JOURNAL OF VIROLOGY, Aug. 1992, p. 5127–5132
0022-538X/92/085127-06$02.00/0

Vol. 66, No. 8

Bipartite Signal for Read-Through Suppression in Murine Leukemia Virus mRNA: an Eight-Nucleotide Purine-Rich Sequence Immediately Downstream of the *gag* Termination Codon Followed by an RNA Pseudoknot

YA-XIONG FENG,[1] HANG YUAN,[2] ALAN REIN,[1] AND JUDITH G. LEVIN[2]*

Laboratory of Molecular Virology and Carcinogenesis, ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, Frederick, Maryland 21702,[1] and Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, Bethesda, Maryland 20892[2]

Received 21 January 1992/Accepted 22 April 1992

The *pol* gene of murine leukemia virus and other mammalian type C retroviruses is expressed by read-through suppression of an in-frame UAG codon which separates the *gag* and *pol* coding regions. In this study, we have analyzed the sequence requirements for read-through suppression by placing different portions of wild-type and mutant viral sequences from the *gag-pol* junction between reporter genes and testing transcripts of these constructs for suppression in reticulocyte lysates. We find that the read-through signal is contained within the first 57 nucleotides on the 3′ side of the UAG codon. Our results indicate that the identities of six conserved bases in the eight-nucleotide, purine-rich sequence immediately downstream of the UAG codon are critical for suppression, as is the existence of a pseudoknot structure spanning the next 49 nucleotides. Thus, read-through suppression depends on a complex, bipartite signal in the mRNA.

In murine leukemia virus (MuLV) and other mammalian type C retroviruses, the *gag* and *pol* coding regions are separated by an in-frame UAG termination codon (8, 13, 19, 28, 32). Read-through suppression of the UAG codon gives rise to a Gag-Pol fusion protein (17, 22, 24, 34), which is subsequently cleaved to give the Pol proteins, i.e., protease, reverse transcriptase, and integrase. The efficiency of suppression is approximately 5% (17).

Studies of this unusual form of translational control have been performed almost entirely with MuLVs, and the mechanism is only partly understood (reviewed in references 12 and 20). The suppression is mediated by a glutamine tRNA (34), which appears to be a normal cellular tRNA (10, 23). Panganiban (23) found that the signals for suppression during translation of MuLV mRNA are completely *cis* acting, with no requirement for viral proteins, and are completely contained within a 300-nucleotide region extending from 100 nucleotides 5′ to 200 nucleotides 3′ of the UAG codon. In a previous study, we observed significant levels of in vitro suppression with a construct whose only viral sequences were 2 codons from the 5′ side of the UAG and 19 codons from the 3′ side (9). Although the efficiency of suppression could not be quantitated in these experiments, the results suggested that the signals for suppression are largely, if not entirely, found within sequences downstream of the UAG codon. In addition, it was shown that the suppression signals do not depend on the presence of a particular termination codon at the *gag-pol* junction. Thus, UAA and UGA codons, like UAG, could be efficiently suppressed in reticulocyte lysates (11) and in infected cells (11, 18). To gain a clearer understanding of the mechanism of read-through suppression, it is obviously of importance to define these signals more precisely.

One possibility is that the signal for suppression is contained in a secondary structure in the viral RNA rather than in specific sequences. In an early proposal, it was suggested that the suppression signal consists of a stem-loop structure in which the UAG codon is in the loop and sequences from *gag* and *pol* are base paired to form the stem (13, 28). However, this putative structure is not found in the mRNAs of other mammalian type C retroviruses (23, 31), and the hypothesis was not supported by tests with viral mutants (14).

In searching for a common structure which might function in suppression, ten Dam et al. (31) noted that the sequence within the first 19 codons on the 3′ side of the UAG codon of MuLV and several other mammalian type C retroviruses can be folded into a pseudoknot structure. Pseudoknots are stem-loop structures in which bases in the loop are paired with bases downstream of the stem to form a second base-paired region (25–27) (see below). Possible involvement of a pseudoknot structure in translational suppression was first deduced from mutational analysis of ribosomal frameshifting in Rous sarcoma virus (16) and was later shown to be required for frameshift suppression in a coronavirus, infectious bronchitis virus (4, 5), a yeast virus (7), and a retrovirus, mouse mammary tumor virus (6). While our work was in progress, two groups reported attempts to test the hypothesis that a putative pseudoknot structure is a crucial element in the suppression signal for Moloney MuLV (Mo-MuLV) and reached opposite conclusions. In one study, Honigman et al. (14) found that mutations disrupting the first stem of the pseudoknot prevent suppression; however, compensatory mutations, which change the sequence but should restore the base-paired structure, failed to restore suppression in vitro. In contrast, the experiments of Wills et al. (33) were strongly supportive of the pseudoknot hypothesis. Their results showed that mutations disrupting either stem of the pseudoknot prevented (or significantly impaired) suppression, and compensatory mutations restored suppression. (The respective mutations made in the two studies were somewhat different.)

* Corresponding author.

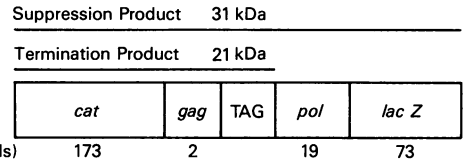

FIG. 1. Schematic diagram of insert sequences in construct with suppressor activity. The basic plasmid, designated pMG_2P_{19}, was a construct containing the *cat* gene up to a 3′ *Nco*I site, Mo-MuLV sequences (28) consisting of 2 codons from the 3′ end of *gag*, TAG, and 19 codons from the 5′ end of *pol*, and vector sequences including those encoding the α peptide of *lacZ*. The construct was made by ligating three fragments: an *Xba*I-*Nco*I *cat* fragment (generated by digestion of the pCAT-Basic plasmid [Promega Biotec] with *Xba*I and *Nco*I), a double-stranded synthetic 70-nucleotide Mo-MuLV oligonucleotide with *Nco*I and *Eco*RI overhangs, and the large *Eco*RI-*Xba*I fragment of pGEM11Zf(−) (Promega Biotec). One strand of the oligonucleotide had the sequence CATG, followed by the Mo-MuLV sequence (28) from bases 2229 to 2294 except that the nucleotide at position 2255 was a G (21) rather than a C as originally reported (28); the other strand began with AATT and continued with the complement of the viral sequence from 2294 to 2229. Synthetic oligonucleotides were provided by Marilyn Powers (PRI/DynCorp, NCI-Frederick Cancer Research and Development Center, Frederick, Md.) or were purchased from Midland Certified Reagent Company (Midland, Tex.). pMG_2P_{19} as well as all other plasmids used in this study were sequenced over a region extending from the last 80 nucleotides of the *cat* fragment to the first 60 nucleotides of the *lacZ* segment by using a Sequenase kit (United States Biochemical Corp.). Translation of the mRNA up to the *gag* termination codon gives a 21-kDa product (termination product); translation beyond the *gag-pol* junction to a termination codon six nucleotides from the end of the region encoding the α peptide of *lacZ* gives a 31-kDa product (suppression product). The diagram is not drawn to scale.

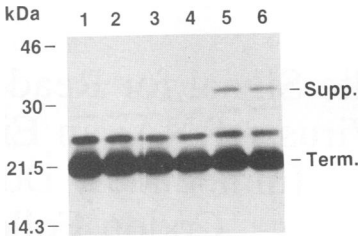

FIG. 2. Minimal nucleotide sequence required for read-through suppression at the Mo-MuLV *gag-pol* junction. The cloned DNAs were digested with *Nde*I [which cleaves in a region downstream of *lacZ* in pGEM11Zf(−)] and were transcribed with SP6 RNA polymerase (Promega Biotec), using the conditions supplied by the company. The resulting mRNAs were translated in rabbit reticulocyte lysates supplemented with 50 μg of calf liver tRNA (Promega Biotec) per ml in the presence of [^{35}S]methionine (1,000 Ci/mmol; New England Nuclear Corp.) as described previously (10). Reactions were incubated for 20 min at 37°C. The protein products were immunoprecipitated with a polyclonal anti-CAT serum (5 Prime-3 Prime, Inc.) as described previously (15) and were analyzed in SDS–15% polyacrylamide gels. The relative intensities of the bands were estimated by scanning the gels with a laser densitometer. The intensity of the suppression product (31-kDa band; see Fig. 1) obtained with the basic plasmid was designated 100%; the lower limit of detection was approximately 5% of this level. Since the suppression product has only one more methionine residue than does the 21-kDa termination product, no correction was made for methionine content when the efficiency of suppression was estimated. The products shown in lanes 1 to 5 were each synthesized from constructs containing two codons from *gag*. The number of *pol* codons was as follows: lane 1, 2 codons, nt +1 to +6 (this construct was designated pMG_2P_2); lane 2, 5 codons, nt +1 to +15; lane 3, 10 codons, nt +1 to +30; lane 4, 15 codons, nt +1 to +45; and lane 5, 19 codons, nt +1 to +57 (pMG_2P_{19}). Lane 6 shows products from a construct containing a deletion of the penultimate *gag* codon (nt −6 to −4), a mutation of the terminal *gag* codon GAC (nt −3 to −1) to GUG, and 19 codons from *pol*. In the constructs with 2, 10, and 15 *pol* codons, the last nucleotide in the sense strand of the viral sequence was changed to a G residue to provide an *Eco*RI overhang at the 3′ end of the oligonucleotide. Numbering of nucleotide positions is as shown in Fig. 3A. The positions of coelectrophoresed molecular size markers are indicated on the left. Supp., suppression product; Term., termination product.

In an effort to clarify the role of the proposed pseudoknot structure and of other sequences or structures near the termination codon in suppression, we have analyzed a series of mutant and wild-type constructs containing various amounts of viral sequence near the Mo-MuLV *gag* termination codon. Our results, like those of Wills et al. (33), support the idea that the pseudoknot is an important part of the suppression signal. In addition, however, our data demonstrate that the signal is complex and includes the eight bases preceding the pseudoknot, immediately downstream of the UAG codon.

Location of read-through signals in viral mRNA. As a first step in identifying *cis*-acting signals within viral mRNA which control suppression, it was important to determine the minimal mRNA sequence which is required. Our goal was to isolate for mutational analysis a viral sequence which could act autonomously to induce suppression in a nonviral context. In this study, we made constructs which permit measurement of the protein product terminating at the UAG codon (termination product) as well as the longer product which is synthesized by suppression of the UAG codon (suppression product). All of the plasmids contained the SP6 promoter, 173 codons from the 5′ end of the chloramphenicol acetyltransferase (*cat*) gene, various amounts of Mo-MuLV sequence from the *gag-pol* junction, and sequences encoding the α peptide of *lacZ* (e.g., the construct shown in Fig. 1). The cloned DNAs were transcribed with SP6 RNA polymerase, the resulting mRNAs were translated in rabbit reticulocyte lysates, and the protein products were immunoprecipitated with a polyclonal anti-CAT serum prior to electrophoresis in sodium dodecyl sulfate (SDS)–15% polyacrylamide gels (10; also see the legend to Fig. 2).

Figure 2 shows that translation of mRNAs containing 2 codons from the 3′ end of *gag* and either 2, 5, 10, or 15 codons from the 5′ end of *pol* (lanes 1 to 4, respectively) gave only the 21-kDa termination product; however, a construct (pMG_2P_{19}; Fig. 1) having the 2 *gag* codons plus 19 codons from *pol* (lane 5) gave both the termination product and a 31-kDa suppression product. (The additional band present in all lanes could be the result of initiation at codons other than the AUG codon at the beginning of the *cat* gene.) The ratio of the termination product to suppression product was about 70:1. In subsequent experiments, pMG_2P_{19} was used as a positive control; for purposes of comparison, the amount of read-through obtained with this construct was designated 100%. To test whether the *gag* sequence is part of the suppression signal, we made a construct with a deletion of the penultimate *gag* codon and a mutation in the last *gag* codon (see the legend to Fig. 2) but having 19 codons from the 5′ end of *pol* (lane 6). This construct showed a high level of read-through suppression (approximately 90% of that seen with pMG_2P_{19}).

The amount of suppression obtained with pMG_2P_{19} was

similar to that obtained earlier with a 5-kb Mo-MuLV *gag-pol* construct containing all of *gag* and two-thirds of *pol* (10) and with the ratio of $Pr65^{gag}$ to $Pr200^{gag\text{-}pol}$ observed in vivo (17). We infer from this observation that the sequence in pMG_2P_{19} contains the entire signal for suppression and is functional within the context of flanking nonviral mRNA sequences. The results are in accord with previous findings (9, 33) and demonstrate that more than 45 and perhaps as many as 57 *pol* nucleotides immediately 3′ of the UAG termination codon are both necessary and sufficient for read-through suppression. Sequences from the *gag* region do not appear to be required.

Possible requirement for an RNA pseudoknot structure in read-through suppression. The fact that such a large sequence is required for suppression is striking and suggests the possibility that a large secondary (or tertiary) structure in the mRNA is part of the read-through signal. As one way of searching for sequences or structures which might be important in suppression, we compared the Mo-MuLV sequence present in pMG_2P_{19} (Fig. 1) with the corresponding sequences of baboon endogenous virus (BaEV) (19) and spleen necrosis virus (SNV) (32), two other members of the mammalian type C retrovirus family using read-through suppression for *pol* gene expression. The results of this comparison are shown in Fig. 3A. Nucleotides in *pol* downstream of the UAG codon are designated by positive numbers, and nucleotides in *gag* are designated by negative numbers; conserved nucleotides are underlined. Inspection of Fig. 3A shows that a large fraction (approximately 60%) of the nucleotides on the 3′ side of the UAG codon are identical in the three viruses.

In addition to this sequence relatedness, there is conservation of predicted RNA structure. As pointed out by ten Dam et al. (31), the *pol* sequences near the UAG codon in a number of mammalian type C retroviral genomes can be folded into a pseudoknot structure; Mo-MuLV and BaEV pseudoknots are shown in Fig. 3B and C, respectively. In both viral mRNAs, one loop portion of the structure contains a stretch of six C residues, beginning 19 nucleotides downstream of the UAG codon, which can interact with a run of six G residues in the 18th and 19th codons 3′ of the UAG codon, to form the second base-paired region (stem 2) of the pseudoknot. These runs of C's and G's are highly conserved (Fig. 3A). The 18 nucleotides between the two stems form loop 2 of the structure. Loop 1 consists of only two nucleotides, A (nucleotide +17 [nt +17]) and G (nt +18).

Experiments were undertaken to test the possibility that sequences in stems 1 and 2 of the proposed Mo-MuLV pseudoknot structure (Fig. 3B) are required for read-through (Fig. 4). If these base-paired structures are essential, changes which destroy the base pairing should prevent read-through, but compensatory changes (which change the sequence but restore the base pairing) should restore activity. Figure 4A shows that mutation either of six bases on the left side of stem 1 (nt +10 to +15 [GGUCAG]) (lane 3) or of six bases (excluding the bulged A) on the right side of stem 1 (nt +26 to +28 [CUG] and +30 to +32 [ACC]) (lane 4) to the complementary bases completely abolished read-through. In contrast, translation of an mRNA with the compensatory changes, i.e., mutation of both sides of stem 1 to the respective complementary bases (lane 5), restored activity. The level of suppression was about 60% of that obtained with pMG_2P_{19} (lane 2). Substitution of the BaEV stem-loop, including the bases in stem 1 of the proposed pseudoknot (Fig. 3C), for the corresponding nucleotides in Mo-MuLV (nt +9 to +33) (lane 6) also resulted in significant

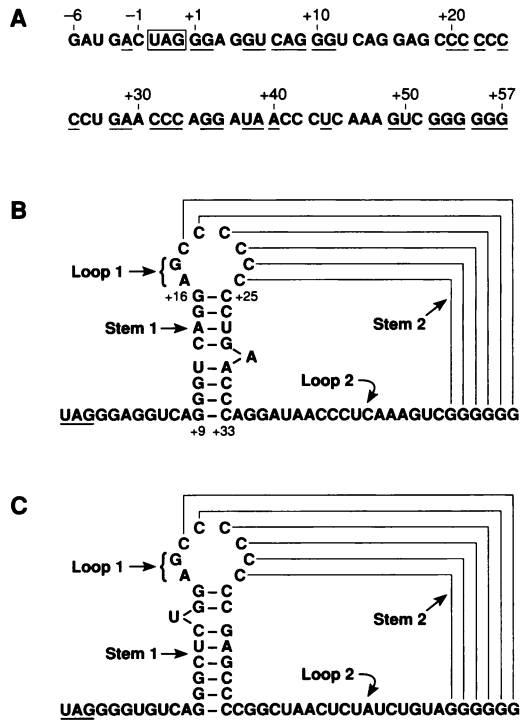

FIG. 3. Nucleotide sequence around the UAG termination codon in Mo-MuLV and proposed pseudoknot structures in Mo-MuLV and BaEV mRNAs. (A) Nucleotide sequence 6 bases 5′ of the UAG termination codon and 57 bases 3′ of the UAG codon in Mo-MuLV (28). The UAG codon is boxed; nucleotide positions 5′ of the UAG are designated by negative numbers, and those 3′ of the UAG are designated by positive numbers. Nucleotides which are identical in Mo-MuLV (28), BaEV (19), and SNV (32) are underlined. Proposed pseudoknot structures and the preceding eight nucleotides are shown for (B) Mo-MuLV and (C) BaEV mRNAs. Stem 1 includes nt +9 to +16 and +25 to +33; stem 2 includes nt +19 to +24 and +52 to +57; loop 1 includes nt +17 and +18; and loop 2 includes nt +34 to +51.

read-through (approximately 50% compared with suppression by pMG_2P_{19}).

Tests of constructs with changes in stem 2 (Fig. 4B) showed that when five of the six G residues (nt +52 to +56) were changed to C's (lane 3), no suppression product could be detected. Mutation of five of the six C's (nt +20 to +24) to G's (lane 4) resulted in a considerable reduction in read-through (fivefold decrease compared with pMG_2P_{19} [lane 2]). When the compensatory changes were made (lane 5), the level of suppression was about 50% of that observed with pMG_2P_{19}.

These results indicate that maintenance of the base-paired structures in stems 1 and 2 is important for suppression and thus provide strong evidence for the participation of a pseudoknot structure.

To determine whether the activity associated with the pseudoknot structure can be preserved if the length of loop 2 (Fig. 3B) is shortened, two constructs with deletions in this region were assayed for suppressor activity (Fig. 4C). In one, the AGG codon (nt +34 to +36) immediately 3′ of stem 1 was deleted (lane 3), and suppression was reduced about fivefold compared with that observed with pMG_2P_{19} (lane 2). Deletion of the GUC codon (nt +49 to +51) immediately 5′ of the stretch of G residues (lane 4) abolished read-through

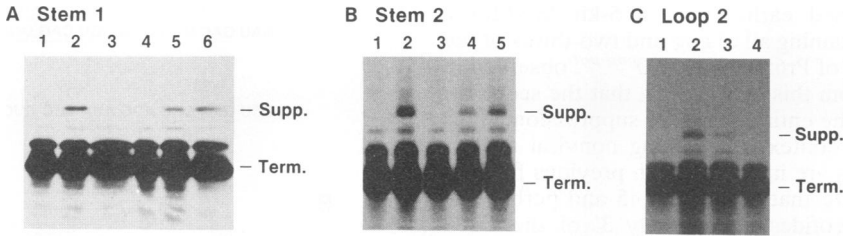

FIG. 4. Mutation of nucleotides in stems 1 and 2 and loop 2 of proposed Mo-MuLV pseudoknot structure. In all panels, lanes 1 and 2 are the protein products from pMG_2P_2 (negative control) and pMG_2P_{19} (positive control), respectively. (A) Stem 1. The mutations in the constructs used for lanes 3 to 6 were as follows: lane 3, nt +10 to +15 changed to complementary bases; lane 4, nt +26 to +28 and +30 to +32 changed to complementary bases; lane 5, nt +10 to +15, nt +26 to +28, and nt +30 to +32 changed to the complementary bases (compensatory mutation); and lane 6, substitution of the BaEV stem-loop containing the bases in stem 1 of the proposed pseudoknot structure (see Fig. 3C) for the corresponding nucleotides in Mo-MuLV (nt +9 to +33). (B) Stem 2. The mutations in the constructs used for lanes 3 to 5 were as follows: lane 3, nt +52 to +56, consisting of a run of G's changed to C's; lane 4, nt +20 to +24, consisting of a run of C's changed to G's; and lane 5, nt +20 to +24 and nt +52 to +56 changed to the complementary bases (compensatory mutation). (C) Deletion of codons in loop 2 of proposed Mo-MuLV pseudoknot. The constructs used for lanes 3 and 4 contain deletions of an AGG codon (nt +34 to +36) and a GUC codon (nt +49 to +51), respectively. Numbering of nucleotide positions is as shown in Fig. 3A. Supp., suppression product; Term., termination product.

completely. These data show that some reduction in the size of loop 2 is compatible with a low level of suppression but that part or all of the GUC codon (nt +49 to +51) is a crucial element in the read-through signal.

Mutational analysis of nucleotides immediately downstream of the UAG termination codon. There are eight nucleotides between the UAG codon at the *gag-pol* junction and the beginning of stem 1 of the pseudoknot structure (Fig. 3B and C), six of which are identical in Mo-MuLV, BaEV, and SNV (Fig. 3A). Mutational analysis of these six bases in MuLV constructs was undertaken to assess their functional importance in read-through suppression (Fig. 5; Table 1 summarizes the results). Changing the first nucleotide (G) downstream of the UAG codon to either A (Fig. 5A, lane 3) or C (Fig. 5A, lane 4) prevented read-through. Similarly, no suppression product could be detected with changes of CA (nt +7 and +8) to GU (Fig. 5B, lane 3), G (nt +2) to C (Fig. 5B, lane 5), GG (nt +1 and +2) to AA (Fig. 5B, lane 6), G (nt +5) to C (Fig. 5C, lane 3), and G (nt +5) to A (Fig. 5C, lane 4) or with deletion of the second *pol* codon, GGU (Fig. 5C, lane 6). The only mutations which gave any suppression were G (nt +2) to A (Fig. 5B, lane 4) or U (nt +6) to A (Fig. 5C, lane 5), and in both cases, the level of read-through was very low (approximately 20 and 15%, respectively, of that observed with pMG_2P_{19}, [lane 2]).

In summary, every alteration which was tested in the six conserved bases in the eight-nucleotide spacer region either eliminated or significantly inhibited suppression. Thus, this sequence constitutes an additional component of the suppression signal.

Conclusions. In this study, we have investigated the nature of the *cis*-acting signal for MuLV read-through suppression. Using an in vitro assay to measure suppression (10), we found that the signal is complex and consists of two distinct elements: a highly conserved, eight-nucleotide purine-rich sequence immediately downstream of the UAG *gag* termination codon, followed by a pseudoknot structure beginning at nt +9 (Fig. 3). The signal is contained within the first 19 *pol* codons, and these 57 nucleotides are both necessary and sufficient for suppression (Fig. 2) (9, 33). The data also show that *gag* sequences upstream of the terminator are not required. A small stimulation (50%) of read-through in the presence of three *gag* codons and a twofold inhibition with 12 *gag* codons have been reported (33).

Our evidence indicating that a pseudoknot structure is required for suppression is based on mutational analysis of stems 1 and 2. Mutations which destroy base pairing in the stems prevent or significantly impair suppression, whereas compensatory mutations which restore base pairing restore

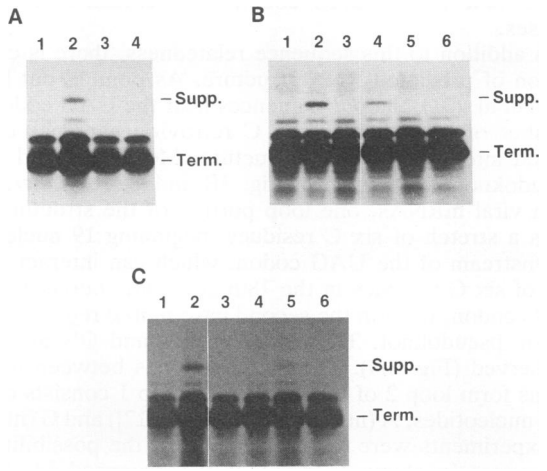

FIG. 5. Mutation of nucleotides within the eight-nucleotide sequence immediately downstream of the UAG codon. In all panels, lanes 1 and 2 are the protein products from pMG_2P_2 (negative control) and pMG_2P_{19} (positive control), respectively. (A) Mutation of the first nucleotide (G) downstream of the UAG codon (nt +1). The mutations in the constructs used for lanes 3 and 4 were a change from G to A and from G to C, respectively. (B) Mutation of CA (nt +7 and +8), the second G downstream of the UAG codon, and both the first and second G residues downstream of the UAG codon. The mutations in the constructs used for lanes 3 to 6 were as follows: lane 3, change of CA (nt +7, +8) to GU; lane 4, change of G (nt +2) to A; lane 5, change of G (nt +2) to C; and lane 6, change of both G residues (nt +1 and +2) to AA. (C) Mutation of the fifth (G) and sixth (U) nucleotides and deletion of the second codon (GGU) downstream of the UAG codon. The mutations in the constructs used for lanes 3 to 6 were as follows: lane 3, change of G (nt +5) to C; lane 4, change of G (nt +5) to A; lane 5, change of U (nt +6) to A; and lane 6, deletion of GGU (nt +4 to +6). Numbering of nucleotide positions is as shown in Fig. 3A. Supp., suppression product; Term., termination product.

TABLE 1. Summary of mutational analysis of the eight-nucleotide sequence preceding the pseudoknot

Mutation[a]	Suppression (% of control)[b]
G1 → A	<5
G1 → C	<5
G2 → A	20
G2 → C	<5
G1, G2 → AA	<5
G5 → A	<5
G5 → C	<5
U6 → A	15
C7, A8 → GU	<5
ΔG4, G5, U6	<5

[a] The bases which were mutated are numbered as shown in Fig. 3A.
[b] The control plasmid was pMG_2P_{19} (see Fig. 1).

suppression (Fig. 4). These results demonstrate that it is the mRNA structure, i.e., maintenance of the base-paired character of the stems, and not the identity of the bases within the stems which is important. This conclusion is further strengthened by the observation that a high level of suppression was obtained when stem 1 of the BaEV pseudoknot (Fig. 3C) was substituted for the corresponding region of Mo-MuLV (Fig. 4A), even though the sequences in the two viruses are not identical (Fig. 3).

In the course of the analysis of stem 2, we found that mutation of five of the six C residues to G's significantly reduced suppression but did not abolish it completely (Fig. 4B). Wills et al. (33) made a similar observation with a stem 2 mutant in which four C's were changed to G's. The reason why mutation of the C residues causes only partial inhibition of read-through is not clear. It is possible that having a base-paired stem 2 in a pseudoknot structure facilitates suppression but is not absolutely required. Alternatively, the structure of the RNA in stem 2 may have some unusual features. Finally, in our mutant, in which nt +20 to +24 are G's (rather than C's), an alternate pseudoknot structure may form, with base pairing between these G's and nt +41 to +45 (CCCUC) (Fig. 3B). The resulting pseudoknot would have a shorter loop 2 than does the wild-type structure.

In addition to the pseudoknot, the importance of the immediate context downstream of the Mo-MuLV termination codon, i.e., the bases between the UAG codon and the pseudoknot (Fig. 3A), has also been investigated. In earlier work, Honigman et al. (14) noted that the first five bases following the Mo-MuLV UAG codon are purines and that similar sequences are common to the genomes of other mammalian type C retroviruses. From results showing that mutation of the first four nucleotides GGAG to ACGC abolished suppression, these investigators concluded that the stretch of five purines is important for MuLV suppression. In this study, we have focused on the eight nucleotides immediately downstream of the UAG codon which precede the pseudoknot. Deletion of either the first (33) or second (Fig. 5C) codons 3′ of the terminator abolished read-through. In a detailed analysis involving single- or double-base changes in the six conserved nucleotides, we also found that every mutation tested either eliminated or significantly reduced suppression (Fig. 5; Table 1). These results demonstrate that the identities of the conserved nucleotides upstream of the pseudoknot structure are crucial for suppression and constitute an important element of the read-through signal.

As noted above, two of the eight bases between the UAG codon and the pseudoknot are not conserved among the mammalian type C retroviruses (AG, nt +3 and +4; Fig. 3A). The fact that suppression occurs with each of these viruses suggests that the identities of these two bases are not crucial for suppression. On the other hand, these bases might interact with other elements in the respective suppression signals; for example, it is possible that the AG sequence is required for suppression in Mo-MuLV because of an interaction with downstream Mo-MuLV sequences, while the corresponding GU sequence in BaEV might interact with BaEV-specific sequences to form an essential part of the BaEV suppression signal. However, the results presented in Fig. 4A, lane 6, argue against this proposal, since suppression is occurring here with the Mo-MuLV sequence at nt +3 and +4 but with the BaEV sequence at nt +9 to +33. From these considerations, it seems likely that the identities of bases at positions +3 and +4 are not important for suppression.

It is of interest to compare the MuLV results with those obtained with other eukaryotic systems in which translational suppression is used. Tobacco mosaic virus has an in-frame UAG codon within the replicase gene which is normally suppressed by a tyrosine tRNA (1, 3). Relative to MuLV, however, the signal for suppression in tobacco mosaic virus is very simple and comprises only the six nucleotides downstream of the termination codon (29). In a quite different case, certain mammalian proteins are known to contain selenocysteine at a position corresponding to an in-frame UGA codon in the message (30). Translation of one such protein, type 1 iodothyronine 5′-deiodinase, was recently shown to require the presence of a 200-nucleotide segment located in the 3′ untranslated region of the mRNA, more than 1 kb downstream of the UGA codon; this segment apparently represents a complex stem-loop structure (2). The influence of nucleotides immediately adjacent to the UGA codon was not investigated.

Among the viruses which exhibit ribosomal frameshift suppression, three (infectious bronchitis virus [4, 5], L-A yeast virus [7], and mouse mammary tumor virus [6]) have a *cis*-acting signal remarkably similar to that of MuLV. The signal is bipartite and consists of a heptanucleotide sequence at the frameshift site and a pseudoknot structure beginning six or seven nucleotides to the 3′ side of the frameshift site. Interestingly, the entire mouse mammary tumor virus pseudoknot can be substituted for the Mo-MuLV pseudoknot, but the level of read-through is extremely low (33). The inefficiency of suppression seen with the chimeric construct suggests that there is some specificity in the structural requirements for the MuLV pseudoknot. Nevertheless, the observation that the different pseudoknot structures are at least to some extent interchangeable, and can serve as elements of signals for frameshifting or read-through suppression, suggests that they may perform a common function in the two types of suppression. The role of the pseudoknot may simply be to cause ribosomes to pause, thereby facilitating ribosome slippage during translation of the heptanucleotide sequence at a frameshift site (16) or, together with the proximal eight-nucleotide component of the read-through signal, allowing the suppressor tRNA to successfully compete with release factor at the read-through site.

The mechanism for recognition of the read-through signal is not yet known. If ribosomal pausing does indeed occur, it would presumably be mediated by protein factors (ribosomal and/or nonribosomal soluble factors) which interact with the RNA signal. In addition, there may be a helicase activity required in a subsequent step to permit continued movement

of the ribosomes and translation of sequences in the pseudoknot structure. Clearly, further investigation is needed to elucidate the precise reactions which are involved in the process of read-through suppression.

We thank Robert Gorelick for assistance with densitometry; Nancy Rice for fruitful discussions; Dolph Hatfield for helpful comments on the manuscript; Ian Brierley and Stephen Inglis, Alik Honigman and Amos Panet, Norma Wills and Ray Gesteland, and Mario Chamorro and Harold Varmus for communicating data prior to publication; and Carol Shawver and Cheri Rhoderick for preparation of the manuscript.

This research was sponsored in part by the National Cancer Institute under contract NO1-CO-74101 with ABL.

REFERENCES

1. **Beier, H., M. Barciszewska, G. Krupp, R. Mitnacht, and H. J. Gross.** 1984. UAG readthrough during TMV RNA translation: isolation and sequence of two $tRNAs^{Tyr}$ with suppressor activity from tobacco plants. EMBO J. **3**:351–356.
2. **Berry, M. J., L. Banu, Y. Chen, S. J. Mandel, J. D. Kieffer, J. W. Harney, and P. R. Larsen.** 1991. Recognition of UGA as a selenocysteine codon in Type I deiodinase requires sequences in the 3′ untranslated region. Nature (London) **353**:273–276.
3. **Bienz, M., and E. Kubli.** 1981. Wild-type $tRNA_G^{Tyr}$ reads the TMV RNA stop codon, but Q base-modified $tRNA_Q^{Tyr}$ does not. Nature (London) **294**:188–190.
4. **Brierley, I., P. Digard, and S. C. Inglis.** 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell **57**:537–547.
5. **Brierley, I., N. J. Rolley, A. J. Jenner, and S. C. Inglis.** 1991. Mutational analysis of the RNA pseudoknot component of a coronavirus ribosomal frameshifting signal. J. Mol. Biol. **220**: 889–902.
6. **Chamorro, M., N. Parkin, and H. E. Varmus.** 1992. An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger RNA. Proc. Natl. Acad. Sci. USA **89**:713–717.
7. **Dinman, J. D., T. Icho, and R. B. Wickner.** 1991. A −1 ribosomal frameshift in a double-stranded RNA virus of yeast forms a *gag-pol* fusion protein. Proc. Natl. Acad. Sci. USA **88**:174–178.
8. **Etzerodt, M., T. Mikkelsen, F. S. Pedersen, N. O. Kjeldgaard, and P. Jorgensen.** 1984. The nucleotide sequence of the Akv murine leukemia virus genome. Virology **134**:196–207.
9. **Feng, Y.-X., T. D. Copeland, S. Oroszlan, A. Rein, and J. G. Levin.** 1990. Identification of amino acids inserted during suppression of UAA and UGA termination codons at the *gag-pol* junction of Moloney murine leukemia virus. Proc. Natl. Acad. Sci. USA **87**:8860–8863.
10. **Feng, Y.-X., D. L. Hatfield, A. Rein, and J. G. Levin.** 1989. Translational readthrough of the murine leukemia virus *gag* gene amber codon does not require virus-induced alteration of tRNA. J. Virol. **63**:2405–2410.
11. **Feng, Y.-X., J. G. Levin, D. L. Hatfield, T. S. Schaefer, R. J. Gorelick, and A. Rein.** 1989. Suppression of UAA and UGA termination codons in mutant murine leukemia viruses. J. Virol. **63**:2870–2873.
12. **Hatfield, D. L., J. G. Levin, A. Rein, and S. Oroszlan.** 1992. Translational suppression in retroviral gene expression. Adv. Virus Res. **41**:193–239.
13. **Herr, W.** 1984. Nucleotide sequence of AKV murine leukemia virus. J. Virol. **49**:471–478.
14. **Honigman, A., D. Wolf, S. Yaish, H. Falk, and A. Panet.** 1991. *cis* acting RNA sequences control the Gag-Pol translation readthrough in murine leukemia virus. Virology **183**:313–319.
15. **Hu, S. C., D. L. Court, M. Zweig, and J. G. Levin.** 1986. Murine leukemia virus *pol* gene products: analysis with antisera generated against reverse transcriptase and endonuclease fusion proteins expressed in *Escherichia coli*. J. Virol. **60**:267–274.
16. **Jacks, T., H. D. Madhani, F. R. Masiarz, and H. E. Varmus.** 1988. Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region. Cell **55**:447–458.
17. **Jamjoom, G. A., R. B. Naso, and R. B. Arlinghaus.** 1977. Further characterization of intracellular precursor polyproteins of Rauscher leukemia virus. Virology **78**:11–34.
18. **Jones, D. S., F. Nemoto, Y. Kuchino, M. Masuda, H. Yoshikura, and S. Nishimura.** 1989. The effect of specific mutations at and around the *gag-pol* gene junction of Moloney murine leukaemia virus. Nucleic Acids Res. **17**:5933–5945.
19. **Kato, S., K. Matsuo, N. Nishimura, N. Takahashi, and T. Takano.** 1987. The entire nucleotide sequence of baboon endogenous virus DNA: a chimeric genome structure of murine type C and simian type D retroviruses. Jpn. J. Genet. **62**:127–137.
20. **Levin, J. G., D. L. Hatfield, S. Oroszlan, and A. Rein.** Mechanisms of translational suppression used in the biosynthesis of reverse transcriptase. *In* A. M. Skalka and S. Goff (ed.), Reverse transcriptase, in press. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Miller, A. D., and I. M. Verma.** 1984. Two base changes restore infectivity to a noninfectious molecular clone of Moloney murine leukemia virus (pMLV-1). J. Virol. **49**:214–222.
22. **Murphy, E. C., Jr., N. Wills, and R. B. Arlinghaus.** 1980. Suppression of murine retrovirus polypeptide termination: effect of amber suppressor tRNA on the cell-free translation of Rauscher murine leukemia virus, Moloney murine leukemia virus, and Moloney murine sarcoma virus 124 RNA. J. Virol. **34**:464–473.
23. **Panganiban, A. T.** 1988. Retroviral *gag* gene amber codon suppression is caused by an intrinsic *cis*-acting component of the viral mRNA. J. Virol. **62**:3574–3580.
24. **Philipson, L., P. Andersson, U. Olshevsky, R. Weinberg, D. Baltimore, and R. Gesteland.** 1978. Translation of MuLV and MSV RNAs in nuclease-treated reticulocyte extracts: enhancement of the *gag-pol* polypeptide with yeast suppressor tRNA. Cell **13**:189–199.
25. **Pleij, C. W. A.** 1990. Pseudoknots: a new motif in the RNA game. Trends Biochem. Sci. **15**:143–147.
26. **Pleij, C. W. A., K. Rietveld, and L. Bosch.** 1985. A new principle of RNA folding based on pseudoknotting. Nucleic Acids Res. **13**:1717–1731.
27. **Schimmel, P.** 1989. RNA pseudoknots that interact with components of the translation apparatus. Cell **58**:9–12.
28. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukaemia virus. Nature (London) **293**:543–548.
29. **Skuzeski, J. M., L. M. Nichols, R. F. Gesteland, and J. F. Atkins.** 1991. The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. J. Mol. Biol. **218**:365–373.
30. **Stadtman, T. C.** 1991. Biosynthesis and function of selenocysteine-containing enzymes. J. Biol. Chem. **266**:16257–16260.
31. **ten Dam, E. B., C. W. A. Pleij, and L. Bosch.** 1990. RNA pseudoknots: translational frameshifting and readthrough on viral RNAs. Virus Genes **4**:121–136.
32. **Weaver, T. A., K. J. Talbot, and A. T. Panganiban.** 1990. Spleen necrosis virus *gag* polyprotein is necessary for particle assembly and release but not for proteolytic processing. J. Virol. **64**:2642–2652.
33. **Wills, N. M., R. F. Gesteland, and J. F. Atkins.** 1991. Evidence that a downstream pseudoknot is required for translational read-through of the Moloney murine leukemia virus *gag* stop codon. Proc. Natl. Acad. Sci. USA **88**:6991–6995.
34. **Yoshinaka, Y., I. Katoh, T. D. Copeland, and S. Oroszlan.** 1985. Murine leukemia virus protease is encoded by the *gag-pol* gene and is synthesized through suppression of an amber termination codon. Proc. Natl. Acad. Sci. USA **82**:1618–1622.