Proteins Encoded by the Long Terminal Repeat Region of Mouse Mammary Tumor Virus: Identification by Hybrid-Selected Translation

JANIS RACEVSKIS* AND OM PRAKASH

Memorial Sloan-Kettering Cancer Center, New York, New York 10021

Received 14 March 1984/Accepted 15 May 1984

The long terminal repeat (LTR) region of mouse mammary tumor virus (MMTV) is known to contain an open reading frame of sufficient length to code for a protein of 36,000 M_r . The coding capacity of the 3' sequences of MMTV genomic RNA has been demonstrated by *in vitro* translation studies, which have reported the synthesis of four related proteins: p36, p24, p21, and p18. These proteins are overlapping translation products of the same open reading frame, with the smaller ones initiating at internal methionine codons. From the predicted amino acid sequence of the LTR protein, we have selected a region likely to be antigenic, obtained a synthetic peptide of that region, and raised antiserum to the peptide. The antipeptide serum specifically immunoprecipitated all four proteins from *in vitro* translated genomic 3' MMTV RNA, plus an additional one of 32,000 M_r . Published sequence data of MMTV LTRs show an internal AUG codon at a position which could initiate a protein of 32,000 M_r . The three smaller *in vitro* translation products (p24, p21, and p18) were consistently synthesized in much greater amounts than the p36 or p32 protein. The relative amount of each *in vitro* synthesized protein from genomic MMTV RNA could be predicted and was in good agreement with the postulated effect of flanking nucleotides on the efficiency of the respective AUG initiation codon. Polyadenylated RNAs, isolated from various mouse tissues, were selected by hybridization to plasmid DNA containing MMTV LTR sequences immobilized on nitrocellulose. *In vitro* translation of hybrid-selected mRNAs isolated from BALB/c mouse lactating mammary glands and carcinogen-induced mammary tumors, followed by immunoprecipitation with antipeptide serum, revealed that only one polypeptide was synthesized by the MMTV LTR-specific mRNA, the 36,000 M_r species.

The long terminal repeat (LTR) of mouse mammary tumor virus (MMTV) is unique among retroviral LTRs in that it contains a long open reading frame in its U3 region (30). The presence of an open reading frame in the MMTV LTR was discovered by DNA sequencing of the proviral LTR (7) and by *in vitro* translation of the 3' end of genomic MMTV RNA (4, 28). The open reading frame in the MMTV LTR has been found to be conserved in all MMTV strains analyzed to date (6, 7, 10, 13, 17) and appears to have a coding capacity sufficient to code for a 36,000 M_r protein. *In vitro* translation studies (4, 22, 28) have reported that four polypeptides, unrelated to any MMTV structural proteins, could be synthesized from the 3' end of genomic MMTV RNA: polypeptides of 36,000, 24,000, 21,000, and 18,000 M_r . The same four polypeptides were also observed after translation of RNA transcribed from cloned MMTV LTR sequences (5), thus firmly establishing their viral origin. The *in vitro* protein products of the LTR appear to be overlapping translation products of the same open reading frame (4, 22, 28), with the smaller polypeptides probably initiating at internal AUG codons. In the *in vitro* translation systems, the three smaller polypeptides (24,000, 21,000, and 18,000 M_r) are always synthesized in greater amounts than the 36,000 M_r protein (4, 22, 28).

Search for expression of the MMTV LTR gene *in vivo* has revealed the presence, in certain mouse tissues, of a spliced 1.6-kilobase (kb) mRNA species hybridizable with MMTV LTR sequences (33). We have raised antisera to the MMTV LTR gene translation products, using a synthetic peptide

predicted by the DNA sequence, and have found that a 36,000 M_r protein precipitable with the antiserum is synthesized *in vitro* from hybrid-selected MMTV LTR-specific, 1.6-kb mRNA isolated from lactating mammary glands of BALB/c mice and 7,12-dimethylbenz(a)anthracene (DMBA)-induced mammary tumors of BALB/c mice.

MATERIALS AND METHODS

Antipeptide serum. The 15-amino acid-long synthetic peptide was purchased from Peninsula Laboratories and was coupled to the carrier protein keyhole limpet hemocyanin with the coupling reagent *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester according to the procedures of Lerner et al. (15). Rabbits were immunized with the peptide-keyhole limpet hemocyanin conjugate according to the following schedule (15): 500 μ g emulsified in complete Freund adjuvant subcutaneously on day 0; 500 μ g in incomplete Freund adjuvant on days 14, 21, and 91. Animals were bled 15 weeks after the first injection and at 2-month intervals after booster injections.

Activity of the antiserum was tested by immunoprecipitation of [35 S]methionine-labeled *in vitro* translation products of C3H MMTV genomic RNA. MMTV C3H virus, isolated from cell line Mm5mt/c1 (20), was used as the source of viral RNA and was provided by the Intramural Viral Resources Program, National Cancer Institute. Sucrose gradient-purified 35S viral RNA was a gift from P. Etkind of this institute. Viral RNA was translated *in vitro* in a rabbit reticulocyte lysate system (New England Nuclear) in the presence of [35 S]methionine (New England Nuclear). The immunoglobulin G (IgG) fraction of the highest-titer antiserum was

* Corresponding author.

isolated by affinity chromatography on a column of the synthetic peptide coupled to CH-Sepharose 4B (Pharmacia Fine Chemicals, Inc.), all as described by Walter et al. (32).

Immunoprecipitation. Initial screening of antisera was performed by adding 5 to 10 μ l of serum to an aliquot of in vitro translated MMTV proteins in reticulocyte lysate, diluted in lysis buffer (26): 20 mM Tris-hydrochloride (pH 7.5), 50 mM NaCl, 0.5% Nonidet P-40, 0.5% sodium deoxycholate. Immune complexes were precipitated with protein A-Sepharose (Pharmacia) as described previously (26). Immunoprecipitates were analyzed by electrophoresis on 5 to 20% exponential gradient polyacrylamide gels followed by fluorography of the gels (24).

A more quantitative recovery of MMTV LTR-specific proteins was obtained by using an indirect immunoprecipitation technique (2). Antiserum- or affinity-purified IgGs were initially preadsorbed to a 10% (vol/vol) Sepharose-protein A suspension in phosphate-buffered saline (pH 7.2)–0.5% Tween 20–0.1% bovine serum albumin (PTA) for 2 h at 4°C with rotation. The Sepharose-protein A was then washed with PTA and aliquoted into samples to be analyzed, which were then allowed to react overnight at 4°C with continual rotation. The Sepharose-protein A antigen complexes were then washed with PTA and analyzed as described above.

RNA purification. RNA was isolated from lactating mammary glands (LMG) and livers of BALB/c mice (Jackson Laboratories), LMG of GR mice (Institute for Medical Research), transplantable DMBA-induced mammary tumors of BALB/c mice (29) (obtained from N. Telang of this institute), and transplantable leukemic cells of DBA/2 mice (24). RNA was extracted from tissues by the guanidinium isothiocyanate-cesium chloride method (18). Tissues were frozen in liquid nitrogen, suspended in 8 volumes of 6 M guanidinium isothiocyanate–5 mM sodium citrate (pH 7.0)–0.1 M β -mercaptoethanol–0.5% Sarkosyl, and then homogenized in a Sorvall Omnimixer. Addition of cesium chloride to the homogenates, layering on cesium chloride cushions, centrifugation, and RNA extraction were all performed as described before (18). Polyadenylated [poly(A)⁺] RNA was selected by passage over an oligodeoxythymidilic acid-cellulose column as previously described (28). Analysis of poly(A)⁺ RNAs was performed with formaldehyde-containing agarose gels (27), with modifications according to Prakash et al. (23). RNAs were electrophoresed in 1.4% agarose gels containing MAE buffer (20 mM MOPS [morpholinepropanesulfonic acid], pH 7.0, 5 mM sodium acetate, 1 mM EDTA), 2.2 M formaldehyde, and 0.5 μ g of ethidium bromide per ml. RNA samples, 10 μ g, were denatured in 25 μ l of MAE buffer containing 5% formamide, 2.2 M formaldehyde, and 0.02% bromophenol blue by heating at 60°C for 2 min before loading on the gel. The electrophoresis was carried out for 3 h at 150 V, using MAE buffer in the reservoirs. The reservoir buffer was circulated to eliminate concentration gradients. The gel was soaked for 30 min in 20 $\times$ SSC (SSC = 0.15 M NaCl plus 0.015 M sodium citrate) before photographing under UV light, and then the RNA was transferred to nitrocellulose (31). Hybridizations were carried out in 50% formamide-containing solutions at 42°C according to Thomas (31). Labeled, nick-translated probe was prepared from the cloned 1.4-kb *Pst*I DNA fragment of the 5' MMTV LTR region (16) (kindly provided by J. Majors and H. Varmus).

Hybrid selection. Hybrid selection was performed essentially as described by Pachl et al. (21). The plasmid DNA used for preparation of filters was the pBR322 clone containing the 1.4-kb *Pst*I fragment from the left end of the C3H

MMTV DNA (16). The plasmid DNA was linearized by cleavage with *Eco*RI, extracted with phenol-chloroform, and ethanol precipitated. A 4- μ g amount of linearized, denatured plasmid DNA was spotted per 6-mm nitrocellulose filter circle. After washing and baking the filters, they were placed in hybridization solution: 50% formamide, 0.4 M NaCl, 10 mM PIPES [piperazine-*N,N'*-bis(2-ethanesulfonic acid)], pH 6.4, 0.2% sodium dodecyl sulfate, 100 μ g of calf liver tRNA per ml. Poly(A)⁺ RNA was added, and hybridization was carried out in a volume of 0.1 ml at 42°C for 18 h. The filters were washed, and the RNA was eluted as described previously (21). The RNA was coprecipitated from ethanol with 5 μ g of calf liver tRNA, washed, and translated in a rabbit reticulocyte cell-free system.

RESULTS

Antipeptide serum. From the predicted amino acid sequence of the C3H exogenous MMTV LTR open reading frame (7), we chose a 14-amino acid-long region as having a high probability of being antigenic on the basis of its high average amino acid hydrophilicity value. In choosing this region we followed the method of Hopp and Woods (12) in which likely antigenic sites are identified by locating regions of high hydrophilicity in the amino acid chain, by assigning numerical hydrophilicity values to all constituent amino acids. Regions containing the hexapeptide with the highest average hydrophilicity value were found to be antigenic in 12 of 12 proteins studied (12). The 14-amino acid sequence is specified by codons 165 through 178 (numbering the codons starting with the AUG at the beginning of the open reading frame [17]; Fig. 1) and has the sequence: Ile-Glu-Asn-Arg-Lys-Arg-Arg-Ser-Thr-Ser-Ile-Glu-Glu-Gln. This peptide is encoded by the sequence beginning two codons downstream from the fourth internal AUG codon (6), a region that we can deduce codes for the amino-terminal end of the 18,000 M_r in vitro translation product (Fig. 1). As it turns out, this sequence is highly conserved and the hydrophilic hexapeptide 166–171 (Glu-Asn-Arg-Lys-Arg-Arg) is identical in all four known MMTV sequences (6, 7, 10, 13, 17). Only single amino acid differences, all in the carboxy-terminal half of the chosen 14-amino acid sequence, are present in the four sequenced MMTV strains. An equally hydrophilic hexapeptide is located at position 267–272; however, the adjacent amino acids are less hydrophilic than those flanking peptide 166–171.

A 15-amino acid-long synthetic peptide containing the chosen 14-amino acid sequence plus a noncoded cysteine at the carboxy terminus, for attachment to a carrier protein, was obtained. The synthetic peptide was coupled to keyhole limpet hemocyanin via the cysteine residue with the bifunctional reagent *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester, following the published procedures of Lerner et al. (15). Rabbits were immunized with the peptide carrier protein conjugate according to the protocol of Lerner et al. (15),

FIG. 1. Schematic representation of the 319-codon-long open reading frame of the C3H exogenous MMTV LTR, according to Donehower et al. (6). Numbers on bottom denote codon number, and arrows indicate position of AUG codons. In vitro LTR translation products are marked above their putative initiation codons. Shaded box indicates location of synthetic peptide (codons 165 to 178).

and sera were tested for activity by immunoprecipitation of the in vitro translation products of exogenous C3H MMTV genomic RNA. After a 15-week immunization protocol, all three rabbits developed some precipitating antibody activity against the in vitro LTR proteins (pLTRs). One of the rabbits developed a considerably stronger response than the other two, and all subsequent experiments, including affinity purification of IgGs, were performed with serum from this one animal. Further booster injections were also given, and higher-activity serum was obtained. In vitro translations were performed in a rabbit reticulocyte lysate system, with [35 S]methionine labeling and using a 35S fraction of genomic MMTV RNA as a template. The viral RNA used in these studies was isolated from exogenous C3H strain virus produced by cell line Mm5mt/c1 (20). A whole range of proteins were synthesized from this template (Fig. 2, lane A), including the high-molecular-weight *gag-pol* and *gag* precursors (4, 28) of 160,000, 105,000, and 75,000 M_r , which were precipitated with anti-MMTV core protein p28 serum (Fig. 2, lane D). Immunoprecipitation with an anti-pLTR serum from an early bleeding showed that at least four proteins were

FIG. 2. [35 S]methionine-labeled in vitro translation products of C3H MMTV 35S genomic RNA (lane A), separated on a polyacrylamide gel, immunoprecipitated with: (B) nonimmune serum; (C) anti-pLTR peptide serum; (D) anti-MMTV core protein p28 serum. Molecular weights of proteins (in thousands) are marked on sides.

FIG. 3. [35 S]methionine-labeled in vitro translation products of MMTV 35S genomic RNA (lane A), immunoprecipitated with: (B) nonimmune serum; (C) anti-pLTR peptide serum; (D) anti-pLTR peptide serum in the presence of 5 μ g of synthetic peptide; (E) affinity-purified IgG fraction of anti-pLTR peptide serum.

specifically precipitated (Fig. 2, lane C). Immunoprecipitation of the in vitro products with a later, higher-titer serum as well as with affinity-purified IgG of this serum, conclusively demonstrated the specific precipitation of the four polypeptides, plus that of a fifth one of 32,000 M_r (Fig. 3, lanes C and E). In general, it was observed that equally good results were obtained whether unfractionated serum or a purified IgG fraction was used for immunoprecipitation with the indirect immunoadsorption technique (2). The precipitation of these five polypeptides could be completely abolished by the addition of a few micrograms of the synthetic peptide to act as a competitive inhibitor (Fig. 3, lane D). Addition of synthetic peptide had no effect on immunoprecipitation with anti-core protein p28 serum (data not shown). The synthesis of the 32,000 M_r protein was probably not previously reported (4, 22, 28) because it is synthesized in relatively low levels. Examination of the published MMTV LTR sequence data (6, 10, 13) shows that an AUG codon is present in all four MMTV strains at a position (codon 38) which could initiate a polypeptide of 32,000 M_r . The one exception is the

sequence of the C3H virus published by Majors and Varmus (17), which shows a valine codon (GUG) at position 38 of the LTR open reading frame. This result (17) is at variance with the sequence published by Donehower et al. (6) since both studies were performed with exogenous C3H virus isolated from the same cell line: Mm5mt/c1 (20). The C3H MMTV LTR open reading frame contains four additional AUG codons located near the 3' end of the open reading frame. Any peptides initiated at these codons would not be detected with our antipeptide serum since these codons are located downstream from the chosen peptide sequence. The relative amounts of each of the five polypeptides synthesized in vitro were observed to be consistent from experiment to experiment. The three smaller polypeptides (24,000, 21,000, and 18,000 M_r) were always synthesized in greater quantities than the 36,000 M_r protein and the 32,000 M_r species, which was always the least abundant (Fig. 4). Taking into consideration the fact that the densitometer scan (Fig. 4) is a measure of incorporated [35 S]methionine, the actual disparity in the relative amounts of synthesis of these proteins is even greater, since the smaller polypeptides have fewer methionine residues. In addition, since the proteins were separated on a polyacrylamide gradient gel, there is a direct relationship between the molecular weight of a protein and the efficiency of detecting a fluorographic signal from the labeled protein, because the smaller proteins are embedded in the thicker part of the dried gel where signals are quenched to a greater degree (11). The ratio of the five LTR proteins to each other is the same in the translation mix (Fig. 3, lane A) as it is in the immunoprecipitates (Fig. 3, lanes C and E), demonstrating that the antiserum has equal avidity for all five polypeptides. An attractive explanation for the different levels of synthesis of the five LTR proteins is the postulated effect of flanking nucleotides on the efficiency of an AUG codon to act as an initiation site (14). Analysis of the published sequence of MMTV LTRs shows that the putative initiative codon for the 32,000 M_r protein, which is least abundant, is in the least favored of all possible arrangements, with a pyrimidine in the -3 position: $^{-3}YXXAUGX^{-4}$.

Hybrid-selected translation. Various mouse tissues were analyzed for the presence of the reported MMTV LTR-specific, spliced, 1.6-kb mRNA species (33) by hybridization of Northern blots of extracted poly(A)⁺ RNAs with cloned MMTV LTR probes. A specific band migrating faster than 18S rRNA could be detected in blots of RNAs extracted from lactating BALB/c mouse mammary glands (Fig. 5, lane B) and DMBA-induced mammary tumors of BALB/c mice (Fig. 5, lane C). No other MMTV-specific mRNAs could be detected in these tissues, which is consistent with the report that carcinogen DMBA-induced mammary tumors of BALB/c mice do not express detectable levels of any MMTV structural proteins (9). No MMTV structural pro-

FIG. 4. Densitometer tracing of lane C, Fig. 3. Immunoprecipitate of MMTV RNA translation products with anti-pLTR peptide serum.

FIG. 5. Northern blot analysis of poly(A)⁺ RNAs extracted from: (A) BALB/c mouse liver; (B) BALB/c LMG; (C) DMBA-induced BALB/c mammary tumor; (D) DBA/2 T-cell lymphoma; (E) GR LMG. A 10- μ g portion of poly(A)⁺ RNA (passed once through an oligodeoxythymidylate column) was applied to each lane. Hybridization was done with the 1.4-kb *Pst*I fragment containing the MMTV LTR sequence. The film was exposed for 48 h, except for lane D, which was exposed for 12 h [the RNA in lane D was twice poly(A) selected].

teins could be detected in the DMBA-induced mammary tumors or in LMG of BALB/c mice by the technique of protein blotting and reaction with anti-MMTV serum (data not shown). Analysis of GR mouse LMG RNA (Fig. 5, lane E) revealed the presence of the two RNA species usually found in MMTV-producing cells (25): the genome-length, 35S, 7.8-kb-long RNA, and the 24S, 3.8-kb-long envelope gene product. No additional, smaller, MMTV-specific RNAs could be detected in the GR LMG, even after prolonged exposures. A number of MMTV-specific RNAs were observed in the sample extracted from the MMTV antigen-expressing (24) T-cell lymphoma of DBA/2 mice (Fig. 5, lane D). In addition to the 35S RNA and a band migrating slightly faster than 24S, there appear to be two lower-molecular-weight species: one of about 1.6 kb and another still smaller (Fig. 5, lane D).

MMTV-specific mRNAs were selected by hybridization to cloned MMTV LTR sequences immobilized on nitrocellulose filters as described in Materials and Methods. Hybridized mRNAs were eluted by boiling and then ethanol precipitated, washed, and translated in a rabbit reticulocyte lysate in vitro protein translation system. The [35 S]methionine-labeled products of in vitro translated mRNAs were immunoprecipitated with antipeptide serum as well as a polyclonal

anti-MMTV serum, and the precipitates were analyzed on polyacrylamide gels (Fig. 6). In addition to the DMBA-induced mammary tumors and LMG of BALB/c mice, we also analyzed LMG of GR mice and T-cell lymphomas of DBA/2 mice. LMG of the high mammary tumor incidence strain GR produce large quantities of virus (1), and the T-cell lymphomas of DBA/2 mice express high levels of MMTV antigens, although they do not secrete mature virions (24). The most striking result was that only one LTR-specific protein, the 36,000 M_r species, was synthesized from the mRNAs of the BALB/c LMG and mammary tumors (Fig. 6, lanes C and G). Precipitation of the 36,000 M_r translation product was completely abolished by the addition of a few micrograms of synthetic peptide (Fig. 7, lanes D and F). Synthesis of other MMTV-specific proteins could not be detected in the translation mixtures of mRNAs isolated from BALB/c tissues (Fig. 6, lanes D and H). The pLTR pattern obtained from the translation of mRNAs of GR LMG (Fig. 6, lane A; Fig. 7, lane A) was very similar to patterns obtained from translation of genomic RNA, probably reflecting the fact that most if not all of the MMTV LTR translation occurred from genomic RNA. The "p36" LTR protein of GR virus appeared to have a slightly higher molecular weight than the analogous protein of BALB/c or C3H. An interesting result was obtained from translation of mRNAs isolated from DBA/2 T-cell lymphomas: the predominant LTR-specific protein was the 32,000 M_r polypeptide (Fig. 6, lane E; Fig. 7, lane G). Immunoprecipitation with polyclonal anti-MMTV serum showed that a whole range of MMTV structural proteins and precursors were synthesized by the poly(A)⁺ RNAs selected from GR LMG and DBA/2 lymphomas (Fig. 6, lanes B and F).

FIG. 6. Analysis of in vitro translation products encoded by hybrid-selected RNAs of GR LMG (lanes A and B); DMBA-induced mammary tumor of BALB/c mouse (lanes C and D); transplantable T-cell lymphoma of DBA/2 mice (lanes E and F); LMG of BALB/c mice (lanes G and H). Translation products were immunoprecipitated with anti-pLTR peptide serum (lanes A, C, E, and G) and polyclonal anti-MMTV serum (lanes B, D, F, and H). Lane I is an aliquot of in vitro translation products.

FIG. 7. Analysis of in vitro translation products encoded by hybrid-selected RNAs of: GR LMG (lanes A and B); DMBA-induced mammary tumor of BALB/c mouse (lanes C and D); LMG of BALB/c mice (lanes E and F); T-cell lymphoma of DBA/2 mice (lanes G and H). The translation products were immunoprecipitated with anti-pLTR peptide serum (lanes A, C, E, and G) and anti-pLTR peptide serum in the presence of 5 μ g of the synthetic peptide (lanes B, D, F, and H).

Immunoprecipitation analysis, with anti-p-LTR serum, of extracts of in vitro labeled primary cultures of LMG and mammary tumors, or freshly excised minced tissues, yielded equivocal results because of high background levels and cross-reactions with unrelated proteins. Cross-reactions with otherwise unrelated proteins, especially with abundant ones, have been observed when antipeptide sera are used to analyze whole-cell extracts (19). This phenomenon is probably because antibodies to a single linear sequence will also recognize with some affinity a closely related linear sequence which might be present by chance on proteins that are otherwise unrelated (19). Antisera prepared against synthetic peptides have weak precipitating activity and do not effect quantitative recovery of antigens, even in antibody excess; thus, resolution of this problem will have to await the production of hyperimmune anti-pLTR polyclonal serum.

Hybrid-selected translation analysis was also performed on RNAs extracted from tissue culture cells of cell lines of normal BALB/c mammary glands as well as mammary tumors; no LTR proteins could be detected.

DISCUSSION

To search for expression of the MMTV LTR gene product, we raised antiserum against a synthetic peptide predicted by the DNA sequence of the C3H exogenous MMTV LTR open reading frame (6, 7, 17). The chosen 14-amino

acid sequence lies near the presumed amino terminus of the 18,000 M_r in vitro translation product (4, 22, 28). The antiserum was found to immunoprecipitate all four previously reported (4, 22, 28) LTR translation products, plus low levels of a fifth one of 32,000 M_r .

In vitro translation of genomic MMTV 35S RNA always produced the five LTR proteins in the same quantitative ratio, with the three smaller polypeptides (p24, p21, and p18) being the most abundant. In terms of total amount of synthesis, as measured by [35 S]methionine incorporation, the five LTR polypeptides were always synthesized in the following relative order: p24 > p18 > p21 > p36 > p32. The difference between the most abundant protein (p24) and the least abundant one (p32) was at least 20-fold (Fig. 4). If the observed quantitative ratios of the in vitro LTR products are caused by RNA degradation, then one would expect to observe a linear inverse relationship between the molecular weight of each LTR protein and the quantity synthesized; this is not the case. RNA analysis and in vitro synthesis of Pr75^{gag} and Pr180^{gag-pol} indicate that degradation of RNA is minimal. Furthermore, the same pattern was also previously observed by Peters et al. (22) upon in vitro translation of three different viral RNAs as well as four different plasmid-derived cRNAs. One of the most attractive explanations for the observed variation in rates of translation is the postulated effect of flanking nucleotides on the efficiency of a given AUG codon to act as an initiation signal of eucaryotic translation, as put forth by Kozak (14). Kozak (14) has shown that nucleotides in the -3 and +4 positions relative to an AUG (XXXAUGX) codon exert a pronounced effect on the efficiency of that codon and the frequency with which that codon is used for eucaryotic initiation of translation. In Table 1, the seven arrangements of initiation codons and nucleotides -3 and +4 are listed in decreasing frequency (top to bottom) of utilization as initiation signals in eucaryotic mRNAs (14). Listed alongside the appropriate codon arrangement is the MMTV LTR translation product which is assumed to be initiated on that particular codon arrangement (6, 7, 17). Comparison of Table 1 and Fig. 4 demonstrates that there is a very good correlation between the relative amount of in vitro synthesis of each LTR protein from a genomic RNA template and the relative efficiency of its respective initiation codon. The proposed initiation codon hierarchy is especially successful at predicting the very low level of synthesis of p32, whose initiation codon is in the least favored of all arrangements. The selective importance of the initiation codon, however, holds only for the in vitro translation of genomic MMTV RNA. If the natural mRNA is used as template, however, then only the 36,000 M_r polypeptide is synthesized, as demonstrated by using hybrid-selected LTR-specific mRNAs of BALB/c LMG and DMBA-induced mammary tumors. This result confirms previous

observations (14) that the primary determinant of a functional initiation site of an mRNA is its proximity to the 5' end of the message. Although we cannot discount the possibility that the activated LTR in BALB/c tissues only contains the AUG for p36, this seems rather remote since the other four AUG codons have been conserved in all four sequenced MMTV strains; furthermore, only one LTR product appears to be synthesized by the DBA/2 LTR mRNA as well.

The transcription of the MMTV LTR gene into a spliced mRNA (33), which we have demonstrated to be a functional message coding for a 36,000 M_r protein, in two types of BALB/c mouse tissues indicates the probable activation of an endogenous provirus. The expression of MMTV LTR sequences in BALB/c mouse LMG was foretold by earlier hybridization studies which detected an elevated level of 3'-end MMTV RNA expression in these tissues (8). Since BALB/c mice do not carry a milk-borne or exogenous virus, a possible source of the LTR products might be the incomplete endogenous provirus unit I (3). In vitro translation of hybrid-selected mRNAs from MMTV-producing tissues such as GR mouse LMG does not give an answer as to whether functional LTR mRNAs are present, since what we observe is probably primarily translation of genomic MMTV RNA. In addition, no LTR-specific, small mRNA species could be detected in the GR tissues by Northern blot analysis. A similar analysis of RNAs isolated from a T-cell lymphoma of DBA/2 mice indicates the presence of two small, LTR-specific RNAs and yields the interesting observation that only the 32,000 M_r LTR translation product is produced. The synthesis of a single product suggests the presence of a functional LTR mRNA. This lymphoma cell line produces high levels of both MMTV *gag* and *env* gene products (24); it does not, however, produce mature virions. The synthesis of the 32,000 M_r polypeptide perhaps indicates that its initiation codon in the DBA/2 provirus differs or that the pLTR gene in the activated DBA/2 provirus is shorter or lacks the first AUG codon.

The possible function of the MMTV LTR gene product can only be the subject of speculation at this time; however, a tantalizing clue is our failure to detect LTR proteins in tissue culture cells of BALB/c mouse mammary glands. One of the alterations in cells of mouse mammary gland tissue culture cell lines as compared to cells in the original tissue is loss of the ability to respond to certain hormones, including prolactin (34).

ACKNOWLEDGMENTS

We are grateful to P. Etkind for her gift of the purified viral RNA and to N. Telang for the transplantable mouse mammary tumors. We thank R. Kopelman for expert technical assistance.

This work was supported by Public Health Service grant CA-16599 from the National Cancer Institute.

LITERATURE CITED

1. Bentvelzen, P., and J. Hilgers. 1980. The murine mammary tumor virus, p. 311-355. In G. Klein (ed.), *Viral oncology*. Raven Press, New York.
2. Bunol, T. F., and R. A. Reisfeld. 1982. Unique glycoprotein-proteoglycan complex defined by monoclonal antibody on human melanoma cells. *Proc. Natl. Acad. Sci. U.S.A.* **79**:1245-1249.
3. Cohen, J. C., J. E. Majors, and H. E. Varmus. 1979. Organization of mouse mammary tumor virus-specific DNA endogenous to BALB/c mice. *J. Virol.* **32**:483-496.
4. Dickson, C., and G. Peters. 1981. Protein-coding potential of mouse mammary tumor virus genome RNA as examined by in vitro translation. *J. Virol.* **37**:36-47.

TABLE 1. Initiation codon arrangements in descending order of efficiency^a and corresponding MMTV LTR proteins^b

Protein(s)	Codon arrangement ^c
p21, p24 p18	AXX AUG G
	AXX AUG A
	AXX AUG Y
	GXX AUG G
p36 p32	GXX AUG A
	GXX AUG Y
	YXX AUG X

^a Kozak (14).

^b Donehower et al. (6).

^c Y = pyrimidine; X = any one of the four ribonucleotides.

5. Dickson, C., R. Smith, and G. Peters. 1981. *In vitro* synthesis of polypeptides encoded by the long terminal repeat region of mouse mammary tumor virus DNA. *Nature (London)* **291**:511–513.
6. Donehower, L. A., B. Fleurdelys, and G. L. Hager. 1983. Further evidence for the protein coding potential of the mouse mammary tumor virus long terminal repeat: nucleotide sequence of an endogenous proviral long terminal repeat. *J. Virol.* **45**:941–949.
7. Donehower, L. A., A. L. Huang, and G. L. Hager. 1981. Regulatory and coding potential of the mouse mammary tumor virus long terminal redundancy. *J. Virol.* **37**:226–238.
8. Dudley, J. P., J. M. Rosen, and J. S. Butel. 1978. Differential expression of poly(A)-adjacent sequences of mammary tumor virus RNA in murine mammary cells. *Proc. Natl. Acad. Sci. U.S.A.* **75**:5797–5801.
9. Dusing-Swartz, D., D. Medina, J. S. Butel, and S. H. Socher. 1979. Mouse mammary tumor virus genome expression in chemical carcinogen induced mammary tumors in low- and high-tumor-incidence mouse strains. *Proc. Natl. Acad. Sci. U.S.A.* **76**:5360–5364.
10. Fasel, N., K. Pearson, E. Buetti, and H. Diggelmann. 1982. The region of mouse mammary tumor virus DNA containing the long terminal repeat includes a long coding sequence and signals for hormonally regulated transcription. *EMBO J.* **1**:3–7.
11. Harding, C. R., and J. R. Scott. 1983. Fluorography limitations on its use for quantitative detection of ^3H and ^{14}C -labeled proteins in polyacrylamide gels. *Anal. Biochem.* **129**:371–376.
12. Hopp, T. P., and K. R. Woods. 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. U.S.A.* **78**:3824–3828.
13. Kennedy, N., G. Knedlitschek, B. Groner, H. E. Hynes, P. Herrlick, R. Michalides, and A. J. J. van Ooyen. 1982. Long terminal repeats of endogenous mouse mammary tumor virus contain a long open reading frame which extends into adjacent sequences. *Nature (London)* **295**:622–624.
14. Kozak, M. 1981. Possible roles of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res.* **9**:5233–5252.
15. Lerner, R. A., N. Green, H. Alexander, F. T. Liu, J. G. Sutcliffe, and T. M. Shinnick. 1981. Chemically synthesized peptides predicted from the nucleotide sequence of the hepatitis B virus genome elicit antibodies reactive with the native envelope protein of Dane particles. *Proc. Natl. Acad. Sci. U.S.A.* **78**:3403–3407.
16. Majors, J. E., and H. E. Varmus. 1981. Nucleotide sequences at host-proviral junctions for mouse mammary tumor virus. *Nature (London)* **289**:253–258.
17. Majors, J. E., and H. E. Varmus. 1983. Nucleotide sequencing of an apparent proviral copy of *env* mRNA defines determinants of expression of the mouse mammary tumor virus *env* gene. *J. Virol.* **47**:495–504.
18. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. Nigg, E. A., G. Walter, and S. J. Singer. 1982. On the nature of crossreactions observed with antibodies directed to defined epitopes. *Proc. Natl. Acad. Sci. U.S.A.* **79**:5939–5943.
20. Owens, R. B., and A. J. Hackett. 1972. Tissue culture studies of mouse mammary tumor cells and associated viruses. *J. Natl. Cancer Inst.* **49**:1321–1332.
21. Pachl, C., B. Biegalka, and M. Linial. 1983. RNA and protein encoded by MH2 virus: evidence for subgenomic expression of *v-myc*. *J. Virol.* **45**:133–139.
22. Peters, G., R. Smith, S. Brookes, and C. Dickson. 1982. Conservation of protein coding potential in the long terminal repeats of exogenous and endogenous mouse mammary tumor viruses. *J. Virol.* **42**:880–888.
23. Prakash, O., R. V. Guntaka, and N. H. Sarkar. 1983. Evidence for a prokaryotic promoter in the murine mammary tumor virus long terminal repeat. *Gene* **23**:117–130.
24. Racevskis, J., and N. H. Sarkar. 1982. ML antigen of DBA/2 mouse leukemias: expression of an endogenous murine mammary tumor virus. *J. Virol.* **42**:804–813.
25. Robertson, D. L., and H. E. Varmus. 1981. Dexamethasone induction of the intracellular RNAs of mouse mammary tumor virus. *J. Virol.* **40**:673–682.
26. Schultz, A. M., E. H. Rabin, and S. Oroszlan. 1979. Post-translational modification of Rauscher leukemia virus precursor polypeptides encoded by the *gag* gene. *J. Virol.* **30**:255–266.
27. Schwinghamer, M. W., and R. J. Shepherd. 1980. Formaldehyde-containing slab gels for analysis of denatured tritium-labeled RNA. *Anal. Biochem.* **103**:426–434.
28. Sen, G. C., J. Racevskis, and H. N. Sarkar. 1981. Synthesis of murine mammary tumor viral proteins in vitro. *J. Virol.* **37**:963–975.
29. Telang, N. T., M. R. Banerjee, A. P. Iyer, and A. B. Kundu. 1979. Neoplastic transformation of epithelial cells in whole mammary gland in vitro. *Proc. Natl. Acad. Sci. U.S.A.* **76**:5886–5890.
30. Temin, H. M. 1981. Structure, variation and synthesis of retrovirus long terminal repeat. *Cell* **27**:1–3.
31. Thomas, P. S. 1983. Hybridization of denatured RNA transferred or dotted to nitrocellulose paper. *Methods Enzymol.* **100B**:255–266.
32. Walter, G., M. A. Hutchinson, T. Hunter, and W. Eckhart. 1981. Antibodies specific for the polyoma virus middle-sized tumor antigen. *Proc. Natl. Acad. Sci. U.S.A.* **78**:4882–4886.
33. Wheeler, D. A., J. S. Butel, D. Medina, R. D. Cardiff, and G. L. Hager. 1983. Transcription of mouse mammary tumor virus: identification of a candidate mRNA for the long terminal repeat gene product. *J. Virol.* **46**:42–49.
34. Yang, J., J. Enami, and S. Nandi. 1977. Regulation of mammary tumor virus production by prolactin in BALB/cfC3H mouse normal mammary epithelial cells in vitro. *Cancer Res.* **37**:3644–3647.