The I2LS mutation contains both a 2-nucleotide insertion and a 3-nucleotide substitution. For simplicity, we will refer to this mutation as primarily an insertion.

In every instance, except for the I4US mutation, the mutagenesis oligonucleotide was of the same orientation as the viral RNA. The mutations were cloned in pDC101B by using the *Rsr*II site in U5 (nucleotide 47) and the *Sac*I site in the leader (nucleotide 259) to create plasmids pAAS4US, pAAI4US, pAAI2LS, pAAI4LS, pAAΔ3LS, pAAS5LS, pAALys3PBS, pAAΔ3PBS, and pAAS3PBS, respectively. A *Bst*EII (cuts at position 103)-to-*Sac*I fragment from pAAI2LS was replaced with the corresponding fragment from pAAΔ3PBS to create the double mutation containing plasmid pAAI2LS, Δ3PBS.

Production of synthetic templates. pDC101B (wild type), pDC103B(Δ3US), pAAS4US, pAAI4US, pAAI2LS, pAAI4LS, pAAΔ3LS, pAAS5LS, pAALys3PBS, pAAΔ3PBS, pAAI2LS-Δ3PBS, and pAAS3PBS (100 ng each) were linearized with *Eco*RI. Then 20 ng of linearized DNA was used as the template for individual PCR amplification reaction with the following primers:

T71–18	5′CCCTAATACGACTCACTATAGCCATTTTACCATTCACC3′
ASLV400–380	5′CACCTTTATGACGGCTTCCAT3′

The T71–18 primer contains a T7 promoter immediately preceding nucleotides 1 to 18 of the direct repeat, R, sequence in the plus orientation. The ASLV400–380 primer is the sequence from nucleotides 380 to 400 in the minus orientation. The thermostable error-correcting polymerase, Vent, was used for the amplification reaction to minimize errors during DNA synthesis. The PCRs were fractionated by 2% agarose gel electrophoresis, and the products were purified by transfer to DEAE-cellulose membranes and eluted with high salt (22). The purified DNA was then used as a template for T7 RNA polymerase as previously described (1). The transcription reactions were stopped by incubation with 10 U of RQ1DNase for 1 h at 37°C. Protein was removed by extraction with phenol-chloroform, and RNA was precipitated with 2 volumes of ethanol and collected. The RNA was suspended in 10 mM Tris · HCl (pH 7.6)–20 mM NaCl–2 mM EDTA. Each transcription reaction typically yielded about 10 μg of RNA for every 1 μg of DNA template used. RNAs were brought to a concentration of 0.3 μg/ml and stored at −20°C. Template RNAs were stored in 1.2-μg aliquots. Immediately prior to use, RNAs were heated to 80°C in a dry bath and slowly cooled to room temperature to permit the formation of stable RNA secondary structures.

Production of synthetic RNA primers. RNA primers were synthesized by the technique of Milligan et al. (15) with the modifications as described by Aiyar et al. (1). Briefly, RNA transcripts were prepared by in vitro transcription catalyzed by T7 RNA polymerase, using the duplex oligodeoxynucleotides, which were purified by 20% polyacrylamide gel electrophoresis (22). Care was taken to avoid contamination by incomplete products of synthesis. After transcription, a 1/20 aliquot of the transcription reaction mixture was incubated with shrimp alkaline phosphatase followed by T4 polynucleotide kinase plus [γ-^{32}P]ATP. The 5′ [^{32}P]RNA was added to the total transcription reaction mixture and subjected to gel electrophoresis through 15% polyacrylamide gels under denaturing conditions (1). Approximately 50% of the RNA was found to be of the correct size. This band was cut from the gel and eluted with 50% efficiency overnight in 1 ml of 0.5% ammonium sulfate–100 mM EDTA at 37°C. The RNA was precipitated with 3 volumes of ethanol, collected by centrifugation, and suspended in RNase-free distilled water. The concentration of RNA was estimated by measurement of its A_{260}. Unlike the template RNA synthesis, RNA primer synthesis was not very efficient. We obtained 1.5 nmol of RNA primer from 0.5 nmol of duplex deoxynucleotide. Primers were stored in RNase-free distilled water at a concentration of 50 pmol/μl (50 μM). Primers were used directly in the reconstitution reactions. The following RNA primers were synthesized:

Trp20	5′GAAUCACGUCGGGGUCACCA3′
Trp28	5′GCGUGUUCGAAUCACGUCGGGGUCACCA3′
Trp28S3PBS	5′GCGUGUUCGACGUACGUCGGGGUCACCA3′
Trp55	5′GCGCGUCUGACUCCAGAUCAGAAGGCUGCGUGUUCGAAUCACGUCGGGGUCACCA3′
Lys20	5′AAGUCCCUGUUCGGGCGCCA3′
Lys28	5′CAGGGUUCAAGUCCCUGUUCGGGCGCCA3′
LysTrp28	5′GCGUGUUCAAGUCCCUGUUCGGGCGCCA3′
Lys55	5′GAGCAUCAGACUUUUAAUCUGAGGGUCCAGGGUUCAAGUCCCUGUUCGGGCGCCA3′
LysTrp55	5′GAGCAUCAGACUUUUAAUCUGAGGGUCGCGUGUUCAAGUCCCUGUUCGGGCGCCA3′

The nomenclature for the primers is as follows. Trp20 and Lys20 are the terminal 20 nucleotides of $tRNA^{Trp}$ and $tRNA_3^{Lys}$, respectively. Trp28 and Lys28 are the 28 nucleotides of $tRNA^{Trp}$ and $tRNA_3^{Lys}$, respectively, which includes the TψC stem-loop of the tRNAs. Trp28S3PBS is the 28 terminal nucleotides of $tRNA^{Trp}$ but containing a 3-base substitution that allows the primer to anneal to a PBS containing the complement of the underlined sequences. LysTrp28 contains the TψC sequences of $tRNA^{Trp}$ juxtaposed to the terminal 20 nucleotides of $tRNA_3^{Lys}$. The 55-nucleotide primers can form the anticodon stem-loop in addition to the TψC loop of the respective tRNA primers. LysTrp55 is another mosaic primer that contains the $tRNA^{Trp}$ TψC loop sequences embedded into a background of the $tRNA_3^{Lys}$ sequences.

Reconstitution of reverse transcription in vitro. Reverse transcription reactions were reconstituted in vitro by a modification of the procedure described by Prats et al. (19) and Barat et al. (3). Various synthetic template RNAs (1.2 μg) were each mixed with 50 pmol of unlabeled RNA primer (as

indicated) and 25 pmol of unlabeled long terminal repeat (LTR) DNA primer in a 10-μl annealing reaction that contained 40 mM Tris · HCl (pH 8.0), 60 U AMV RT, and 60 mM NaCl. The reaction mixture was incubated at 40°C for 30 to 40 min, at which time it was brought to a final volume of 20 μl containing 60 mM NaCl; 6.7 mM $MgCl_2$; 5 mM dithioerythritol; 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 10 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 (22). Newly synthesized cDNAs were detected by autoradiography with Kodak XR-OMAT film. Each experiment was repeated three times with similar results. The amount of product DNA was estimated by densiometric tracing analysis of the autoradiograms with a SciScan 5000 (U.S. Biochemical Corp.).

RESULTS

The RSV RNA sequence extending from nucleotides 56 to 130 is shown in Fig. 1A. The U5-IR and U5-leader stem structures and the U5-TψC interaction region are as depicted. Using a reconstituted in vitro system, we previously demonstrated that all of these structural interactions were required for efficient initiation of reverse transcription (1). Additionally, disruption of either the U5-leader stem or the U5-IR stem abrogated the U5-TψC interaction (1). Lying between the U5-IR and U5-TψC region and between the latter and the U5-leader stem are 4- and 5-nucleotide non-base-paired sequences, respectively. These sequences will be referred to as the upper and lower spacers, respectively. We have previously introduced a 3-nucleotide deletion into the upper spacer and shown that it decreased viral replication in vivo as a result of a defect in initiation of reverse transcription (1, 7). To investigate whether this type of defect was caused by a distortion of a specific conformation or orientation between the different secondary structures, we have introduced a series of insertions, deletions, or substitutions into the upper and lower spacers as well as the 3′ end of the PBS.

Reconstitution of reverse transcription in vitro. To analyze the effect of these mutations, a series of T7 transcripts that represent the 5′ nontranslated U5 RSV template RNA and the $tRNA^{Trp}$ primer were prepared as described in Materials and Methods. The viral RNA templates begin at nucleotide 1 in R (numbering of Schwartz et al. [23]) and extend to nucleotide 400 in the *gag* gene. The RNA primers represent the 3′-terminal 20, 28, and 55 nucleotides, respectively, of $tRNA^{Trp}$. The templates and primers were annealed in a molar ratio of 1:5 (template to primer) in the presence of AMV RT under conditions described in Materials and Methods. As a control for the amount of RNA added to a reaction, an excess of an LTR deoxynucleotide primer was also added. Both annealed primers were then extended by the addition of deoxynucleotides (one of which was α-^{32}P labeled) and an additional 30 U of AMV RT (1). Following extension, the protein was removed by phenol extraction, and reaction products were treated with RNase and separated on 8% polyacrylamide gels under denaturing conditions. A DNA extension from the RNA primer annealed at the PBS produces a runoff cDNA transcript of 101 deoxynucleotides for the wild type. Extension from the DNA primer plus the primer itself is 40 deoxynucleotides. This is schematically depicted in Fig. 1B.

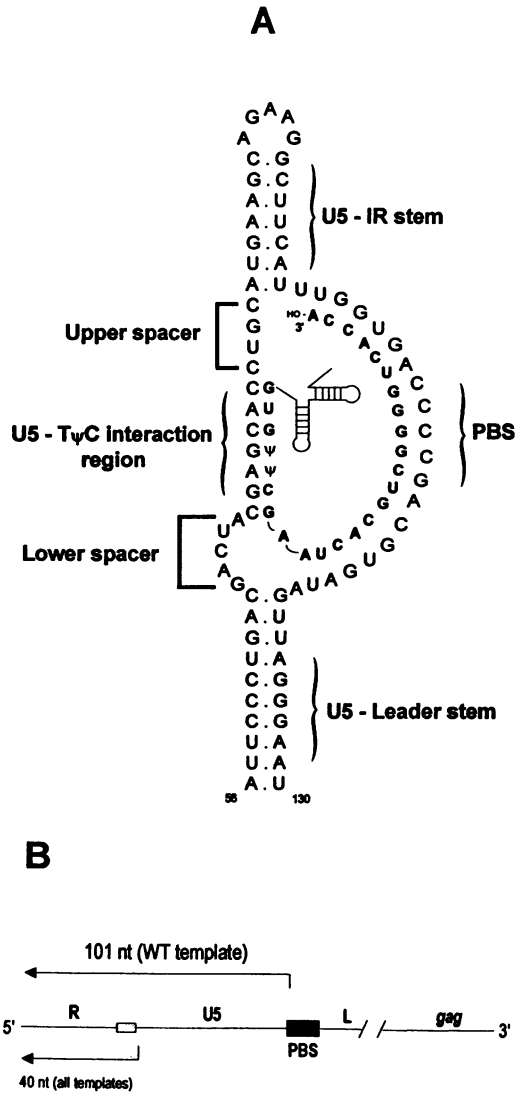

FIG. 1. Predicted RNA secondary structure for the 5′ noncoding region of wild-type RSV (SR-A) RNA near the PBS. (A) Numbers indicate nucleotides from the viral RNA 5′ end. U5, nucleotide sequences derived from the 5′ end of viral RNA; PBS, minus-strand primer-binding site. The U5-leader stem and the U5-IR stem are as shown. The acceptor stem and the TψC stem-loops of the $tRNA^{Trp}$ have been unwound to permit interactions with the PBS and U5 RNA sequences. Also shown are non-base-paired regions between the secondary-structure elements referred to as the Upper Spacer and Lower Spacer, respectively. (B) Schematic representation of the reconstituted reverse transcription system. Template RNA was annealed to the RNA primers (as described in the text) and the LTR DNA primer in the presence of AMV RT as in Materials and Methods. The DNA products from primer-dependent reverse transcription were treated with RNase, and the newly synthesized DNAs were analyzed on a 8% denaturing polyacrylamide gel. The extension product from the wild-type (WT) RNA primer is 101 nucleotides long; the extension product from the LTR DNA primer is 40 nucleotides long. L, nucleotide sequence derived from the leader region of viral RNA. Symbols: ■, PBS; ▩, control DNA primer-annealing site.

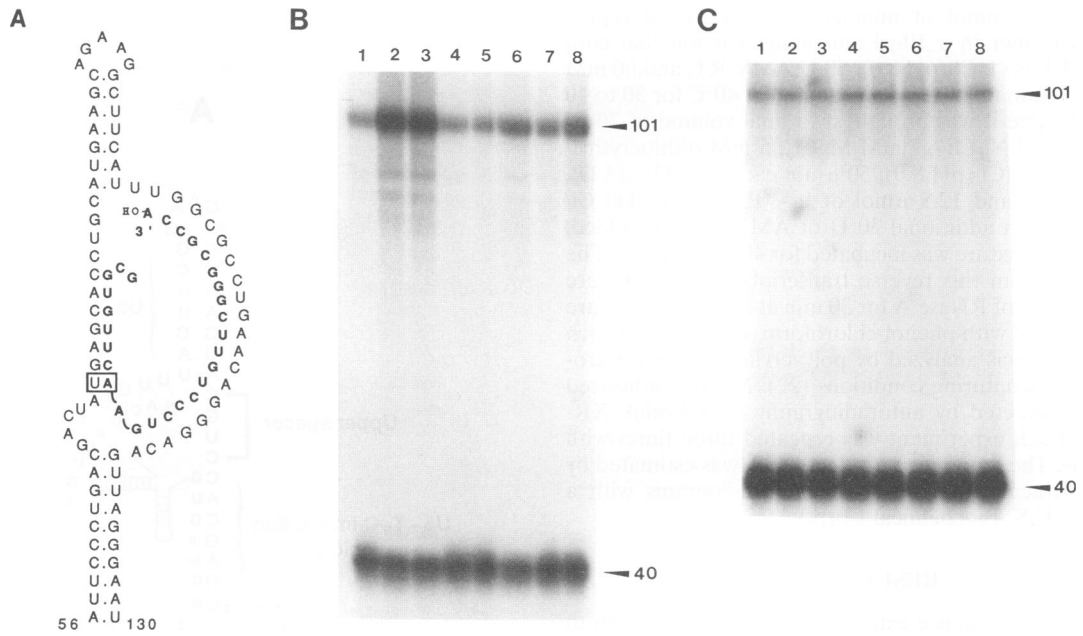

FIG. 2. (A) Schematic representation of the RSV Lys3 mosaic RNA annealed to the LysTrp28 primer. The RSV Lys3 RNA contains an additional mutation at position 71, a C-to-T nucleotide change (boxed). (B) DNA products from reconstituted reverse transcription reactions in which wild-type RSV RNA (lanes 1 to 3) and RSV Lys3 RNA (lanes 4 to 8) were used as template RNAs with RNA primers Trp20 (lane 1), Trp28 (lane 2), Trp55 (lane 3), Lys20 (lane 4), Lys28 (lane 5), LysTrp28 (lane 6), Lys55 (lane 7), and LysTrp55 (lane 8). AMV RT was used during the annealing and extension reactions. The cDNA synthesized from the PBS RNA primer is 101 nucleotides long, whereas that from the LTR DNA primer is 40 nucleotides long. (C) RNA templates and primers were as in part B except that the HIV-1 RT was used during the annealing and extension reactions.

The effect of size of the RNA primer on reconstitution of reverse transcription with wild-type RSV RNA is presented in Fig. 2. As previously shown, the primer (a 28-mer) containing the TψC sequence of $tRNA^{Trp}$, which can anneal to U5 RNA in addition to the PBS, is approximately fivefold more efficient than a primer (20-mer) which can anneal only to the PBS (Fig. 2B, lanes 1 and 2). The use of a longer primer (55-mer), which includes the anticodon stem-loop, does not increase the efficiency of synthesis of the 101-deoxynucleotide cDNA over that observed with the 28-mer (lane 3). Differences in cDNA synthesis observed with these various primers do not reflect differences in primer annealing to template RNA (data not shown). On the basis of these results, all subsequent experiments were performed with the 28-mer primer unless otherwise indicated.

Mutations in the upper spacer region. The series of insertions, deletions, and base substitutions in the upper spacer analyzed are summarized in Fig. 3A. The effect of placing a 4-nucleotide insertion mutation into the upper spacer is shown in Fig. 3B (lane 3). For reference, the wild-type RNA (Fig. 3B, lane 1) and the RNA template containing the 3-nucleotide deletion (lane 3), which we previously characterized (1), were included in this analysis. The cDNA product expected for a 4-nucleotide insertion is 105 deoxynucleotides, whereas that for the 3-nucleotide deletion is 98 deoxynucleotides. As shown in Fig. 3B (lane 3), the presence of the 4-nucleotide insertion caused a two- to threefold decrease in the amount of the 105-deoxynucleotide cDNA synthesized compared with the wild-type 101-deoxynucleotide cDNA (lane 1). The magnitude of the decrease was similar to that observed for the 3-nucleotide deletion (lane 2). In contrast, a 4-nucleotide substitution in the same region of the RNA had no detectable effect on 101-deoxynucleotide cDNA synthesis (lane 4). As a control, the amount of the DNA-primed 40-deoxynucleotide cDNA product was found to be the same for all reactions. The size of this product does not change since the LTR DNA primer anneals 5′ to the U5 RNA mutations. These results indicate that the sequences in the upper spacer are not recognized in a sequence-specific manner during cDNA synthesis. However, changes in the size of upper spacer induced by either deletions or insertions of sequence reduce the efficiency of reverse transcription.

Mutations in the lower spacer region. A similar set of mutations were introduced into the lower spacer (Fig. 4A) and analyzed for cDNA synthesis in vitro (Fig. 4B). These mutations include a 2-nucleotide (I2LS) and a 4-nucleotide (I4LS) insertion, a 3-nucleotide deletion (Δ3LS), and a 5-nucleotide substitution (S5LS). The sizes of the cDNA products expected for I2LS, I4LS, Δ3LS, and S5LS mutations are 103, 105, 98, and 101 deoxynucleotides, respectively. As in the upper spacer, perturbations that alter the length of the lower spacer result in defects in strong stop DNA synthesis (Fig. 4B). The magnitudes of the decrease for the 2-nucleotide insertion and the 3-nucleotide deletion were similar to one another (Fig. 4B, lanes 2 and 3) and to that observed for the comparable insertion and deletion in the upper spacer (Fig. 3B, lanes 2 and 3). The 4-nucleotide insertion (Fig. 4B, lane 4), however, had a much stronger inhibitory effect on reverse transcription than did the 2-nucleotide insertion (lane 3). This result suggests that there is a relationship between the size of the insertion in the lower spacer and the severity of the effect on reverse transcription. In contrast to these results, the 5-nucleotide substitution had no apparent effect on initiation of reverse transcription (lane 5).

Mutations in the PBS. The results shown in Fig. 3 and 4 suggest that the RT-RNA initiation complex has a defined

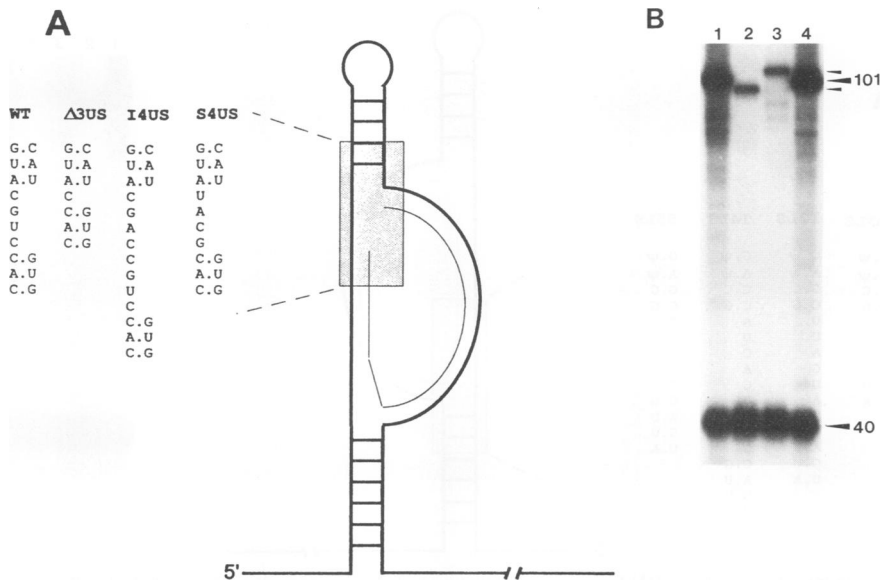

FIG. 3. (A) Schematic representation of the RNA secondary structures of the 5′ noncoding RSV RNA showing mutations to the upper spacer sequence. The thin line represents the tRNA primer, and the thick line represents the RNA template. The expanded region shows the sequence of the mutations which delete 3 nucleotides (Δ3US), insert 4 nucleotides (I4US), or substitute 3 nucleotides (S3US) of template RNA. WT, wild type. (B) cDNA products from reconstituted reverse transcription reactions with wild-type (lane 1), Δ3US (lane 2), I4US (lane 3), or S3US (lane 4) template RNA. The Trp28 RNA primer was used in all reactions. Extension from the wild-type and the substituted RNA templates from the RNA primer is 101 nucleotides long. Extension from the template RNA containing the 3-nucleotide deletion or the 4-nucleotide substitution are accordingly smaller or larger (small arrowheads). The extension product from the LTR DNA primer is 40 nucleotides long.

tertiary structure, which, when distorted, results in a decrease in the efficiency of initiation of reverse transcription. If this is correct, one would predict that defects in reverse transcription could be caused by distortions of structure introduced into the PBS region. To test this hypothesis, a 3-nucleotide deletion was placed at the 3′ end of the PBS (Δ3PBS), reducing its ability to anneal to the $tRNA^{Trp}$ from 18 to 15 complementary nucleotides. As a control, a 3-nucleotide substitution was also placed in the same region (S3PBS). Both mutations are depicted in Fig. 5A. As with the deletions in the upper and lower spacers, the deletion in the PBS reduced the in vitro efficiency of initiation of reverse transcription two- to threefold (Fig. 5C, lane 2) compared with the wild type (lane 1). In contrast, the 3-nucleotide substitution had little or no effect on reverse transcription (Fig. 5C, lanes 3 and 4). This was the case whether the wild-type 28-mer primer (lane 3) or a mutant primer, which contains a 3-nucleotide substitution permitting it to fully anneal to the modified PBS (lane 4), was used. Note that PBS mutations will not change the size of the cDNA products. These results indicate that the loss in the base pairing at the 3′ end of the PBS is not responsible for the loss in efficiency of reverse transcription. This conclusion is further substantiated by combining the 3-nucleotide PBS deletion mutation with the 2-nucleotide substitution in the lower spacer (Fig. 5B). The combination of these two mutations, which would restore the primer RNA base pairing to the wild-type level, did not restore the efficient reverse transcription (Fig. 5C, lanes 5 and 6). Thus the reverse transcription complex is sensitive to perturbations that alter the relative orientation of any of the duplex regions.

Substitution of the RSV PBS for that from HIV-1. We and others have shown that RT is required to efficiently anneal a primer tRNA to the PBS (1, 4, 24). In an effort to further elucidate the role of RT in this process, a mosaic RNA containing U5 and leader sequences from RSV were combined with the $tRNA_3^{Lys}$ PBS of HIV-1 (Fig. 2A). A set of primers which correspond to the 3′-terminal 55 (Lys55), 28 (Lys28), or 20 (Lys20) nucleotides, respectively, of $tRNA_3^{Lys}$ were synthesized as T7 transcripts. The name of the primer is indicated in parentheses. Additionally, two mosaic primers of 28 and 55 nucleotides, in which the TψC loop sequences in Lys28 and Lys55 were replaced with those from $tRNA^{Trp}$, were synthesized. These hybrid primers are called LysTrp28 and LysTrp55, respectively. Each of the primer-template combinations was examined in vitro by using either AMV RT (Fig. 2B) or HIV-1 RT (Fig. 2C) during the annealing and extension reactions. In contrast to what is observed with the RSV template and RSV primer (Fig. 2B, lanes 2 and 3), none of the mosaic RNAs supported efficient initiation of reverse transcription (lanes 4 to 8). There was a difference in the small amount of DNA synthesis observed with the various $tRNA_3^{Lys}$ primers. Those that could anneal to the RSV TψC interaction region were slightly more efficient than those that could not (compare lanes 6 and 8 with lanes 4, 5, and 7). Taken together, these results suggest that there must be sequence-specific interactions between RT and the duplex regions of the primer-template combination in the initiation complex which are lost when the sequence of the PBS is changed.

If the PBS-primer duplex is the primary site for specific RT interaction, one might reconstitute the efficient initiation of reverse transcription by addition of the HIV-1 RT to the in vitro reaction. However, even though equivalent units of HIV-1 RT were added, the amount of DNA synthesis observed with the same primer-templates was small (Fig. 2C). This suggests that efficient initiation of reverse transcription requires multiple specific interactions between the PBS and the U5 RNA TψC region. The latter is consistent with the higher activity observed with the LysTrp28 primer and AMV RT (Fig. 2B, lane 6).

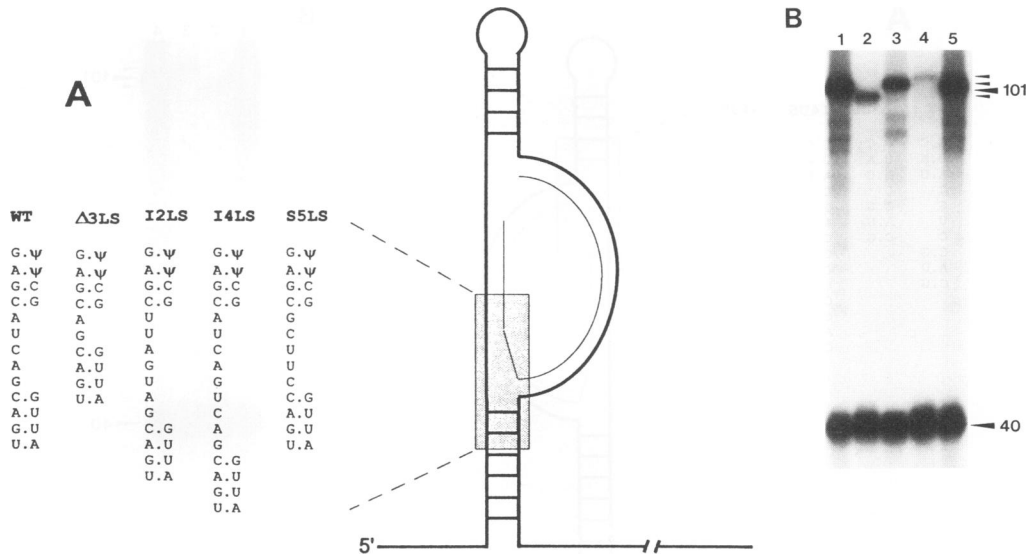

FIG. 4. (A) Schematic representation of the RNA secondary structures of the 5′ noncoding RSV RNA showing mutations to the lower spacer sequence. Primer and template are as indicated in the legend to Fig. 3. The expanded region shows the sequence of the mutations which delete 3 nucleotides (Δ3LS), insert 2 or 4 nucleotides (I2LS and I4LS, respectively), or substitute 5 nucleotides (S5LS) of the template RNA. WT, wild type. (B) cDNA products from reconstituted reverse transcription reactions with wild-type (lane 1), Δ3LS (lane 2), I2LS (lane 3), I4LS (lane 4), or S5LS (lane 5) template RNA. The Trp28 RNA primer was used in all reactions. Extension from the wild-type and the substituted RNA templates from the RNA primer is 101 nucleotides long. Extension from the template RNA containing the 3-nucleotide deletion or the 2- or 4-nucleotide substitution are accordingly smaller or larger (small arrowheads). The extension product from the LTR DNA primer is 40 nucleotides long.

DISCUSSION

The initiation complex formed with avian sarcoma/leukosis virus U5 RNA and the $tRNA^{Trp}$ primer in the presence of avian sarcoma/leukosis virus RT appears to be complex. As well as the interaction of the PBS with the tRNA primer, there are a series of secondary structures, U5-IR stem, U5-leader stem, and U5 RNA-TψC interaction region, all required for efficient initiation of reverse transcription (1, 7, 8). Moreover, these duplex structural elements appear to be maintained in a specific orientation relative to one another for optimum initiation of reverse transcription. Insertions or deletions of similar size placed in spacer regions between these elements result in similar reductions in cDNA synthesis. Furthermore, the magnitude of the reduction in cDNA synthesis is related to the size of the insertion. In contrast, nucleotide substitutions introduced into the same regions have no effect on cDNA synthesis. The simplest explanation for these data is that insertions or deletions in the spacer regions structurally distort the RT-RNA complex, reducing the efficiency of initiation of reverse transcription. This interpretation is further supported by the observation that a deletion in the 3′ end of the PBS also causes a reduction in reverse transcription. This is in contrast to a 3-base substitution, which has little effect, irrespective of whether the primer can anneal to the substituted bases. The changes in synthesis observed here are due to defects in initiation rather than DNA elongation synthesis since we have previously shown that mutations that disrupt the duplex regions result in a decrease in incorporation of the first deoxynucleotide, dATP (1, 7, 8). For increased sensitivity, however, we have monitored strong stop DNA synthesis in these experiments.

Characteristic 5′ RNA structures and primer interactions are conserved in other type C and D retroviruses, suggesting that these viruses use a common mechanism for the initiation of reverse transcription (13). Direct evidence for the U5-IR and U5-leader stem structures, in the absence of primer, has recently been obtained by nuclease sensitivity studies of Baudin et al. (5) for murine leukemia virus and HIV-1 RNAs. Under these conditions, the PBS is annealed to U5 RNA as predicted from studies of RSV (1). The mutational studies of Rizvi and Panganiban (21) have led to similar conclusions for simian immunodeficiency virus RNA. Berkhout and Schoneveld (6) have extended the in vitro observations to HIV-2 RNA annealed to $tRNA_3^{Lys}$. In this case, both nuclease sensitivity studies and structural interference with DNA synthesis initiated from a primer annealed downstream 3′ to the PBS signal the presence of the U5-IR stem, U5-leader stem, and U5 RNA-TψC interaction regions.

The deletion mutagenesis studies of Rhim et al. (20) also provide evidence for defined spatial relationships between the structural elements of the PBS region of HIV-1 (HXB2 strain) RNA. These investigators deleted between 9 and 12 nucleotides from the 3′ end of the PBS. Upon passage of viruses containing these deletions in cell cultures, wild-type growth revertants were rapidly recovered. These stable revertants were cloned and sequenced. Of 43 clones examined, all but 1 possessed 18-nucleotide PBS regions, suggesting that there was a selection against an alteration in the size of the PBS. In contrast to the conservation of size, several stable revertants possessed sequence changes at two or three positions within the PBS that would reduce base pairing to the primer. The fact that these are stable revertants indicates that the loss of base pairing does not interfere with reverse transcription. These results are similar to those described here for the S3PBS mutation, which had wild-type levels of initiation of reverse transcription when using an 18-nucleotide PBS but with reduced base pairing to the primer. Smaller but functional PBS regions, 10 or 12 nucleotides long, are also found in retroelements such as *Drosophila* copia (12). Together these results suggest that changes in the normal size of the PBS cause

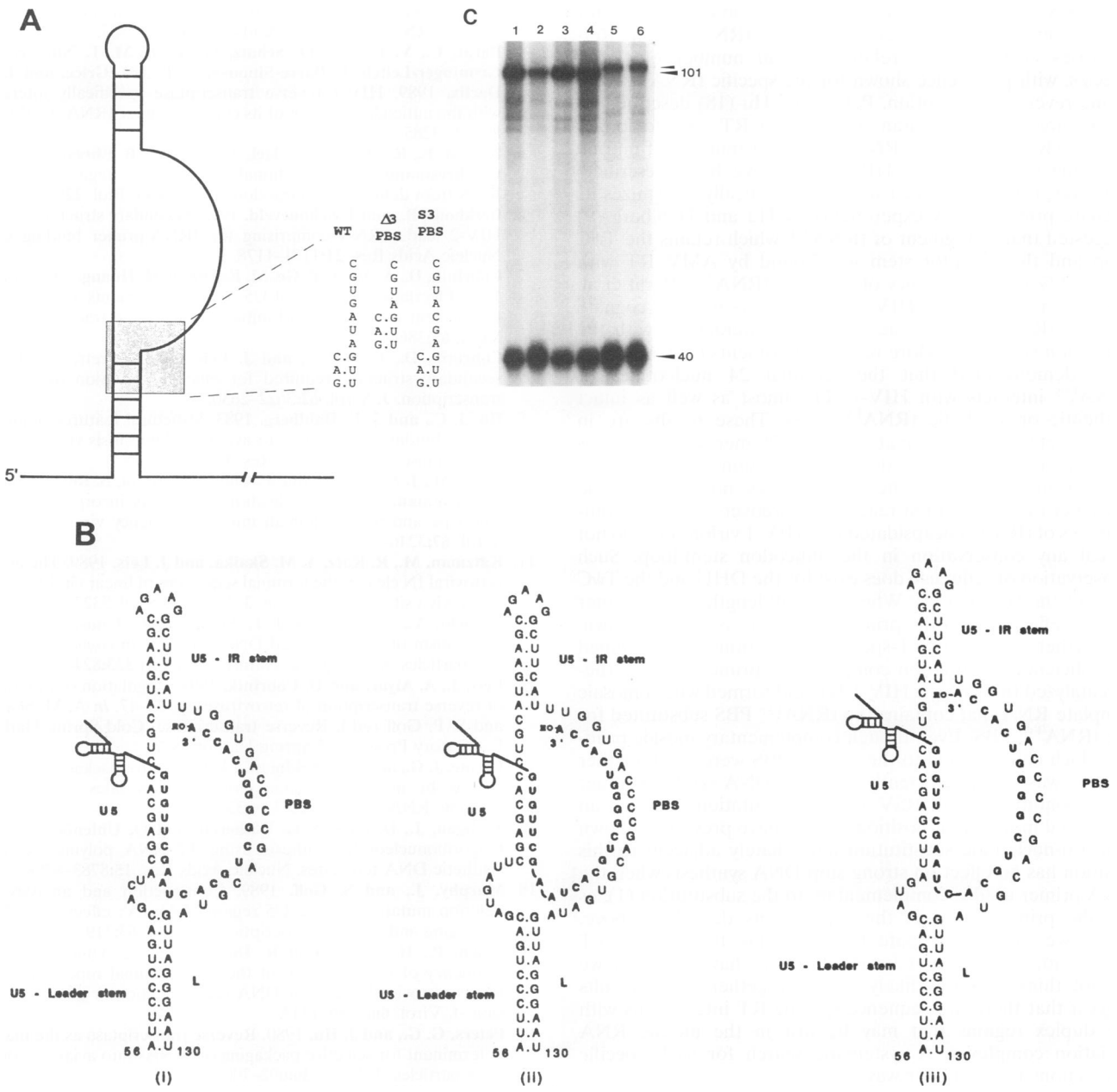

FIG. 5. (A) Schematic representation of the RNA secondary structures of the 5′ noncoding RSV RNA showing mutations to the PBS sequence. The expanded region shows the sequence of the mutations which delete 3 nucleotides (Δ3PBS) or substitute 3 nucleotides (S3PBS) of the PBS. WT, wild type. (B) Putative RNA secondary structure of the Δ3PBS (i), I2LS (ii), and combined Δ3PBS-I2LS (iii) mutant RNAs showing potential base pairing with the Trp28 primer. All other notations are as in the legend to Fig. 1A. (C) cDNA products from reconstituted reverse transcription reactions with wild-type (lane 1), Δ3PBS (lane 2), S3PBS (lanes 3 and 4), I2LS (lane 5), and Δ3PBS-I2LS (lane 6) template RNAs. The Trp28 RNA primer was used in all reactions except for that shown in lane 4, where the Trp28S3 primer was used. Size of cDNA products were as in Fig. 3 and 4.

structural distortions in the RT-RNA initiation complex that decrease the efficiency of reverse transcription.

In some revertant clones obtained by Rhim et al. (20), deletions 3′ to the PBS of 6 to 11 nucleotides were detected. These deletions remove duplications of sequences that appear to be present in the HXB2 (20) and MAL strains (5) of HIV-1. The duplication of these sequences causes the formation of an additional stem-loop at the 3′ end of the PBS. Removal of this duplication will not interfere with the formation of the U5-leader stem predicted by the structural studies of Baudin et al. (5). Spatial relationships are also supported by the studies of Olsen et al. (17), which introduced small insertions into the lower spacer of a reticuloendotheliosis virus-based vector. The altered lower spacers rapidly reverted to wild-type size when the vector was introduced into cell cultures. In a few revertants, small variations from the wild-type sequence were found.

In addition to the structural requirements for formation of the initiation complex, we have begun to look for specific

RT-RNA interactions. There is considerable evidence for specific interactions between RT and the tRNA primer. Retroviruses encapsidate a relatively small number of tRNA species, with preference shown for the specific tRNA used to prime reverse transcription. Peters and Hu (18) described an avian sarcoma virus mutant that lacked RT and no longer specifically encapsidated $tRNA^{Trp}$. Similar mutants of murine leukemia virus (14) and HIV-1 (10) have been described. However, it is not yet clear how RT specifically recognizes its cognate primer. Early experiments by Hu and Dahlberg (9) suggested that a fragment of $tRNA^{Trp}$ which retains the TψC loop and the acceptor stem was bound by AMV RT with about 80% of the efficiency of an intact $tRNA^{Trp}$. Barat et al. (4) have reported that HIV-1 RT interacts with its cognate primer ($tRNA_3^{Lys}$) by virtue of specific interactions with the anticodon stem-loop. More recent experiments (24), however, have demonstrated that the terminal 24 nucleotides of $tRNA_3^{Lys}$ interacts with HIV-1 RT almost as well as intact authentic or synthetic $tRNA_3^{Lys}$ does. These results are in agreement with our observations that a 28-mer RNA primer is much more efficiently used during initiation than a 20-mer RNA primer and that the 55-mer does not improve the efficiency of the initiation reaction. Moreover, sequence comparisons of tRNAs encapsidated into HIV-1 virions (10) do not reveal any conservation in the anticodon stem-loop. Such conservation of sequence does exist for the DHU and the TψC arms of the tRNAs (2a). Whether a full-length tRNA primer is more efficient than the primers used here is not yet known. To further search for RT-specific interactions, we examined the efficiency of initiation complexes to prime DNA synthesis catalyzed by AMV or HIV-1 RT and formed with a mosaic template RNA that contains the $tRNA_3^{Lys}$ PBS substituted for the $tRNA^{Trp}$ PBS. Even though complementary mosaic primers which could anneal to the altered PBS were used, neither enzyme was able to efficiently catalyze cDNA synthesis under these conditions. The RSV Lys3PBS mutation contains an additional mutation at position 71. We have previously shown that a 6-nucleotide substitution immediately adjacent to this position has no effect on strong stop DNA synthesis when the RNA primer used is complementary to the substitution (1), as are the primers used in the experiments described above. While we cannot eliminate the possibility that the C-to-T substitution is responsible for the effects we have observed, we do not think that this likely. Taken together, these results suggest that there are sequence-specific RT interactions with the duplex regions that may be lost in the mosaic RNA initiation complexes. A systematic search for such specific interactions is now under way.

ACKNOWLEDGMENTS

We thank Ed Houts, Molecular Genetic Resources, for AMV RT and Stuart LeGrice, Case Western Reserve University School of Medicine, for HIV-1 RT.

This work was supported in part by Public Health Service grant CA38046 from the National Cancer Institute and by Cancer Research Center grant P30 CA 43703.

REFERENCES

1. **Aiyar, A., D. Cobrinik, Z. Ge, H.-J. Kung, and J. Leis.** 1992. Interaction between retroviral U5 RNA and the TψC loop of the $tRNA^{Trp}$ primer is required for efficient initiation of reverse transcription. J. Virol. **66**:2464–2472.
2. **Aiyar, A., and J. Leis.** 1993. Modification of the megaprimer method of PCR mutagenesis: improved amplification of the final product. BioTechniques **14**:366–367.

2a. **Aiyar, A., and J. Leis.** Unpublished data.

3. **Barat, C., S. LeGrice, and J.-L. Darlix.** 1991. Interaction of the HIV-1 reverse transcriptase with a synthetic form of its replication primer, $tRNA^{Lys,3}$. Nucleic Acids Res. **19**:751–757.
4. **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.
5. **Baudin, F., R. Marquet, C. Isel, J.-L. Darlix, B. Ehresmann, and C. Ehresmann.** 1993. Functional sites in the 5′ region of HIV-1 RNA from defined structural domains. J. Mol. Biol. **229**:382–397.
6. **Berkhout, B., and I. Schoneveld.** 1993. Secondary structure of the HIV-2 leader RNA comprising the tRNA-primer binding site. Nucleic Acids Res. **21**:1171–1178.
7. **Cobrinik, D., A. Aiyar, Z. Ge, M. Katzman, H. Huang, and J. Leis.** 1991. Overlapping retroviral U5 sequence elements are required for efficient integration and initiation of reverse transcription. J. Virol. **65**:3864–3872.
8. **Cobrinik, D., L. Soskey, and J. Leis.** 1988. A retroviral RNA secondary structure required for efficient initiation of reverse transcription. J. Virol. **62**:3622–2630.
9. **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.
10. **Jiang, M., J. Mak, A. Ladha, E. Cohen, M. Klein, B. Rovinski, and L. Kleinman.** 1993. Identification of tRNAs incorporated into wild-type and mutant human immunodeficiency virus type 1. J. Virol. **67**:3246–3253.
11. **Katzman, M., R. Katz, A. M. Skalka, and J. Leis.** 1989. The avian retroviral IN cleaves the terminal sequences of linear viral DNA at the in vivo sites of integration. J. Virol. **63**:5319–5327.
12. **Kikuchi, Y., Y. Ando, and T. Shiba.** 1986. Unusual priming mechanism of RNA-directed DNA synthesis in *copia* retrovirus-like particles in *Drosophila*. Nature (London) **323**:824–826.
13. **Leis, J., A. Aiyar, and D. Cobrinik.** 1993. Regulation of initiation of reverse transcription of retroviruses, p. 33–47. *In* A. M. Skalka and S. P. Goff (ed.), Reverse transcriptase. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
14. **Levine, J. G., and J. G. Seidman.** 1979. Selective packaging of host tRNAs by murine leukemia virus particles does not require genomic RNA. J. Virol. **29**:328–335.
15. **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.
16. **Murphy, J., and S. Goff.** 1989. Construction and analysis of deletion mutations in the U5 region of MuLV: effects on RNA packaging and reverse transcription. J. Virol. **63**:319–327.
17. **Olsen, P., H. Temin, and R. Dornburg.** 1992. Unusually high frequency of reconstitution of the long terminal repeats in U3-minus retrovirus vectors by DNA recombination or gene conversion. J. Virol. **66**:1336–1343.
18. **Peters, G. G., and J. Hu.** 1980. Reverse transcriptase as the major determinant for selective packaging of tRNAs into avian sarcoma virus particles. J. Virol. **36**:692–700.
19. **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.
20. **Rhim, H., J. Park, and C. D. Marrow.** 1991. Deletions in the $tRNA^{Lys}$ primer-binding site of human immunodeficiency virus type 1 identify essential regions for reverse transcription. J. Virol. **65**:4555–4564.
21. **Rizvi, T. A., and A. T. Panganiban.** 1993. Simian immunodeficiency virus RNA is efficiently encapsidated by human immunodeficiency virus type 1 particles. J. Virol. **67**:2681–2688.
22. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
23. **Schwartz, D. E., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell **22**:853–869.
24. **Weiss, S., B. Konig, H. J. Muller, H. Seidel, and R. S. Goody.** 1992. Synthetic human $tRNA^{Lys,3}$ and natural bovine $tRNA^{Lys,3}$ interact with HIV-1 reverse transcriptase and serve as primers for retroviral cDNA synthesis. Gene **111**:183–197.