JOURNAL OF VIROLOGY, Sept. 1985, p. 870–873
0022-538X/85/090870-04$02.00/0

Vol. 55, No. 3

Translational Readthrough of an Amber Termination Codon During Synthesis of Feline Leukemia Virus Protease

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

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

Received 18 March 1985/Accepted 21 May 1985

Feline leukemia virus contains a protease which apparently has the same specificity as murine leukemia virus protease. It cleaves in vitro the Pr65^{gag} of Gazdar-mouse sarcoma virus into the constituent p15, p12, p30, and p10 proteins. We purified the protease and determined its NH_2-terminal amino acid sequence (the first 15 residues). Alignment of this amino acid sequence with the nucleotide sequence (I. Laprevotte, A. Hampe, C. H. Sherr, and F. Galibert, J. Virol. 50:884–894, 1984) reveals that the protease is a viral-coded enzyme and is located at the 5′ end of the *pol* gene. As previously found for murine leukemia virus (Y. Yoshinaka, I. Katoh, T. D. Copeland, and S. Oroszlan, Proc. Natl. Acad. Sci. U.S.A. 82:1618–1622, 1985), feline leukemia virus protease is synthesized through in-frame suppression of the *gag* amber termination codon by insertion of a glutamine in the fifth position, and the first four amino acids are derived from the *gag* gene.

Feline leukemia virus (FeLV) is a non-genetically transmitted exogenous retrovirus shown to be associated with disease in domestic cats (4). As is characteristic of most retroviruses, FeLV genomic RNA contains three genes, *gag*, *pol*, and *env*, which are necessary for viral replication. The *gag* gene is translated into the polyprotein precursor Pr65^{gag}, which is processed to the structural proteins p15, p12, p30, and p10; the *pol* gene encodes an RNA-dependent DNA polymerase, and the *env* gene encodes the envelope glycoproteins of the virion surface (2). Based on the order of the *gag* gene-coded structural proteins, protein sequence homology, and immunological relatedness with murine leukemia virus (MuLV), FeLV is classified as type C, subgroup 1 (15). Further, Laprevotte et al. (10) reported a nucleotide sequence of 2,565 base pairs which includes a portion of the 5′ long terminal repeat, the *gag* leader, the complete *gag* gene, and 389 base pairs of the *pol* gene. Their data indicated that FeLV *gag* and *pol* genes are translated in different reading frames. Recently, we purified and sequenced Moloney (Mo)-MuLV protease responsible for the proteolytic processing of precursor polyprotein Pr65^{gag} (18). The results showed that this protease is encoded by the *gag-pol* gene and synthesized within Pr180$^{gag\text{-}pol}$ through suppression of the amber termination codon located at the end of the *gag* gene.

In this report we describe the purification and partial sequence of a protease from FeLV, its location in the viral genome, and the translational control for its synthesis. FeLV (Rickard strain AB) was grown in feline lymphoblasts (16) and purified by sucrose density gradient centrifugation (13). In earlier studies we demonstrated Mo-MuLV protease activity under assaying conditions which involved endogenous substrate i.e., uncleaved Pr65^{gag} of Mo-MuLV (19) released by disruption of the virus by Nonidet P-40. Subsequently, we adopted for routine analysis a method of assaying Mo-MuLV protease (20) with an exogenous substrate, Gazdar-mouse sarcoma virus (Gz-MSV) Pr65^{gag} (6). Attempts to detect FeLV protease activity by the former method were unsuccessful because of the extremely low levels of uncleaved Pr65^{gag} in purified FeLV. Therefore, in the present studies designed to purify FeLV proteolytic enzyme, we used Gz-MSV Pr65^{gag} to assay FeLV protease activity. The suitability of Gz-MSV Pr65^{gag} as a substrate was expected because the cleavage sites in the mouse and feline mature *gag* proteins are very similar based on protein and nucleotide seuqence data available for both systems (10, 11, 13, 17). The previously described methods for protease assay (18) and purification were used without major modifications.

To purified FeLV (158 mg) suspended in 2 ml of STE buffer (0.13 M NaCl, 0.01 M Tris hydrochloride [pH 7.2], and 0.001 M EDTA), 20 volumes of cold acetone (−70°C) was added, and the suspension was centrifuged at 4,000 × *g* for 10 min at 4°C. The precipitate was dried in vacuo. To solubilize the protease, the acetone powder was extracted (at 4°C for 30 min, stirred continuously) with 4 ml of TD buffer (0.02 M Tris hydrochloride [pH 7.0], 5 mM dithiothreitol [Sigma Chemical Co., St. Louis, Mo.]) containing 1.0 M NaCl. The extract was centrifuged at 20,000 × *g* for 20 min at 4°C. The supernatant was then fractionated on a Sephacryl S-200 column (2.5 by 90 cm) with TD buffer, and the protease activity was determined as described above by adding 100 μl of each fraction to 15 μg of Gz-MSV substrate in 1% Nonidet P-40. The protease-active fractions were pooled, lyophilized, and then further fractionated by reverse-phase high-pressure liquid chromatography (RP-HPLC) on a Bondapak C_{18} column (0.39 by 30 cm) (Waters Associates, Inc., Milford, Mass.). The protease activity was eluted with about 33% acetonitrile (Fig. 1A; fraction 30) and was detected by assaying lyophilized 5% aliquots of the fractions. When fractions 26 to 35 were incubated with disrupted Gz-MSV, fractions 27 to 33 cleaved Pr65^{gag} into what appears to be the mature proteins p30, p15, p12, and p10 (Fig. 1B). The peak activity appeared in fraction 30. In addition, in fractions 28 and 32 intermediate cleavage products, presumably Pr40^{gag} (p30 plus p10) and Pr27^{gag} (p15 plus p12), were produced as observed in previous studies with Mo-MuLV protease (20). The purified protein, the majority of which eluted in fractions 29 to 31, showed a single band in sodium dodecyl sulfate-polyacrylamide gel electrophoresis. From this analysis, the total protein was estimated to be approximately 7 μg.

* Corresponding author.

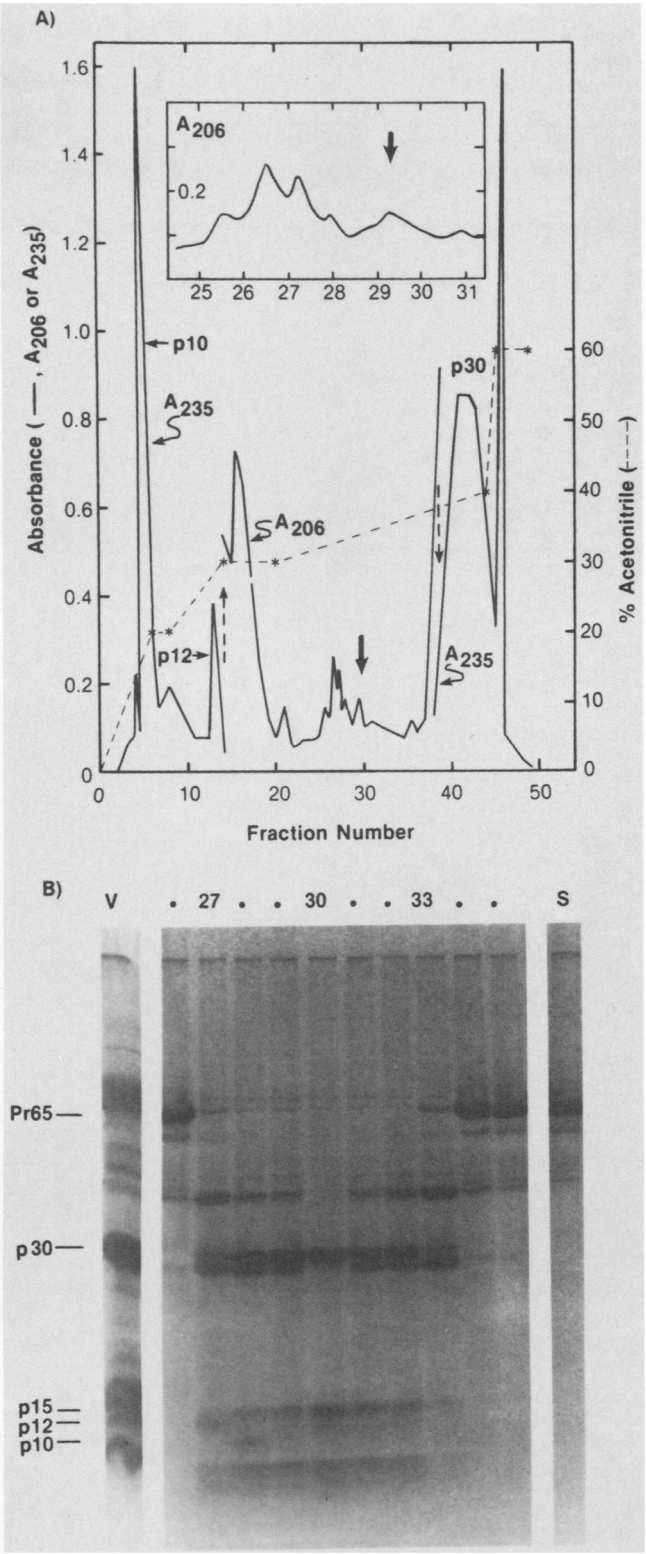

FIG. 1. Purification of protease by RP-HPLC. (A) Absorbance profile. Sephacryl S-200 chromatography fractions were applied to a Bondapak C_{18} column and then eluted with increasing concentration of acetonitrile as follows: 0 to 20% acetonitrile over 20 min, 20 to 30% acetonitrile over 30 min, and 30 to 60% acetonitrile over 90 min at a constant flow rate of 1.0 ml/min. Absorbance was measured at 235 nm or at 206 nm as indicated. (B) Assay of RP-HPLC fractions for protease activity. One-twentieth of each fraction was lyophilized and assayed for protease activity as described in the text. Proteins were visualized by staining with Coomassie brilliant blue R-250.

To determine the NH_2-terminal amino acid sequence of the protease, approximately 50 pmol of RP-HPLC-purified protein was subjected to automated Edman degradation in a gas-phase sequenator (9) with the program supplied by the manufacturer (Applied Biosystems, Inc). Conversion of the anilinothiazolinone amino acids to the phenylthiohydantoin amino acids was accomplished with 25% trifluoroacetic acid in water. Phenylthiohydantoin amino acids were identified and quantitated by RP-HPLC (8). The first 15 residues of the NH_2-terminal amino acid sequence were determined. The residue assignments together with the quantitative recoveries are given in Fig. 2A.

To examine whether the protease protein is viral coded, we aligned the experimentally determined sequence with the amino acid sequence deduced from DNA nucleotide sequence (10). The protease amino acid sequence begins with asparagine coded by triplet 2072 to 2074 and overlaps with the last four amino acids of the *gag* region (Fig. 2B). However, the third amino acid residue of protease was glycine and not glutamic acid, predicted from nucleotide sequence as shown in the alignment. The fifth amino acid, glutamine, corresponds to the *gag* termination codon TAG positioned at nucleotides 2084 to 2086. This is followed by a glutamic acid residue coded by the first triplet of the *pol* gene. The nucleotide sequence then continues in the *gag* reading frame and matches the protein sequence Thr-Gln-Gly-Gln-Asp-Pro-Pro-Pro-. At nucleotide 2113, however, we encounter a TGA stop codon in the DNA sequence. But if we remove one of the eight consecutive C residues from the sequence occupying positions 2104 through 2111, the nucleotide sequence matches both the FeLV protease sequence as determined here as well as the previously reported MuLV protease sequence (18) beyond nucloeotide 2111. The result is that FeLV protease is now in the same frame as *gag* and reverse transcriptase. This suggests that the DNA clone of FeLV strain B reported by Laprevotte et al. (10) is a noninfectious clone which, like many cloned DNAs of retroviruses, is defective. The protease NH_2-terminal amino acid sequence was found to be different from the nucleotide sequence of strain B at position 3 (Gly-Glu) and at position 7 (Thr-Ser). This may indicate strain differences and suggests that the major component of FeLV(AB) is strain A; the NH_2-terminal amino acid sequence analysis data of p10 (3) and p12 (unpublished data) also show differences from the strain B nucleotide sequence. In conclusion, these results show that the FeLV protease is a viral-coded enzyme and that it is synthesized by readthrough of the amber termination codon as in the murine system (18).

When we align the amino acid sequence of FeLV and MuLV protease deduced from the DNA sequence, 80% homology (25 different of 125 residues) is observed. Furthermore, the FeLV protease cleavage products of $Pr65^{gag}$ from Gz-MSV made sense when we compared the feline cleavage sites with the murine cleavage site between the *gag* proteins (p15, p12, p30, and p10); each feline cleavage site is very similar to the corresponding murine site (Fig. 2C). In both FeLV and Mo-MuLV systems (18), a single protease is responsible for the complete proteolytic processing of $Pr65^{gag}$ which is temporally linked to virus maturation. This conclusion regarding specificity is supported by the similarity in the chemical structure of all the cleavage sites (Fig. 2C). Although the peptide bonds cleaved (viz., tyrosyl-proline between p15 and p12, phenylalanyl-leucyl–proline between p12 and p30, leucyl-alanine between p30 and p10, and leucyl–asparagine-threonine between p10 and protease) are not identical, the carboxyl-terminal amino acid se-

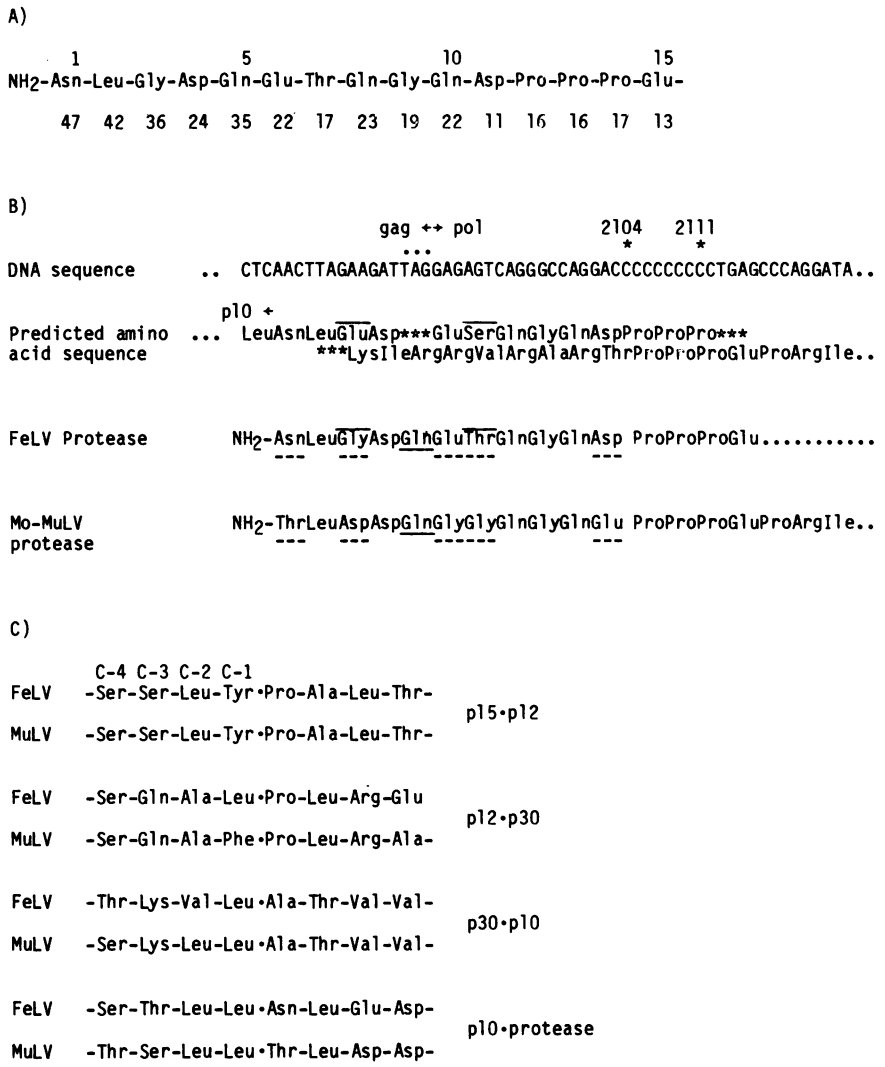

FIG. 2. (A) NH_2-terminal sequence of FeLV protease. Number below each residue is the yield (in picomoles) of the phenylthiohydantoin amino acid. (B) Alignment of the NH_2-terminal amino acid sequence with the DNA sequence of FeLV (10). The amber codon UAG is translated into glutamine (underline). Lines above amino acids indicate differences between deduced and determined sequences. Dotted lines under amino acids indicate differences between FeLV and Mo-MuLV protease (20). (C) Comparison of *gag* cleavage site sequences of FeLV and Mo-MuLV.

quences of the cleavage products are strikingly similar. They suggest a consensus (14) as follows: the penultimate residue (C-2) to the newly generated carboxyl terminus (C-1) is always a hydrophobic amino acid. The residue (C-3) next to it always has a polar side chain, charged or uncharged (serine, glutamine, and lysine), and the C-4 residue is either serine or threonine, both of which are known to initiate β-turns.

The exact mechanism of suppression is not clear. However, it is quite possible that the insertion of glutamine results from the misreading of the termination codon (UAG) by normal glutamyl tRNA as we proposed for Mo-MuLV (18). Other translational control mechanisms, such as suppression by nonsense suppressor tRNA (7), by splicing and by frame shift suppression (1, 5), have also been observed in both procaryotic and eucaryotic cell systems. The effects of the surrounding sequences on the suppression of a nonsense codon have also been investigated in bacteria (12). Retroviruses provide a useful model system for studying translational control in eucaryotic cells. It will be of interest to compare the influence of neighboring sequences on the suppression of certain amber termination codons in retroviral mRNAs.

We thank Young Kim for excellent technical assistance, and Jeannie Clarke and Sally Lockhart for preparation of the manuscript.

This research was sponsored by the National Cancer Institute, Department of Health and Human Services, under contract NO1-CO-23909 with Litton Bionetics, Inc., Kensington, Md.

LITERATURE CITED

1. **Atkins, J. F., and R. F. Gesteland.** 1983. Resolution of the discrepancy between a gene translation-termination codon and the deduced sequence for release of the encoded polypeptide. Eur. J. Biochem. **137**:509–516.
2. **Baltimore, D.** 1975. Tumor viruses. Cold Spring Harbor Symp. Quant. Biol. **39**:1187–1200.
3. **Copeland, T. D., M. A. Morgan, and S. Oroszlan.** 1984. Com-

plete amino acid sequence of the basic nucleic acid binding protein of feline leukemia virus. Virology **133**:137–145.
4. **Essex, M.** 1980. Feline leukemia and sarcoma viruses, p. 205–229. *In* G. Klein (ed.), Viral oncology. Raven Press, New York.
5. **Fox, T. D., and B. Weiss-Brummer.** 1980. Leaky 1^+ and 1^- frameshift mutations at the same site in a yeast mitochondrial gene. Nature (London) **288**:60–63.
6. **Gazdar, A. F., L. A. Phillips, P. S. Sarma, P. T. Peebles, and H. C. Chopra.** 1971. Presence of sarcoma genome in a "non-infectious" mammalian virus. Nature (London) New Biol. **324**:69–72.
7. **Hatfield, D. L., B. S. Dudock, and F. C. Eden.** 1983. Characterization and nucleotide sequence of a chicken gene encoding an opal suppressor tRNA and its flanking DNA segments. Proc. Natl. Acad. Sci. U.S.A. **80**:4940–4944.
8. **Henderson, L. E., T. D. Copeland, and S. Oroszlan.** 1980. Separation of all amino acids phenylthiohydantoins by high-performance liquid chromatography on phenylalkyl support. Anal. Biochem. **102**:1–7.
9. **Hewick, R. M., M. M. Hunkapiller, L. E. Hood, and W. J. Dreyer.** 1981. A gas-liquid solid phase peptide and protein sequenator. J. Biol. Chem. **254**:7990–7997.
10. **Laprevotte, I., A. Hampe, C. H. Sherr, and F. Galibert.** 1984. Nucleotide sequence of *gag* gene and *gag-pol* junction of feline leukemia virus. J. Virol. **50**:884–894.
11. **Miller, A. D., and I. M. Verma.** 1984. Two base changes restore infectivity to a noninfectious molecular clone of Moloney murine leukemia virus (pMLV-1). J. Virol. **49**:214–222.
12. **Miller, J. H., and A. Albertini.** 1983. Effects of surrounding sequence on the suppression of nonsense codons. J. Mol. Biol. **164**:59–71.
13. **Morgan, M. A., T. D. Copeland, and S. Oroszlan.** 1983. Structural and antigenic analysis of the nucleic acid-binding proteins of bovine and feline leukemia viruses. J. Virol. **46**:177–186.
14. **Oroszlan, S., and T. D. Copeland.** 1985. Primary structure and processing of *gag* and *env* gene products of human T-cell leukemia viruses HTLV-1_{CR} and HTLV-I_{ATK}. Curr. Top. Microbiol. Immunol. **115**:221–233.
15. **Oroszlan, S., and R. V. Gilden.** 1980. Primary structure analysis of retrovirus proteins, p. 299–344. *In* J. R. Stephenson (ed.), Molecular biology of RNA tumor viruses. Academic Press, Inc., New York.
16. **Rickard, C. G., J. E. Pest, F. Noronha, and L. M. Barr.** 1969. A transmissible virus-induced lymphocytic leukemia of the cat. J. Natl. Cancer Inst. **42**:987–1014.
17. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukemia virus. Nature (London) **293**:543–548.
18. **Yoshinaka, Y., I. Katoh, T. D. Copeland, and S. Oroszlan.** 1985. Murine leukemia virus protease is encoded by the *gag-pol* gene and is synthesized through suppression of an amber termination codon. Proc. Natl. Acad. Sci. U.S.A. **82**:1618–1622.
19. **Yoshinaka, Y., and R. B. Luftig.** 1977. Properties of a p70 proteolytic factor of murine leukemia viruses. Cell **12**:709–719.
20. **Yoshinaka, Y., and R. B. Luftig.** 1982. P65 of Gazdar murine sarcoma viruses contains antigenic determinants from all four of the murine leukemia virus (MuLV) *gag* polyproteins (p15, p12, p30 and p10) and can be cleaved *in vitro* by the MuLV proteolytic activity. Virology **118**:380–388.