A Unique Sequence in Murine Leukemia Virus Long Terminal Repeat Functions as a Termination Signal for Transcription in Escherichia coli

ALIK HONIGMAN* and AMOS PANET

Departments of Microbiological Chemistry1 and Virology,2 The Hebrew University—Hadassah Medical School, Jerusalem, Israel

Received 1 June 1982/Accepted 27 August 1982

Moloney murine leukemia virus DNA fragments were subcloned into two plasmids downstream of active promoters to detect termination sequences for transcription initiated at these promoters. Two of the viral DNA fragments (2,050 and 595 base pairs) were able to block transcription when inserted in one orientation but not when inserted in the other. These two viral DNA fragments contained long terminal repeat sequences. Deletion of 84 base pairs from the U-3 and R regions in the long terminal repeat removed the transcription termination activity.

Transcription of retrovirus 35S RNA appears to be initiated within the left-hand long terminal repeat (LTR) and terminated at the right-hand LTR of the integrated proviral DNA (24). Sequence analysis of the LTR region of murine leukemia virus (MuLV) and murine sarcoma virus has revealed a TATA-like sequence (AATAAA) which appears to be part of the promoter region (1, 5, 19, 22, 26). This sequence is located 24 nucleotides upstream of the site coding for the 5’ end of the 35S viral RNA. Another adenine-plus-thymine (A+T)-rich sequence in the LTR (AATAAAA), 16 nucleotides upstream from the site representing the 3’ end of the viral RNA, is the putative polyadenylic acid addition signal (1, 5). Although a large body of evidence is available on the DNA structures required for the initiation of RNA transcription, very little is known about transcription termination in eucaryotic systems. Studies on the regulation of procaryotic gene expression have indicated that termination occurs at unique sequences which may form stem and loop structures. Models have been put forward to explain the function of such structures in transcription termination in Escherichia coli (20, 21).

Several studies have indicated that some promoters of eucaryotic genes are recognized by the RNA polymerase of E. coli for the initiation of transcription (8, 9, 16). To determine whether specific DNA sequences of Moloney MuLV (M-MuLV) cause termination of transcription in a procaryotic system, we introduced the viral DNA fragments into two types of plasmids which were constructed to select for transcription termination sequences. One plasmid, pKG1900 (15), was constructed from part of pBR322 and two DNA fragments from the E. coli gal operon containing the gal promoter and the galK gene (Fig. 1). When this plasmid is introduced into an E. coli strain defective in the expression of the galK gene, it converts the Gal− bacteria (white colonies on MacConkey galactose plates) to Gal+ (red colonies on MacConkey galactose plates; Table 1).

The plasmid pBR-MuLV, which was used in this work as a source of M-MuLV DNA, contains a complete copy of M-MuLV unique sequences and two tandem copies of the LTR region. The viral DNA portion of the plasmid was cut into four fragments with XmaI and ligated to pKG1900, which was cleaved with the restriction enzyme Avai. After transformation of E. coli C600 and E. coli N100 cells, both GalK− (15), the cells were plated on LB agar containing ampicillin (100 μg/ml). Bacterial clones carrying viral DNA sequences were recognized by hybridization to 32P-labeled cDNA from M-MuLV (17, 18) by the colony hybridization technique (7). Each of the bacterial clones that hybridized to the M-MuLV cDNA was tested on MacConkey galactose agar plates containing ampicillin (100 μg/ml), and their plasmids were analyzed by restriction enzyme mapping. All white colonies on MacConkey galactose plates carried hybrid plasmids composed of pKG1900 and one of two viral DNA fragments: the 2,050-base-pair (bp) fragment A, which includes 472 bp of the adjacent LTR, or the 595-bp fragment B, which includes the entire LTR.
FIG. 1. Schematic illustration of the complete M-MuLV DNA and the construction of hybrid plasmids. The heavy line on each circle represents the pBR322 portion in each hybrid plasmid. For detailed description of each plasmid vector, see references 13 and 15. The source of M-MuLV DNA was pBR-M-MuLV (clone 1), originated from λ c1387, a gift from D. Linemeyer. The numbers in parentheses at each restriction endonuclease site represent the position of the cleavage site relative to the 5' end of the viral RNA, as published by Shinnick et al. (22). The two adjacent tandem LTRs (white and black boxes on the straight line) are drawn as arranged in the plasmid pBR-MuLV. Each of the viral fragments A and B were inserted separately and in both possible orientations (O and R) in pKO1 and pKG1900. Fragment C was inserted in both orientations in pHA10.
sequence in a permuted order (114 bp usually located at the 3' end of the LTR are situated in this fragment at the 5' end [Fig. 1]).

Restriction enzyme analysis of plasmids purified from GalK⁻ bacterial clones (about 100 white colonies were tested [data not shown]) indicated that the two viral fragments containing the LTR sequences were inserted in the right-hand (R) orientation, i.e., in the polarity to give viral plus-strand transcript initiated at the gal promotor (Fig. 1). Of 100 colonies which gave a red color on MacConkey galactose plates, 80 carried one of these two viral DNA fragments in the opposite (O) orientation. The remaining 20 colonies carried the 328-bp fragment located within the unique viral sequences (6213 in Fig. 1). Viral DNA fragments were also identified by hybridization to viral [³²P]cDNA by the technique of Southern (23) (data not shown).

For more precise localization of the region responsible for the transcription termination in the GalK⁻ clones, plasmid pKU3 (Table 1) was isolated, and 84 bp were removed from the LTR region. The deletion was achieved by digestion with the restriction enzyme FnuDII. The FnuDII viral DNA fragment, containing 390 nucleotides from the LTR region and 1,251 nucleotides of unique viral sequences, was purified from low-temperature-gelling agarose after fractionation by electrophoresis of the pKU3 digest. The purified DNA fragment was ligated into the Smal site of pKG1900. Before transfection of E. coli C600 (galK⁻) cells, the ligation mixture was recut with SmaI to linearize any remaining original plasmid sequences. Clones containing the FnuDII M-MuLV DNA fragments were identified by colony hybridization to M-MuLV cDNA and analyzed by restriction enzyme mapping. Two of the clones, each carrying the FnuDII MuLV DNA fragment in a different orientation, pKF12 in the R and pKF4 in the O orientation, were further analyzed on MacConkey galactose plates. Both colonies were red on these plates (Table 1), indicating that the sequence responsible for the transcription termination in E. coli lies within the 84 bp located between the FnuDII and SmaI cleavage sites in the LTR region (Fig. 1 and 2). Since digestion with FnuDII removed 84 bp from the 3' end and 355 bp from the 5' end of fragment A (Fig. 1), the termination signal could therefore be at either end of fragment A. However, since deletion of 282 bp located between the FnuDII and SmaI sites at the 3' end of fragment A also removed the termination signal (pKP1, Table 1), the termination signal must be located within the 84 bp at the 3' end of the LTR. Moreover, galactokinase expression was blocked in pKL1, which carries viral fragment B (Fig. 1 and Table 1). Fragment B consists of only the permuted LTR, which indicates that termination occurs within the LTR region.

Efficiencies of termination in the different bacterial clones were compared by measuring

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevant properties</th>
<th>Color on MacConkey galactose plates</th>
<th>Survival rate at 40°C</th>
<th>Galactokinase relative activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>pHA10</td>
<td>Amp⁺, λ immune, kills at 40°C</td>
<td>Red 1</td>
<td>3 x 10⁻⁵</td>
<td>1</td>
</tr>
<tr>
<td>pHM6</td>
<td>pHA10 plus MuLV BglII fragment C, Kil⁻ at 40°C, orientation R</td>
<td>Red 2</td>
<td>2 x 10⁻¹</td>
<td>0.02</td>
</tr>
<tr>
<td>pHM18</td>
<td>pHA10, MuLV BglII fragment C, Kil⁻ at 40°C orientation O</td>
<td>Red 3</td>
<td>3 x 10⁻⁴</td>
<td>0.07</td>
</tr>
<tr>
<td>pKG1900</td>
<td>Amp⁺, galK, gal promoter</td>
<td>White</td>
<td>10⁻⁴</td>
<td>0.03</td>
</tr>
<tr>
<td>pKU3</td>
<td>pKG1900, MuLV Smal fragment A, orientation R</td>
<td>White</td>
<td>10⁻⁴</td>
<td>0.09</td>
</tr>
<tr>
<td>pKU8</td>
<td>pKG1900, MuLV Smal fragment A, orientation O</td>
<td>White</td>
<td>10⁻⁴</td>
<td>0.07</td>
</tr>
<tr>
<td>pKL1</td>
<td>pKG1900, MuLV Smal fragment B, orientation R</td>
<td>White</td>
<td>10⁻⁴</td>
<td>0.07</td>
</tr>
<tr>
<td>pKL2</td>
<td>pKG1900, MuLV Smal fragment B, orientation O</td>
<td>White</td>
<td>10⁻⁴</td>
<td>0.07</td>
</tr>
<tr>
<td>pKF12</td>
<td>pKU3 with 439-bp deletion, orientation R</td>
<td>Red</td>
<td>1.15</td>
<td>0.77</td>
</tr>
<tr>
<td>pKF4</td>
<td>pKU3 with 439-bp deletion, orientation O</td>
<td>Red</td>
<td>1.15</td>
<td>0.77</td>
</tr>
<tr>
<td>pKP1</td>
<td>pKU3 with 282-bp deletion, orientation R</td>
<td>Red</td>
<td>1.15</td>
<td>0.77</td>
</tr>
<tr>
<td>pOU3</td>
<td>pKO1, MuLV Smal fragment A, orientation O</td>
<td>Pink</td>
<td>2.22</td>
<td>0.15</td>
</tr>
<tr>
<td>pOU5</td>
<td>pKO1, MuLV Smal fragment A, orientation O</td>
<td>White</td>
<td>0.02</td>
<td>0.16</td>
</tr>
<tr>
<td>pOL1</td>
<td>pKO1, MuLV Smal fragment B, orientation O</td>
<td>Pink</td>
<td>0.16</td>
<td>0.16</td>
</tr>
</tbody>
</table>

a Assay for λ kil gene expression was carried out with E. coli (recA supE44) carrying pHA10 or its derivatives. The survival rate was determined as the ratio of the number of surviving colonies at 40°C to the number surviving at 30°C (11, 13).

b Galactokinase assays were performed in E. coli C600 (galK recA) cells carrying the plasmids listed as described by McKenney et al. (15) (Fig. 3). The values for relative activity of galactokinase were calculated from the conversion values of galactose to galactose 1-phosphate for each clone relative to the values obtained for C600(pKG1900) after 10 min of enzyme reaction. The values for the last three hybrid plasmids represent transcription initiation activity.
enzyme activity by phosphorylation of [\(^{14}\)C]galactose (15). The enzymatic activity of galactokinase dropped dramatically (more than 10-fold) once the LTR was introduced into pKG1900 in the R orientation (pKU3) (Fig. 3 and Table 1). When 84 bp were deleted from the viral DNA (pKF12), the galactokinase activity was restored to the level found in pKG1900-carrying bacteria.

The DNA sequence responsible for reduction in enzyme activity is thus located within the 84 bp which map between -122 to -38 nucleotides from the 3' end of the viral 35S RNA within the LTR region (Fig. 2).

Expression of galactokinase activity in the plasmids carrying the LTR in the O orientation was reduced by 20% (pKU8 and pKL2, Table 1). The same small reduction in galactokinase activity was also observed after removal of the 84 nucleotides responsible for termination (pKF4, Table 1). This reduction in enzyme activity, therefore, appeared to be dependent on a DNA sequence within the viral DNA plus strand which is not complementary to the transcription termination signal in the viral DNA strand with the polarity to give viral plus-strand transcript.

To confirm the results described above, we used another plasmid (pHA10) to select for transcription termination signals in M-MuLV DNA. pHA10 was constructed from pBR322 and the EcoRI L\(_D\) fragment (Fig. 1). One of the genes carried by this DNA fragment codes for a temperature-sensitive \(\lambda\) repressor, c1857. Inactivation of the \(\lambda\) repressor at 40°C allows transcription to start at the \(\lambda p_L\) promoter, resulting in the expression of the \(\lambda\) kil gene, which kills the bacterial host. Introduction of a transcription termination signal between the \(\lambda p_L\) promoter and the \(\lambda\) kil gene prevents killing of the host at 40°C (11).

The 4.3-kilobase M-MuLV DNA fragment (cleavage product of \(Bg\)III restriction enzyme) was ligated into the BamHI cleavage site of

![FIG. 2. Terminator-like secondary structure in M-MuLV LTR. The RNA sequence is that described by Shinnick et al. (22). The calculated free energy of the possible secondary structure is \(\Delta G (25^\circ C) = -17.4\) kcal, and for the cloned LTR, starting at the \(Sm\)al cleavage site, it is \(\Delta G (25^\circ C) = -12\) kcal. The underlined nucleotides between position -98 to -92 point out the proposed LTR TATA box. The underlined region between -22 to -17 indicates the proposed polyadenylic acid addition signal. Arrows marked 1 and 2 represent restriction sites on the DNA plus strand for \(Sm\)al and \(Fnu\)DII, respectively.](http://jvi.asm.org/)

![FIG. 3. In vitro phosphorylation of [\(^{14}\)C]galactose by cell lysates of various clones. Enzymatic reactions were carried out as described by McKenney et al. (15). The reaction was initiated by the addition of tolenuen
ized bacteria, grown to a cell density of 0.6 at 650 nm. The reaction mixture (100 \(\mu\)l) was composed of 1 mM dithiothreitol, 3 mM NaF, 4 mM MgCl\(_2\), 50 mM Tris-hydrochloride (pH 7.9), 1.6 mM ATP, and 1 mM [\(^{14}\)C]-galactose (65,000 cpm/0.1 \(\mu\)mol). Portions (25 \(\mu\)l) were withdrawn for analysis of [\(^{14}\)C]galactose phosphorylation on DE81 filters.](http://jvi.asm.org/)
pHA10 (Fig. 1 and Table 1). The ampicillin-resistant colonies of E. coli C600 (recA-) grown at 30°C were hybridized to cDNA from M-MuLV; about 15% of these colonies hybridized to the cDNA. The colonies containing the viral DNA insert were tested at 30 and 40°C for heat resistance. All heat-resistant clones (80% of the total viral DNA-containing clones) acquired the viral DNA fragment in orientation R, whereas in the clones that remained heat sensitive (20%), the viral DNA was always in the O orientation. The enrichment for heat-resistant clones (80%) could be explained by the fact that transformation with pHA10 at 30°C resulted in the killing of transformed bacteria, in which the expression of the λ kil gene occurs before synthesis of the λ repressor (similar to λ infection). However, when the expression of the λ kil gene is blocked, all transformed cells survive (11, 13).

Two clones, each carrying the viral DNA insert in a different orientation (pHM6 in the R and pHM18 in the O orientation), were tested for growth at 40 and 30°C. In pHM6, expression of the λ kil gene was blocked up to 90%, whereas in pHM18, the killing rate was about the same as it was in pHA10 (Table 1).

The expression of a possible promoter located in the LTR of M-MuLV was tested in E. coli cells by inserting the viral DNA fragments A and B into the pKO1 plasmid (Fig. 1). Plasmid pKO1 was designed to identify promoter sequences. The galK gene was inserted into a pBR322 derivative missing the entire tetracycline resistance (Tc) region. No promoter precedes the galK gene, and thus galK is not transcribed unless a promoter is inserted into the plasmid (15). In plasmids pOU3 and pOL1, the viral LTR DNA fragments A and B, respectively, are inserted in the same orientation relative to the galK gene as is the promoting LTR relative to the viral genes in the integrated provirus (the left-end LTR). The LTR promoted expression of the galK gene at a level 10-fold above the levels obtained in bacteria containing either pKO1 or pOU5 (Table 1). In pOU5 the LTR fragment is inserted in the O orientation.

The experimental approach taken in this work is based upon the application of current models and research systems developed for the study of transcription control in procaryotes. As shown by McKenney et al. (15), control of transcription in these plasmids is accurately measured by the galactokinase assay. This research strategy was used to overcome the complexity involved in the genetic study of eucaryotic systems. However, as long as each of the genetic signals defined in E. coli is not tested in its natural background, the function attributed to the DNA sequence may not represent its role in eucaryotic cells.

In this work, we showed that the LTR DNA of M-MuLV can terminate transcription when introduced into E. coli cells. An 84-nucleotide segment of the LTR contained a DNA sequence between −38 and −122 nucleotides from the 3' end of the 35S RNA which appeared to signal transcription termination in E. coli cells.

Promotion of transcription from the LTR has been extensively studied in both eucaryotic (2, 4, 6) and procaryotic systems (8, 9, 16). At present, transcription termination in procaryotic cells is well understood, and numerous signals for termination of E. coli mRNAs have been sequenced (for a review, see reference 21). Little is known, however, about transcription termination in eucaryotic cells.

While it is not yet clear whether eucaryotic transcription signals are recognized as such in E. coli, several lines of evidence indicate some common features between eucaryotic and procaryotic signals. (i) Promoters from either source contain an A+T-rich region (TATA box). In fact, E. coli RNA polymerase can initiate transcription at the promoter site of Rous sarcoma virus LTR DNA (8, 9, 16). (ii) Stem and loop structures (hyphenated dyad symmetry) in the late mRNA of simian virus 40 have been implicated in transcription termination and attenuation during its lytic cycle in monkey cells (N. Hay, H. Scollnick-David, and Y. Aloni, Cell, in press).

One hypothesis suggests that a region adjacent to the 3' end of the 35S RNA of MuLV may also be involved in transcription termination (1). This sequence can be arranged to form a stem and loop structure (Fig. 2). Deletion of the DNA fragment which included the proposed stem and loop structure eliminated the termination function. Several characteristics of this stem and loop structure are notable. (i) The guanine plus cytosine content in this region was high (69%), and the calculated free energy contribution of the base-paired region and loops was −17.4 kcal/mol (see the legend to Fig. 2), which is consistent with a duplex form for this sequence. Even a shorter stem and loop starting at the Smal site (with calculated free energy of −12 kcal) still abolished transcription. (ii) Comparison of the nucleotide sequences of seven isolates of MuLV (1, 5, 19, 22, 26) indicated a strong conservation in the sequence forming the stem and loop, with only two substitutions in the 100-nucleotide sequence adjacent to the 3' end of the virus RNA. (iii) The distances from the stem and loop to the 3' terminus of the RNA (34 nucleotides) and to the presumed polyadenylation signal (12 nucleotides) were identical for all MuLV genomes analyzed. (iv) No stretch of polyuridine followed the stem and loop structure, as it does in some procaryotic transcription terminators (21). (v) The termination signal located in
the M-MuLV was resistant to the antitermination function of the λ N gene product. Very few N-resistant termination signals have been described previously (11, 21).

Our results indicate that the MuLV LTR could also serve as a promoter for transcription in E. coli cells when introduced upstream of the galactokinase structural gene in plasmid pKO1. The dual activity of the LTR in promotion and termination of transcription in E. coli is reminiscent of the function of the LTR in expression of the integrated proviral genes during virus replication. The left-end LTR serves mainly as a promoter, while the right-end LTR downstream of the active promoter serves as a termination signal (24). In our experimental system, the LTR can initiate transcription when it is not transcribed from an upstream promoter (when inserted in pKO1). However, when inserted downstream of the gal promoter (in pKG1900), it functions as a terminator.

For the analysis of transcription termination activity of the LTR in eucaryotic cells, new constructions will have to be designed. One such construction which we are pursuing involves the insertion of the LTR and its deleted derivatives (described in this work) downstream of the promoter of the late gene of simian virus 40 (3). Premature termination of transcription at the inserted LTR region may serve as an assay for termination in mammalian cells.

This work was supported by grants from the Israel Academy of Sciences (to A.H.), the Stiftung Volkswagenwerk, and the Bergman Foundation (to A.P.).

The excellent technical assistance of H. Falk is acknowledged. We thank H. Hyman for helpful discussions and D. Linemeyer and L. McKenney for supplying the plasmids.

LITERATURE CITED