JOURNAL OF VIROLOGY, Apr. 1995, p. 2214–2222
0022-538X/95/$04.00+0

Vol. 69, No. 4

An Internal Ribosomal Entry Mechanism Promotes Translation of Murine Leukemia Virus *gag* Polyprotein Precursors

CLARISSE BERLIOZ AND JEAN-LUC DARLIX*

LaboRétro, Unité de Virologie Humaine (INSERM-ENS U412), Ecole Normale Supérieure de Lyon, 69364 Lyon Cedex 07, France

Received 26 September 1994/Accepted 10 January 1995

The genomic retroviral RNA is the messenger for the translation of the *gag* and *pol* genes encoding the precursors to the major structural proteins and enzymes, respectively, of the virion core. The long 5′ untranslated region, the leader, is formed of independent well-structured domains involved in key steps of the viral life cycle such as the initiation of proviral DNA synthesis, genomic RNA dimerization and packaging, and the initiation of *gag* translation. These functional features and the presence of stable secondary structures between the cap and the *gag* initiation codon suggested that translation initiation of *gag* might proceed through a mechanism different from the canonical ribosome scanning process. Interestingly enough, murine leukemia viruses code also for a glycosylated *gag* precursor, named *glyco-gag*, initiated at a CUG codon upstream and in the same open reading frame as the AUG^{gag}. We have investigated the translation initiation of *gag* and *glyco-gag* precursors of Friend murine leukemia virus (F-MLV) in the rabbit reticulocyte lysate system and in murine cells. Through site-directed mutagenesis of *gag* and *glyco-gag* initiation codons, we show that initiation of *gag* and *glyco-gag* synthesis does not utilize the classical ribosome scanning. When poliovirus protease 2A is coexpressed in murine cells, expression of MLV-*lacZ* RNA is not modified, indicating that translation initiation of MLV *gag* precursors is a cap-independent mechanism. In addition, the F-MLV leader was inserted between two genes in a dicistronic *neo*-MLV-*lacZ* mRNA, and its ability to promote expression was examined in vitro and in vivo. Results obtained demonstrate that an internal ribosome entry mechanism promotes translation of F-MLV *gag* precursors. This finding led us to construct a new dicistronic retroviral vector in which the F-MLV leader can promote both packaging of recombinant genomic RNA and expression of the 3′ gene.

Cells infected by a retrovirus synthesize virus-specific RNAs by transcription of the integrated proviral DNA. The primary transcript is 5′ capped and 3′ polyadenylated and can be either spliced to give the subgenomic RNAs or transported to the cytoplasm where it serves as an mRNA for the synthesis of *gag* and *gag-pol* polyproteins and/or as the pregenomic RNA. The *gag* and *gag-pol* polyproteins are the precursors to the major proteins and enzymes, respectively, of the virion core (10).

The genetic organization of the 5′ sequences of the genomic RNA appears to be original and complex in that a long multifunctional 5′ untranslated domain, the leader, precedes the *gag* gene. In fact, the 5′ leader is formed of independent well-structured domains involved in key steps of the viral life cycle such as the initiation of proviral DNA synthesis (the primer binding site [PBS] and R sequences), genomic RNA dimerization and encapsidation (the dimerization-encapsidation signal [E/DLS]), and the initiation of *gag* precursor translation (AUG-619 *gag* initiation codon) (Fig. 1). The genomic RNA of murine leukemia virus (MLV) directs the synthesis of two *gag*-related proteins: Pr65^{gag}, the precursor to the virion core structural proteins, and Pr75$^{glyco\text{-}gag}$, the precursor to the glycosylated *gag* protein found on the surface of MLV-infected cells (*glyco-gag* is also called Gross cell surface antigen) (16, 17, 36, 46). Translation of Pr65^{gag} starts at an AUG (position 619 for Friend murine leukemia virus [F-MLV]) while Pr75$^{glyco\text{-}gag}$ is initiated at a CUG (position 355 for F-MLV) located within the 5′ leader and in frame with AUG^{gag} (43, 44). The functional characteristics of the leader and the identification of two initiation codons in the same open reading frame in the 5′ leader of MLV RNA are two features which could be expected to inhibit the scanning ribosome from reaching the correct initiation sites. According to the ribosome scanning model, most likely used by most 5′ capped eukaryotic mRNAs (29, 32, 34, 39), the ribosomal 40S subunit, carrying Met-tRNA and initiation factors, binds to the 5′ cap and migrates along the 5′ leader until it encounters an AUG codon in a favorable context, i.e., (A/G)CCAUGG (30, 34). The presence of stable secondary structures ($\Delta G = -50$ kcal [ca. -200 kJ]/mol) between the cap and the initiation codon was shown to strongly inhibit the initiation of translation by a 40S ribosomal subunit scanning stop (29). The 5′ MLV leader has been shown to be very structured (40, 50) and thus could interfere with the scanning of 40S ribosomal subunits along the 5′ leader. These structural features and the functional characteristics of the 5′ leader described above suggested that translation initiation of *gag* might proceed through a mechanism different from the canonical ribosome scanning process.

In this study, initiation of F-MLV *gag* and *glyco-gag* translation was investigated in the rabbit reticulocyte lysate system (RRL) and in murine cells. Results show that initiation of F-MLV *gag* and *glyco-gag* translation does not utilize the classical ribosome scanning or leaky scanning. Rather, results indicate that initiation of *gag* precursor translation is most probably controlled by an internal ribosome entry mechanism.

MATERIALS AND METHODS

Plasmid construction. All Moloney, Friend, and AKR strains of MLV code for *glyco-gag* and *gag* polyproteins (16, 17, 36, 46). To study the translation initiation of these two *gag* precursors, we used F-MLV 57 provirus (F-MLV 57), which is responsible for the development of hemolytic anemia and erythroleukemia in

* Corresponding author. Mailing address: LaboRétro, Unité de Virologie Humaine (INSERM-ENS U412), Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France. Phone: (33) 72.72.81.69/(33) 72.72.80.51. Fax: (33) 72.72.80.80.

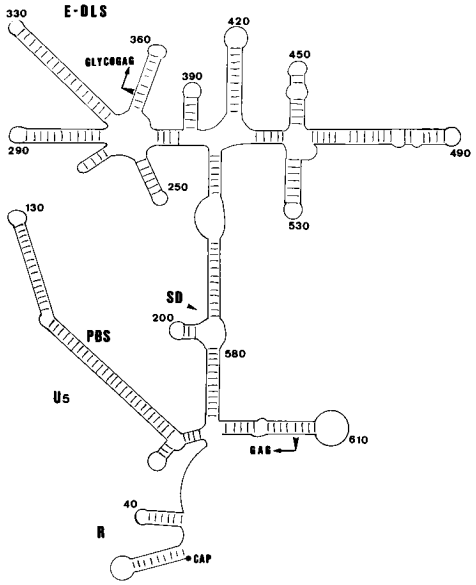


FIG. 1. A secondary structure model of F-MLV 5′ leader RNA based on chemical probing and computer analysis. Numbering is with respect to the genomic RNA cap site (position +1). R, U5, PBS, SD, and E-DLS are the repeat, the untranslated 5′ region, the primer $tRNA^{Pro}$ binding site, the donor splice site, and the packaging-dimerization sequence, respectively. CUG-355 and AUG-619 are the *glyco-gag* and *gag* initiation codons, respectively.

mice (48). F-MLV 57 encodes $Pr65^{gag}$, initiated at an AUG at position 619, and $Pr75^{glyco\text{-}gag}$, initiated at a CUG at position 355 (11, 44).

Standard procedures were used for restriction endonuclease digestion and plasmid DNA construction (45). *Escherichia coli* HB101 strain 1035 (RecA$^-$) was used for plasmid DNA amplification. Details of the constructions are given below. All numbering is with respect to the genomic RNA cap site (position +1).

(i) pMLV-CB1. The Moloney MLV DNA fragment *Nhe*I-*Sac*I from pMLV-SI4 (12) containing the 5′ leader and *glyco-gag* and *gag* coding sequences (positions 1 to 2558) under the control of the T7 RNA polymerase promoter was ligated to the pMLV-CG19 DNA fragment [positions 8233 to 8333 plus 23-nucleotide poly(A) tail] obtained by digestion with *Sac*I and *Spe*I (3).

(ii) pMLV-CB2. The F-MLV DNA fragment *Kpn*I-*Sac*I (positions 36 to 2559) from p57A (gift of M. Sitbon, Paris, France) corresponding to the 5′ leader, *glyco-gag*, and *gag* coding sequences was inserted into pMLV-CB1 digested with *Kpn*I and *Sac*I (see Fig. 2).

(iii) pPolio-CB27. The pLNPOZ DNA fragment *Eco*RI-*Nhe*I (2) (position 1025 upstream of *neo* gene to position 6035 downstream of *lacZ* gene; see Fig. 9A) was cloned into Bluescript II KS$^+$ digested with *Eco*RI and *Spe*I to generate pCB25. The *Hin*dIII-*Xba*I fragment of pCB25 containing the *neo*-poliovirus internal ribosome entry site (IRES)-*lacZ* sequences was cloned into the pRc/CMV vector (Invitrogen), lacking positions 1284 to 3253, digested with *Hin*dIII and *Xba*I. The *neo*-IRES-*lacZ* sequences are under the control of the T7 RNA polymerase promoter for in vitro expression and the cytomegalovirus early promoter for expression in eukaryotic cells.

(iv) pMLV-CB28. The F-MLV 57 DNA fragment 1-651 was generated by PCR. This PCR fragment was digested with *Xho*I and *Bam*HI and inserted between *neo* and *lacZ* of pPolio-CB27 partially digested with *Xho*I and *Bam*HI. In this plasmid, initiation of β-Gal translation was under the control of F-MLV AUG^{gag} (position 619).

(v) pMLV-CB39. The F-MLV DNA fragment 212-651 was generated by PCR and digested with *Nhe*I. This fragment was inserted in the antisense orientation (positions 651 to 212) between *neo* and *lacZ* of pMLV-CB28 digested with *Nhe*I. In this construct, the initiation of β-Gal translation was under the control of an AUG in a favorable context, i.e., (A/G)CCAUGG (30), which was created by PCR at position 212 (see Fig. 5).

(vi) pMLV-CB50. The F-MLV DNA fragment 1-565 was generated by PCR. The 280-565 fragment was recovered after digestion with *Spe*I (position 280) and *Nhe*I (inserted after position 565) and inserted between *neo* and *lacZ* of pMLV-CB28 digested with *Nhe*I. In this construct and in pMLV-CB61 described below, the initiation of β-Gal translation was under the control of an AUG in a favorable context, i.e., (A/G)CCAUGG (30), which was created by PCR at position 565 (see Fig. 5).

(vii) pMLV-CB61. The F-MLV DNA fragment 1-565 was generated by PCR and digested with *Nhe*I. This fragment was inserted between *neo* and *lacZ* of pMLV-CB28 digested with *Nhe*I (see Fig. 5).

(viii) pMLV-CB63. The F-MLV DNA fragment 1-620 was generated by PCR. The PCR fragment was digested by *Nhe*I and inserted between *neo* and *lacZ* of pMLV-CB28 digested by *Nhe*I. In this construct and in pMLV-CB94 and pMLV-CB98, the initiation of β-Gal translation was under the control of F-MLV AUG^{gag} (position 619) (see Fig. 5).

(ix) pMLV-CB71. The Cla12-AP *Eco*RI-*Xba*I DNA fragment (21a) containing the alkaline phosphatase gene was cloned into Bluescript II KS$^+$ digested with *Eco*RI and *Xba*I to generate pCB70. The *Eco*RV-*Sal*I fragment of pCB70 containing the alkaline phosphatase gene and the *neo* gene obtained by *Sal*I-*Bam*HI digestion of the PCR-amplified DNA 1659-2499 of pLNPOZ were cloned into pLNPOZ (see Fig. 9A) digested with *Bal*I and *Bam*HI by a trimolecular ligation. A *Spe*I site was inserted during the PCR of the *neo* fragment between the *Sal*I site and position +4 of the *neo* gene.

(x) pMLV-CB74. The F-MLV DNA fragment 212-651 was generated by PCR and digested with *Nhe*I. This fragment was inserted in the antisense orientation (positions 651 to 212) between *phosphatase* and *neo* of pMLV-CB71 digested with *Spe*I. In this construct, the initiation of *neo* translation was under the control of an AUG in a favorable context, i.e., (A/G)CCAUGG (30), which was created by PCR at position 212 (see Fig. 9A).

(xi) pMLV-CB78. The F-MLV DNA fragment 212-620 was generated by PCR and digested with *Nhe*I. This fragment was inserted between *phosphatase* and *neo* of pMLV-CB71 digested with *Spe*I. In this construct, the initiation of *neo* translation was under the control of F-MLV AUG^{gag} (position 619) (see Fig. 9A).

(xii) pMLV-CB92. The *Sma*I-*Xba*I fragment of pMLV-CB61 containing the F-MLV deleted leader and the *lacZ* gene was ligated into pMLV-CB28 digested with *Eco*RV and *Xba*I. This construct contains the F-MLV leader upstream of the *lacZ* gene under the control of the cytomegalovirus early promoter and the T7 promoter (see Fig. 4A).

(xiii) pMLV-CB93. The *Sma*I-*Xba*I fragment of pMLV-CB63 containing the entire F-MLV leader and the *lacZ* gene was ligated into pMLV-CB28 digested with *Eco*RV and *Xba*I. This construct contains the F-MLV leader upstream of the *lacZ* gene under the control of the cytomegalovirus early promoter and T7 promoter (see Fig. 4A).

(xiv) pMLV-CB94. The F-MLV DNA fragment 280-620, obtained by *Spe*I-*Nhe*I digestion of the PCR-amplified DNA 1-620, was inserted between *neo* and *lacZ* of pMLV-CB28 digested with *Nhe*I (see Fig. 5).

(xv) pHMSV-LacZ. The Harvey murine sarcoma virus DNA fragment 1-380 was generated by PCR and digested with *Nhe*I (position +1) and *Hin*dIII (position 380). This fragment was inserted in the antisense orientation (positions 380 to 1) upstream of the *lacZ* gene of pMLV-CB28 digested with *Hin*dIII and *Nhe*I. This construct contains the Harvey murine sarcoma virus sequence (positions 380 to 1) upstream of the *lacZ* gene under the control of the cytomegalovirus early promoter and T7 promoter. The initiation of β-Gal translation was under the control of an AUG in a favorable context, i.e., (A/G)CCAUGG, which was created by PCR at position +1.

(xvi) pMLP-P2A and pMLP-RP2A. Plasmids pMLP-P2A and pMLP-RP2A contain the poliovirus protease 2A coding sequence derived from poliovirus type 1 (Mahomey strain). In pMLP-P2A, the protease coding sequence was inserted downstream of the adenovirus major late promoter and its tripartite leader. The pMLP-P2AR plasmid containing the insert in the reverse orientation served as a control. These plasmids were gifts of N. Fouillot (18).

(xvii) pβ-actin-LacZ. The β-actin-LacZ plasmid contains the *lacZ* gene under the control of the β-actin promoter. This plasmid was a gift of P. Savatier (47) (see Fig. 4A).

Site-directed mutagenesis of $Pr75^{glyco\text{-}gag}$ and $Pr65^{gag}$ initiation codons. Mutagenesis was carried out with the Muta-gene M13 in vitro mutagenesis protocol (Bio-Rad Laboratories) based on the method described by Kunkel (35). The *Kpn*I-*Eco*RI F-MLV DNA fragment (positions 36 to 2852) containing the initiation codons of *glyco-gag* and *gag* was ligated to M13mp19 DNA digested with *Eco*RI and *Kpn*I (pMLV-CB4). The following oligonucleotides were used to mutate the CTG-355 codon to CAG (5′-GACGTCTCCCTGGGTTGCGGC-3′) or to ATG (5′-GACGTCTCCCATGGTTGCGGC-3′) (mutated nucleotides are underlined). The oligonucleotide used to alter the ATG-619 codon to AGC was 5′-AGTCTGGCCGCTGTTTTCAGA-3′. The following oligonucleotide was used to mutate the GTG-319 codon to ATG (5′-CCGTCAGTTCCATC ACGGGTCCGCC-3′). The *Kpn*I-*Sac*I DNA fragment (positions 36 to 2559) containing one mutation or a combination of two or three mutations was reinserted into pMLV-CB2. The resulting plasmids were named pMLV-CB6 (CTG-355 to CAG), pMLV-CB7 (CTG-355 to ATG), pMLV-CB16 (CTG-355 to ATG; ATG-619 to AGC), pMLV-CB82 (GTG-319 to ATG), pMLV-CB84 (GTG-319 to ATG; CTG-355 to CAG), and pMLV-CB86 (GTG-319 to ATG; CTG-355 to ATG). Mutant clones were selected upon sequencing by the dideoxyribonucleotide chain termination method using DNA primers, modified T7 DNA polymerase (Sequenase), and ^{35}S-dATP.

In vitro-generated RNA. pMLV-CB2, -CB6, -CB7, -CB16, -CB82, -CB84, and -CB86 plasmids were digested with *Sac*I to generate the template used to synthesize wild-type and mutated F-MLV RNAs (positions 1 to 2559) with T7 RNA polymerase. pMLV-CB39, pMLV-CB50, pMLV-CB61, pMLV-CB63, pMLV-CB92, pMLV-CB93, and pMLV-CB94 plasmids were digested with *Ssp*I (position 1240 in *lacZ* gene) to obtain the template used to synthesize dicistronic RNAs with T7 RNA polymerase.

Five micrograms of linearized plasmid DNA was transcribed in 0.1 ml of 40 mM Tris-HCl (pH 7.5)–6 mM $MgCl_2$–2 mM spermidine–10 mM dithiothreitol–10 mM NaCl–0.5 mM (each) ribonucleoside triphosphate with 40 U of T7 RNA polymerase and 80 U of the RNase inhibitor RNasin for 3 h at 37°C. Following DNase treatment (RQ1 DNase), the RNA was extracted with phenol-chloroform and precipitated with ethanol. The pellet was washed with 70% ethanol, and the RNA was dissolved in sterile bidistilled water. RNA was analyzed by electrophoresis through a 0.7% (wt/vol) agarose gel in 50 mM Tris-borate (pH 8.3)–0.1 mM EDTA at 5 V/cm and visualized by ethidium bromide staining (1 μg/ml for 5 min). To synthesize capped RNAs, the reaction was performed with the same conditions as described above, except that ATP, CTP, and UTP were added at a final concentration of 0.5 mM with 0.05 mM GTP and 0.5 mM $m^7G(5')ppp(5')Gm$ (Pharmacia).

In vitro translation. Wild-type and mutated MLV RNAs were translated in the nuclease-treated RRL (Promega) at 50% of the original concentration with 10 μg of RNA per ml and 1 mCi of [^{35}S]methionine per ml (Amersham) at 31°C for 1 h. The assays were supplemented in potassium acetate to a final concentration of 130 mM. The samples were then boiled in 62.5 mM Tris-HCl (pH 6.8)–2% sodium dodecyl sulfate (SDS)–10% glycerol–5% β-mercaptoethanol–0.02% bromophenol blue, and the ^{35}S-labelled proteins were analyzed by 7% (wt/vol) polyacrylamide–0.2% SDS gel electrophoresis. $Pr65^{gag}$ and $Pr75^{gag}$ products were quantitated by means of scanning with Cirrus 1.0.1 (Softhansa GmbH) and Scan Analysis 2.0 (Biosoft).

pMLV-CB39, pMLV-CB50, pMLV-CB61, pMLV-CB63, pMLV-CB92, pMLV-CB93, and pMLV-CB94 RNAs were translated with the RRL assay as described above. The ^{35}S-labelled proteins were analyzed by 12% (wt/vol) polyacrylamide–0.2% SDS gel electrophoresis either directly or after immunoprecipitation. Efficiencies of neomycin and β-Gal translation were quantitated by means of scanning with Cirrus 1.0.1 (Softhansa GmbH) and Scan Analysis 2.0 (Biosoft). The efficiencies of IRES-mediated translation initiation were determined by the ratio of truncated β-Gal expression to neomycin expression.

Immunoprecipitation. Anti-β-galactosidase mouse monoclonal antibody (4 μg) (Promega) and 1× TNE (90 μl) (150 mM NaCl, 50 mM Tris HCl [pH 7.4], 1 mM EDTA) were added to the translation products. After 2 h at 4°C, immunoprecipitates were captured by addition of 10% (vol/vol) protein A-Sepharose (Pharmacia) and incubated for 1 h at 4°C. Then, immunoprecipitates were centrifuged and the pellets were washed three times with TNE. After a 10-min incubation at 100°C in the presence of protein sample buffer (62.5 mM Tris-HCl [pH 6.8], 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.02% bromophenol blue), the supernatants containing the immunoprecipitated proteins were resolved on an SDS–12% polyacrylamide gel. After electrophoresis, the gel was fixed with acetic acid-methanol (10%/30% [vol/vol]) and dried, and the results were visualized by autoradiography.

Cell culture and DNA transfection. NIH 3T3 cells were grown in Dulbecco modified Eagle medium containing 5% newborn calf serum at 37°C in the presence of 5% CO_2. For DNA transfections, cells were seeded at 5×10^5 cells per 100-mm plate 24 h prior to transfection. Cells were transfected by the calcium phosphate procedure exactly as described previously (9). Transfections were performed in triplicate for each plasmid with 6 μg of the pMLV DNA construct. Two days later, the transfected cells were washed. For each triplicate, 24 h after, transfected cells were either fixed to detect β-galactosidase expression by histochemical staining with X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) (20), pooled to measure β-galactosidase enzyme activity, or lysed to extract the RNAs for Northern (RNA) analysis. Efficiencies of IRES-mediated translation initiation were determined by the ratio of β-galactosidase activity to level of arbitary units of scanned mRNA *lacZ* per blue-stained cell. Cotransfection experiments with the protease 2A were performed in duplicate for each plasmid combination with 5 μg of pMLP-P2A or pMLP-P2AR plasmid and 7 μg of pMLV, or 10 μg of pβ-actin-LacZ plasmids by plate. Three days later, proteins of the transfected cells were extracted to assay β-galactosidase activity, and cellular RNAs were subjected to Northern analysis. Efficiencies of β-Gal expression were determined by the ratio of β-galactosidase activity to level of arbitary units of scanned mRNA *lacZ*.

β-Galactosidase enzyme assays were performed on cell extracts by the β-galactosidase enzyme assay system (Promega).

Transfection and infection with dicistronic retroviral vectors. An ecotropic retroviral packaging cell line, GP-E+86 (38), was grown in Dulbecco modified Eagle medium containing 5% newborn calf serum at 37°C in the presence of 5% CO_2 and was seeded at 5×10^5 cells per 100-mm plate 24 h prior to transfection. The cells were transfected by the calcium phosphate procedure (9) with 18 μg of pLNPOZ (2), pMLV-CB74, or pMLV-CB78 plasmid. The transfected cells were incubated 3 days before the virus-containing medium was harvested and were either fixed for alkaline phosphatase or β-galactosidase histochemical staining or placed under G418 selection at 0.8 mg/ml for 2 weeks.

All retrovirus infections were performed by adding 4 ml of virus-containing medium to the cells. One day prior to infection, NIH 3T3 cells were seeded at 3×10^5 cells per 60-mm plate. Freshly harvested viruses from transient expression of helper cells were filtered (0.45-μm-pore-size filter), and Polybrene was added at a concentration of 8 μg/ml. Infected cells were grown for 24 h. Before their passage, infected cells were either placed under G418 selection at 0.8 mg/ml for 2 weeks or stained for alkaline phosphatase or β-galactosidase expression detection.

Preparation of cellular RNAs and Northern analysis. Extraction of cellular RNAs from transfected cells and Northern analysis were performed on transfected cells as described previously by Khandjian and Meric (28). A ^{32}P-labelled probe complementary to the *lacZ* gene (positions 23 to 436) (27) was labelled by the NonaPrimer Kit I (Appligene). The membrane was autoradiographed overnight at −80°C. RNAs were quantitated by means of scanning with Cirrus 1.0.1 (Softhansa GmbH) and Scan Analysis 2.0 (Biosoft). The level of *lacZ* mRNA expressed was given in arbitary units by fixed area scanning on the same autoradiography.

Alkaline phosphatase histochemical staining. After transfection or infection, cells were fixed with 2% formaldehyde and 0.2% glutaraldehyde, washed twice with phosphate-buffered saline, and incubated for 30 min in phosphate-buffered saline at 65°C. Then, cells were washed twice with phosphatase buffer containing 100 mM Tris-HCl (pH 9.5), 100 mM NaCl, and 50 mM $MgCl_2$ and incubated for 3 h in a staining solution containing 0.1 mg of BCIP (5-bromo-4-chloro-3-indolylphosphate toluidinium) per ml, 1 mg of nitroblue tetrazolium salt per ml, and 1 mM levamisole in phosphate buffer.

RESULTS

Effect of the 5′ cap on *gag* precursor synthesis. To investigate whether *gag* and *glyco-gag* translation initiation is dependent on ribosome scanning, we analyzed the effect of the 5′ RNA cap on the efficiency of *glyco-gag* and *gag* translation. In vitro-generated capped or uncapped RNAs containing the 5′ leader and the *glyco-gag* and *gag* coding sequences of F-MLV (pMLV-CB2) (see Fig. 2 and Materials and Methods) were translated in the RRL at 50% (79 mM potassium acetate) or 70% (110 mM potassium acetate) of the original concentration. Two major products, $Pr75^{glyco\text{-}gag}$ and $Pr65^{gag}$, were obtained, and no change in the level of in vitro *glyco-gag* or *gag* synthesis was observed when uncapped RNA was used (data not shown), whatever the dilution of the RRL. Synthesis of F-MLV *glyco-gag* and *gag* was found to be independent of the 5′ cap in the RRL at both dilutions, suggesting that a mechanism different from the canonical ribosome scanning could be involved (51). It should be noted that, under these translation conditions, a four- to sevenfold increase in β-Gal synthesis was observed with capped versus uncapped control Harvey murine sarcoma virus-LacZ RNA containing the Harvey murine sarcoma virus sequence (positions 1 to 380) in an antisense orientation upstream of the *lacZ* gene (pHMSV-LacZ) (data not shown). To further investigate the mechanism of translation initiation of *gag* protein, we mutated the *glyco-gag* and *gag* initiation codons and analyzed their influence on the synthesis of the *gag* precursors.

Influence of *glyco-gag* initiation codon mutations on the level of *gag* precursor synthesis. Several other cap-dependent mechanisms, such as termination-reinitiation, promoter switching, or leaky ribosome scanning, have been proposed to account for the initiation of translation of bifunctional mRNAs with initiation codons in the same open reading frame as in F-MLV genomic RNA (33, 34). Termination-reinitiation is not thought to occur with F-MLV RNA since no additional cistron was found upstream of the *gag* open reading frame (31). In addition, promoter switching is very unlikely because only one F-MLV genomic RNA has been identified (10). On the other hand, leaky ribosome scanning, in which some 40S ribosomes bypass the first initiation codon and start translation at the second initiation codon, may explain translation initiation at both the $CUG^{glyco\text{-}gag}$ and AUG^{gag} (8, 34).

To investigate this possibility, the CUG-355 initiation codon of *glyco-gag*, which is in a favorable context, i.e., ACCCUGG (30), was mutated to either CAG to prevent *glyco-gag* synthesis (pMLV-CB6) or to AUG in order to overexpress *glyco-gag* (pMLV-CB7). In addition, a double mutation was made in which $AUG^{glyco\text{-}gag}$ was substituted for CUG and AUG^{gag} was changed to AGC in order to abolish *gag* synthesis (pMLV-CB16) (Fig. 2; see also Materials and Methods). Subsequently,

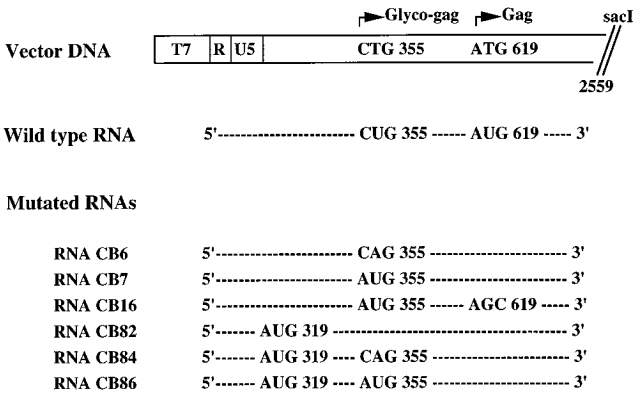


FIG. 2. F-MLV mutant RNAs generated in vitro. Wild-type (pMLV-CB2) and mutated (pMLV-CB6, -CB7, -CB16, -CB82, -CB84, and -CB86) plasmids containing the *glyco-gag* and *gag* coding sequences (in the same open reading frame) under the control of the T7 RNA polymerase promoter were digested with *Sac*I (+2559) to generate the templates used to synthesize F-MLV wild-type and mutated RNAs.

wild-type and mutated in vitro-generated RNAs were translated in the nuclease-treated RRL supplemented in potassium acetate at the final concentration of 130 mM (Fig. 3) (51). The positive control was RNA coding for luciferase (65 kDa; lane 1), and the negative control corresponded to translation without RNA (lane 2). Translation of wild-type RNA containing the 5′ first 2,558 nt of F-MLV yielded two major products, $Pr65^{gag}$ and $Pr75^{glyco\text{-}gag}$ (lane 3). An additional band with a mobility of 55 kDa was seen. This band may well correspond to a premature stop of $Pr65^{gag}$ translation since it disappeared when AUG^{gag} was mutated to AGC (lane 6). Mutation of CUG-355 to CAG resulted in a complete inhibition of *glyco-gag* synthesis while a slight increase in $Pr65^{gag}$ synthesis was observed (~30%) (lane 4). When CUG-355 was mutated to AUG, $Pr75^{glyco\text{-}gag}$ synthesis was increased 1.6-fold while $Pr65^{gag}$ synthesis was decreased 1.3-fold (lane 5 versus lane 3). On the other hand, only $Pr75^{glyco\text{-}gag}$ expression was observed (compared with 30% for the wild type) when CUG-355 was mutated to AUG and AUG^{gag} was changed to AGC (lane 6 versus lane 3). The persistent expression of $Pr65^{gag}$ (~75% of that observed for the wild type) even when expression of the upstream reading frame $Pr75^{glyco\text{-}gag}$ was increased 1.6-fold suggests that a small percentage of *gag* synthesis may be due to leaky scanning. However, this would account for at most a small percentage of *gag* translation initiation.

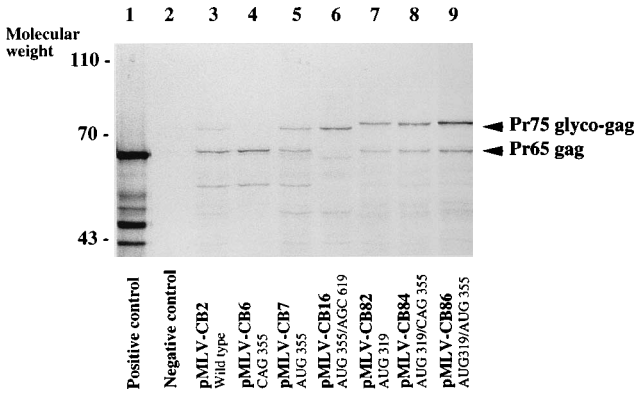


FIG. 3. Translation of F-MLV RNA in the RRL. Wild-type (lane 3) and mutant MLV (lanes 4 to 9) RNAs were translated in the RRL. After heat denaturation, ^{35}S-labelled proteins were analyzed by 7% (wt/vol) polyacrylamide–0.2% SDS gel electrophoresis. The positions of F-MLV $Pr65^{gag}$ and $Pr75^{glyco\text{-}gag}$ are indicated. Lane 1, luciferase RNA. Lane 2, no RNA. Molecular weight is in thousands.

To further rule out the possibility that translation was due to scanning from the 5′ end, the GUG-319 codon located in a good translation context, i.e., GUGGUGG, and in frame with CUG-355 was mutated to AUG (pMLV-CB82). The effect of this mutation was analyzed when the *glyco-gag* initiation codon was changed to either CAG (pMLV-CB84) or AUG (pMLV-CB86) (Fig. 2). Translation of these RNAs in the RRL could potentially yield three products, $Pr65^{gag}$, $Pr75^{glyco\text{-}gag}$, and a higher-molecular-weight protein (Pr77) corresponding to initiation at the new AUG-319. *glyco-gag* synthesis was abolished when AUG-319 was introduced, even when CUG-355 was mutated to AUG (lanes 7 and 9). The strong negative effect on *glyco-gag* translation resulting from the insertion of AUG-319 upstream of CUG-355 indicates that the ribosomes scan between positions 319 and 355 of the F-MLV leader. In addition, regardless of the level of Pr77 synthesis, $Pr65^{gag}$ was always expressed at 56% of the wild-type level (lanes 7 to 9 versus lane 3). It should be noted that the level of Pr77 synthesis in relation to *gag* synthesis was similar for the three constructs tested (lanes 7 to 9) as determined by scan analysis. These results are consistent with those described above when CUG-355 was mutated and support the idea that a minor part of *gag* translation initiation may result from leaky scanning of 40S ribosomal subunits between the $CUG^{glyco\text{-}gag}$ and the AUG^{gag}. However, the majority of *gag* translation initiation does not depend on leaky ribosome scanning between these two codons.

Influence of poliovirus protease 2A on the expression of monocistronic vectors containing the F-MLV leader 5′ to the *lacZ* gene. In poliovirus, ribosomes bind directly within the 5′ nontranslated region of the uncapped viral RNA and scan to the initiation codon while translation of cellular mRNAs is inhibited by inactivation of the cap-binding complex eIF-4F, a process mediated by poliovirus protease 2A (23). In a previous study of the translation of the hepatitis B virus polymerase gene, Fouillot et al. have shown that transient expression of protease 2A was sufficient to inhibit cap-dependent hepatitis B virus P gene translation (18). To examine whether *gag* and *glyco-gag* translation is a cap-dependent mechanism, monocistronic vectors with the F-MLV leader upstream of the *lacZ* gene were constructed (Fig. 4A) and cotransfected into NIH 3T3 cells with a plasmid expressing poliovirus protease 2A. In the latter plasmid, it should be noted that the poliovirus protease 2A-coding region was inserted downstream of the adenovirus major late promoter and its tripartite leader (18). The adenovirus tripartite leader confers on this mRNA the ability to be translated independently of the formation of the cap-binding complex and allows its translation in poliovirus-infected cells (14). The plasmid pMLP-P2A contains the insert in the correct orientation, allowing the expression of protease 2A, while the plasmid pMLP-RP2A contains the insert in the reverse orientation, serving as a control. In pMLV-CB93 and pMLV-CB92, the F-MLV leader (positions 28 to 620 or positions 28 to 565, respectively) was inserted upstream of the *lacZ* reporter gene (see Materials and Methods). The pβ-actin-LacZ plasmid corresponded to a positive control for a cap-dependent translation process. Cotransfection experiments with protease 2A were performed in duplicate for each plasmid combination. After 48 h, β-galactosidase activity was assayed on transfected cells and cellular RNAs were extracted to perform Northern analysis. Efficiencies of translation were measured as units of β-galactosidase per arbitary unit of mRNA β-Gal.

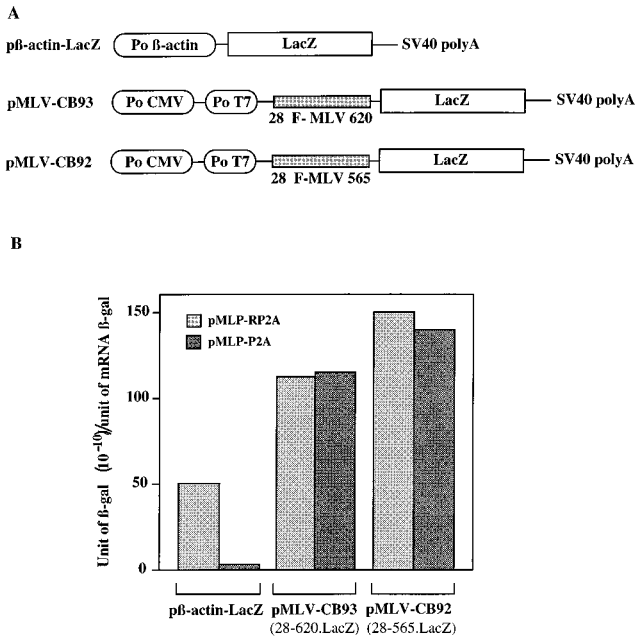


FIG. 4. Effect of poliovirus protease 2A on the translation of monocistronic MLV β-galactosidase mRNA. (A) Plasmids pMLV-CB93 and pMLV-CB92 contain the F-MLV leader (positions 28 to 620 or 28 to 565) 5′ to the *lacZ* gene. In pβ-actin-LacZ, β-galactosidase expression is driven by the rat β-actin promoter. (B) Plasmid pMLP-P2A expresses poliovirus protease 2A under the control of the adenovirus major late promoter; pMLP-RP2A has the p2A sequence in the reverse orientation as a control plasmid. Cotransfection experiments with protease 2A were performed in duplicate for each plasmid combination. After 48 h, β-galactosidase activity was assayed on transfected cells, and cellular RNAs were extracted to perform Northern analysis. Efficiencies of translation are given as units of β-galactosidase per arbitrary unit of mRNA β-Gal. Results are the average of two independent experiments. SV40, simian virus 40; CMV, cytomegalovirus.

The results of the cotransfection experiments are reported in Fig. 4B. The level of β-Gal in cells cotransfected with pβ-actin-LacZ and pMLP-P2A was 20 times lower than in cells cotransfected with pMLP-P2AR (Fig. 4B). These results indicate that the poliovirus protease 2A encoded by the pMLP-P2A plasmid was capable of strongly inhibiting the cap-dependent synthesis of β-galactosidase. When protease 2A was coexpressed with pMLV-CB93 or pMLV-CB92 in NIH 3T3 cells, the level of β-galactosidase expression was very similar to that in the presence of pMLP-P2AR. Since protease 2A had no effect on gene expression directed by the F-MLV leader, it can be concluded that *gag* precursor translation initiation is most probably directed by a cap-independent mechanism.

In vitro translation of dicistronic *neo*-MLV-*lacZ* mRNAs. In picornaviruses, positive-strand RNA viruses (7, 21, 25, 26, 42), and in the human immunoglobulin heavy-chain-binding protein mRNAs (24, 37), translation initiation occurs by an internal ribosome entry mechanism. Defined structural elements and pyrimidine-rich regions compose the IRES and permit ribosome scanning-independent translation by direct binding of ribosomes to the leader sequences of these mRNAs (15, 25). Similar structured and pyrimidine-rich regions are located in the leader of F-MLV RNA upstream of $CUG^{glyco\text{-}gag}$ and of AUG^{gag} (40, 50). In addition, the F-MLV 5′ first 205 nucleotides corresponding to R, U5, and PBS sequences have a potential RNA secondary structure with a stability in the range of −50 kcal (ca. −200 kJ)/mol (16a) (Fig. 1). The presence of these stable secondary structures between the cap and the

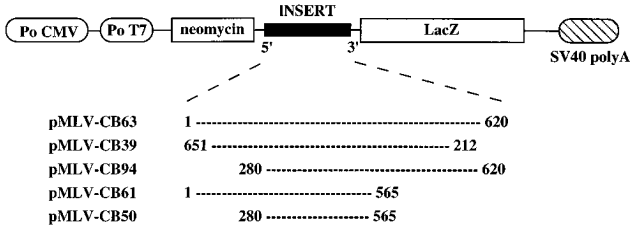


FIG. 5. Dicistronic Neo-MLV-LacZ plasmid DNA constructs used in vitro and in vivo. Dicistronic plasmid DNAs contain the leader sequence of F-MLV between the neomycin and β-galactosidase genes under the control of the T7 RNA polymerase promoter (Po T7) and the cytomegalovirus early promoter (Po CMV). The viral DNA fragments cloned in the plasmids are shown (numbering is with respect to the +1 cap site of the viral RNA). SV40, simian virus 40.

$CUG^{glyco\text{-}gag}$ and AUG^{gag} initiation codons could strongly inhibit the initiation of *glyco-gag* and *gag* translation by prematurely stopping the scanning of 40S ribosomal subunits. Indeed, the presence of stable secondary structures in the 5′ leader of mRNAs appears to strongly inhibit the initiation of translation (29). Thus, these features suggest that initiation of *glyco-gag* and *gag* protein translation might occur by direct binding of ribosomes to the F-MLV 5′ leader sequence.

To investigate this possibility, the complete F-MLV leader sequence was inserted between the neomycin (*neo*) and the β-galactosidase (*lacZ*) reporter genes (Fig. 5; see Materials and Methods). The recombinant pMLV plasmids allowed the in vitro synthesis of dicistronic RNAs ranging from 2,400 to 2,800 nucleotides in length (Fig. 5). Agarose gel electrophoresis of the in vitro-generated RNAs confirmed that they were intact and unique (data not shown).

Subsequently, dicistronic RNAs were translated in the RRL. These RNAs encoded *neo* protein, directing standardization of the level of translation, and a C-terminally truncated β-galactosidase with an average molecular mass of 46 kDa ($\beta\text{-Gal}^{46}$). This system allows β-galactosidase expression only if an IRES is present in the F-MLV RNA leader (Fig. 6). The positive control employed RNA coding for luciferase (65 kDa; lane 1), and the negative control corresponded to translation without RNA (lane 2). Insertion of the F-MLV leader in the antisense orientation (pMLV-CB39) resulted in the synthesis of a very low level of $\beta\text{-Gal}^{46}$ (lane 5), corresponding to a basal level of *lacZ* initiation resulting most probably from termination-reinitiation between *neo* and *lacZ*. Translation of the dicistronic RNA in which the F-MLV 5′ leader (positions 1 to 620) was inserted between the *neo* and *lacZ* genes (pMLV-CB63) produced a high level of $\beta\text{-Gal}^{46}$ (lane 4). This indicates that ribosomes can recognize sequences within the F-MLV leader to initiate β-Gal translation. Two forms of β-Gal were obtained with pMLV-CB63 as well as with the monocistronic construct pMLV-CB93 (Fig. 4): one initiated at $CUG^{glyco\text{-}gag}$ (molecular mass, 56 kDa), and the other one initiated at AUG^{gag} (molecular mass, 46 kDa) (lanes 3 and 4). When the 5′ first 280 nucleotides of the F-MLV leader were deleted (pMLV-CB94), synthesis of the β-Gal form initiated at the $CUG^{glyco\text{-}gag}$ was abolished (compare higher-molecular-weight band corresponding to $\beta\text{-Gal}^{56}$ in lane 4 versus lane 6). It should be noted that no $\beta\text{-Gal}^{56}$ was observed with pMLV-CB94 even after overexposure of the gel (data not shown). However, initiation at the AUG^{gag} was decreased by only about 10% (lane 6). To determine if the region 565-620 of the F-MLV leader is important, a construct containing only the region 1-565 was tested (pMLV-CB61). In this construct, the $CUG^{glyco\text{-}gag}$ is no longer in phase with β-Gal. However, an AUG initiation codon was

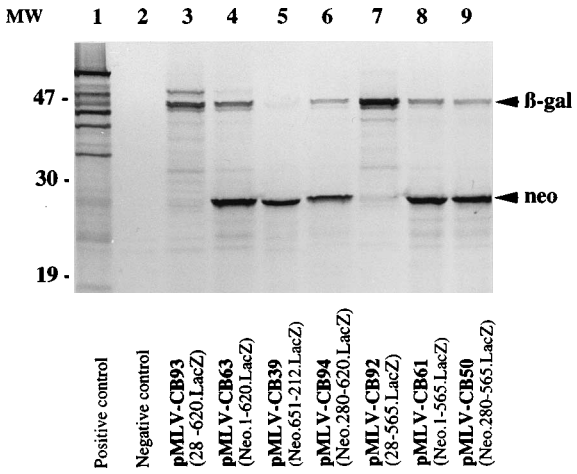


FIG. 6. Translation of monocistronic and dicistronic RNAs in RRL. Lanes 1 and 2 are positive (luciferase plasmid DNA [Promega]) and negative (no plasmid DNA) controls, respectively. Lanes 3 to 9 are pMLV-CB93, pMLV-CB63, pMLV-CB39, pMLV-CB94, pMLV-CB92, pMLV-CB61, and pMLV-CB50 RNAs, respectively, expressed in the RRL (Promega). After heat denaturation, ^{35}S-labelled proteins were analyzed by 12% (wt/vol) polyacrylamide–0.2% SDS gel electrophoresis. Positions of *neo* (28 kDa) and C-terminally truncated β-Gal proteins (46 kDa) are indicated. Lower-molecular-weight (MW) (in thousands) bands corresponded to premature translation stops.

introduced just before the β-Gal coding region. pMLV-CB61 was able to efficiently direct synthesis of β-Gal[46] (lane 8), indicating that this region is not important for the internal ribosome entry mechanism. Finally, deletion of the first 280 nucleotides of the F-MLV 1-565 fragment (pMLV-CB50) still permitted the synthesis of β-Gal[46] with a ratio of translation corresponding to 85% of that obtained with pMLV-CB61 (lane 9 versus lane 8). Identical results were obtained with dicistronic RNAs coding for *neo* and the entire *lacZ* sequence (data not shown). These results strongly suggest that an internal ribosome entry mechanism promotes translation initiation of *gag* precursors. Results obtained with pMLV-CB94 indicate that the 5′ first 280 nucleotides are important for the translation of the β-Gal form initiated at the CUG$^{glyco\text{-}gag}$ since the synthesis of this β-Gal form was abolished when this region was deleted (lane 6 versus lane 4). Moreover, it appears that the region 280-565 of F-MLV is still able to promote internal translation initiation of β-Gal in vitro.

Expression of the dicistronic *neo*-MLV-*lacZ* DNA constructs in NIH 3T3 cells. In order to confirm the in vitro data, the dicistronic pMLV DNA plasmids were assayed in cell culture by means of DNA transfection, and β-galactosidase expression was examined by histochemical staining of the cells with X-Gal (20), quantitation of β-galactosidase activity, and Northern analysis of dicistronic RNAs.

To rule out that a cryptic promoter and/or cryptic splice sites could function with the dicistronic MLV plasmids used, cellular RNAs were extracted 2 days after transfection and subjected to Northern analysis using a ^{32}P-labelled *lacZ* probe (see Materials and Methods). As shown in Fig. 7, only one major RNA was detected at the expected size by Northern analysis in NIH 3T3 cells transfected with the dicistronic pMLV DNA constructs. If a cryptic promoter and/or cryptic splice site were to be used, a mRNA of at least 3,000 nucleotides (*lacZ* gene is 3,078 bp) should be present and thus should migrate between the 18S (1,900-nucleotide) and the 28S (4,700-nucleotide) RNAs. Since no RNA was observed in this region or outside of the region presented (Fig. 7), the *neo*-MLV-*lacZ* RNAs were dicistronic in our in vivo assay conditions.

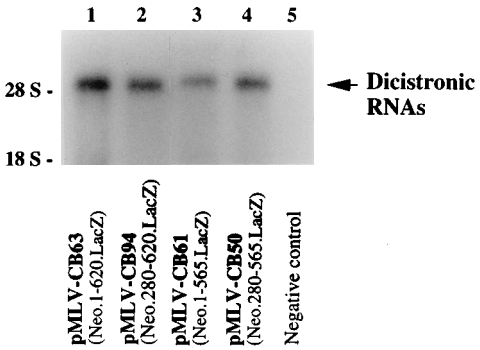


FIG. 7. Synthesis of MLV-derived dicistronic mRNAs in NIH 3T3 cells. Total RNA was extracted from NIH 3T3 cells transfected with 6 μg of plasmid DNA pMLV-CB63, pMLV-CB94, pMLV-CB61, and pMLV-CB50. RNAs were subjected to electrophoresis in an 0.8% agarose–formaldehyde gel in morpholinopropanesulfonic acid buffer and transferred to nitrocellulose. Hybridization was conducted with a ^{32}P-labelled probe corresponding to bases 23 to 436 of *lacZ*. The 18S and 28S rRNAs were revealed by ethidium bromide staining prior to RNA transfer and hybridization. Positions of 18S and 28S rRNAs and of dicistronic *neo*-MLV-*lacZ* RNAs are indicated.

NIH 3T3 cells were transfected with monocistronic and dicistronic pMLV plasmids, and the results of these experiments are reported as units of β-galactosidase per arbitrary unit of dicistronic mRNA β-Gal per blue-stained cell (see Materials and Methods) (Fig. 8). As a positive control of β-galactosidase expression, we used pMLV-CB93 in which the cytomegalovirus early promoter directed the synthesis of a monocistronic 5′ MLV leader *lacZ* mRNA with the F-MLV AUG^{gag} as the β-Gal initiation codon (Fig. 4A). All transfected pMLV plasmids produced about 20 to 100 times more β-galactosidase than the negative control corresponding to the MLV leader inserted in the antisense orientation (pMLV-CB39) (Fig. 8).

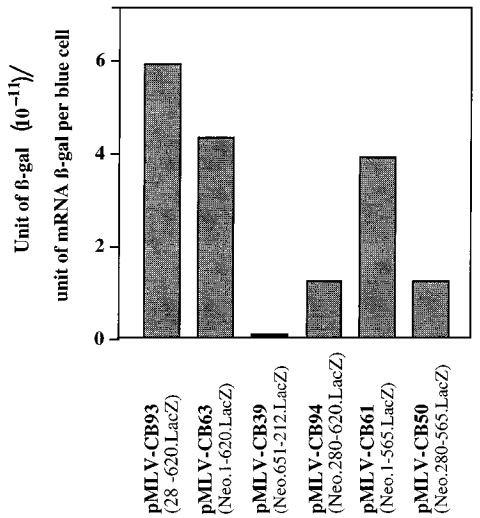
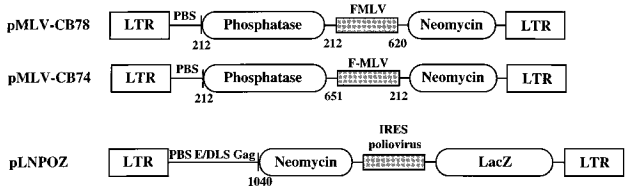


FIG. 8. β-Galactosidase expression of F-MLV dicistronic DNA constructs in transfected cells. NIH 3T3 cells (5×10^5 cells per 100-mm plate) were transfected in triplicate with pMLV-CB93, pMLV-CB63, pMLV-CB94, pMLV-CB61, pMLV-CB50, and pMLV-CB39 by the calcium phosphate procedure. Two days after transfection, transfected cells of one plate were either fixed to reveal β-galactosidase expression by histochemical staining with X-Gal, pooled to measure β-galactosidase activity, or lysed to extract the RNAs for Northern analysis. Efficiencies of IRES-mediated translation initiation are given as units of β-galactosidase per arbitary unit of mRNA *lacZ* per blue-stained cell. Results are the averages of three experiments.



B

Plasmids	% of transfected cells giving Neo^R foci	Relative viral titers (normalized to pMLV-CB78)	% of infected cells foci expressing the two genes
pMLV-CB78	2.1	100	91
pMLV-CB74	0.2	0.06	0
pLNPOZ	2.3	1400	80

FIG. 9. Dicistronic MLV-derived retroviral vectors with the F-MLV leader controlling packaging and internal initiation of translation. (A) A pMLV dicistronic retroviral vector was constructed with the leader sequence of F-MLV between the phosphatase and neomycin cistrons flanked by the MLV long terminal repeats (LTRs). The MLV DNA fragments used in the constructs are shown. In these vectors, the 5′ sequences contained the R, U5, PBS, and donor splice elements while the dimerization-encapsidation signal (E/DLS) from positions 212 to 620 was located between the two reporter cistrons. The pLNPOZ dicistronic vector was used as a positive control and contained the poliovirus IRES between the neomycin and β-galactosidase cistrons. The E/DLS or Psi$^+$ (5) is located at the 5′ end of this vector and extended to position 1040. (B) Data presented are from two independent experiments. The relative titer of each recombinant virus was normalized to that of pMLV-CB78. The average titer of pMLV-CB78 was 1,400 CFU/ml. The percentage of transfected cells giving rise to G418-resistant foci is expressed as the ratio of the number of G418^r foci to the number of cells transiently expressing the indicator gene (*lacZ* or phosphatase). The percentage of infected foci expressing the two genes was calculated as the ratio of infected G418^r foci expressing the indicator gene (*lacZ* or phosphatase) after 2 weeks of G418 selection to the number of infected G418^r foci.

The low level of β-Gal expression in the negative control is most likely due to termination-reinitiation between *neo* and *lacZ*. Furthermore, NIH 3T3 cells containing the negative control were much less intensely stained (data not shown). The dicistronic vector containing the entire F-MLV leader (pMLV-CB63) between the two reporter genes produced almost the same quantity of β-Gal as the monocistronic plasmid containing the F-MLV 5′ leader upstream of the *lacZ* gene (pMLV-CB93; Fig. 8). This result is in complete agreement with the in vitro translation data and confirms the presence of an internal ribosome entry in the F-MLV leader. Deletion of nucleotides 565 to 620 containing AUG^{gag} caused only a slight reduction in the level of β-Gal observed (Fig. 8, pMLV-CB61). In contrast, deletion of region 1-280 resulted in a fourfold decrease in β-Gal (Fig. 8, pMLV-CB94 versus pMLV-CB63 and pMLV-CB61 versus pMLV-CB50). These effects correlate with those observed in vitro but are amplified. Moreover, these results support the idea that the region 1-280 is important for the internal ribosome entry function of the F-MLV leader.

A new retroviral vector with the MLV leader inserted between two genes. The finding that the F-MLV leader contains a functional internal ribosome entry prompted us to construct a new dicistronic retroviral vector. In this DNA construct, the F-MLV leader (positions 212 to 620) was inserted between the phosphatase and neomycin genes (Fig. 9A, pMLV-CB78) flanked by the MLV 5′ long terminal repeat and PBS sequences, and by the 3′ long terminal repeat, respectively. Insertion of the F-MLV leader in the antisense orientation between the phosphatase and neomycin genes represented a negative control. As a positive control, we used the pLNPOZ dicistronic vector (2) in which the poliovirus IRES (42) was inserted between the neomycin and β-galactosidase genes. It should be noted that the Psi$^+$ sequence (also called E/DLS) (positions 212 to 1040) is present in pLNPOZ and is known to promote a 10- to 30-fold increase in MLV-derived RNA packaging compared with RNA containing only the leader (5).

Ecotropic GP-E+86 helper cells were transfected with these MLV-derived retroviral vectors, and short-term expression of β-Gal or phosphatase as well as long-term expression of neomycin by the transfected helper cells was monitored (see Materials and Methods). The number of transfected helper cells was measured by means of the short-term expression of β-galactosidase or phosphatase (data not shown). After 2 weeks of G418 selection (see Materials and Methods), the percentage of phosphatase-expressing transfected cells giving Neo^r foci was 2.1% with the pMLV-CB78 dicistronic vector and thus similar to the 2.3% of β-Gal-expressing transfected cells obtained with the pLNPOZ vector. The number of Neo^r foci was only 0.2% with the negative control where the F-MLV leader was in the antisense orientation (Fig. 9B).

The 5′ leader of MLV contains the packaging Psi (or E/DLS) sequence, which is a *cis*-acting element necessary for the specific encapsidation of the genomic RNA and MLV-derived RNA in MLV virions (1). To investigate the ability of the F-MLV leader inserted between two genes in a retroviral vector to direct RNA packaging, the recombinant viruses transiently expressed by the helper cells were used to infect fresh 3T3 cells. The titer of the dicistronic pMLV-CB78 vector was found to be about 1.4×10^3 CFU/ml (Fig. 9B). As expected, the titer of the pLNPOZ vector was about 14 times higher since it contains the Psi$^+$ packaging signal extending into *gag* known to enhance by 10- to 30-fold the encapsidation of recombinant MLV-derived RNA (Fig. 9B) (5). The titer of the dicistronic pMLV-CB74 vector with the leader in the antisense orientation was very low (Fig. 9B).

The infected 3T3 cells were put under G418 selection to monitor the long-term expression of neomycin and phosphatase or β-Gal. As shown in Fig. 9B, 91 and 80% of the infected Neo^r foci also expressed phosphatase and β-Gal, respectively. In conclusion, these results clearly show that the F-MLV leader inserted between two genes can promote both translation as an IRES and RNA packaging into MLV virions as Psi$^+$ (E/DLS). Further improvements of this dicistronic MLV vector are presently under way by inserting the leader and the 5′ *gag* sequences between the two genes (Psi$^+$) (5).

DISCUSSION

One important feature of retroviruses like MLV, Rous sarcoma virus, and human immunodeficiency virus type 1 (4, 13, 40, 50) is the presence of a long and well-structured leader at the 5′ end of the genomic RNA that is involved in key steps of the viral life cycle such as the initiation of *gag* precursor translation, the dimerization and encapsidation of the unspliced genomic RNA, and the initiation of reverse transcription (Fig. 1). The presence of a long stem-loop structure (ΔG of −50 kcal [ca. −200 kJ]/mol [16a]) as well as other structures between the 5′ cap of F-MLV RNA and the *glyco-gag* and *gag* initiation codons suggested that the initiation of translation of these polyprotein precursors might not proceed through the classical ribosome scanning (Fig. 1). In favor of this notion, it has been shown that stable secondary structures (ΔG of −50 kcal [ca. −200 kJ]/mol) inserted between the 5′ cap and the initiation codon can strongly inhibit the initiation process, probably by interfering with the scanning of 40S ribosomal subunits (29).

Through site-directed mutagenesis of the initiation codons of F-MLV *glyco-gag* and *gag* polyprotein precursors, we have

shown that the majority of initiation at the downstream AUG^{gag} did not result from leaky ribosome scanning between the CUG$^{glyco\text{-}gag}$ and the AUG^{gag} (Fig. 3). These results are in contrast to those obtained for *tat* and *rev* translation in equine infectious anemia virus. These two regulatory proteins are synthesized from the same viral subgenomic mRNA in which AUG^{rev} is downstream from CUG^{tat}. Translation of *rev* appears to rely on a leaky scanning since changing CUG^{tat} to AUG resulted in a drastic reduction of *rev* translation (8).

To examine whether the *gag* and *glyco-gag* translation is a cap-dependent or independent process, poliovirus protease 2A, which inhibits cap-dependent translation by inactivating the cap-binding complex eIF-4F (23), was coexpressed in NIH 3T3 cells with pMLV plasmids. The protease 2A had no effect on gene expression directed by the F-MLV leader (Fig. 4). These results strongly suggest that expression of *glyco-gag* and *gag* is initiated by a cap-independent mechanism and are in agreement with the in vitro translation data obtained with capped and uncapped F-MLV RNAs.

Translation initiation of picornavirus uncapped genomic RNA and human immunoglobulin heavy-chain-binding protein capped cellular mRNAs occurs by an internal ribosome entry mechanism (7, 21, 24–26, 37, 42). To investigate the possible presence of an IRES in the F-MLV leader, dicistronic vectors were constructed and assayed in vitro and in cell culture. The in vitro experiments clearly showed that sequences within the F-MLV 5′ leader between nucleotides 1 and 620 can direct the efficient internal initiation of protein synthesis (Fig. 6). In addition, preliminary electron microscopic observations showed that ribosomes directly bind to internal sequences of the 5′ leader of MLV RNA (43). The results obtained in cell culture are in agreement with in vitro experiments and confirm that the F-MLV leader possesses an internal entry for ribosomes. In picornavirus, the internal ribosome entry domain is a long, highly structured RNA element containing pyrimidine-rich regions (15, 25). Similar structured and pyrimidine-rich regions are located in the leader of F-MLV RNA upstream of CUG$^{glyco\text{-}gag}$ (positions 210 to 280) and AUG^{gag} (positions 510 to 580; Fig. 1) (40, 50). Removal of this sequence (positions 1 to 280) had a negative effect on translation (Fig. 6, lane 6 versus lane 4), indicating that this region is implicated in F-MLV internal ribosomal initiation. It is possible that, as was observed with picornavirus, the pyrimidine-rich region is an important element of the IRES (15, 25). It is also possible that removal of this region perturbs the secondary structure or alters the spacing of conserved elements in the IRES. On the basis of these results, we can envisage two possible models for the initiation of *gag* and *glyco-gag* synthesis. Translation of *glyco-gag* starts by the direct binding of the 40S ribosomal subunit within the leader upstream of CUG$^{glyco\text{-}gag}$ and necessitates the presence of structural elements and/or the pyrimidine-rich sequence within the region 1-280 (Fig. 1). Subsequently, the 40S ribosomal subunit scans to reach the CUG$^{glyco\text{-}gag}$, which is in a favorable context (34), as described for translation initiation of poliovirus (25). This model is in total agreement with the results of in vitro translation since we show that the 40S ribosomal subunits scan between positions 319 and 355 (Fig. 3, lanes 7 to 9). Initiation of *gag* synthesis appears to be much less clear since leaky ribosome scanning between the CUG$^{glyco\text{-}gag}$ and AUG^{gag} codons is unlikely to account for the majority of *gag* synthesis as indicated by our data (Fig. 3). It could be envisioned that the 40S ribosomal subunit, bound upstream of the CUG$^{glyco\text{-}gag}$, would be directly transferred (shunted) to the F-MLV 540-580 sequence, which is probably base paired to the ribosome binding region as indicated by the secondary structure model of the F-MLV leader (Fig. 1), and initiate the synthesis of *gag* protein (13, 19, 49). Or, it is possible that the 40S ribosomes bind to the RNA between the two initiation codons, a process facilitated by the presence of a highly structured RNA element containing pyrimidine-rich regions, and scan until they encounter the AUG^{gag} initiation codon. This model is supported by results obtained with pMLV-CB50 containing the region 280-565 of the leader (Fig. 6 and 8) in which synthesis of β-Gal was observed. Both models are in agreement with the observation of Evans et al., who showed that the levels of *glyco-gag* and *gag* proteins may vary independently according to the cell growth phase (17).

Initiation of picornavirus RNA translation clearly requires cell-specific factors that are thought to be RNA-binding proteins (7, 22). Interestingly, β-galactosidase expression was clearly lower in HeLa cells than in NIH 3T3 cells (data not shown). Thus, it is likely that in murine retroviruses the internal initiation of translation promoted by the 5′ leader requires specific cellular factors. This is presently under investigation.

Identification of an IRES in the F-MLV leader has resulted in the development of a new dicistronic MLV vector for gene transfer in which the F-MLV leader has been inserted between two genes. In this retroviral vector, the F-MLV IRES promotes efficient expression of the second gene in almost the same number of transfected or infected cells as the poliovirus IRES. Furthermore, insertion of the F-MLV leader between two genes did not strongly impair the packaging function (Fig. 9). New dicistronic MLV-derived vectors are being constructed to improve the packaging function by extending the leader into the 5′ *gag* sequences (5).

The presence of an IRES in the F-MLV leader adjacent to or overlapping other elements essential for the viral life cycle such as the packaging sequence (Psi or E/DLS) raises important questions. These elements are implicated in the control of *gag* translation and genomic RNA dimerization-packaging, two steps which are necessary for virus replication but which appear to be mutually exclusive (10, 43, 44). In hepatitis B virus, it has been shown that the binding of 80S ribosomes to the encapsidation signal of the genomic RNA prevents packaging (41). Therefore, one function of the IRES in the retroviral leader could be to regulate the balance between *gag* protein synthesis and genomic RNA packaging in the course of virion formation. This is presently under investigation.

ACKNOWLEDGMENTS

Thanks are due to M. Sitbon for the gift of F-MLV plasmid p57A, to N. Fouillot for the gift of pMLP-P2A and pMLP-RP2A, and to P. Savatier for the gift of pβ-actin-LacZ. We thank M. Lapadat-Tapolsky for a critical reading of the manuscript and C. Ehresmann for the secondary structure model of the F-MLV leader. Particular thanks are due to C. Torrent for encouragement and helpful discussions during this work.

This work was supported by a grant from the Fondation Merieux, by the French Association against AIDS (ANRS), and by the French Association against Cancer (ARC).

REFERENCES

1. **Adam, M. A., and A. D. Miller.** 1988. Identification of a signal in a murine retrovirus that is sufficient for packaging of nonretroviral RNA into virions. J. Virol. **62:**3802–3806.
2. **Adam, M. A., N. Ramesh, A. D. Miller, and W. R. Osborne.** 1990. Internal initiation of translation in retroviral vectors carrying picornavirus 5′ nontranslated regions. J. Virol. **65:**4985–4990.
3. **Allain, B., M. Lapadat-Tapolsky, C. Berlioz, and J.-L. Darlix.** 1994. Transactivation of the minus-strand DNA transfer by nucleocapsid protein during reverse transcription of the retroviral genome. EMBO J. **13:**973–981.
4. **Baudin, F., R. Marquet, C. Isel, J.-L. Darlix, B. Ehresmann, and C. Ehresmann.** 1993. Functional sites in the 5′ region of human immunodeficiency virus type 1 RNA form defined structural domains. J. Mol. Biol. **229:**382–397.

5. **Bender, M. A., T. D. Palmer, R. E. Gelinas, and A. D. Miller.** 1987. Evidence that the packaging signal of Moloney murine leukemia virus extends into the *gag* region. J. Virol. **61:**1639–1646.
6. **Borman, A., M. T. Howell, J. G. Patton, and R. J. Jackson.** 1993. The involvement of a spliceosome component in internal initiation of human rhinovirus RNA translation. J. Gen. Virol. **74:**1775–1788.
7. **Borman, A., and R. J. Jackson.** 1992. Initiation of translation of human rhinovirus RNA: mapping the internal ribosome entry site. Virology **188:** 685–696.
8. **Carroll, R., and D. Derse.** 1993. Translation of equine infectious anemia virus bicistronic *tat-rev* mRNA requires leaky ribosome scanning of the *tat* CTG initiation codon. J. Virol. **67:**1433–1440.
9. **Chen, C., and H. Okayama.** 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. **7:**2745–2752.
10. **Coffin, J.** 1984. Structure of the retroviral genome, p. 261–368. *In* R. Weiss, N. Teich, H. Warmus, and J. Coffin (ed.), RNA tumor viruses, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
11. **Corbin, A., A.-C. Prats, J.-L. Darlix, and M. Sitbon.** 1994. A nonstructural *gag*-encoded glycoprotein precursor is necessary for efficient spreading and pathogenesis of murine leukemia viruses. J. Virol. **68:**3857–3867.
12. **Cornille, F., Y. Mely, D. Ficheux, I. Savignol, D. Gerard, J.-L. Darlix, M. C. Fournié-Zaluski, and B. P. Roques.** 1990. Solid phase synthesis of the retroviral nucleocapsid protein NCp10 of moloney murine leukemia virus and related zinc fingers in free SH forms: influence of zinc chelation on structural and biochemical properties. Int. J. Pept. Protein Res. **36:**551–558.
13. **Darlix, J.-L., M. Zuker, and P.-F. Spahr.** 1982. Structure-function relationship of rous sarcoma virus leader RNA. Nucleic Acids Res. **17:**5183–5196.
14. **Dolph, P. J., J. Huang, and R. J. Schneider.** 1990. Translation by the adenovirus tripartite leader: elements which determine independence from cap-binding protein complex. J. Virol. **64:**2669–2677.
15. **Duke, G. M., M. A. Hoffman, and A. C. Palmenberg.** 1992. Sequence and structural elements that contribute to efficient encephalomyocarditis virus RNA translation. J. Virol. **66:**1602–1609.
16. **Edwards, S. A., and H. Fan.** 1979. *gag*-related polyproteins of Moloney murine leukemia virus: evidence for independent synthesis of glycosylated and unglycosylated forms. J. Virol. **30:**551–563.

16a. **Ehresmann, C.** Personal communication.

17. **Evans, L., S. Dresler, and D. Kabat.** 1977. Synthesis and glycosylation of polyprotein precursors to the internal core proteins of Friend murine leukemia virus. J. Virol. **24:**865–874.
18. **Fouillot, N., S. Tlouzeau, J. M. Rossignol, and O. Jean-Jean.** 1993. Translation of the hepatitis B virus P gene by the ribosomal scanning as an alternative to internal initiation. J. Virol. **67:**4886–4895.
19. **Fütterer, J., Z. Kiss-Laszio, and T. Hohn.** 1993. Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA. Cell **73:**789–802.
20. **Ghattas, I., J. Sanes, and J. Majors.** 1991. The encephalomyocarditis virus internal ribosome entry site allows efficient coexpression of two genes from a recombinant provirus in cultured cells and in embryos. Mol. Cell. Biol. **11:**5848–5859.
21. **Glass, M. J., X.-Y. Jia, and D. F. Summers.** 1993. Identification of the hepatitis A virus internal ribosome entry site: in vivo and in vitro analysis of bicistronic RNAs containing the HAV 5′ noncoding region. Virology **193:** 842–852.

21a. **Goff, D.** Unpublished data.

22. **Hellen, C. U., G. W. Witherell, M. Schmid, S. H. Shin, T. V. Pestova, A. Gil, and E. Wimmer.** 1993. A cytoplasmic 57-kDa protein that is required for translation of picornavirus RNA by internal ribosomal entry is identical to the nuclear pyrimidine tract-binding protein. Proc. Natl. Acad. Sci. USA **90:**7642–7646.
23. **Hellen, C. U. T., M. Fäcke, H.-G. Kräusslich, C.-K. Lee, and E. Wimmer.** 1991. Characterization of poliovirus 2A proteinase by mutational analysis: residues required for autocatalytic activity are essential for induction of cleavage of eukaryotic initiation factor 4F polypeptide p220. J. Virol. **65:** 4226–4231.
24. **Jackson, R. J.** 1991. Initiation without an end. Nature (London) **353:**14–15.
25. **Jackson, R. J., M. T. Howell, and A. Kaminski.** 1990. The novel mechanism of initiation of picornavirus RNA translation. Trends Biochem. Sci. **15:**477–483.
26. **Jang, S. K., H. Kräusslich, M. J. H. Nicklin, G. M. Duke, A. C. Palmenberg, and E. Wimmer.** 1988. A segment of the 5′ nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. J. Virol. **62:**2636–2643.
27. **Kalnins, A., K. Otto, U. Rütter, and B. Müller-Hill.** 1983. Sequence of the *LacZ* gene of Escherichia coli. EMBO J. **2:**593–597.
28. **Khandjian, E., and C. Meric.** 1986. A procedure for northern blot analysis of native RNA. Anal. Biochem. **159:**80–85.
29. **Kozak, M.** 1986. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. Proc. Natl. Acad. Sci. USA **83:**2850–2854.
30. **Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44:**283–292.
31. **Kozak, M.** 1987. Effects of intercistronic length on the efficiency of reinitiation by eukaryotic ribosomes. Mol. Cell. Biol. **7:**3438–3445.
32. **Kozak, M.** 1989. Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. Mol. Cell. Biol. **9:**5134–5142.
33. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell Biol. **108:**229–241.
34. **Kozak, M.** 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. J. Cell Biol. **115:**887–903.
35. **Kunkel, T.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82:**488–492.
36. **Ledbetter, J., and R. C. Nowinski.** 1977. Identification of the Gross cell surface antigen associated with murine leukemia virus-infected cells. J. Virol. **23:**315–322.
37. **Macejak, D., and P. Sarnow.** 1991. Internal initiation of translation mediated by the 5′ leader of a cellular mRNA. Nature (London) **353:**90–94.
38. **Markowitz, D., S. P. Goff, and A. Bank.** 1988. A safe packaging line for gene transfer: separating viral genes on two different plasmids. J. Virol. **62:**1120–1124.
39. **Merrick, W.** 1992. Mechanism and regulation of eukaryotic protein synthesis. Microbiol. Rev. **56:**291–315.
40. **Mougel, M., N. Tounekti, J.-L. Darlix, J. Paoletti, B. Ehresmann, and C. Ehresmann.** 1993. Conformational analysis of the 5′ leader and the *gag* initiation site of Mo-MuLV RNA and allosteric transitions induced by dimerization. Nucleic Acids Res. **21:**4677–4684.
41. **Nassal, M., M. Junker-Niepmann, and H. Schaller.** 1990. Translational inactivation of RNA function: discrimination against a subset of genomic transcripts during HBV nucleocapsid assembly. Cell **63:**1357–1363.
42. **Pelletier, J., G. Kaplan, V. R. Racaniello, and N. Sonenberg.** 1988. Cap-independent translation of poliovirus mRNA is conferred by sequence elements within the 5′ noncoding region. Mol. Cell. Biol. **8:**1103–1112.
43. **Prats, A. C.** 1988. Thesis. P. Sabatier University, Toulouse, France.
44. **Prats, A. C., G. De Billy, P. Wang, and J.-L. Darlix.** 1989. CUG initiation codon used for the synthesis of the cell surface antigen coded by the murine leukemia virus. J. Mol. Biol. **205:**363–372.
45. **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.
46. **Saris, C., J. van Eenbergen, R. Liskamp, and H. Bloemers.** 1983. Structure of glycosylated and unglycosylated *gag* and *gag-pol* precursor proteins of Moloney murine leukemia virus. J. Virol. **46:**841–859.
47. **Savatier, P., J. Morgenstern, and R. S. P. Beddington.** 1990. Permissiveness to murine leukemia virus expression during preimplantation and early postimplantation mouse development. Development **109:**655–665.
48. **Sitbon, M., B. Sola, L. Evans, J. Nishio, S. F. Hayes, K. Nathanson, C. F. Garon, and B. Chesebro.** 1986. Hemolytic anemia and erythroleukemia, two distinct pathogenic effects of friend MuLV: mapping of the effects to different regions of the viral genome. Cell **47:**851–859.
49. **Sonenberg, N.** 1993. Remarks on the mechanism of ribosome binding to eukaryotic mRNAs. Gene Expr. **3:**317–323.
50. **Tounekti, N., M. Mougel, C. Roy, R. Marquet, J.-L. Darlix, J. Paoletti, B. Ehresmann, and C. Ehresmann.** 1992. Effect of dimerization on the conformation of the encapsidation Psi domain of moloney murine leukemia virus RNA. J. Mol. Biol. **223:**205–220.
51. **Weber, L. A., E. D. Hickey, and C. Baglioni.** 1978. Influence of potassium salt concentration and temperature on inhibition of mRNA translation by 7-methylguanosine 5′-monophosphate. J. Biol. Chem. **253:**178–183.