JOURNAL OF VIROLOGY, Feb. 1984, p. 629–632
0022-538X/84/020629-04$02.00/0

Vol. 49, No. 2

Nucleotide Sequences of the Envelope Genes of Two Isolates of Feline Leukemia Virus Subgroup B

JACK H. NUNBERG,* MARK E. WILLIAMS, AND MICHAEL A. INNIS

Department of Molecular and Biological Research, Cetus Corporation, Emeryville, California 94608

Received 11 July 1983/Accepted 23 September 1983

We determined the nucleotide sequences of the envelope genes of the Snyder-Theilen and Gardner-Arnstein isolates of feline leukemia virus subgroup B. Comparison of the deduced amino acid sequences of the envelope gene products revealed regions of sequence divergence, which we relate to structural features of the viral protein. We also examined nucleotide sequences within the long terminal repeats of these related isolates of feline leukemia virus subgroup B.

The feline leukemia viruses (FeLV) comprise a group of horizontally transmitted retroviruses associated with a variety of naturally occurring malignant and degenerative diseases in domestic cats (5, 7, 13). Three subgroups of FeLV have been defined on the basis of viral interference (15, 16). Evidence exists to correlate subgroup type(s) with differences in host range and determinants of antibody-mediated virus neutralization, as well as with differences in natural infectivity and disease progression (7, 8, 17). As in other retrovirus systems, it is likely that many of these phenomena involve the products of the viral envelope (*env*) gene, specifically the surface glycoprotein gp70 . To initiate studies to determine the structure of FeLV gp70 and to correlate structural features with functional differences, we determined the nucleotide sequences of the *env* genes of two molecularly cloned isolates of FeLV of the B subgroup (FeLV-B).

A molecularly cloned Gardner-Arnstein FeLV-B (GA-FeLV) isolate (λHF60) (12) was kindly provided by A. Roach and N. Davidson in the form of the pKC7-derived plasmid pKHR-1. We confirmed that this GA-FeLV proviral isolate is infectious in transfection experiments (12) with CCC clone 81 (S^+L^-) cells (3; J. Gilbert, unpublished data). A molecularly cloned Snyder-Theilen FeLV-B (ST-FeLV) isolate (λST-FeLV) (18) was kindly provided by C. Sherr and was subcloned into the *Eco*RI site of pACYC184 (pST-FeLV). The method originally used to molecularly clone this virus has resulted in the deletion of some long terminal repeat (LTR) sequences, and this isolate is not amenable to infectivity studies by a transfection assay.

The complete DNA sequence of the viral *env* and LTR regions was obtained by Maxam-Gilbert (10) and M13-dideoxynucleotide (11, 14) methods. The nucleotide sequence, as well as the deduced amino acid sequence of the *env* gene products, is presented in Fig. 1. The deduced amino acid sequences of FeLV gp70 and p15E were determined by homology to known murine leukemia virus *env* proteins (19). Since the time this work was initiated, Elder and Mullins (2) have published the nucleotide sequence of the same GA-FeLV isolate; our sequences are identical.

Comparison of the ST- and GA-FeLV nucleotide sequences demonstrates that this ST-FeLV sequence was derived from a viral genome defective in *env*; an adenosine has been inserted into the ST-FeLV *env* gene at position 370 to 376, generating a frameshift mutation within gp70 and subsequent premature termination. (The deduced amino acid sequence shown in Fig. 1 assumes deletion of the inserted adenosine.) The *env* gene is additionally defective as a result of the deletion of the C-terminal portion of p15E. The production of defective viral genomes is known to accompany retroviral replication.

The deletion within p15E extends to the putative inverted repeat of the ST-FeLV LTR. The inverted repeat, identical to that found in the LTRs of GA-FeLV (2), Gardner-Arnstein feline sarcoma virus (GA-FeSV), and ST-FeSV (4), is intact. The ST-FeLV LTR contains a series of direct repeats derived from, and identical to, a 50-nucleotide sequence found, in single copy, in the LTR of ST-FeSV (4). The structure of these repeats is shown in Fig. 2. A homologous single-copy sequence is found in the LTRs of GA-FeLV (2) and GA-FeSV (4). These repeats are different from the putative enhancer sequences previously identified in the LTRs of feline retroviruses (4) and show no significant homology to enhancer sequences found in direct repeats of Moloney murine sarcoma virus (9). Unanswered is the question of whether this ST-FeLV LTR, containing the direct repeats, is functional or is derived from a provirus additionally defective in the LTR region.

Because of the defective nature of the ST-FeLV *env* gene, any conclusions as to the functional significance of structural features of the ST-FeLV gp70 or comparisons with the GA-FeLV gp70 remain tentative. Comparison of the nucleotide sequences within the *env* genes of GA- and ST-FeLV, however, reveals extensive homology with only single nucleotide changes. Other than the single adenosine insertion at position 370 to 376, no other insertions, deletions, or rearrangements are observed within the ST-FeLV gp70. Homology between the nucleic acid sequences of ST- and GA-FeLV ranges from greater than 99% in the N-terminal half of gp70 to ca. 85% in the C-terminal half. This difference in sequence conservation is reflected in the deduced amino acid sequences of these proteins. Host range recognition functions are believed to reside within N-terminal gp70 sequences (1), and the extreme conservation of these sequences between these FeLV-B may reflect the functional relatedness of these isolates. The relative divergence of C-terminal sequences may reflect polymorphisms permitted within this portion of gp70. Conservation of ST- and GA-FeLV DNA and protein sequences resumes in p15E.

Amino acid differences within the ST- and GA-FeLV *env*

* Corresponding author.

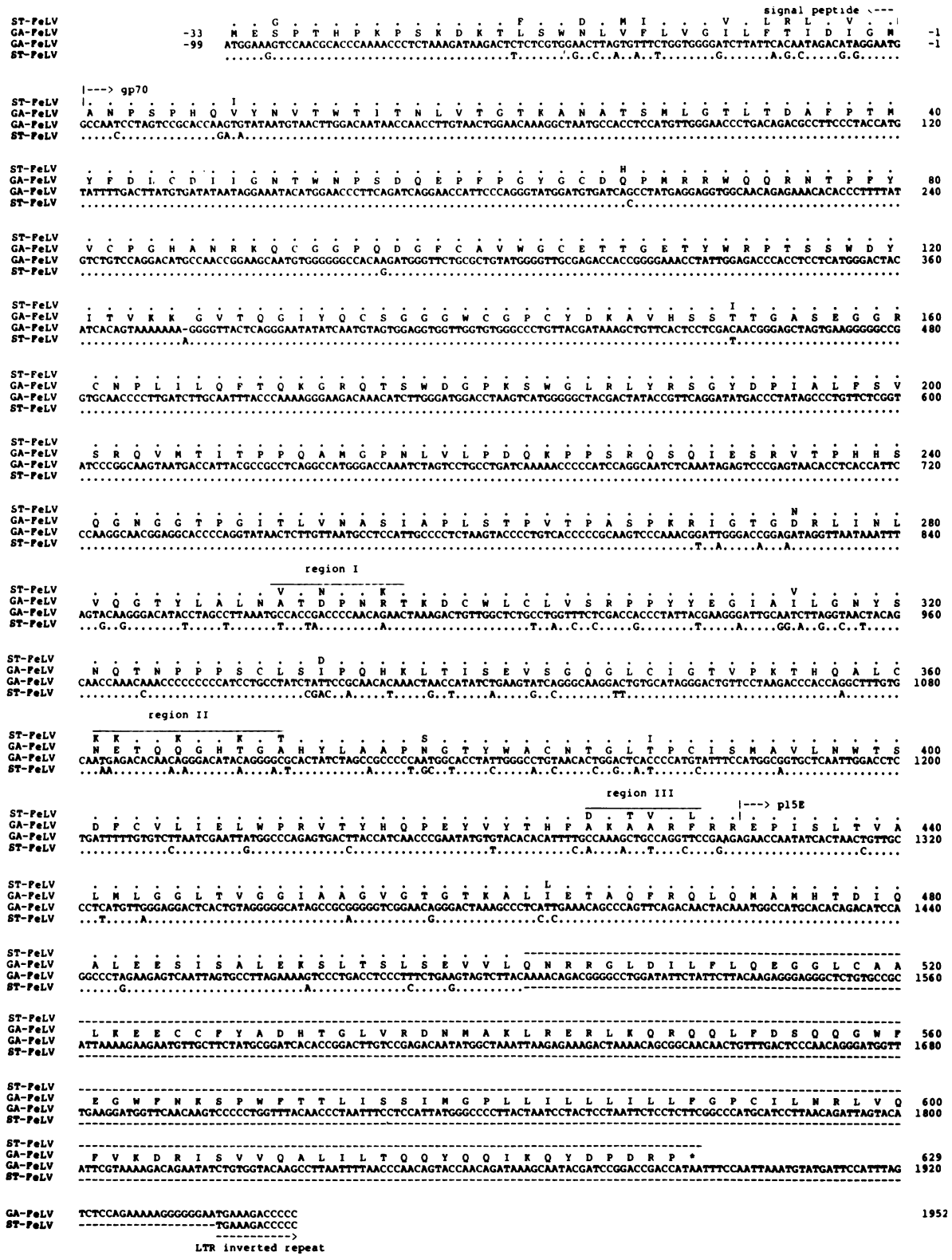

FIG. 1. Nucleotide and deduced amino acid sequences of ST-FeLV and GA-FeLV *env* genes. Nucleotide and amino acid sequences are as indicated. The nucleotide sequence extends to the LTR inverted repeat (indicated). Deletions in ST-FeLV are indicated by (—); conserved residues are indicated by (·). The additional adenosine resulting in the ST-FeLV frameshift is indicated between nucleotides 375 and 376 in the GA-FeLV sequence. The overlined regions in gp70 represent regions of clustered divergence between ST-FeLV gp70 and GA-FeLV gp70. The asterisk denotes the translational termination codon TAA.

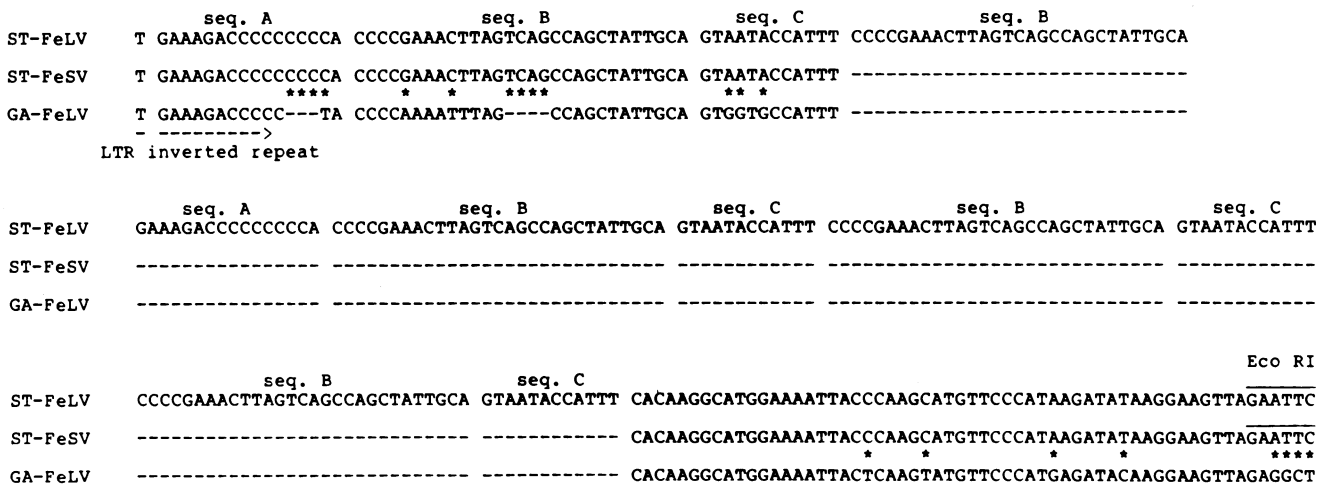

FIG. 2. Comparison of LTR sequences in ST-FeLV, ST-FeSV, and GA-FeLV. Homologous sequences between LTR inverted repeats and the *Eco*RI sites in the LTRs of ST-FeLV and ST-FeSV are compared. Sequence differences between adjacent lines of sequence are indicated by (*); deletions relative to ST-FeLV sequences are indicated by (—). Sequences (seq.) A, B, and C represent blocks repeated within the ST-FeLV LTR. The ST-FeSV LTR sequence is from Hampe et al. (4).

genes were found to be clustered in four specific regions of the protein. To further analyze these regions, the algorithm of Hopp and Woods (6) was used to compute hydrophilicity profiles of the FeLV *env* proteins. These results and those of a comparable analysis of Moloney murine leukemia virus gp70 (19) and Moloney mink cell focus-forming murine leukemia virus gp70 (1) are shown in Fig. 3. As expected from conservation of the primary amino acid sequences, substantial structural homology exists among the C-terminal gp70 and p15E regions of these murine and feline proteins.

Three of the four regions of divergence between the ST- and GA-FeLV *env* proteins occur at hydrophilicity peaks within the C-terminal half of gp70 (as indicated in Fig. 1 and 3). Region I contains two potential glycosylation sites that are maintained; conservative amino acid changes occur within the two N-X-T sequences. Region II contains predominantly nonconservative amino acid changes, one of which destroys a potential glycosylation site present in GA-FeLV gp70. Region III covers the site at which the *env* precursor is cleaved during virus maturation to yield gp70 and p15E; amino acid changes within this region are predominantly conservative. These regions are likely present on the surface of the protein, as judged both by hydrophilicity and by the presence of either potential glycosylation sites or protein cleavage sites. The fourth, and perhaps most striking, region of divergence between ST- and GA-FeLV *env* proteins is located within the signal peptide region of the initial *env* translation product. Although the amino acid sequence has diverged to 70% identity, changes are for the most part conservative, and the regions maintain the characteristic features of eucaryotic signal peptides (20).

We interpret these regions of amino acid differences between ST- and GA-FeLV to identify regions of the protein in which protein structure, and not specific amino acid sequence per se, is important for function. The divergence observed within the putative signal peptide suggests that substantial flexibility exists within primary amino acid sequences compatible with signal peptide function. Models utilizing structural requirements for signal peptide function have been described by others (20). Analogously, the divergent regions of gp70 may also serve a structural, rather than sequence-specific, role in gp70 function. An alternative possibility, that these sequence differences delineate some as yet undefined functional difference between ST- and GA-FeLV, cannot be excluded at this time.

We have presented here the nucleotide sequences of the envelope genes of two isolates of FeLV-B and have compared the deduced amino acid sequences of the *env* gene products. The molecularly cloned ST-FeLV sequence used in this study was derived from a viral genome defective in *env*. We believe, however, that the described amino acid changes between the related ST- and GA-FeLV are significant and relate to structural and functional features of the *env* gene products. Analysis of additional nondefective isolates of FeLV-B and of other FeLV subgroups will generate more extensive correlates of FeLV *env* structure and FeLV viral function.

We thank C. Sherr, J. Mullins, A. Roach, and N. Davidson for providing us with the molecularly cloned FeLV sequences used in this study and for useful discussions.

This work was partly supported by funds from TechAmerica Group, Inc.

LITERATURE CITED

1. **Bosselman, R. A., F. van Straaten, C. Van Beveren, I. M. Verma, and M. Vogt.** 1982. Analysis of the *env* gene of a molecularly cloned and biologically active Moloney mink cell focus-forming proviral DNA. J. Virol. **44**:19–31.
2. **Elder, J. H., and J. I. Mullins.** 1983. Nucleotide sequence of the envelope gene of Gardner-Arnstein feline leukemia virus B reveals unique sequence homologies with a murine mink cell focus-forming virus. J. Virol. **46**:871–880.
3. **Fischinger, P. J., C. S. Blevins, and S. Nomura.** 1974. Simple, quantitative assay for both xenotropic murine leukemia and ecotropic feline leukemia viruses. J. Virol. **14**:177–179.
4. **Hampe, A., M. Gobet, J. Even, C. J. Sherr, and F. Galibert.** 1983. Nucleotide sequences of feline sarcoma virus long terminal repeats and 5′ leaders show extensive homology to those of other mammalian retroviruses. J. Virol. **45**:466–472.
5. **Hardy, W. D., Jr., L. J. Old, P. W. Hess, M. Essex, and S. Cotter.** 1973. Horizontal transmission of feline leukaemia virus. Nature (London) **244**:266–269.

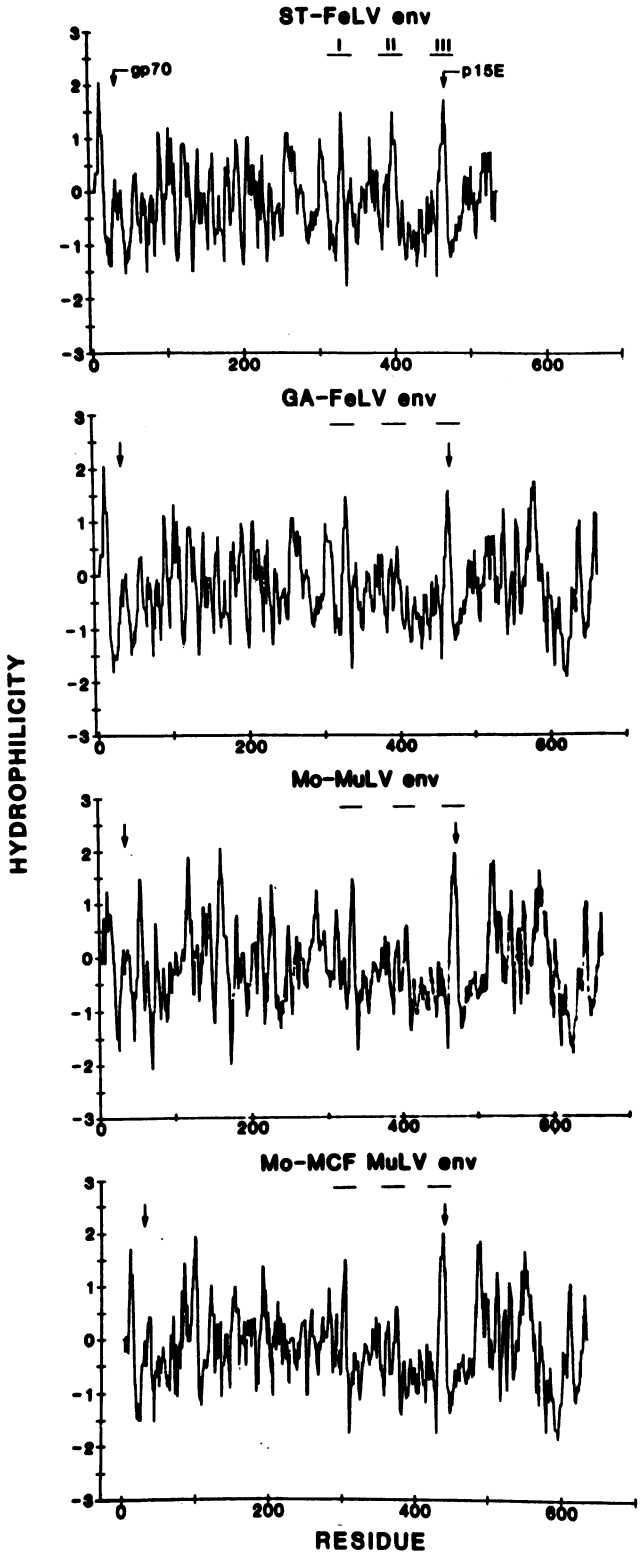

FIG. 3. Comparison of hydrophilicity profiles of *env* gene products of ST-FeLV, GA-FeLV, Moloney murine leukemia virus, and Moloney mink cell focus-forming murine leukemia virus. Analysis, using the algorithm developed by Hopp and Woods (6), is shown. Profiles are aligned at the gp70-p15E junction (indicated). Regions of clustered divergence between ST-FeLV gp70 and GA-FeLV gp70 are indicated. Moloney murine leukemia virus sequences are derived from Shinnick et al. (19); Moloney mink cell focus-forming murine leukemia virus sequences are derived from Bosselman et al. (1).

6. **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.
7. **Jarrett, O., W. D. Hardy, Jr., M. C. Golder, and D. Hay.** 1978. The frequency of occurrence of feline leukaemia virus subgroups in cats. Int. J. Cancer. **21**:334–337.
8. **Jarrett, O., and P. H. Russell.** 1978. Differential growth and transmission in cats of feline leukaemia viruses of subgroups A and B. Int. J. Cancer. **21**:466–472.
9. **Levinson, B., G. Khoury, G. Vande Woude, and P. Gruss.** 1982. Activation of SV40 genome by 72-base pair tandem repeats of Moloney sarcoma virus. Nature (London) **295**:568–572.
10. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labelled DNA with base-specific chemical cleavage. Methods Enzymol. **65**:499–560.
11. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
12. **Mullins, J. I., J. W. Casey, M. O. Nicolson, K. Bauman Burck, and N. Davidson.** 1981. Sequence arrangement and biological activity of cloned feline leukemia virus proviruses from a virus-productive human cell line. J. Virol. **38**:688–703.
13. **Pedersen, N. C., G. Theilen, M. A. Keane, L. Fairbanks, T. Mason, B. Orser, C.-H. Chen, and C. Allison.** 1977. Studies of naturally transmitted feline leukemia virus infection. Am. J. Vet. Res. **38**:1523–1531.
14. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5467.
15. **Sarma, P. S., and T. Log.** 1971. Viral interference in feline leukemia-sarcoma complex. Virology **44**:352–358.
16. **Sarma, P. S., and T. Log.** 1973. Subgroup classification of feline leukemia and sarcoma viruses by viral interference and neutralization tests. Virology **54**:160–169.
17. **Sarma, P. S., T. Log, D. Jain, P. R. Hill, and R. J. Heubner.** 1975. Differential host range of viruses of feline leukemia-sarcoma. Virology **64**:438–446.
18. **Sherr, C. J., L. A. Fedele, M. Oskarsson, J. Maizel, and G. Vande Woude.** 1980. Molecular cloning of Snyder-Theilen feline leukemia and sarcoma viruses: comparative studies of feline sarcoma virus with its natural helper virus and with Moloney murine sarcoma virus. J. Virol. **34**:200–212.
19. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukemia virus. Nature (London) **293**:543–548.
20. **Steiner, D. F., P. S. Quinn, C. Patzelt, S. J. Chan, J. Marsh, and H. S. Tager.** 1980. Proteolytic cleavage in the posttranslational processing of proteins, p. 175–202. *In* D. M. Prescott and L. Goldstein (ed.), Cell biology, vol. 4. Academic Press, Inc., New York.