Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 1618–1622, March 1985
Biochemistry

Murine leukemia virus protease is encoded by the *gag–pol* gene and is synthesized through suppression of an amber termination codon

(retroviral protease/amino acid sequence/translational control)

YOSHIYUKI YOSHINAKA, IYOKO KATOH, TERRY D. COPELAND, AND STEPHEN OROSZLAN

Litton Bionetics, Inc., Basic Research Program, Laboratory of Molecular Virology and Carcinogenesis, National Cancer Institute–Frederick Cancer Research Facility, Frederick, MD 21701

Communicated by David Baltimore, November 7, 1984

ABSTRACT **We have purified from Moloney murine leukemia virus (Mo-MuLV) a protease that has the capacity of accurately cleaving the polyprotein precursor Pr65^{gag} into the mature viral structural proteins. Both the NH_2- and COOH-terminal amino acid sequences have been determined and aligned with the amino acid sequence deduced from the DNA sequence of Mo-MuLV by other workers. The results show that: (*i*) the protease is located at the 5′ end of the *pol* gene, and the first four amino acids are overlapped with the 3′ end of the *gag* gene; (*ii*) the fifth amino acid residue is glutamine, which is inserted by suppression of the UAG termination codon at the *gag–pol* junction; and (*iii*) the protease is composed of 125 amino acids with calculated M_r = 13,315, and the COOH terminus of the protease is adjacent to the NH_2 terminus of reverse transcriptase. The map order of the *gag–pol* gene is proposed to be 5′-p15–p12–p30–p10–protease–reverse transcriptase–endonuclease-3′.**

The internal structural proteins of murine leukemia virus (MuLV) are encoded by the group-specific antigen (*gag*) gene and synthesized as a precursor polyprotein designated Pr65^{gag}. In addition to the *gag* gene, all replication-competent retroviruses possess a polymerase (*pol*) and an envelope (*env*) gene which have been mapped as 5′-*gag-pol-env*-3′. Although the *gag* and *pol* genes are separated by an amber termination codon (UAG), translation of the genome-size mRNA yields, in addition to Pr65^{gag}, a larger precursor designated Pr180$^{gag\text{-}pol}$ (for review see ref. 1). Jamjoom *et al.* (2) suggested that the synthesis of this *gag-pol* polyprotein, which is made in amounts 4–10% of those for Pr65^{gag}, may be translationally controlled. Using an *in vitro* translational system and yeast suppressor tRNA, Philipson *et al.* (3) provided evidence that synthesis of Pr180$^{gag\text{-}pol}$ was enhanced by suppression of an amber termination codon.

During virus maturation Pr65^{gag} is proteolytically cleaved into the final products designated p15, p12, p30, and p10. The processing is accomplished by a virion-associated protease (4), which first cleaves Pr65^{gag} into Pr27^{gag} (p15 + p12) and Pr40^{gag} (p30 + p10), the two major intermediate cleavage products (5, 6). However, the origin of protease (viral or cellular) remained unknown. In avian retroviruses a protein designated p15 and encoded by the 3′ end of the *gag* gene has been shown to have associated protease activity (7, 8), but the *gag* proteins of MuLV were not found to cleave the precursor. Genetic studies by Traktman *et al.* (9) with conditional maturation mutants of MuLV have indicated the importance of Pr180$^{gag\text{-}pol}$ for the proteolytic processing of Pr65^{gag}. Levin *et al.* (10), who studied a natural *pol* frameshift mutant, confirmed and extended these observations and predicted the map position of a putative virally coded protease to be 5′ to the reverse transcriptase coding region. We have determined the NH_2-terminal sequence of the 80-kilodalton reverse transcriptase (11) derived from Pr180^{gag} *pol* and located its genetic locus to begin 360 nucleotides downstream from the amber termination codon positioned at the end of the *gag* gene (12). This finding suggested to us that the *pol* gene segment upstream to the codon specifying the NH_2 terminus of reverse transcriptase may code for an approximately 14-kilodalton additional polypeptide, and we hypothesized that it may be the protease since the deduced primary and more importantly the predicted secondary structure of the putative protein resembled those of avian myeloblastosis virus and avian sarcoma virus protease (13, 14).

In this communication we report the purification and primary structure analysis of a protease from Moloney (Mo)-MuLV. These data provide the evidence that the protease is encoded by the *gag–pol* gene and is synthesized by a translational readthrough of the amber termination codon for the *gag* gene.

MATERIALS AND METHODS

Viruses. Mo-MuLV was grown in BALB/c mouse bone marrow JLS-V9 cells (MJD-54 cells) kindly supplied by K. Manly (Roswell Park Memorial Institute, Buffalo, NY). Gazdar murine sarcoma virus (Gz-MSV) was grown in HTG-2 cells (15). The viruses were purified by sucrose density gradient centrifugation and obtained from the Biological Products Laboratory, Program Resources, Inc., National Cancer Institute–Frederick Cancer Research Facility (Frederick, MD).

Assay of Protease. Gz-MSV, which itself has no protease activity and contains uncleaved Pr65^{gag} as its major core protein, was the source of the polyprotein substrate. Protease activity was assayed as previously described (6).

Extraction of Protease Activity from Virus. To 50 mg of purified Mo-MuLV suspended in 2 ml of 0.13 M NaCl/0.01 M Tris·HCl, pH 7.2/0.001 M EDTA (STE buffer), 20 vol of cold acetone (−70°C) was added, and then the suspension was centrifuged at 5000 rpm for 10 min at 4°C in a Sorvall SS-24 rotor. The precipitate was dried under reduced pressure. To solubilize the protease, extraction of the acetone powder (4°C, 30 min with constant stirring) was done stepwise first with 10 ml of 0.02 M piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (Pipes), pH 7.0/5 mM dithiothreitol (PD buffer) alone, PD buffer plus 0.1 M NaCl, PD buffer plus 0.5 M NaCl, and finally PD buffer plus 2.0 M NaCl. Each aqueous extract was centrifuged at 10,000 rpm for 10 min at 4°C in a Sorvall SS-24 (SS-1) rotor. Aliquots were dialyzed against PD buffer and assayed for protease activity. Extracts

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: MuLV, murine leukemia virus; Mo-MuLV, Moloney MuLV; Gz-MSV, Gazdar murine sarcoma virus; AKV, AKR mouse leukemia virus; FeLV, feline leukemia virus; BaEV, baboon endogenous virus; RP-HPLC, reversed-phase high-performance liquid chromatography.

having protease activity were pooled, concentrated by lyophilization, and saved for further purification.

NaDodSO$_4$/PAGE. Various protein materials were analyzed by discontinuous NaDodSO$_4$/PAGE (16). Specifically, to separate low molecular weight proteins, a 32-cm 8–18% polyacrylamide gradient gel was used as described (4). Visualization of proteins was by staining with either Coomassie brilliant blue R-250 or silver (17).

Ion-Exchange and Gel-Permeation Chromatography. Fractionation of protease by phosphocellulose and Sephadex G-75 chromatography was done as previously described (4).

Reversed-Phase High-Performance Liquid Chromatography (RP-HPLC). Lyophilized samples shown to have protease activity were dissolved in saturated guanidine·HCl (Gdn·HCl), and further fractionated by RP-HPLC (18) on a μBondapak C_{18} column (Waters Associates). Five-milliliter fractions were collected and aliquots were taken for protein composition analysis on NaDodSO$_4$/PAGE and protease activity measurement.

NH$_2$-Terminal Microsequence Analysis. Semi-automated microsequence analysis was performed with a Beckman sequencer model 890C equipped with a cold trap accessory as described (19). Phenylthiohydantoin derivatives of amino acids were identified and quantitated by HPLC (20).

COOH-Terminal Sequence Analysis. Protein samples were digested with carboxypeptidase Y (21) for various time intervals and the released amino acids were quantitated on the Durrum 500 analyzer.

RESULTS

Purification of Protease. In initial studies designed to purify the Mo-MuLV protease, the proteolytic activity was first concentrated and fractionated by phosphocellulose chromatography using stepwise elution with increasing concentration of KCl. The activity that eluted at 0.3 M KCl was further fractionated by gel filtration on Sephadex G-75. Each fraction was assayed for protease activity by incubating aliquots with disrupted Gz-MSV and subsequently determining the protein pattern by NaDodSO$_4$/PAGE. Shown in Fig. 1 are the electrophoretic profiles of Sephadex G-75 fractions 20–23 after they were incubated alone or with the substrate (Gz-MuLV Pr65^{gag}). These four fractions were found to have protease activity as judged by the decrease in band intensity of Gz-MSV Pr65^{gag} and concomitant appearance of Pr40^{gag}, Pr27^{gag}, p30, and p10 bands, which were readily detectable. It is also seen by this semiquantitative assay that fractions 21 and 22 had the peak activity inasmuch as Pr40^{gag}, the proximal precursor for p30 and p10 (4), completely disappeared and was further cleaved into the final products p30 and p10. It was observed previously as well as in the present studies that, in contrast to Pr40^{gag}, the intermediate cleavage product, Pr27^{gag}, is relatively difficult to process *in vitro* to the constituent p15 and p12.

Attempts to purify the protease to homogeneity by conventional methods as described were unsuccessful. Although the protease could be concentrated considerably (in some cases 500-fold on protein basis), it was difficult to identify which protein was responsible for proteolytic activity. The protein patterns of concentrated phosphocellulose/Sephadex G-75 fractions made visible by staining after NaDodSO$_4$/PAGE (see Fig. 1) were very complex. More than 10 proteins were detected in the 10- to 20-kilodalton region of NaDodSO$_4$/PAGE, where the protease itself migrates (4).

To purify the protease in sufficient amounts and purity for structure analysis we utilized RP-HPLC. For these studies we first prepared an acetone powder from purified virus. From this powder we solubilized the protease by stepwise extraction with PD buffer having increasing concentrations

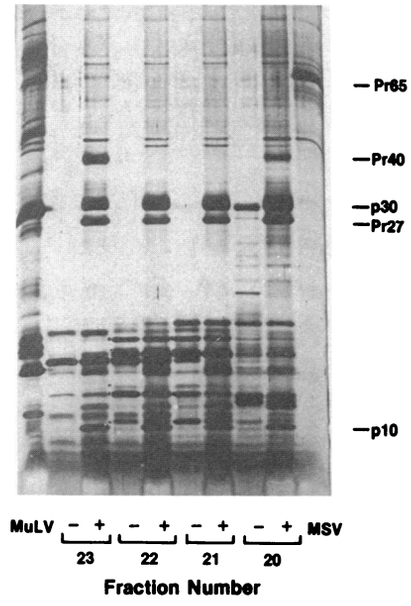

FIG. 1. Cleavage of Gz-MSV Pr65^{gag} with partially purified protease of Mo-MuLV. Purified disrupted virus (50 mg) was fractionated by a combination of phosphocellulose and Sephadex G-75 chromatography, and the protease activity of each fraction was determined as described in the text. Protein composition and protease activity of Sephadex G-75 fractions 20–23 are shown: −, fraction alone; +, fraction plus Gz-MSV Pr65^{gag}. The gel was stained with silver nitrate.

of NaCl. Much of the protease activity was extracted with 0–0.5 M NaCl in PD buffer, while most of the membrane proteins, including p15E and p15, stayed in the insoluble residue. The protease-active extracts were pooled, lyophilized, and dissolved in 3 ml of PD buffer, then fractionated on a Sephacryl S200 column in the cold and further purified by RP-HPLC using a μBondapak C_{18} column as shown in Fig. 2*A*. Fractions (5 ml) were collected and lyophilized to recover proteins. Purity and protease activity were determined by NaDodSO$_4$/PAGE analysis as shown in Fig. 2 *B* and *C*, respectively. The peak activity was eluted at about 33% acetonitrile (fraction 24 of Fig. 2) and clearly separated from p30, p12, p10, and other low molecular weight proteins. The purified protein showed a single band in NaDodSO$_4$/PAGE (Fig. 2*B*). When incubated with disrupted Gz-MSV it cleaved Pr65^{gag} to produce Pr40^{gag}, Pr27^{gag}, and p30 (Fig. 2*C*). The total protein recovered in RP-HPLC fractions 22–24 was 14 μg.

In the absence of a quantitative assay for the protease the determination of the recovery of enzymatic activity is difficult. If we define a unit as the activity (per unit volume) capable of 50% reduction of Pr65^{gag} band intensity after 16-hr incubation (see *Materials and Methods*), we can estimate that we extracted a total of 110 units of activity from the virus and found 66 units in HPLC fractions (Fig. 2). This corresponds to 60% overall recovery of protease activity. The possibility for the actual protease being a minor component copurifying with the protein peak cannot be completely excluded. However, this is unlikely since in NIH/3T3 cells transfected with cloned viral DNA having deletions only in the protease region, Pr65^{gag} is synthesized but not processed into mature protein components (unpublished observations, and S. Crawford and S. P. Goff, personal communication).

NH$_2$-terminal Amino Acid Sequence of the Protease. To determine the NH$_2$-terminal amino acid sequence of purified protease recovered from fraction 24, 0.5 nmol of protein was degraded in a single microsequence analysis. The amino

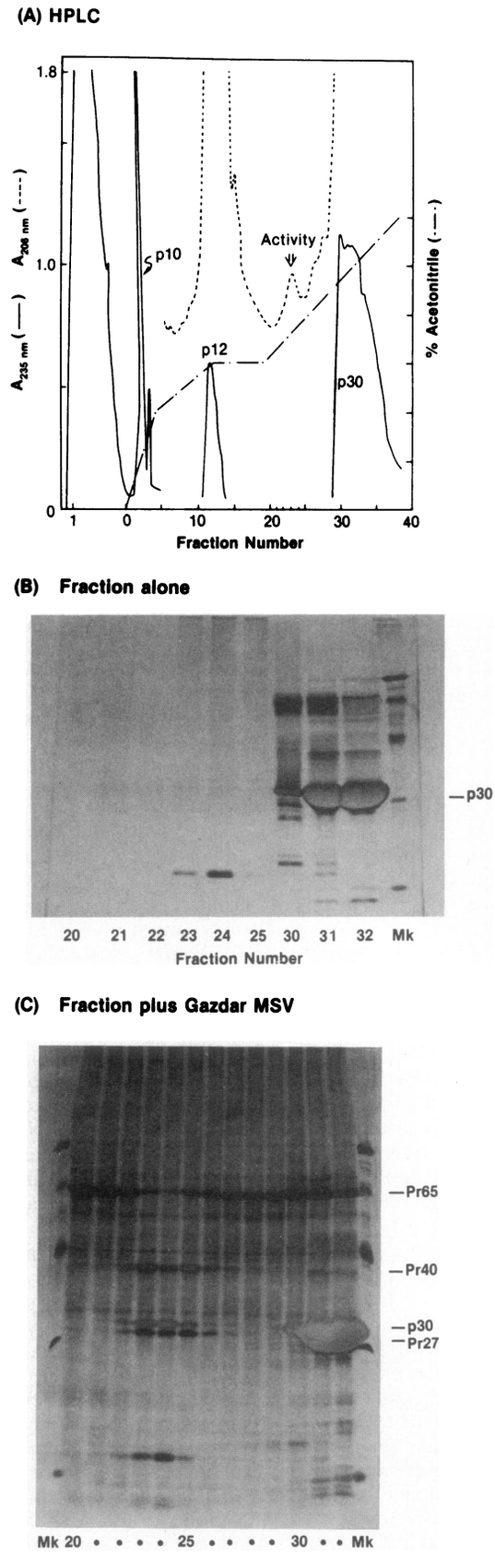

FIG. 2. Purification of protease by RP-HPLC. Purified virus (50 mg) was extracted with acetone. Protease activity extracted from acetone powder was further fractionated by Sephacryl S200 and then by RP-HPLC (μBondapak C_{18} column 0.39 × 30 cm, Waters Associates). (*A*) Absorbance profile. Gradient conditions were 0–20% (vol/vol) acetonitrile over 20 min; 20–30% acetonitrile over 40 min; isocratic at 30% acetonitrile for 30 min, and 30–60% acetonitrile over 90 min at a constant flow rate of 1.0 ml/min. (*B*) Purity of RP-HPLC-separated proteins by NaDodSO$_4$/PAGE. One-twentieth of each fraction was lyophilized and analyzed; staining was with silver nitrate. Lane Mk contained molecular weight markers: phosphorylase b, 92,000; bovine serum albumin, 68,000; ovalbumin, 46,000; carbonic anhydrase, 29,000; lysozyme, 14,400. (*C*) Assay of RP-HPLC fractions for protease activity. One-twentieth of each fraction was lyophilized and assayed for protease activity as described in text. Proteins are visualized by silver staining. Lanes Mk as in *B*.

```
                    5                                  10
Thr-Leu-Asp-Asp-Gln-Gly-Gly-Gln- X -Gln-
240 575 300 490 240 140 220 230     195

                   15                                  20
Glu-Pro-Pro-Pro-Glu- X -Arg-Ile-Thr-Leu-
180  65  90  85 110     135  50  50  70
```

FIG. 3. NH_2-terminal sequence of Mo-MuLV protease. The number under each assigned residue is the recovery (in pmol) of that residue. X, an unidentified residue.

acids identified at each cycle (first 20) are shown in Fig. 3 together with the quantitative yields.

COOH-Terminal Sequence Analysis. Purified protease was digested for various time periods with carboxypeptidase Y and the amino acids released were determined on the analyzer. The data shown in Table 1 allow us to conclude that the COOH terminus of the protease is leucine. Other amino acids released in smaller quantities were valine, glutamine (or serine), and proline. The kinetic analysis data do not by themselves define an accurate sequence, but they could be interpreted by comparison with the amino acid sequence deduced from the DNA sequence (12) as will be discussed below.

Position of Protease on Viral Genome and Suppression of Amber Codon into Glutamine. To determine whether the protease protein is virus encoded or not, we aligned the NH_2- and COOH-terminal amino acid sequences with nucleotide sequences. As shown in Fig. 4, the protease amino acid sequence starts with threonine encoded by triplet 2223–2225 and includes the last four amino acids of the *gag* region. Furthermore, the amber codon (UAG) is translated as glutamine, which is residue five of the protease. This is followed by a glycine residue encoded by the first triplet of the *pol* gene, indicating that translation continues in the same reading frame. These results clearly show that the $Pr65^{gag}$-specific protease is a virus-encoded enzyme and that it is synthesized by reading through the termination codon.

It is also seen that the determined NH_2-terminal sequence for the protease (15 residues are shown) matches the amino acid sequence predicted from the nucleotide sequence of proviral DNA designated pMLV-1 (12) except at position 11, where the protein has Glu instead of Asp. At this juncture it is important to point out that the pMLV-1 clone is not infectious and that the infectious clone (pMLV-48) of Miller and Verma (22), like the protein, also has Glu in position 11. A single base change, C → G, in the codon accounts for this difference. A comparison of the COOH-terminal sequence analysis results with the translated sequence of proviral DNA indicates that the COOH-terminal Leu must be the codon 2595–2597, which is adjacent to the NH_2-terminal Thr of reverse transcriptase (11). This determines that the COOH-terminal sequence is Pro-Leu-Gln-Val-Leu-OH and that the protease is composed of 125 amino acids (Figs. 4 and 5).

DISCUSSION

We have succeeded in purifying from Mo-MuLV a virus-encoded protease that is capable of processing *in vitro* the *gag* precursor polyprotein $Pr65^{gag}$ into the constituent struc-

Table 1. COOH-terminal amino acid sequence analysis of Mo-MuLV protease

Digestion time, min	Amino acids released,* pmol			
	Leu	Val	Gln	Pro
0.5	70	29	0	0
1.0	101	30	33	0
2.0	123	31	45	30
5.0	140	48	49	31
10.0	158	67	55	48
20.0	191	72	62	60

*Each sample analyzed at each time point had 200 pmol of protein.

tural proteins, which appear to be the same as those occurring in mature virions. We have definitely identified cleavage products p30, p10, and p15. The identification of p15 was aided by the high sensitivity of detection after its specific labeling with [^{3}H]myristate (unpublished data). The fourth *gag*-gene encoded protein, p12, however, could not be identified with absolute certainty in our assay system due to its poor affinity for Coomassie blue and to the apparent lack of specificity of the more sensitive silver stain (nucleic acids are stained equally well as proteins).

The NH_2- and COOH-terminal sequences of the Mo-MuLV protease as determined in this study and the availability of DNA sequences now present an opportunity for deducing the complete primary structure of this virus-encoded proteolytic enzyme. The complete amino acid sequence of the Mo-MuLV protease aligns (Fig. 5) without gaps with the amino acid sequences of putative proteases as inferred from DNA sequences of AKR mouse leukemia virus (AKV) (23), feline leukemia virus (FeLV) (24), and baboon endogenous virus (BaEV) (25). It is seen from this alignment that there are only four amino acid differences between the two mouse proteases compared. With respect to the Mo-MuLV sequence, FeLV and BaEV proteases have 25 (20%) and 39 (31.2%) changes, respectively, indicating highly conserved primary structures. From the combined results it appears that there are three variable regions: one at the NH_2-terminal region (residues 1–12), the second in the middle part of the molecule (residues 60–77), and the third at the COOH-terminal region (residues 108–120). The last five residues of the MuLV and FeLV sequences are identical, while in the BaEV sequence there are two substitutions, Leu → Ile and Val → Ile. These changes, however, would not significantly alter the nature of the cleavage site between the protease and reverse transcriptase. The regions involving residues 29–41 are identical among the sequences, and another highly conserved long stretch is present, extending from residue 78 to residue 101. Little is known about the active site of the retroviral protease, but inhibition studies done with avian protease (8) suggested that cysteine may be involved. Furthermore, it was also shown that the mouse protease is inhibited by tosyllysyl chloromethyl ketone (TLCK) (4). It is known that TLCK can inhibit thiol proteases or similar enzymes just as well as serine proteases, and on occasions it has been effectively used to identify cysteines at the active site (26). Interestingly, with the exception of BaEV, each of the viral proteases, including those of avian myeloblastosis virus and avian sarcoma virus, has only a single cysteine (residue 88 in the alignment of Fig. 5), which is preceded by Asp or Glu. It will be important to develop quantitative assays, perhaps utilizing synthetic peptide substrates (27), for the retroviral proteases to characterize them more completely.

Our results, together with the previously determined NH_2- and COOH-terminal sequences of reverse transcriptase (11) and their alignments with nucleotide sequences, suggest the map order for $Pr180^{gag\text{-}pol}$ of the mouse retrovirus to be 5′-p15–p12–p30–p10–protease–reverse transcriptase–endonuclease-3′. The polyprotein itself most likely has no proteolytic activity. It remains to be seen how the active protease is generated. Autocatalysis or an initial cleavage by another enzyme (probably cellular) may be responsible.

The most significant result reported in this study relevant to virus replication is the finding that *in vivo* translation of the *pol* gene resulting in the synthesis of the precursor polyprotein $Pr180^{gag\text{-}pol}$ occurs through in-frame readthrough of the amber termination codon. While we do not know the exact mechanism by which glutamine is inserted at the termination site in the translation process taking place in mouse fibroblasts, we can assume that this insertion is accomplished via the misreading of the UAG codon by normal $tRNA^{Gln}$ due to the wobble in the 3′ position of the anticodon. A more remote possibility is suppression by a specific nonsense suppressor tRNA. Such tRNAs have been identified not only in prokaryotes but also recently in eukaryotes (28, 29). Suppression of termination codons has been proposed to occur in plant viruses (30–32) and alphaviruses (33, 34). It appears that plant and animal viruses are capable of effectively utilizing the translational readthrough mechanism to produce from a single initiation site different amounts of proteins and polyproteins required for specific functions. The importance of this translational control for virus replication, infectivity, and pathogenicity could be directly tested by utilizing mutants in which the respective termination codons are eliminated.

In Mo-MuLV the *gag* and *pol* genes are in the same reading frame. However, available nucleic acid sequences indicate that apparently this is not true for Rous sarcoma virus (14), FeLV (24), human T-cell leukemia virus (35), and bovine leukemia virus (36). As with the Mo-MuLV study, protein sequencing will reveal whether in-frame suppres-

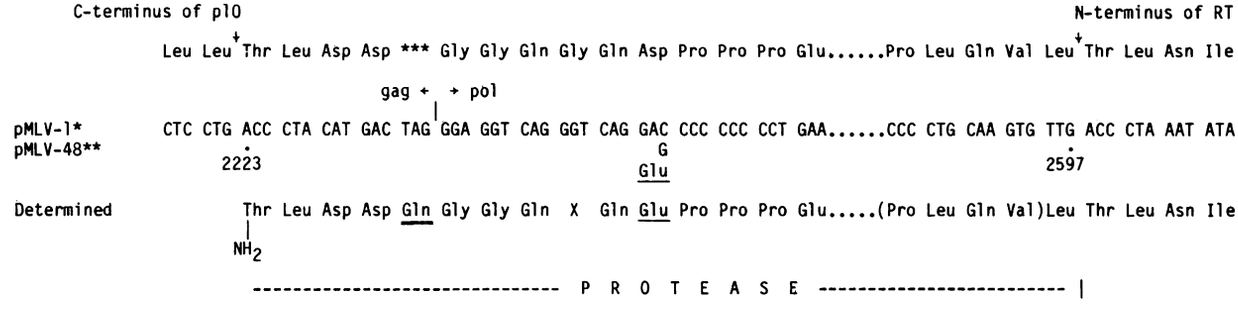

FIG. 4. Alignment of NH_2- and COOH-terminal amino acid sequences with DNA sequences of pMLV-1 (12) and pMLV-48 (21). The amber codon UAG is translated as glutamine (double underline). RT, reverse transcriptase.

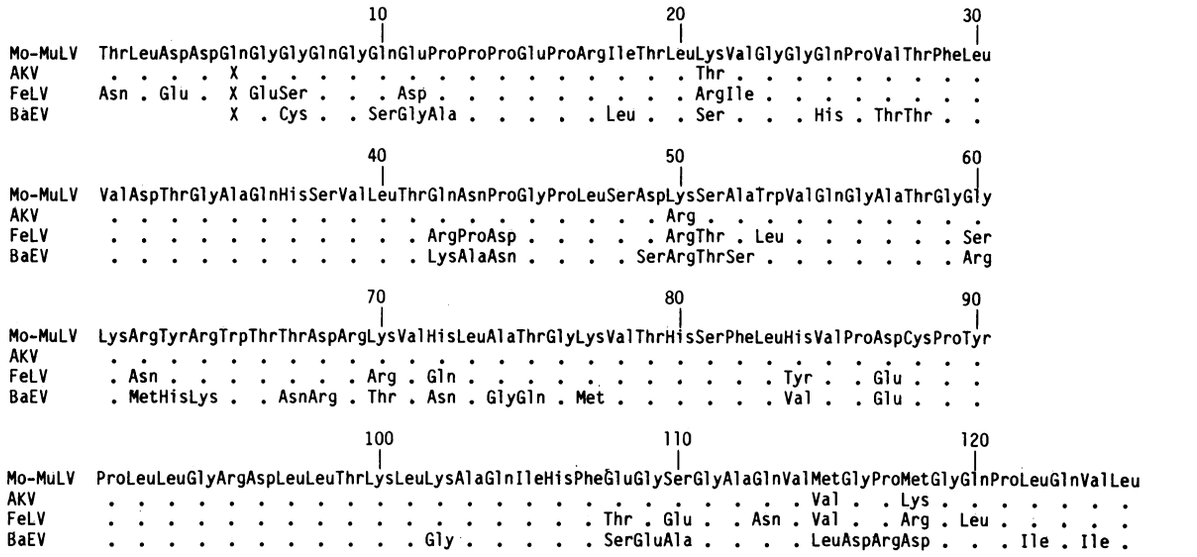

FIG. 5. Amino acid sequence alignment of Mo-MuLV protease with corresponding sequences of AKR mouse leukemia virus (AKV) (23), feline leukemia virus (FeLV) (24), and baboon endogenous virus (BaEV) (25) inferred from nucleotide sequences. Corresponding to position 5 (Gln in Mo-MuLV and X in the other sequences) the genes have an amber termination codon. In this alignment the predicted NH_2-terminal sequence (positions 1–12) of the putative FeLV protease is translated from the nucleotide sequence of Laprevotte *et al.* (24) in the *gag* reading frame. The remainder of the sequence, starting with Pro in position 12, is in a reading frame as published (24). The DNA sequence between codon GAC (Asp in position 11) and CCC (Pro in position 12) has an additional C which was not decoded for the purposes of this alignment. It remains to be seen whether the FeLV suppression occurs in frame, as in MuLV, or requires a frameshift. The third possibility of course is splicing, as proposed by Laprevotte *et al.* (24).

sion, frameshift suppression, or splicing is involved in these latter cases. In all cases the elucidation of actual mechanism responsible for suppression will require further biochemical experiments with purified tRNAs.

We thank Cathy Hixson and Young Kim for their excellent technical assistance and Jeannie Clarke for preparation of the manuscript. This research was sponsored in part by the National Cancer Institute under contract no. N01-C0-23909 with Litton Bionetics, Inc.

1. Dickson, C., Eisenman, R., Fan, H., Hunter, E. & Teich, N. (1982) in *Molecular Biology of Tumor Viruses*, eds. Weiss, R. A., Teich, N. M., Varmus, H. E. & Coffin, J. M. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 513–648.
2. Jamjoom, G. A., Naso, R. B. & Arlinghaus, R. B. (1977) *Virology* **78,** 11–34.
3. Philipson, L. P., Andersson, V., Olshevsky, R., Weinberg, D., Baltimore, D. & Gestland, R. (1978) *Cell* **13,** 189–199.
4. Yoshinaka, Y. & Luftig, R. B. (1977) *Cell* **12,** 709–719.
5. Witte, O. N. & Baltimore, D. (1978) *J. Virol.* **26,** 750–761.
6. Yoshinaka, Y. & Luftig, R. B. (1982) *Virology* **118,** 380–388.
7. Von der Helm, K. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 911–915.
8. Dittmar, K. J. & Moelling, K. (1978) *J. Virol.* **28,** 106–118.
9. Traktman, P. & Baltimore, D. (1980) in *Biosynthesis, Modification, and Processing of Cellular and Viral Proteins*, eds. Koch, G. & Richter, D. (Academic, New York), pp. 301–306.
10. Levin, J. G., Hu, S. C., Rein, A., Messer, L. I. & Gerwin, B. (1984) *J. Virol.* **51,** 470–478.
11. Copeland, T. D., Gerard, G. F., Hixson, C. G. & Oroszlan, S. (1985) *Virology*, in press.
12. Shinnick, T. M., Lerner, R. A. & Sutcliffe, J. G. (1981) *Nature (London)* **293,** 543–548.
13. Sauer, R. T., Allen, D. D. W. & Niall, H. D. (1981) *Biochemistry* **20,** 3784–3791.
14. Schwartz, D. E., Tizard, R. & Gilbert, W. (1983) *Cell* **32,** 853–869.
15. Gazdar, A. F., Phillips, L. A., Sarma, P. S., Peebles, P. T. & Chopra, H. C. (1971) *Nature (New Biol.) (London)* **234,** 69–72.
16. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
17. Wray, W., Boulikas, T., Wray, V. P. & Hancock, R. (1981) *Anal. Biochem.* **118,** 197–203.
18. Henderson, L. E., Sowder, R., Copeland, T. D., Smythers, G. & Oroszlan, S. (1984) *J. Virol.* **52,** 492–500.
19. Copeland, T. D., Grandgenett, D. P. & Oroszlan, S. (1980) *J. Virol.* **36,** 115–119.
20. Henderson, L. E., Copeland, T. D. & Oroszlan, S. (1980) *Anal. Biochem.* **102,** 1–7.
21. Hayashi, R. (1976) *Methods Enzymol.* **45,** 568–587.
22. Miller, A. D. & Verma, I. M. (1984) *J. Virol.* **49,** 214–222.
23. Herr, W. (1984) *J. Virol.* **49,** 471–478.
24. Laprevotte, I., Hampe, A., Sherr, C. J. & Galibert, F. (1984) *J. Virol.* **50,** 884–894.
25. Tamura, T. (1983) *J. Virol.* **47,** 137–145.
26. Whitaker, J. R. & Perez-Villasenor, J. (1968) *Arch. Biochem. Biophys.* **124,** 70–76.
27. Copeland, T. D. & Oroszlan, S. (1982) in *Peptides: Synthesis, Structure, Function*, eds. Rich, D. H. & Gross, E. (Pierce, Rockford, IL), pp. 497–500.
28. Diamond, A., Dudock, B. & Hatfield, D. (1981) *Cell* **25,** 497–506.
29. Hatfield, D. L., Dudock, B. S. & Eden, F. C. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 4940–4944.
30. Pelham, H. R. B. (1978) *Nature (London)* **272,** 469–471.
31. Goelet, P., Lomossoff, G. P., Butler, P. J. G., Akam, M. E., Gait, M. J. R. & Karn, J. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 5818–5822.
32. Morch, M. D., Drugon, G. & Benicourt, C. (1982) *Virology* **119,** 193–198.
33. Strauss, E. G., Rice, C. M. & Strauss, J. H. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 5271–5275.
34. Strauss, E. G., Rice, C. M. & Strauss, J. H. (1984) *Virology* **133,** 92–110.
35. Seiki, M., Hattori, S., Hirayama, Y. & Yoshida, M. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3618–3622.
36. Rice, N. R., Stephens, R. M., Burny, A. & Gilden, R. V. (1985) *Virology*, in press.