JOURNAL OF VIROLOGY, Sept. 1983, p. 413–420
0022-538X/83/090413-08$02.00/0

Vol. 47, No. 3

Nucleotide Sequence of the gp70 Gene of Murine Retrovirus MCF 247

CHRISTIE A. HOLLAND, JOHN WOZNEY, AND NANCY HOPKINS*

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received 14 March 1983/Accepted 19 May 1983

We determined the nucleotide sequence and predicted the amino acid sequence of the gp70 gene of MCF 247, a recombinant murine retrovirus isolated from an AKR mouse. Information specifying the first 286 amino acids of the protein was probably derived from the presumptive nonecotropic parent of MCF 247, whereas the C-terminal 154 amino acids were probably derived from the ecotropic parent Akv. The nonecotropic sequences at the amino terminus of MCF 247 show only 38% homology, at the amino acid level, to those of Akv. In contrast, these sequences are strikingly similar (99% homologous) to those reported for another MCF virus, Moloney MCF, which was isolated from a BALB/c mouse. Moloney MCF also has ecotropic-derived sequences encoding the C-terminal portion of its gp70 protein; however, the recombination event that introduced these sequences occurs 213 nucleotides further towards the C terminus of gp70 than it does in MCF 247.

MCF 247 is the prototype of a class of dual-tropic retroviruses isolated from thymic lymphomas of AKR mice (7, 9). These viruses are XC^-, induce foci of morphological alteration on mink cells (MCF^+), are thymotropic, and accelerate leukemia when injected into newborn AKR mice (3). They arise by recombination between inherited ecotropic Akv viruses and other nonecotropic endogenous viruses. It has been suggested that their novel envelopes play a role in the leukemogenic process (6, 22).

The envelope (*env*) gene product of murine leukemia viruses is translated as a glycosylated precursor (Prgp85), which is proteolytically cleaved to yield gp70 (4, 26) located on the virion surface; Prp15E, which in turn is cleaved to yield p15E (18) anchored to the viral membrane; and a small peptide, 16 to 21 amino acids long, of unknown function (8).

To determine the structure of the gp70 gene of MCF 247, we determined a 1,320-nucleotide-long sequence encompassing the gp70 gene of an infectious DNA clone of MCF 247. We compared this sequence and the predicted amino acid sequence of gp70 with those of another MCF virus, Moloney (Mo)-MCF (1), as well as with the ecotropic progenitors of these MCF viruses, Akv (13) and Moloney virus (25), respectively.

MATERIALS AND METHODS

Molecular clones. The MCF 247 clone (λ247-9) was generated essentially as described by Kelly et al. (12), except that a λ vector, NM788, was used. When the DNA insert of λ247-9 is transfected in SC-1 or NIH 3T3 cells, infectious viral particles are produced which are XC^- and MCF^+ and accelerate leukemia when injected into newborn AKR mice. The gp70-coding sequences of λ247-9 contain all of the restriction endonuclease sites characterized by Chattopadhyay et al. (2). A subclone (pgp70-1) of the gp70-coding sequences of λ247-9 was prepared by isolating a 3.2-kilobase fragment from the unique *Xho*I site in the polymerase gene to the *Xba*I site in the p15E gene. The 3.2-kilobase viral fragment and *Sal*I-digested XF3 plasmid DNA (a derivative of pBR322 constructed by D. Hannahan) were ligated, which resulted in linear molecules joined at the *Xho*I and *Sal*I sites. These were filled by using the large fragment of DNA polymerase I, ligated to *Sal*I linkers, recleaved with *Sal*I, cyclized, and used to transform the bacterial strain DH1 (D. Hannahan, J. Mol. Biol., in press).

Sequence determination. The subclone (pgp70-1) or the *Xho*I-to-*Xba*I 3.2-kilobase fragment isolated directly from λ247-9 was cleaved with restriction endonucleases. Fragments isolated from either preparative agarose or acrylamide gels were labeled at their 5′ ends with [γ-^{32}P]ATP and polynucleotide kinase. Labeled fragments were either recleaved and isolated on agarose gels or strand separated on acrylamide gels, and their sequence was determined by using the procedures of Maxam and Gilbert (16).

RESULTS AND DISCUSSION

Nucleotide sequence of the gp70 gene of MCF 247. The sequence of the gp70 gene of MCF 247 was determined by using the method of Maxam and Gilbert (16) to sequence the fragments shown in Fig. 1. The fragments were generated

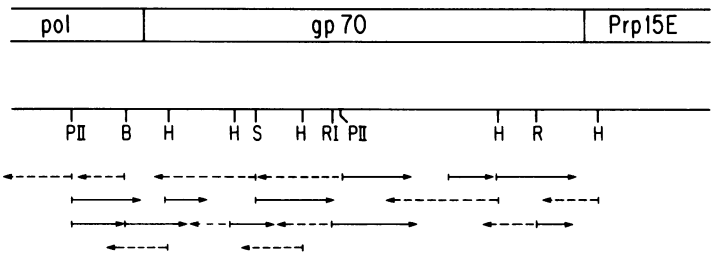


FIG. 1. Strategy for sequencing the gp70 gene of MCF 247. Arrows indicate regions of DNA sequenced from the restriction site corresponding to the tail of the arrow. (←) Regions sequenced from minus-strand DNA; (→) regions sequenced from plus-strand DNA. Restriction enzymes are as follows: PII, *Pvu*II; B, *Bam*HI; H, *Hin*fI; S, *Sma*; RI, *Eco*RI; R, *Rsa*I.

either from an infectious molecular clone of MCF 247 (λ247-9) or from a subclone of the gp70-coding sequences (pgp70-1). The sequence that was determined had one large open reading frame of 1,320 nucleotides contiguous with the sequence of 603 nucleotides encoding the Prp15E of MCF 247 determined by Kelly et al. (12). The termination codon for the open reading frame is located after the Prp15E coding sequences, 34 nucleotides before the long terminal repeat.

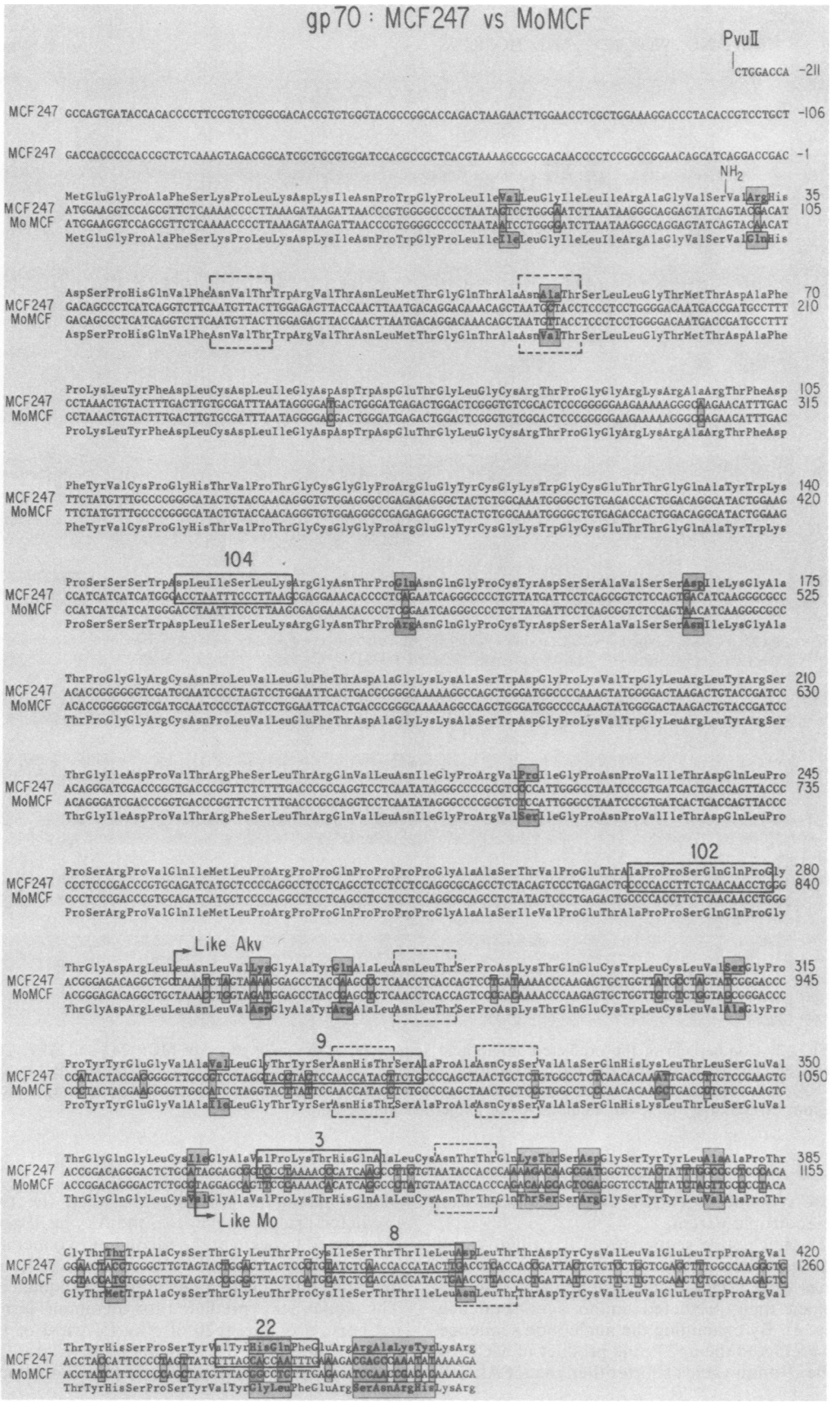
The nucleotide sequence and the predicted amino acid sequence of the gp70 of MCF 247 are shown in Fig. 2. Although the NH_2-terminal amino acids of MCF 247 have not been determined, Schultz et al. (23) have recently determined the sequence of the 25 amino-terminal amino acids of the gp69 protein of a Rauscher-derived MCF virus. An almost identical sequence (1 amino acid change out of 25) occurs in MCF 247, and so we concluded that the NH_2 terminus of mature gp70 is located at the Val residue coded for by nucleotides 97 to 99 in Fig. 2. A single methionine codon is located 96 nucleotides upstream and in phase with this residue. These 96 nucleotides code for 32 predominantly uncharged, hydrophobic amino acids, typical of a leader sequence (5). A 3′ splice consensus sequence (24) (CTCTCCAAG) was found 247 nucleotides 5′-ward of the initiator methionine of gp70 (data not shown).

The carboxyl terminus of gp70 is not yet known for any murine retrovirus. Previously, by analogy with Moloney murine leukemia virus (MoMuLV) we placed the amino terminus of p15E of MCF 247 at the first amino acid after those shown in Fig. 2 (12). It is possible that several amino acids of gp70 are lost during the cleavage from p15E (5, 11), and thus, the carboxyl terminus of the mature gp70 protein might be different from that shown in Fig. 2.

Six sequences which can serve as glycosylation sites (Asn-X-Thr and Asn-X-Ser) (17, 21) were found in the gp70 of MCF 247. Five of these sites are shared with Mo-MCF, Akv, and Mo-MuLV, and four are clustered in the carboxyl-terminal half of gp70 (Fig. 2 and 5). The sixth site is present near the N terminus of both MCF 247 and Mo-MCF but is not present in either of the ecotropic viruses.

Comparison of the nucleotide sequences encoding the gp70 proteins of Akv and MCF 247. We compared the nucleotide sequences of the gp70 genes of MCF 247 with those of its putative ecotropic parent, Akv (13). It was immediately clear that the nucleotide sequences encoding the carboxyl-terminal third of these gp70 proteins were nearly identical. Except for a single nucleotide difference (nucleotide 1,097, Fig. 2), the sequence of the gp70 gene of MCF 247 is identical to that of Akv from the arrow at nucleotide 857 in Fig. 2 to the beginning of the p15E gene. Since the single base change could well be a mutation in one of the clones, we presume that

FIG. 2. Comparison of the nucleotide sequences and predicted amino acid sequences of the gp70 genes of MCF 247 and Mo-MCF. The nucleotide sequence of the gp70 gene of MCF 247 was determined as described in the text with the translation of the *env* gene product presented above. The predicted amino terminus of the protein was identified by comparison with protein sequencing data for a Rauscher MCF virus (23). The nucleotide sequence and predicted amino acid sequence for Mo-MCF are those of Bosselman et al. (1). Nucleotide and amino acid differences between the viruses are accentuated by gray boxes. The points at which the nucleotide sequences of each MCF isolate become like those of their respective ecotropic parents are indicated arrows. The broken lines indicate potential glycosylation sites. Unshaded boxes enclose sequences corresponding to large RNase T1-resistant oligonucleotides analyzed previously in Akv and MCF 247 fingerprints (14, 15, 19, 20).


gp70 : MCF247 vs MoMCF
PvuII
CTGGACCA -211
MCF 247 GCCAGTGATACCACACCCCTTCCGTGTCGGCGACACCGTGTGGGTACGCCCGGCACCAGACTAAGAACTTGGAACCTCGCTGGAAAGGACCCTACACCGTCCTGCT -106
MCF247 GACCACCCCCACCGCTCTCAAAGTAGACGGCATCGCTGCGTGGATCCACGCCCGCTCACGTAAAAGCCGCCGACAACCCCTCCGGCCCGGAACAGCATCAGGACCGAC -1
NH2
MCF247 MetGluGlyProAlaPheSerLysProLeuLysAspLysIleAsnProTrpGlyProLeuIleValLeuGlyIleLeuIleArgAlaGlyValSerValArgHis 35
Mo MCF ATGGAAGGTCCAGCGTTCTCAAAACCCCTTAAAGATAAGATTAACCCGTGGGGCCCCCTAATAGTCCTGGGAATCTTAATAAGGGCAGGAGTATCAGTACGACAT 105
ATGGAAGGTCCAGCGTTCTCAAAACCCCTTAAAGATAAGATTAACCCGTGGGGCCCCCTAATAATCCTGGGGATCTTAATAAGGCCAGGAGTATCAGTACAACAT
MetGluGlyProAlaPheSerLysProLeuLysAspLysIleAsnProTrpGlyProLeuIleIleLeuGlyIleLeuIleArgAlsGlyValSerValGlnHis
MCF247 AspSerProHisGlnValPheAsnValThrTrpArgValThrAsnLeuMetThrGlyGlnThrAlaAsnAlaThrSerLeuLeuGlyThrMetThrAspAlaPhe 70
MoMCF GACAGCCCTCATCAGGTCTTCAATGTTACTTGGAGAGTTACCAACTTAATGACAGGACAAACAGCTAATGCTACCTCCCTCCTGGGGACAATCACCGATGCCTTT 210
GACAGCCCTCATCAGGTCTTCAATGTTACTTGGAGAGTTACCAACTTAATGACAGGACAAACAGCTAATGTTACCTCCCTCCTGGGGACAATGACCGATCCCTTT
AspSerProHisGlnValPheAsnValThrTrpArgValThrAsnLeuMetThrGlyGlnThrAlaAsnValThrSerLeuLeuGlyThrMetThrAspAlaPhe
MCF247 ProLysLeuTyrPheAspLeuCysAspLeuIleGlyAspAspTrpAspGluThrGlyLeuGlyCysArgThrProGlyGlyArgLysArgAlaArgThrPheAsp 105
MoMCF CCTAAACTGTACTTTGACTTGTGCGATTTAATAGGGGATGACTGGGATGAGACTGGACTCGGGTGTCGCACTCCCGGGGGAAGAAAAAGGGCGACAACATTTGAC 315
CCTAAACTGTACTTTGACTTGTGCGATTTAATAGGGGATGACTGGGATGAGACTGGACTCGGGTGTCGCACTCCCGGGGGAAGAAAAAGGGCGACAACATTTGAC
ProLysLeuTyrPheAspLeuCysAspLeuIleGlyAspAspTrpAspGluThrGlyLeuGlyCysArgThrProGlyGlyArgLysArgAlaArgThrPheAsp
MCF247 PheTyrValCysProGlyHisThrValProThrGlyCysGlyGlyProArgGluGlyTyrCysGlyLysTrpGlyCysGluThrThrGlyGlnAlaTyrTrpLys 140
MoMCF TTCTATGTTTGCCCCGGGCATACTGTACCAACAGGGTGTGGAGGGCCGAGAGAGGGCTACTGTGGCAAATGGGGCTGTGAGACCACTGGACAGGCATACTGGAAG 420
TTCTATGTTTGCCCCGGGCATACTGTACCAACAGGGTGTGGAGGGCCGAGAGAGGGCTACTGTGGCAAATGGGGCTGTGAGACCACTGGACAGGCATACTGGAAG
PheTyrValCysProGlyHisThrValProThrGlyCysGlyGlyProArgGluGlyTyrCysGlyLysTrpGlyCysGluThrThrGlyGlnAlaTyrTrpLys
104
MCF247 ProSerSerSerTrpAspLeuIleSerLeuLysArgGlyAsnThrProGlnAsnGlnGlyProCysTyrAspSerSerAlaValSerSerAspIleLysGlyAla 175
MoMCF CCATCATCATCATGGGACCTAATTTCCCTTAAGCGAGGAAACACCCCTCAGAATCAGGGCCCCTGTTATGATTCCTCAGCGGTCTCCAGTGACATCAAGGGCGCC 525
CCATCATCATCATGGGACCTAATTTCCCTTAAGCGAGGAAACACCCCTCAGAATCAGGGCCCCTGTTATGATTCCTCAGCGGTCTCCAGTGACATCAAGGGCGCC
ProSerSerSerTrpAspLeuIleSerLeuLysArgGlyAsnThrProArgAsnGlnGlyProCysTyrAspSerSerAlaValSerSerAsnIleLysGlyAla
MCF247 ThrProGlyGlyArgCysAsnProLeuValLeuGluPheThrAspAlaGlyLysLysAlaSerTrpAspGlyProLysValTrpGlyLeuArgLeuTyrArgSer 210
MoMCF ACACCGGGGGGTCGATGCAATCCCCTAGTCCTGGAATTCACTGACGCCGGCCAAAAAGGCCAGCTGGGATGGCCCCAAAGTATGGGGACTAAGACTGTACCGATCC 630
ACACCGGGGGGTCGATGCAATCCCCTAGTCCTGGAATTCACTGACGCCGGCCAAAAAGGCCAGCTGGGATGGCCCCAAAGTATGGGGACTAAGACTGTACCGATCC
ThrProGlyGlyArgCysAsnProLeuValLeuGluPheThrAspAlaGlyLysLysAlaSerTrpAspGlyProLysValTrpGlyLeuArgLeuTyrArgSer
MCF247 ThrGlyIleAspProValThrArgPheSerLeuThrArgGlnValLeuAsnIleGlyProArgValProIleGlyProAsnProValIleThrAspGlnLeuPro 245
MoMCF ACAGGGATCGACCCGGTGACCCGGTTCTCTTTGACCCGCCAGGTCCTCAATATAGGGCCCCCGCCTCCCCATTGGGCCTAATCCCGTGATCACTGACCAGTTACCC 735
ACAGGGATCGACCCGGTGACCCGGTTCTCTTTGACCCGCCAGGTCCTCAATATAGGGCCCCCGCCTCCCCATTGGGCCTAATCCCGTGATCACTGACCAGTTACCC
ThrGlyIleAspProValThrArgPheSerLeuThrArgGlnValLeuAsnIleGlyProArgValSerIleGlyProAsnProValIleThrAspGlnLeuPro
102
MCF247 ProSerArgProValGlnIleMetLeuProArgProProGlnProProProProGlyAlaAlaSerThrValProGluThrAlaProProSerGlnGlnProGly 280
MoMCF CCCTCCCGACCCGTCAGATCATGCTCCCCAGGCCTCCTCAGCCTCCTCCTCCAGGGCGCAGCCTCTACAGTCCCTGAGACTGCCCCACCTTCTCAACAACCTGGG 840
CCCTCCCGACCCGTCAGATCATGCTCCCCAGGCCTCCTCAGCCTCCTCCTCCAGGGCGCAGCCTCTATAGTCCCTGAGACTGCCCCACCTTCTCAACAACCTGGG
ProSerArgProValGlnIleMetLeuProArgProProGlnProProProProGlyAlaAlaSerIleValProGluThrAlaProProSerGlnGlnProGly
Like Akv
MCF247 ThrGlyAspArgLeuLeuAsnLeuValLysGlyAlaTyrGlnAlaLeuAsnLeuThrSerProAspLysThrGlnGluCysTrpLeuCysLeuValSerGlyPro 315
MoMCF ACGGGAGACAGGCTGCTAAACTCTAGTAAAAGGAGCCTACCAAGGCCTCAACCTCACCAGTCCTGATAAAACCCAAGAGTGCTGGTTATGCCTAGTATCCGGACCC 945
ACGGCAGACAGGCTGCTAAACCTCGTAGATGGAGCCTACCAAGCCCTCAACCTCACCAGTCCGGACAAAACCCAAGAGTGCTGGTTGTGCCTGGTAGCCGGACCC
ThrGlyAspArgLeuLeuAsnLeuValAspGlyAlaTyrArgAlaLeuAsnLeuThrSerProAspLysThrGlnGluCysTrpLeuCysLeuValAlaGlyPro
9
MCF247 ProTyrTyrGluGlyValAlaValLeuGlyThrTyrSerAsnHisThrSerAlaProAlaAsnCysSerValAlaSerGlnHisLysLeuThrLeuSerGluVal 350
MoMCF CCATACTACGAGGGGGTTGCCGTCCTAGGTACGTATTCCAACCATACTTCTGCCCCAGCTAACTGCTCTGCCGTGGCCTCCAACACAAGTTCACCTTGTCCGAAGTC 1050
CCCTACTACGAAGGGGTTGCCATCCTAGGTACTTATTCCAACCATACCTCTGCCCCAGCTAACTGCTCCGTGGCCTCCAACACAAGCTCACCTTGTCCGAAGTC
ProTyrTyrGluGlyValAlaIleLeuGlyThrTyrSerAsnHisThrSerAlaProAlaAsnCysSerValAlaSerGlnHisLysLeuThrLeuSerGluVal
3
MCF247 ThrGlyGlnGlyLeuCysIleGlyAlaValProLysThrHisGlnAlaLeuCysAsnThrThrGlnLysThrSerAspGlySerTyrTyrLeuAlaAlaProThr 385
MoMCF ACCGGACAGGGACTCTGCATAGGAGCGGTCCCTAAAACCCATCAAGCCTTGTGTAATACCCACCCAAAAGACAAGCGATCGGTCCTACTATTTGCCCCCTCCCACA 1155
ACCGGACAGGGACTCTGCGTAGGAGCAGTTCCCAAAACACATCAGGCCCTATGTAATACCCACCCAGACAAGCAGTCGACGGTCCTATTATTTGCCCCCTACA
ThrGlyGlnGlyLeuCysValGlyAlaValProLysThrHisGlnAlaLeuCysAsnThrThrGlnThrSerSerArgGlySerTyrTyrLeuValAlaProThr
Like Mo
8
MCF247 GlyThrThrTrpAlaCysSerThrGlyLeuThrProCysIleSerThrThrIleLeuAspLeuThrThrAspTyrCysValLeuValGluLeuTrpProArgVal 420
MoMCF GGAACTACCTGGCCTTGTAGTACTCGACTTACTCCGTGTATCTCCACCACCATACTTGACCTGACCACCGATTACTGTGTCCTGGTCGAGCTTTGGCCAAGGGTC 1260
GGTACCATGTGGCCTTGTAGTACCGGCTTACTCCATGTATCTCGACCACCATACTGAACCTTACCACCGATTATTGTGTTCTTGTCGAACTGTGGCCAAGAGTC
GlyThrMetTrpAlaCysSerThrGlyLeuThrProCysIleSerThrThrIleLeuAsnLeuThrThrAspTyrCysValLeuValGluLeuTrpProArgVal
22
MCF247 ThrTyrHisSerProSerTyrValTyrHisGlnPheGluArgArgAlaLysTyrLysArg
MoMCF ACCTACCATTCCCCTAGTTATGTTTACCACCAATTTGAAAGACGAGCCAAATATAAAAGA
ACCTATCATTCCCCGAGCTATGTTTACGGCCGCTTTGAGAGATCCAACCGACACAAAAGA
ThrTyrHisSerProSerTyrValTyrGlyLeuPheGluArgSerAsnArgHisLysArg

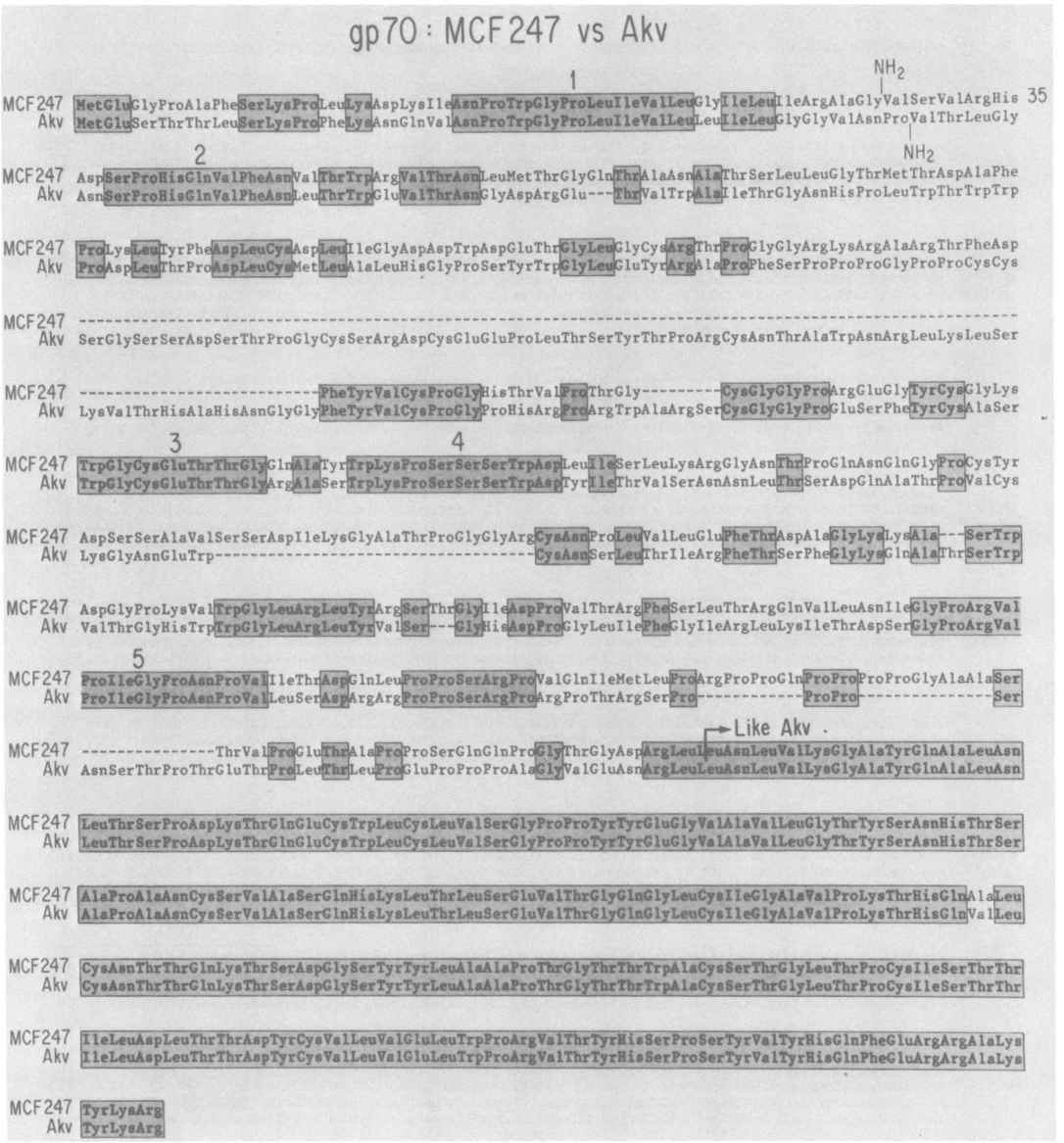


FIG. 3. Comparison of the predicted amino acid sequences of the gp70 genes of MCF 247 and Akv. The predicted amino acid sequence of the gp70 gene of MCF 247 was aligned to that of Akv, determined by Lenz et al. (13), to allow maximum homology of amino acids. Identical amino acids are accentuated by gray boxes. Regions of homology of seven amino acids or longer are numbered.

these sequences of MCF 247 are derived from the ecotropic parent.

The nucleotide sequences of the amino-terminal two-thirds of the Akv and MCF 247 gp70 genes were very difficult to align. Therefore, we aligned their predicted amino acid sequences (Fig. 3). By examining the nucleotide sequences we predicted the gp70 gene product of MCF 247 to be 29 amino acids shorter than that of Akv. To obtain maximum homology between the two predicted protein sequences, the Akv gp70 gene is assumed to begin at the second of two methionine residues identified by Lenz et al. (13), and deletions have been allowed in both sequences. The amino acids predicted for the amino-terminal portion of the gp70 of Akv (5′-ward of the arrow, Fig. 3) are 38% homologous to those of MCF 247. To determine whether the amino acid

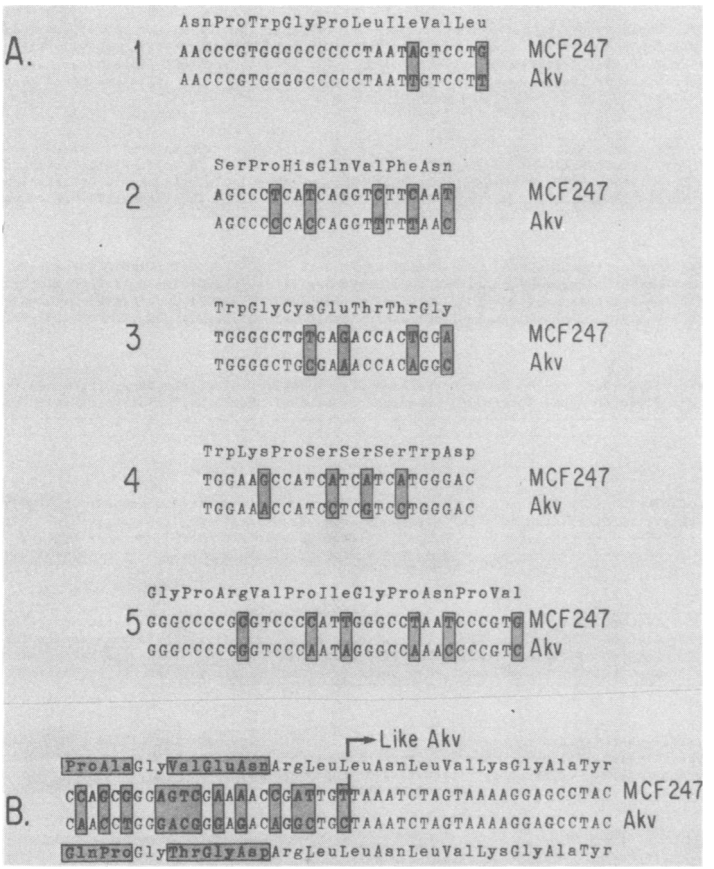
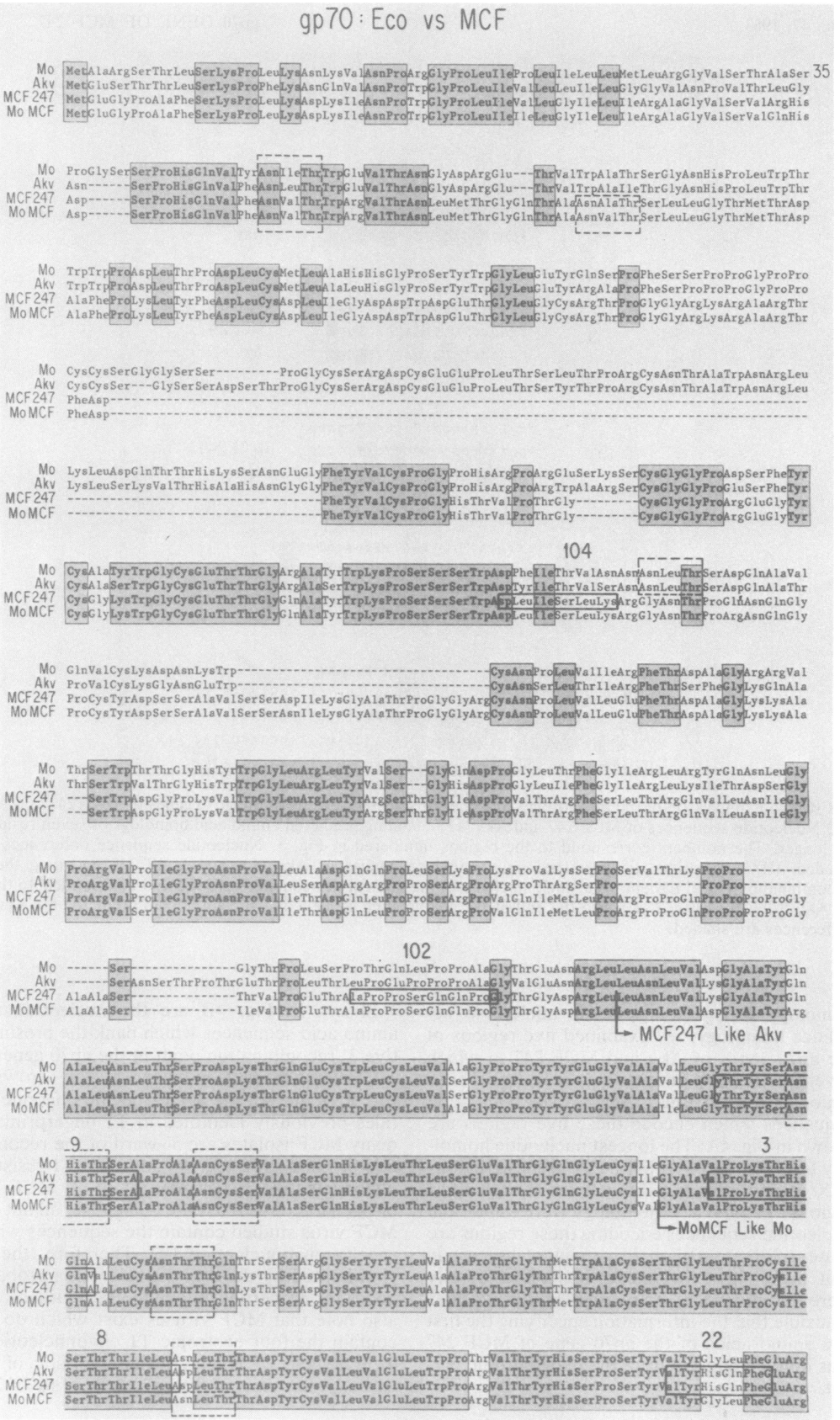


FIG. 4. Comparison of the nucleotide sequences of selected regions of the gp70 genes of MCF 247 and Akv. (A) Nucleotide sequences of MCF 247 and Akv (13) encoding predicted amino acid homologs of seven residues or longer. The numbers correspond to the regions numbered in Fig. 3. Nucleotide sequence differences are shaded. (B) Nucleotide and predicted amino acid sequences of Akv (13) and MCF 247 encoding the 3′ recombination point. Except for a single nucleotide, the nucleotide sequences of MCF 247 are identical to those of Akv from the arrow 3′-ward to the end of the gp70 coding sequences. Nucleotide and predicted amino acid differences are shaded.

homology reflected significant nucleotide sequence homology, we examined five regions of the gp70 genes of Akv and MCF 247 in which seven amino acids or longer are identical. These regions are numbered in Fig. 3. The nucleotide sequences which encode these five regions are shown in Fig. 4A. The longest nucleotide homology between Akv and MCF 247 in these regions is 20 nucleotides long. Thus, even though short regions of amino acid homology were found, the nucleotide sequences encoding these regions are quite different. Although we cannot be certain that these short regions of nucleotide identity were not introduced by recombination, it seems plausible that the information specifying the first 286 amino acids of the gp70 gene of MCF 247 was derived entirely from the nonecotropic parent.

Shown in Fig. 4B are the nucleotide and amino acid sequences which flank the presumptive 3′ recombination point in the gp70 genes of MCF 247 and Akv, indicated by the arrows in Fig. 2, 3, and 4. The MCF-specific oligonucleotides previously identified in T1 fingerprints of many MCF isolates are 5′-ward of the recombination point (13, 15, 19, 20). The T1-resistant oligonucleotide 104 is present in all MCF isolates characterized to date. Likewise, all but one MCF virus studied contain the sequences which specify oligonucleotide 102. Therefore, the region limited by these oligonucleotides probably codes for properties unique to MCF viruses. We also note that MCF viruses exist which do not contain the four ecotropic T1 oligonucleotides, 9, 3, 8, and 22 (Fig. 2), characteristic of the carboxyl terminus of the gp70 gene of MCF 247,


gp70: Eco vs MCF
Mo
Akv
MCF247
Mo MCF
35
104
102
MCF247 Like Akv
MoMCF Like Mo
9
3
8
22

but instead inherit the C-terminal sequences of gp70 from their nonecotropic parent (13). Although the nucleotide sequences of these viruses are therefore different in the region encoding the C terminus of gp70, they could still code for proteins of a very similar amino acid sequence. It is therefore impossible to conclude from these data which region of gp70 specifies the host range.

We have not yet located the nucleotide sequence at which recombination introduces nonecotropic information to the 5′ side of *env*. In an attempt to do this, we compared the nucleotide sequences before the initiator Met codon of MCF 247 shown in Fig. 2 with those of Akv determined by Herr et al. (10). Of 218 nucleotides shown, there are 29 nucleotide differences resulting in six amino acid changes. There is also one deletion of four amino acids in this region. Preliminary sequencing evidence shows that there are still limited nucleotide differences for at least an additional 220 nucleotides 5′-ward of the *Pvu*II site shown in Fig. 2. Therefore, the 5′ recombination point must be more than 435 nucleotides 5′ to the start of the gp70-coding sequences.

Comparison of the nucleotide and predicted amino acid sequences of the gp70 genes of MCF 247 and Mo-MCF. We have compared the nucleotide and amino acid sequences of the gp70 gene of MCF 247 with those determined by Bosselman et al. (1) for an MCF virus isolated from a BALB/Mo mouse (Fig. 2). Although by examining the nucleotide sequences (Fig. 2) we predicted that both genes code for proteins of identical size, the recombination events which generated MCF 247 and Mo-MCF did not occur at the same position within the molecules. Interestingly, although these two MCF viruses were generated by recombination within two different strains of mice (Akv and BALB), the nonecotropic sequences that each has acquired are remarkably similar. Thus, Mo-MCF and MCF 247 have acquired gene substitutions in the amino-terminal two-thirds of gp70 that are 99% homologous even though they arose in different mouse strains by recombination involving different ecotropic parents. The 25 N-terminal amino acids of Rauscher MCF (23) are also highly homologous to those of MCF 247 and Mo-MCF.

Comparison of the predicted amino acid sequences of the gp70 genes of two MCF viruses with their ecotropic progenitors. In Fig. 5, the predicted amino acid sequences of the gp70 genes of MCF 247 and Mo-MCF (1) have been aligned with those of Mo-MuLV (25) and Akv (13) as in Fig. 3, and common amino acids are shaded. The figure emphasizes that the carboxyl-terminal third of all four proteins is conserved. In part, this is because, to the 3′ side of the recombination points, the sequences of each of the MCF-viruses are like those of their ecotropic progenitors. However, the nucleotide sequences between the 3′ recombination points of the two MCF viruses are 90% homologous and the predicted amino acid sequences in this region are 93% homologous, even though in MCF 247 these sequences were inherited from the ecotropic parent, whereas in Mo-MCF they were derived from the nonecotropic parent. Therefore, the nonecotropic parent of Mo-MCF is similar in amino acid sequence to the two ecotropic viruses in this C-terminal region of gp70.

ACKNOWLEDGMENTS

We thank D. Hannahan for providing XF3 and Lynn Enquist for the λ vector (NM 788). We thank Alan Schultz for a preprint of the amino acid sequence of the Rauscher MCF *env* gene.

C.A.H. acknowledges the support of the Muscular Dystrophy Association. J.W. acknowledges the support of the Damon Runyon Walter Winchell Cancer Fund (grant no. DRG-471-F). This work was supported by grants from the National Institutes of Health (Program Project grant no. P01-CA26717 and grant no. R01-CA19308) to N.H. and partially by a Public Health Service core grant (CA14051) to S. E. Luria.

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. **Chattopadhyay, S. K., M. R. Lander, S. Gupta, E. Rands, and D. R. Lowy.** 1981. Origin of mink cytopathic focus-forming (MCF) viruses: comparison with ecotropic and xenotropic murine leukemia virus genomes. Virology **113**:465–483.
3. **Cloyd, M. W., J. W. Hartley, and W. P. Rowe.** 1980. Lymphomagenicity of recombinant mink cell focus-inducing murine leukemia viruses. J. Exp. Med. **151**:542–552.
4. **Dickson, C., R. Eisenman, H. Fan, E. Hunter, and N. Teich.** 1982. Protein biosynthesis and assembly, p. 513–648. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
5. **Engelman, D. M., and T. A. Steitz.** 1981. The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. Cell **23**:411–422.
6. **Famulari, N. G.** 1982. Murine leukemia viruses with

FIG. 5. Comparison of the predicted amino acid sequences of the gp70 genes of MCF 247 and Mo-MCF with that of their ecotropic progenitors, Akv and Mo-MuLV. The predicted amino acid sequences of Mo-MuLV, Mo-MCF, and Akv are those of Shinnick et al. (25), Bosselman et al. (1), and Lenz et al. (13), respectively. The sequences were aligned to allow maximum homology. Amino acids shared in all four viruses are shaded. Unshaded boxes indicate T1 oligonucleotides previously identified in fingerprints of MCF 247 or Akv. Dashed boxes represent potential glycosylation sites.

recombinant *env* genes: a discussion of their role in leukemogenesis. Curr. Top. Microbiol. Immunol. **103**:75–108.

7. **Fischinger, P. J., S. Nomura, and D. P. Bolognesi.** 1975. A novel murine oncornavirus with dual eco- and xenotropic properties. Proc. Natl. Acad. Sci. U.S.A. **72**:5150–5155.
8. **Green, N., T. M. Shinnick, O. White, A. Ponticelli, J. G. Sutcliffe, and R. A. Lerner.** 1981. Sequence-specific antibodies show that maturation of Moloney leukemia virus envelope polyprotein involves removal of a COOH-terminal peptide. Proc. Natl. Acad. Sci. U.S.A. **78**:6023–6027.
9. **Hartley, J. W., N. K. Wolford, L. J. Old, and W. P. Rowe.** 1977. A new class of murine leukemia virus associated with development of spontaneous lymphomas. Proc. Natl. Acad. Sci. U.S.A. **74**:789–792.
10. **Herr, W., V. Corbin, and W. Gilbert.** 1982. The nucleotide sequence of the 3′ half of Akv. Nucleic Acids Res. **10**:6931–6944.
11. **Jou, W. M., M. Verhoeyen, R. Devos, E. Saman, R. Fang, D. Huylebroeck, W. Fiers, G. Threlfall, C. Barber, N. Carey, and S. Emtage.** 1980. Complete structure of the hemagglutinin gene from the human influenza A/Victoria/3/75 (H3N2) strain as determined from cloned DNA. Cell **19**:683–696.
12. **Kelly, M., C. A. Holland, M. L. Lung, S. K. Chattopadhyay, D. R. Lowy, and N. H. Hopkins.** 1983. Nucleotide sequence of the 3′ end of MCF 247 murine leukemia virus. J. Virol. **45**:291–298.
13. **Lenz, J., R. Crowther, A. Straceski, and W. Haseltine.** 1982. Nucleotide sequence of the Akv *env* gene. J. Virol. **42**:519–529.
14. **Lung, M. L., J. W. Hartley, W. P. Rowe, and N. H. Hopkins.** 1983. Large RNase T1-resistant oligonucleotides encoding p15E and the U3 region of the long terminal repeat distinguish two biological classes of mink cell focus-forming type C viruses of inbred mice. J. Virol. **45**:275–290.
15. **Lung, M. L., C. Hering, J. W. Hartley, W. P. Rowe, and N. Hopkins.** 1979. Analysis of the genomes of mink cell focus-inducing murine type C viruses: a progress report. Cold Spring Harbor Symp. Quant. Biol. **44**:1269–1274.
16. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. **74**:560–564.
17. **Montreuil, J.** 1980. Primary structure of glycoprotein glycans. Basis for the molecular biology of glycoproteins. Adv. Carbohydr. Chem. Biochem. **37**:157–223.
18. **Pinter, A., and E. Fleissner.** 1977. The presence of disulfide-linked gp70-p15(E) complexes in AKR murine leukemia virus. Virology **83**:417–422.
19. **Rommelaere, J., D. V. Faller, and N. Hopkins.** 1977. RNase T1-resistant oligonucleotides of Akv-1 and AKV-2 type C viruses of AKR mice. J. Virol. **24**:690–694.
20. **Rommelaere, J., D. V. Faller, and N. Hopkins.** 1978. Characterization and mapping of RNase T1-resistant oligonucleotides derived from the genomes of Akv and MCF murine leukemia viruses. Proc. Natl. Acad. Sci. U.S.A. **75**:495–499.
21. **Rosner, M. R., L. S. Grinna, and P. W. Robbins.** 1980. Differences in glycosylation patterns of closely related murine leukemia viruses. Proc. Natl. Acad. Sci. U.S.A. **77**:67–71.
22. **Rowe, W. P., M. W. Cloyd, and J. W. Hartley.** 1979. Status of the association of mink cell focus-forming viruses with leukemogenesis. Cold Spring Harbor Symp. Quant. Biol. **44**:1265–1268.
23. **Schultz, A., A. Rein, L. Henderson, and S. Oroszlan.** 1983. Biological, chemical, and immunological studies of Rauscher ecotropic and mink cell focus-forming viruses from JLS-V9 cells. J. Virol. **45**:995–1003.
24. **Sharp, P.** 1981. Speculations on RNA splicing. Cell **23**:643–646.
25. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukemia virus. Nature (London) **293**:543–548.
26. **Witte, O. N., and D. F. Wirth.** 1979. Structure of the murine leukemia virus envelope glycoprotein precursor. J. Virol. **29**:735–743.