Terminal Amino Acid Sequences and Proteolytic Cleavage Sites of Mouse Mammary Tumor Virus *env* Gene Products

LOUIS E. HENDERSON,* RAYMOND SOWDER, GARY SMYTHERS, AND STEPHEN OROSZLAN

LBI-Basic Research Program, Laboratory of Molecular Virology and Carcinogenesis, National Cancer Institute-Frederick Cancer Research Facility, Frederick, Maryland 21701

Received 11 May 1983/Accepted 19 July 1983

The mature envelope glycoproteins of mouse mammary tumor virus (gp52 and gp36) were isolated by reversed-phase high-pressure liquid chromatography. The N-terminal amino acid sequence of gp36 was determined for 28 residues. The C-terminal amino acid sequences of gp52 and gp36 were determined by carboxypeptidase digestion. The N-terminal amino acid sequence of gp52 has been reported previously (L. O. Arthur et al., *J. Virol.* **41**:414-422, 1982). These data were aligned with the predicted amino acid sequence of the *env* gene product obtained by translation of the DNA sequence (S. M. S. Redmond and C. Dickson, *Eur. Mol. Biol. Org. J.* **2**:125-131, 1983). The amino acid sequences of the mature viral proteins were in agreement with the predicted amino acid sequence of the *env* gene product over the regions of alignment. This alignment showed the sites of proteolytic cleavages of the *env* gene product leading to the mature viral envelope glycoproteins. The N-terminal amino acid sequence of gp52 starts at residue 99 of the predicted structure indicating proteolytic cleavage of a signal peptide. A dipeptide (Lys-Arg) is excised between the C-terminus of gp52 and the N-terminus of gp36. The C-terminal amino acid sequence of gp36 is identical to the sequence predicted by the codons immediately preceding the termination codon for the *env* gene product. The data show that there is no proteolytic processing at the C-terminal of the murine mammary tumor virus *env* gene product and that the *env* gene coding region extends into the long terminal repeat.

Mouse mammary tumor virus (MMTV) is the prototype for the B-type virus of the RNA retroviral family. The synthesis, structure, and biology of MMTV have recently been reviewed (6). During reverse transcription of viral RNA, regions of the 5' and 3' ends are copied twice, resulting in a long terminal repeat (LTR) on the linear viral DNA transcript and in the integrated proviral DNA. The gene order of MMTV linear viral DNA transcript from 5' to 3' ends is LTR-*gag-pol-env*-LTR where *gag*, *pol*, and *env* are genes coding for known viral proteins. The *gag* gene codes for a precursor polyprotein (Pr73^{gag}) which is post-translationally modified by phosphorylation (19) and myristylation (26). The *pol* gene codes for reverse transcriptase. The *env* gene codes for a precursor polyprotein (Pr69^{env}) which is cotranslationally transported into the endoplasmic reticulum where a signal peptide is removed by proteolysis (4) and post-translational glycosylation is initiated. The glycosylated precursor (gPr73^{env}) is transported to the cell membrane where it participates in viral assembly and is subsequently cleaved by proteolysis to viral gp52 and gp36 (4, 25). These two glycoproteins are associated with the lipid envelope of

the mature virus. Gp36 is believed to be a transmembrane glycoprotein. Gp52 probably determines the tissue tropism of the virus by interacting with specific cell surface receptors during viral infection. In addition, MMTV glycoproteins show immunological cross-reactivity with antigens associated with breast cancer (3, 18, 20).

Co- and post-translational modification of the viral precursor polyprotein may be important controlling events in the viral assembly and maturation process. The nature of some of these modifications can be directly determined by structural studies on the viral proteins. The nature of post-translational proteolytic cleavages can best be determined by comparing amino acid sequences of the mature viral protein to the nucleic acid sequence of the viral genome.

The nucleotide sequence of MMTV (C3H) LTR (9, 16) and of two fragments of MMTV (GR) proviral genome including the entire *env* gene (23) and part of the LTR (10, 23) has been determined. In the LTR region where the known nucleotide sequences of the two strains can be compared, there are only 60 base changes scattered over 1,360 nucleotides suggesting minimal

FIG. 1. Separation of MMTV (C3H) structural proteins by RP-HPLC. A 2.0-mg amount of MMTV (C3H) was suspended in 10 ml of water and adjusted to pH 2.0 with trifluoroacetic acid and added to 20 ml of saturated guanidine-HCl. The mixture was shaken for 1 min and then applied to a column (0.39 by 30 cm) of μ Bondapak C_{18} at a flow rate of 1.0 ml/min. After sample application, the column was washed with 40 ml of 0.1% trifluoroacetic acid at 1.0 ml/min (data not shown). Elution was accomplished at 1.0 ml/min and 23°C with gradients of increasing acetonitrile (0.1% trifluoroacetic acid) (---) as follows: 0 to 20%, 10 min; 20 to 50%, 120 min; 50 to 60%, 10 min. The column temperature was changed to 52°C and the organic solvent changed to 1-propanol (0.1% trifluoroacetic acid), and elution continued with increasing gradients as follows: 30% isocratic, 10 min; 30 to 100%, 60 min. Eluted components were detected by UV absorption at 206 nm. Peaks 1 through 10 were analyzed by SDS-PAGE as shown in Fig. 2.

structural differences between the two strains. The nucleotide sequence shows three possible initiation codons in the same open reading frame coding for the *env* gene product (23). It is not known which of these is the initiation codon for the synthesis of p69^{env} in vitro. Therefore, the DNA sequence was translated into amino acid sequence starting at the first possible initiation codon and the predicted amino acid sequence of p69^{env}, numbered accordingly (23). Structural studies on the viral *env* gene products (gp52 and gp36) from either strain should elucidate the exact coding region for these proteins and define proteolytic processing events that lead to the final viral products. Here, we report the N-terminal amino acid sequence of gp36 (C3H) for the first 28 residues and the C-terminal amino acid sequences of gp52 (GR) and gp36 (GR). The N-terminal amino acid sequence of gp52 (C3H) was also determined in this laboratory and reported previously (1). These data are compared with the amino acid sequence of the MMTV *env* precursor predicted from the nucleotide sequence of the *env* gene (23) and the LTR (9, 16). The comparison reveals the proteolytic cleavage sites in the precursor between gp52 and gp36 and shows that the N-terminal amino acid sequence of gp36 (C3H) is identical to the sequence predicted from the nucleotide sequence of the GR strain. The comparison also shows that the C-terminal region of the viral gp36 is coded by nucleotide sequences which extends 83 bases into the LTR. This is the first observation that

nucleotides in retroviral LTR structures can code for protein synthesis.

MMTV structural proteins were purified by reversed-phase high-pressure liquid chromatography (RP-HPLC) by using techniques similar to those previously described for the purification of type C viral proteins (13, 14). Fig. 1 shows the separation of the MMTV (C3H) structural proteins by RP-HPLC. The viral proteins were separated by gradient elution at pH 2 as indicated in Fig. 1. Column fractions 1 through 6 were eluted with acetonitrile gradients at 23°C. Fractions 7 through 10 were eluted with a 1-propanol gradient at 50°C as indicated in Fig. 1. Portions of each fraction were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (Fig. 2). Column fraction 1 did not contain any protein detectable by Coomassie-stained SDS-PAGE analysis. Fraction 2 contained highly purified gp52 (Fig. 2, lane 2), fraction 4 contained purified p27, fraction 5 contained purified p10, and fraction 6 contained additional p27. Fractions 7 and 10 did not contain any detectable protein. Fraction 8 contained lipids that migrate near the dye front on SDS-PAGE and are detectable with Coomassie staining. Fraction 9 contained gp36 and lipids (Fig. 2, lane 9). The separation of the structural proteins of MMTV (GR) was identical to the separation shown in Fig. 1. In the preparations of MMTV used, both p14 and p21 were present in minor amounts compared with the amount of gp52, p27, p10, and gp36. In other separations by the

same techniques but with larger inputs of virus (ca. 10 mg), p14 and p21 were detected (data not shown). P14 eluted in about 25% acetonitrile and p21 eluted in about 31% acetonitrile.

Amino acid analysis of purified gp36 (GR) is given in Table 1. The data are given as moles of each amino acid per mole of protein, calculated on the assumption that the peptide chain contains 232 amino acid residues as predicted by the nucleotide sequence of the *env* gene product (GR) and the defined coding region for gp36 as deduced from the data given in this report. The predicted amino acid composition of gp36 (GR) is also given in Table 1. The experimentally determined values for each residue are within 10% of the predicted values for all residues except methionine and isoleucine. The low values obtained for methionine and isoleucine are probably due to low recoveries of the residues after acid hydrolysis.

Table 2 gives the results of semiautomated Edman degradation of gp36 (C3H). The N-terminal amino acid sequence was determined for 28 residues (Table 2). No phenylthiohydantoin (PTH) amino acid was detected in cycle 18. All other assignments listed in Table 2 were unambiguous.

The C-terminal amino acid sequence of gp36 (GR) was determined by carboxypeptidase digestion by methods previously described (22). Quantitative data for the amino acids released by digestion with either carboxypeptidase A or

TABLE 1. Amino acid analysis of gp36 (GR)

Amino acid	Residues per mole	
	Predicted from nucleotide sequence	Found ^a
Asp + Asn	26	27.1
Thr	14	15.4
Ser	12	13.5
Glu + Gln	23	23.0
Pro	6	6.6
Gly	9	9.5
Ala	25	25.3
Val	19	17.6
Met	5	3.8
Ile	19	15.1
Leu	33	32.1
Tyr	5	5.5
Phe	9	8.4
His	4	4.1
Lys	10	9.9
Arg	8	8.1
Trp	3	ND
Cys	3	ND

^a Assuming a molecular weight of 26,000 for the peptide chain. Results based on 24-h hydrolysis of the protein and a recovery of 85% for Thr and 80% for Ser. ND, Not determined.

carboxypeptidase B are given in Table 3. These results show that the C-terminal amino acid sequence of gp36 (GR) is -Arg-Val-Ser-Tyr-Thr-OH.

The data given in Tables 2 and 3 were used to establish the gp36 coding region in the predicted amino acid sequence of the *env* gene product (23). The experimentally determined N-terminal amino acid sequence of mature viral gp36 (C3H) given in Table 2 is identical to the amino acid sequence predicted from nucleotide sequences for both C3H strain (H. E. Varmus, personal communication) as well as the GR strain (23). Residue 18 was not identified by Edman degradation of the mature viral protein and is a Ser in the predicted sequence. The low recovery of PTH-Ser or a strain difference may account for the failure to detect a PTH amino acid in Edman cycle 18, but the possibility of a post-translationally modified Ser at this position cannot be excluded. The experimentally determined C-terminal amino acid sequence of mature viral gp36 (GR) is identical to the sequence from residue 684 to the C-terminal end (residue 688) of the predicted sequence of the *env* gene product of MMTV GR (23) and also of MMTV C3H (H. E. Varmus, personal communication). The amino acid composition of viral gp36 (GR) is in good agreement with the amino acid composition of residues 457 through 688 in the predicted sequence of the *env* gene product (GR) (Table 1). The alignment of the viral gp36 amino acid

FIG. 2. Shows the results of SDS-PAGE analysis of RP-HPLC fractions 1 through 10 (lanes 1 through 10) from Fig. 1. Molecular weight standards (Bio-Rad Laboratories) were run in lane S. The analysis was performed by the method of Laemmli on a 10 to 20% gradient gel (15). Proteins were visualized by Coomassie staining.

TABLE 2. Amino terminal amino acid sequence of MMTV gp36 (C3H)^a

Cycle no.	Nanomoles of PTH amino acid	Cycle no.	Nanomoles of PTH amino acid
1	Phe (4.25)	15	Ile (0.90)
2	Val (3.71)	16	Ile (0.92)
3	Ala (3.09)	17	Thr (0.60)
4	Ala (2.46)	18	X
5	Ile (2.45)	19	Phe (0.78)
6	Ile (2.64)	20	Ala (0.70)
7	Leu (1.98)	21	Val (1.20)
8	Gly (1.21)	22	Ala (1.03)
9	Ile (2.02)	23	Thr (0.39)
10	Ser (0.37)	24	Thr (0.39)
11	Ala (1.25)	25	Ala (0.93)
12	Leu (0.85)	26	Leu
13	Ile (1.50)	27	Val
14	Ala (1.19)	28	Lys

^a Sequence was determined by semiautomatic Edman degradation of 4.5 nm of MMTV gp36 (C3H). PTH amino acid was identified by RP-HPLC as previously described (11). X indicates no PTH amino acid identified in cycle 18 (see text).

sequence data with the nucleotide sequence data is shown schematically in Fig. 3. This alignment reveals that 83 nucleotides at the 5' end of the LTR region are within the coding region for the gp36 protein. The nucleotide sequences of MMTV (GR) (10, 23) and MMTV (C3H) LTR (9, 16) segments coding for the C-terminal region of gp36 are identical with the exception of a single base change which does not change the translation of the codon. A long open reading frame in the nucleotide sequence of MMTV LTR has been noted, and it has been suggested that sequences in the LTR may direct protein synthesis (9, 10, 16). The initiation codon for this long open reading frame in the LTR is located 53 nucleotides upstream from the termination codon at the end of the *env* gene. Thus, the long open reading frame in the LTR overlaps the *env* gene but is not in the same reading frame as the *env* gene. Therefore, the *env* gene products discussed here are not directly related to the suggested translational products of the LTR long open reading frame. However, the fact that the C-terminal portion of gp36 is coded by nucleotides in the LTR shows that nucleotides in the LTR can direct protein synthesis.

The first 47 amino acid residues in the amino acid sequence of MMTV gp52 (C3H) have been determined by direct analysis of the protein and have been previously reported (1). This sequence is identical to residues 99 through 145 in the predicted sequence of the *env* gene product (GR) with the exception of two residues. The nucleotide sequence predicts a Ser at position 118 and a Met at position 130 in the *env* gene

product of the GR strain, and the amino acid sequence studies identified Asn and Ile, respectively, at these positions in the C3H strain. Digestion of gp52 (GR) for 1 h with carboxypeptidase B released 0.34 mol of alanine, 0.31 mol of arginine, and 0.26 mol of isoleucine per mol of protein. These results are consistent with a C-terminal sequence Ile-Arg-Ala-OH for viral gp52. This sequence is found at residues 452 through 454 in the predicted amino acid sequence for the *env* gene product (GR) (23). These data, when compared with the nucleotide sequence of the *env* portion of the proviral DNA (23), define the initial and final coding regions for the fully processed gp52 (Fig. 3).

Fig. 3 depicts the known nucleotide sequence of the 3' end of the MMTV genome starting at the first initiation codon in the *env* gene and summarizes the co- and post-translational proteolytic cleavages of the *env* gene product that are shown here to occur during the synthesis and processing of viral envelope proteins.

The biosynthetic process for p69^{env} (7) and gPr73^{env} (5) has been recently reviewed (6). Proteolytic cleavage of a 9,000-dalton signal peptide from the N-terminal end of p69^{env} is thought to be a cotranslational event accompanying insertion of the nascent peptide into the endoplasmic reticulum where glycosylation is initiated. The N-terminal amino acid residues of p69^{env} and gPr73^{env} have not been determined. Therefore, the exact initiation site for p69^{env} synthesis and the nature and number of proteolytic events leading ultimately to the N-terminal end of the fully processed gp52 are not known. The scheme represented in Fig. 3 assumes a common N-terminal residue for gp73^{env} and gp52 (residue 99 in the predicted sequence of p69^{env}). The C-terminal amino acid sequence of gp36 is identical to the predicted C-terminal sequence for p69^{env}; therefore, there is no proteolytic processing at the C-terminal end of the *env* gene products of MMTV. This is in contrast to type C murine viruses in which proteolytic processing at the C-terminal end of the *env* gene products is known to occur (21, 27). Proteolytic processing of gPr73^{env} to gp52 and gp36 involves

TABLE 3. Residues released by carboxypeptidase digestion of MMTV gp36 (GR)

Digestion	Picomoles of amino acid released ^a				
	Arg	Val	Ser	Tyr	Thr
Carboxypeptidase A		25.7	31.6	90.7	92.4
Carboxypeptidase B	9.2	14.1	32.8	91.3	93.5

^a In each experiment 132 pmol of gp36 (GR) digested as previously described (22). The protein concentration was determined by amino acid analysis as in Table 1.

FIG. 3. A schematic representation of the synthesis and proteolytic processing of the MMTV *env* gene product. The nucleotide sequences of the *env* (23) and LTR (9) regions of the genome are represented by the first line in the diagram. An open reading frame extends from the first ATG initiation site through two more ATG codons and ends at a TAG termination codon which is located 84 nucleotides into the LTR as shown. The second line of the diagram represents the predicted amino acid sequence of the *env* gene product (p69^{env}), assuming that protein synthesis initiates at the first ATG codon. Numbers indicate the position of amino acid residue in the predicted sequence of p69^{env}. Arrows indicate proven proteolytic cleavage sites. The predicted amino acid sequences around the cleavage sites and at the C-terminal end of p69^{env} are shown and compared to the N- and C-terminal amino acid sequences of the mature gp52 and gp36. Intervening sequences not shown are indicated by saw-toothed lines.

the removal of a Lys-Arg sequence as indicated in Fig. 3. An identical sequence is found in the *env* precursor of type C murine virus (27) where proteolytic cleavage at the C-terminal side of Arg is known to take place (21). Thus, the cleavage sites between gp52 and gp36 of MMTV bear a strong resemblance to their functional counterpart in murine type C viruses. In the case of Friend murine leukemia virus, the Lys-Arg sequence is retained at the C-terminal end of the glycoprotein (2), whereas in the Rauscher and Moloney strains, most of the isolated glycoprotein has these residues removed, possibly by the action of carboxypeptidases (L. E. Henderson, T. D. Copeland, G. W. Smythers, R. Sowder, and S. Oroszlan, unpublished data).

Gp36 is known to be a hydrophobic membrane-associated protein (8, 17). Its hydrophobic nature is indicated by its strong binding to the reversed-phase support (Fig. 1) which facilitates its isolation by RP-HPLC. Similar strong binding to reversed-phase supports is observed for type C membrane-associated homologs of gp36 (i.e., p15E of murine leukemia viruses) (13, 14). Ongoing work on the type B and type C transmembrane proteins indicates that additional post-translational modification may result in the covalent attachment of strongly hydrophobic moieties to these proteins. The exact chemical nature of these hydrophobic moieties is unknown, but it is probable that they involve fatty acid acylation at serine or threonine hydroxyl groups such as has been proposed for the enve-

lope proteins of vesicular stomatitis and Sindbis virus (24). This type of fatty acid acylation is different from the well-documented covalent attachment of a long-chain fatty acid (myristic acid) to the N-terminal residue of the type C (12) and type B *gag* precursors (26). Continued work on the nature and function of these hydrophobic post-translational modifications of the *gag* and *env* proteins may reveal important steps in the mechanisms of virus assembly and budding. Moreover, these post-translational events may have some relationship to the oncogenic potential of these viruses.

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

LITERATURE CITED

1. Arthur, L. O., T. D. Copeland, S. Oroszlan, and G. Schochetman. 1982. Processing and amino acid sequence analysis of the mouse mammary tumor virus *env* gene product. *J. Virol.* 41:414-422.
2. Chen, R. 1982. Complete amino acid sequence and glycosylation of glycoprotein gp71A of Friend murine leukemia virus. *Proc. Natl. Acad. Sci. U.S.A.* 79:5788-5792.
3. Day, N. K., S. S. Witkin, N. H. Sarkar, D. Kinne, D. J. Jussawalla, A. Levin, C. C. Hsia, N. Geller, and R. A. Good. 1981. Antibodies reactive with murine mammary tumor virus in sera of patients with breast cancer: Geographic and family studies. *Proc. Natl. Acad. Sci. U.S.A.* 78:2483-2487.
4. Dickson, C., and M. Atterwill. 1979. Composition, arrangement and cleavage of the mouse mammary tumor virus polyprotein precursor Pr77^{gag} and p110^{gag}. *Cell* 17:1003-1012.
5. Dickson, C., and M. Atterwill. 1980. Structure and proc-

- essing of the mouse mammary tumor virus glycoprotein precursor Pr73^{env}. *J. Virol.* **35**:349–361.
6. **Dickson, C., R. Eisenman, H. Fan, E. Hunter, and N. Teich.** 1982. Protein biosynthesis and assembly, p. 513–648. *In* R. A. Weiss, N. M. Teich, H. E. Varmus, and J. M. Coffin (ed.), *Molecular biology of tumor viruses*, 2nd ed., RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 7. **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.
 8. **Dion, A. S., C. J. Williams, and A. A. Pomenti.** 1977. The major structural proteins of murine mammary tumor virus: techniques of isolation. *Anal. Biochem.* **82**:18–28.
 9. **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.
 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. *Eur. Mol. Biol. Org. J.* **1**:3–7.
 11. **Henderson, L. E., T. D. Copeland, and S. Oroszlan.** 1980. Separation of all amino acid phenylthiohydantoins by high-performance liquid chromatography on phenylalkyl support. *Anal. Biochem.* **102**:1–7.
 12. **Henderson, L. E., H. C. Krutzsch, and S. Oroszlan.** 1983. Myristyl amino terminal acylation of murine retroviral proteins: an unusual post-translational protein modification. *Proc. Natl. Acad. Sci. U.S.A.* **80**:339–343.
 13. **Henderson, L. E., R. Sowder, and S. Oroszlan.** 1981. Protein and peptide purification by reversed-phase high-pressure chromatography using volatile solvents, p. 251–260. *In* D. T. Liu, A. N. Schechter, R. Henriksson, and P. G. Condliffe (ed.), *Chemical synthesis and sequencing of peptides and proteins*. Elsevier, Amsterdam.
 14. **Henderson, L. E., R. Sowder, and S. Oroszlan.** 1982. Reversed phase (RP) HPLC of proteins and peptides and its impact on protein microsequencing, p. 409–416. *In* M. Elzinga (ed.), *Proceedings of the IV international conference on methods in protein sequence analysis*. The Humana Press, Inc., Clifton, New Jersey.
 15. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685.
 16. **Majors, J., and L. Donehower.** 1982. Appendix D: sequences of retroviral LTRs, p. 1333–1336. *In* R. A. Weiss, N. M. Teich, H. E. Varmus, and J. M. Coffin (ed.), *Molecular biology of tumor viruses*, 2nd Ed., RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 17. **Marcus, S. L., S. W. Smith, J. Racevskis, and N. H. Sarkar.** 1978. The relative hydrophobicity of oncornaviral structural proteins. *Virology* **86**:398–412.
 18. **Masa-Tejada, R., I. Keydar, M. Ramanarayanan, T. Ohno, C. Fenoglio, and S. Spiegelman.** 1979. Detection in human breast carcinomas of an antigen immunologically related to a group-specific antigen of mouse mammary tumor virus. *Proc. Natl. Acad. Sci. U.S.A.* **75**:1529–1533.
 19. **Nusse, R., F. A. M. Asselbers, M. H. Salden, R. J. A. M. Michalides, and H. Bloemendal.** 1978. Translation of mouse mammary tumor virus RNA: precursor polypeptides are phosphorylated during processing. *Virology* **91**:106–115.
 20. **Ohno, T., R. Masa-Tejada, I. Keydar, M. Ramanarayanan, J. Bausch, and S. Spiegelman.** 1979. Human breast carcinoma antigen is immunologically related to the polypeptide of the group-specific glycoprotein of mouse mammary tumor virus. *Proc. Natl. Acad. Sci. U.S.A.* **76**:2460–2464.
 21. **Oroszlan, S., L. E. Henderson, T. D. Copeland, A. M. Schultz, and E. Rabin.** 1980. Processing and structure of murine leukemia virus *gag* and *env* gene encoded polypeptides, p. 219–233. *In* G. Koch, and D. Richter (ed.), *Biosynthesis, modification and processing of cellular and viral polyproteins*. Academic Press, Inc., N.Y.
 22. **Oroszlan, S., L. E. Henderson, J. R. Stephenson, T. D. Copeland, C. W. Long, J. N. Ihle, and R. V. Gilden.** 1978. Amino- and carboxyl-terminal amino acid sequences of proteins coded by *gag* gene of murine leukemia virus. *Proc. Natl. Acad. Sci. U.S.A.* **75**:1404–1408.
 23. **Redmond, S. M. S., and C. Dickson.** 1983. Sequence and expression of the mouse mammary tumor virus *env* gene. *Eur. Mol. Biol. Org. J.* **2**:125–131.
 24. **Schlesinger, M. J., and C. Malfer.** 1982. Cerulenin blocks fatty acid acylation of glycoproteins and inhibits vesicular stomatitis and sindbis virus particle formation. *J. Biol. Chem.* **257**:9887–9890.
 25. **Schochetman, G., S. Oroszlan, L. Arthur, and D. L. Fine.** 1977. Gene order of the mouse mammary tumor virus glycoproteins. *Virology* **83**:72–83.
 26. **Schultz, A. M., and S. Oroszlan.** 1983. *In vivo* modification of retroviral *gag* gene encoded polyproteins by myristic acid. *J. Virol.* **46**:355–361.
 27. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukemia virus. *Nature (London)* **293**:543–548.