B. A. Palevitz and P. K. Hepler [*Chromosoma* **46**, 297 (1974)] and Palevitz *et al.* (6). Slices were viewed on Reichert Zetopan microscopes equipped for epifluorescence, phase, or DIC viewing. Low-light television was performed with double- and triple-intensified video cameras (Venus Scientific Inc., Farmingdale, N.Y.) and ancillary equipment as described in (6).
8. Nanospec models SDP-2000 and 10-0112, Nanometrics Inc., Sunnyvale, Calif.
9. N. V. Raikhel, D. J. O'Kane, B. A. Palevitz, unpublished observations.
10. B. A. Palevitz and P. K. Hepler, *Planta* **132**, 71 (1976).
11. M. T. Davison and P. B. Garland, *J. Gen. Microbiol.* **98**, 147 (1977); B. Grobe and C. G. Arnold, *Protoplasma* **86**, 291 (1975); M. Pellegrini, *J. Cell Sci.* **46**, 313 (1980).
12. Supported by National Science Foundation grant PCM81-04470 to B.A.P. and Public Health Service national research service award to D.J.O.

1 July 1981; revised 11 August 1981

## Complete Nucleotide Sequence and Organization of the Moloney Murine Sarcoma Virus Genome

Abstract. *The complete nucleotide sequence of a mammalian transforming retrovirus, Moloney murine sarcoma virus, has been determined. MSV, a recombinant virus derived of helper viral and cellular sequences, possesses termini resembling prokaryotic transposable elements. The viral genome has the coding capacity for the Moloney murine leukemia virus* gag *gene product and contains large deletions in* pol *and* env *genes. A large open reading frame encompassing its cell-derived sequences codes for its putative transforming protein. The nature of some of the important domains in the viral genome has been established, and their structure is discussed in relation to their function.*

Type C RNA viruses (retroviruses) represent a class of genetic elements capable of neoplastic transformation in their natural host. These viruses can be grouped into two broad classes on the basis of their biological activities. The leukemia or leukosis viruses are replication-competent, cause leukemia in susceptible hosts, but do not transform cells in tissue culture. In contrast, the sarcoma viruses, which are usually replication-defective, cause neoplastic transformation of fibroblasts in vitro and produce solid tumors in vivo.

Moloney murine sarcoma virus (MSV) is a representative of the class of replication-defective sarcoma viruses. This virus arose by recombination of the nondefective Moloney murine leukemia virus (MuLV) and cellular sequences present within the normal mouse genome (*1–3*). These latter sequences are essential for viral transforming activity (*3–5*). According to recent convention (*6*) the cell-derived sequences of MSV are designated v-*mos* and those of the cell are designated c-*mos*. The development of molecular cloning and DNA sequencing techniques has made the detailed analysis of the virus genome structure possible. We present in this report the first complete sequence analysis of a transforming retroviral genome.

Figure 1 provides a detailed restriction map of the 5.8-kbp (kilobase pair) MSV genomic DNA cloned in bacteriophage λ (*2*) as well as our sequencing strategy. Sequence analysis was performed according to the procedures of Maxam and Gilbert (*7*). The entire MSV nucleotide sequence is presented as it occurs in the linear proviral genome in Fig. 2.

An important structural feature of the MSV genome is the occurrence of two large terminal repeats of 585 bases (LTR's) at both 5′ and 3′ ends of the proviral genome. These LTR's bear striking similarities to the terminal repeats of prokaryotic transposable elements (*8, 9*). Previous studies have provided sequence data for this region of the genome (*10–15*). The salient features are summarized below.

1) *Inverted terminal repeats.* An inverted repeat of 11 nucleotides, 5′-TGAAAGACCCC-3′ (T, thymine; G, guanine; A, adenine; C. cytosine), appeared at the termini of each LTR at positions 1 to 11, 575 to 585, 5244 to 5254, and 5818 to 5828. Prokaryotic transposable elements are also flanked by such inverted repeats and are capable of translocation to different positions on the chromosome or to another replicon in the cell (*8, 9*). Like these bacterial elements, retroviruses integrate into the host DNA in a linear orientation with defined endpoints (*10, 11, 15*).

2) *Transcription initiation and termination signals.* The LTR is composed of a track of about 440 nucleotides derived from the 3′ end of the viral RNA, directly followed to the right by a stretch of approximately 145 nucleotides derived from the 5′ end of the viral genome (*16*). This stretch of 585 bases contained sig-

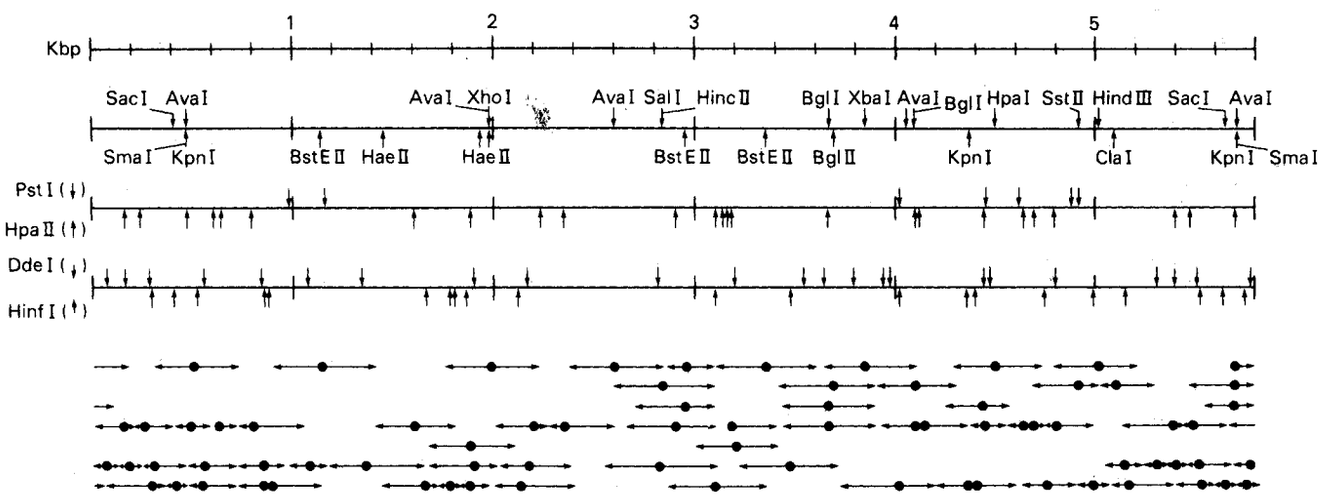
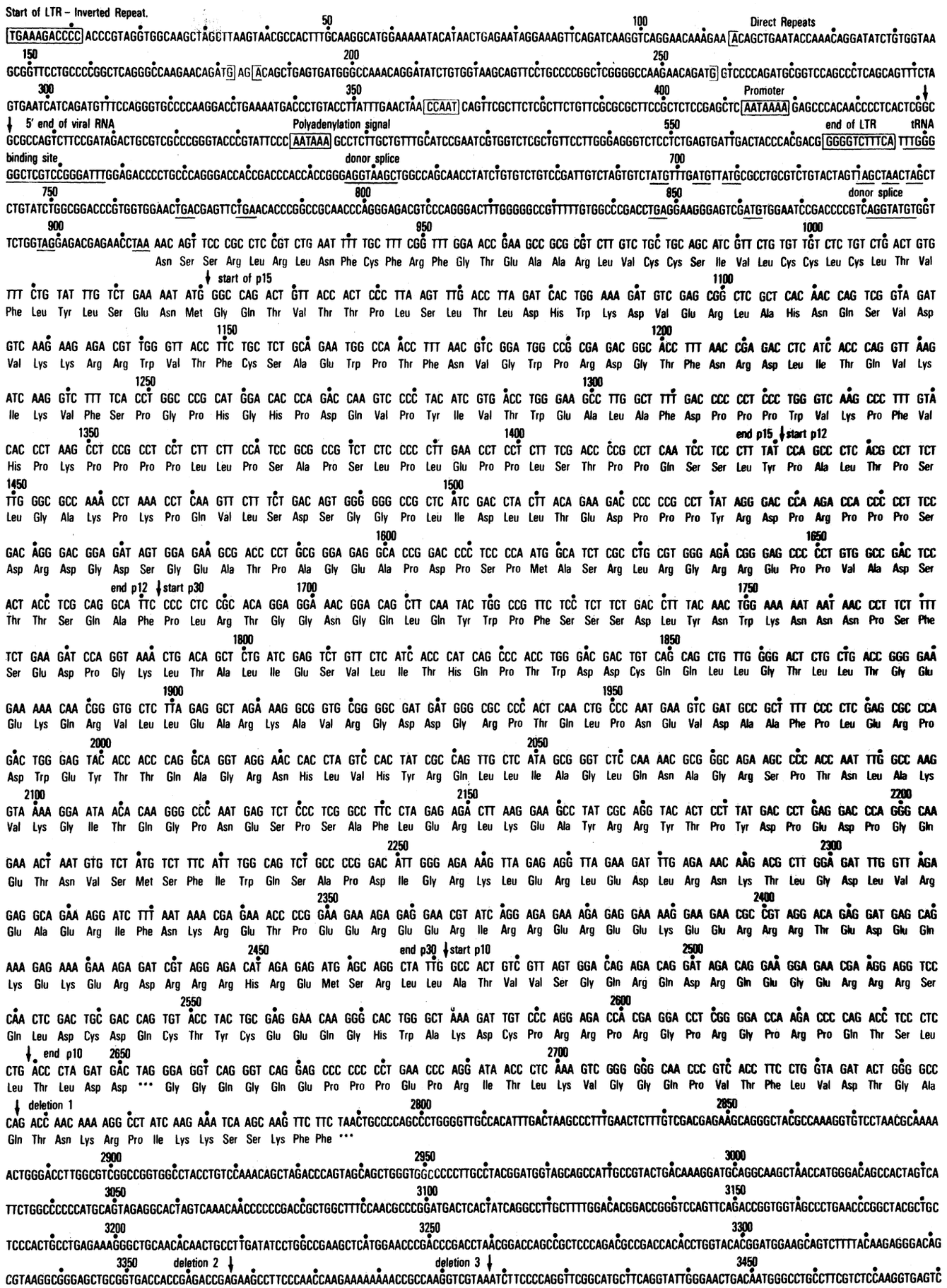



Fig. 1. Restriction enzyme map and strategy for sequencing of the MSV genome. The genome was sequenced with the use of the restriction sites indicated on the diagrammatic map. The 5′ ends were labeled with [γ-$^{32}$P]ATP and $T_4$ polynucleotide kinase (*7*). The 3′ ends were labeled with [$^{32}$P]cordycepin phosphate and terminal transferase (*48*). The labeled end of each fragment is indicated by the filled circle and the extent and direction of sequencing are indicated by the arrows.


0036-8075/81/1023-0445$01.00/0


Start of LTR – Inverted Repeat.
Direct Repeats
Promoter
5′ end of viral RNA
Polyadenylation signal
end of LTR
tRNA
binding site
donor splice
donor splice
start of p15
end p15
start p12
end p12
start p30
end p30
start p10
end p10
deletion 1
deletion 2
deletion 3

nals for RNA transcription as well as messenger RNA (mRNA) capping and polyadenylation. Thus, a promoter-like sequence, 5′-AATAAAA-3′ (*17, 18*), was found at positions 413 to 419. This signal preceded by 22 and 24 nucleotides, respectively, the two GCG triplets likely to constitute the 5′ end of the viral RNA (*11, 13, 19*). Moreover, the sequence 5′-CCAAT-3′ was detected at positions 362 to 366, which is 75 to 77 bases upstream from the 5′ cap structure. An analogous sequence occurs 77 ± 10 bp (base pairs) upstream from the 5′ end of the mRNA capping site of most eukaryotic structural genes (*20*).

A polyadenylation signal (*21*) was found in each LTR at positions 489 to 494 and 5732 to 5737. This signal preceded the dinucleotide 5′-CA-3′, a preferred polyadenylation site (*21*), by 16 bases at positions 510–511 and 5753–5754. It is not known if the polyadenlyation signal at position 489 to 494 functions in the left LTR, but oligonucleotide mapping of the 3′ end of Moloney MuLV RNA (*16*) is consistent with the termination of viral RNA synthesis occurring at the site (5753–5754) in the right LTR.

Genomic RNA contains a direct terminally repeated sequence of 50 to 60 nucleotides termed R or trs (*16*). Thus, if the CA signal at 5753–5754 corresponds to the 3′ end of MSV RNA, the sequence of 68 to 70 nucleotides between the CA and the GCG triplets representing the 5′ end of the viral RNA should constitute the R region of MSV (Fig. 2).

The positioning of LTR's at both ends of the integrated provirus has strongly suggested that the viral genomic RNA is initiated at the promoter of the left LTR and terminated at the polyadenylation signal of the right LTR. The juxtaposition of promoter and termination signals within the LTR might also result in the formation of short RNA transcripts or in transcripts initiating or terminating in flanking cellular DNA. It has been suggested that secondary structure may play a crucial role in preventing premature

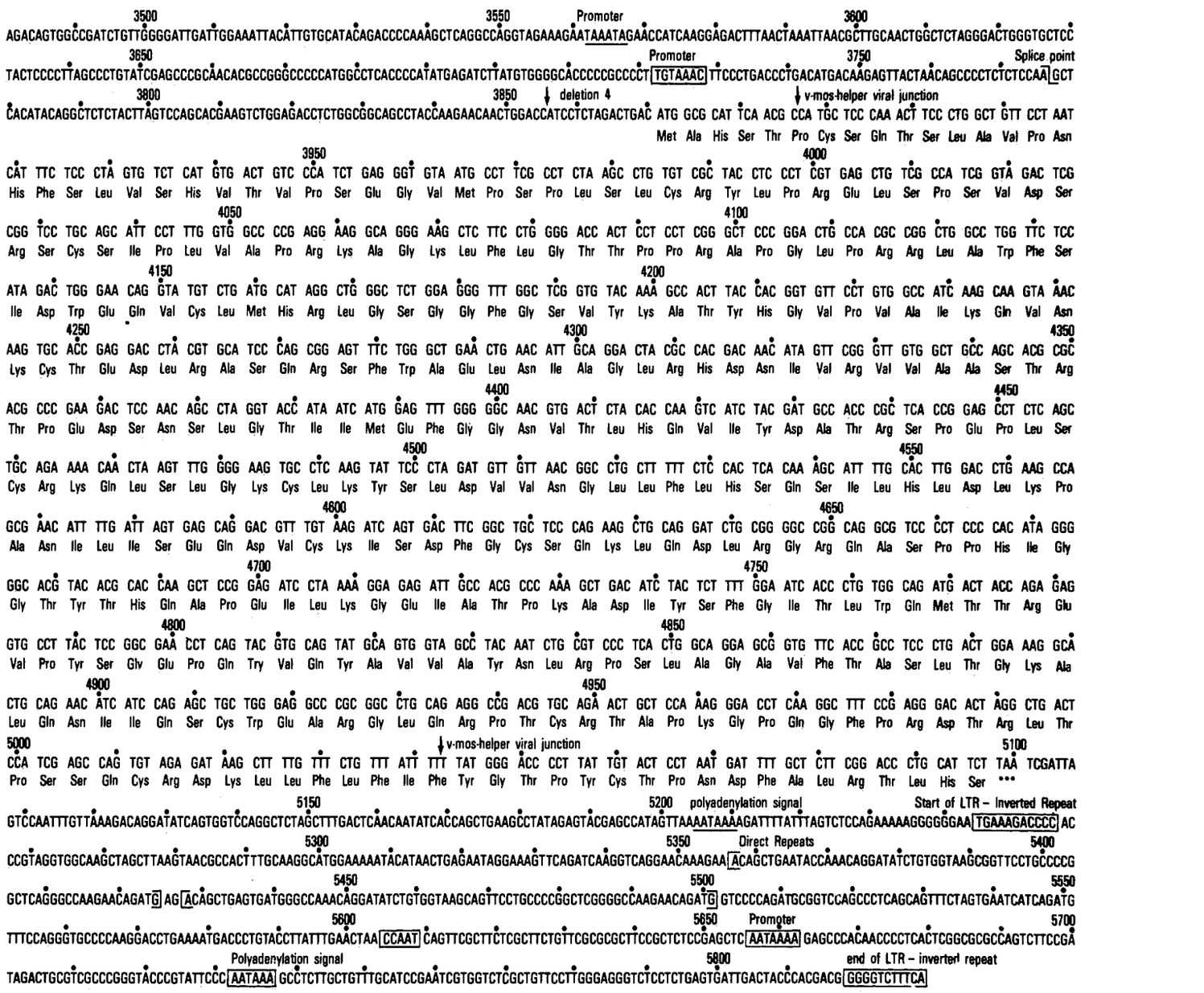



Fig. 2. Complete nucleotide sequence of the MSV genome. The upper line shows the sequence proceeding in the 5′ to 3′ direction and has the same polarity as MSV genomic RNA. The amino acid sequence deduced from the open reading frames is given in the bottom line. The major structural features of the genome are indicated.

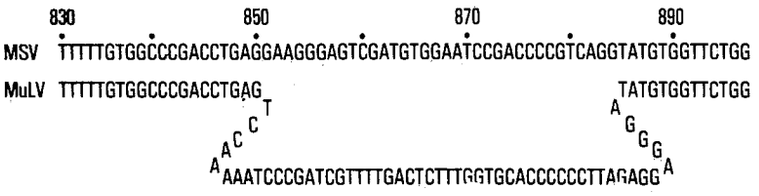



Fig. 3. DNA sequence comparison between Moloney MuLV and MSV genomes near the 5′ end of each molecule. The nucleotide numbers correspond to those numbers for the MSV genome in Fig. 2.

termination at the polyadenlyation site located 70 bases downstream from the promoter in the left LTR (*22*). There is evidence from studies with avian retroviruses that transcripts can be initiated by "downstream" promotion of cellular genes from an LTR (*23*). In fact, these studies suggest that the deregulation of certain cellular genes by this mechanism may contribute to the transforming activity of leukosis viruses (*23*).

3) *Sequence duplications*. The MSV LTR's contained a nearly perfect duplication of 69 and 74 bases, respectively, at positions 114 to 182, 185 to 258, 5357 to 5425, and 5428 to 5501. Although the function of these sequences in the MSV genome is not yet known, similar tandem repeats do occur on the late side of the SV40 DNA replication origins (*24*). These 72 bp repeated sequences are an essential element for promotion of the early transcripts of the SV40 genome (*25*), and the analogous tandem repeats present in MSV may play a similar role.

4) *Transfer RNA binding site*. Proline transfer RNA ($tRNA^{Pro}$) has been reported to be the primer for Moloney MuLV (*26*). A 21-base sequence complementary to the 3′ end of $tRNA^{Pro}$ [5′-CAAATCCCGGACGAGCCCCCA-3′ (*26*)] was localized at position 588 to 608. This sequence includes the complement to the tRNA terminal CCA. Presumably, this sequence represents the primer binding site of the viral genomic RNA. The primer binding site appears to be essential for rescue of biologically active transforming virus, since MSV recombinant DNA clones that have suffered a deletion of this region retain transforming activity but lack the ability to yield rescuable transforming virus (*27*).

Our molecularly cloned MSV genome codes for the Moloney MuLV *gag* gene product. The *gag* gene resides within the 5′ region of the viral RNA [for review see (*28, 29*)]. However, its structure and location with respect to the 5′ end of the viral genome are not known precisely. There is evidence that this region contains sequences that are spliced to the main body of subgenomic mRNA's such as the *env*-mRNA of leukemia viruses and possibly mRNA's of the transforming sarcoma viruses (*30, 31*). Seif *et al.* (*32*) have proposed a general model for RNA splicing, in which the consensus sequence 5′-AGGTAAGT-3′ is a donor for splicing to a receptor sequence located downstream in the primary RNA transcript. Two sequences with seven out of eight nucleotides in common with the model were present at positions 646 to 653 and 881 to 888.

We observed six ATG codons at positions 696 to 698, 703 to 705, 708 to 710, 861 to 863, 885 to 887, and 1038 to 1040. The first five of these were not considered likely to function as initiator codons because they were followed closely by in-phase termination codons. These findings were unusual, since in virtually all eukaryotic mRNA's that have been analyzed, the AUG codon (U, uracil) located closest to the 5′ end has been found to be the one used to initiate translation (*33*). This suggests the possibility that the genomic MSV RNA does not directly act as the functional message for *gag* gene proteins, but that the actual mRNA is generated as a result of processing.

When we compared the DNA sequence of MSV in the region between positions 830 to 890 with that of molecularly cloned Moloney MuLV DNA (Fig. 3), we observed considerable sequence divergence; such divergence has been detected in the intron sequences of genes such as β-globin gene family (*34*). This degree of divergence might also reflect a recombinational event involving MSV. In either case, the striking sequence divergence between MSV and MuLV in this region provides another line of evidence for processing of the mRNA coding for *gag* gene products.

Starting from position 915, a large open reading frame of 1736 bases and ending in a TAG codon at position 2652 to 2654 was detected. We presumed that this sequence coded for the viral *gag* gene precursor polypeptide. Figure 2 shows the predicted amino acid sequence of this polypeptide. This predicted sequence matches well with the partial amino acid sequence data available for Rauscher and Moloney MuLV p15, p12, p30, and p10 polypeptides (*35, 36*). The junctions between p15 and p12, p12 and p30, and p30 and p10 could be localized on the basis of amino acid sequence analysis (*35*). Thus, our nucleotide-sequencing analysis confirms the proposed order of internal structural proteins coded by the viral *gag* gene (*37*). Moreover, the nucleotide sequence data provide direct experimental proof for the hypothesis that the COOH-terminal tyrosine of p15 and $NH_2$-terminal proline of p12 are contiguous, as are the COOH-terminal phenylalanine of p12 and $NH_2$-terminal proline of p30, and the COOH-terminal leucine of p30 and $NH_2$-terminal alanine of p10 (*35*).

The *gag* gene is formed by processing of a larger precursor polypeptide (*38*), which is approximately 6000 daltons larger than the 65,000-dalton protein containing p15, p12, p30, and p10. Synthesis of this precursor protein must be initiated at a point around 180 nucleotides upstream from the amino terminus of p15. Since there is no AUG codon followed by an open reading frame in this region of the viral genome, the *gag* precursor could be made from a processed RNA. The COOH-terminus of p10 demonstrated four additional amino acids that are apparently not present in the processed p10 molecules (*35*). These findings suggest COOH-terminal processing of *gag* gene precursor as well.

The reverse transcriptase is thought to be synthesized via a 180,000-dalton precursor protein by a complicated readthrough mechanism (*39–41*). This readthrough protein, which contains products of the *gag* and reverse transcriptase genes (*39–41*), is present at low level both in vivo and in vitro and can be enhanced in vitro by amber (or to a lesser extent by ocher) suppressor tRNA's (*41*). The *gag* gene ends with an amber codon (TAG), which is directly followed by another open reading frame (Fig. 2). The MSV genome has undergone four deletions at positions 2746, 3366, 3402, and 3885 in the helper virus region that codes for reverse transcriptase. We mapped these deletions by sequencing the corresponding regions of the Moloney MuLV genome (Fig. 2). These deletions render the polymerase gene nonfunctional within the MSV genome.

Heteroduplex analysis has revealed that the MSV genome has undergone a sequence substitution of approximately 1.2 kilobases in the region coding for the envelope protein, where *env* gene sequences have been replaced by cellular sequences (*42*). In order to localize the exact point of this recombinational event, we sequenced the corresponding

regions of the Moloney MuLV genome. It was thus possible to localize junction points between v-*mos* and helper viral sequences (Fig. 2) (*13*). Examination of the viral RNA strand of v-*mos* (Fig. 2) revealed an open reading frame starting with an initiation codon ATG at position 3871 and terminating with the triplet TAA at position 5098. This stretch of 1227 nucleotides began 16 bases to the left of the v-*mos*-helper viral junction and ended 61 nucleotides into helper viral sequences to the right of v-*mos*.

The region of the MSV genome between positions 4890 and 4920 has been the most difficult to sequence because of the occurrence of short direct and inverted repeats in close proximity. Sequencing of either strand did not appear to give the complete sequence of this region; analysis of both strands gave a consensus sequence which incorporated three additional nucleotides, not revealed by analysis of either strand alone (Fig. 4). The consensus sequence deduced from these sequencing procedures is given in Fig. 2. The sequence of this region differs from our earlier results by three nucleotides but does not change the reading frame (*13*). This sequence also differs from that of Van Beveran *et al.* (*43*) by one nucleotide at position 4943. The sequence of Van Beveran *et al.* alters the reading frame of the MSV putative transforming gene product, such that it terminates within v-*mos* sequences to the left of the Hind III site (*43*).

The MSV transforming region as determined here has the coding capacity for a protein of 409 amino acids and a molecular size of 44,000 daltons. The functional product of the MSV transforming gene has yet to be identified. However, in vitro translation of MSV RNA has revealed a family of proteins with molecular sizes of 43, 40, 31, and 24 kilodaltons, whose syntheses were specifically inhibited when MSV genomic RNA was first annealed with v-*mos*-specific DNA fragments (*44*). Analyses by cyanogen bromide cleavage has further revealed that these proteins are related and that their respective sizes correspond reasonably well with the locations of initiator codons within the 1227-nucleotide open reading frame (*44*).

The mechanism by which the mRNA coding for the putative transforming protein is generated is not known. However, it is known that splicing acceptor sites at the 3′ end of intervening sequences contain a pyrimidine-rich nucleotide track followed by the dinucleotide 5′-AG-3′ (*32*). The MSV genome was found to contain such a possible acceptor splice point at position 3776. In addition, two promoter-like sequences were found at positions 3562 to 3567 and 3722 to 3728. Thus the functional mRNA coding for the MSV transforming protein could be independently transcribed or could be a product of splicing. Findings that the MSV transforming gene can function effectively in subgenomic MSV DNA fragments (*3–5*) have indicated that transcription can be initiated in the absence of the 5′ LTR. Thus, by whichever mechanism it is normally transcribed, the MSV transforming gene may not require a spliced message for it to function.

Toward the right end of the MSV genome, we observed a cluster of sequences rich in A + T which included eight base pairs at positions 5204 to 5211, followed by an inverted repeat of the same sequence at positions 5214 to 5222. The inverted repeats contained the polyadenylation signal 5′-AATAAA-3′. This signal at position 5205 to 5210 preceded the dinucleotide CA in position 5228 by 18 base pairs, which is a preferred site for polyadenylation. These signals might serve as terminators for the transcription of v-*mos* mRNA. This CA dinucleotide was followed 14 bases later by the right LTR of MSV genome whose features have been described above.

The complete nucleotide sequence of the MSV genome presented here, taken together with available information concerning viral transcripts and proteins, makes possible a much more detailed understanding of the structure and function of the transforming retrovirus genome. We presented the sequence of the MSV genome for the purpose of convenience as it occurs in the linear proviral genome. However, the genome that we analyzed was a covalently closed circle, whose two LTR's appeared in tandem. The sequence at the LTR junction site was determined to be GGGGTC-TTTCATTTAATGAAAGACCCC. Comparison of this sequence with that of the viral-cellular junction sequence of integrated MSV (*11*) indicates that during integration there is a loss of nucleotides TTTAA from the termini of LTR's during proviral integration. This loss of sequences in the process of integration demonstrates another important structural analogy of retroviruses to transposable elements, which also suffer a deletion of a few dinucleotides during integration (*8*). Their similarity to transposable elements is consistent with the possibility that retroviruses have evolved from small movable cellular genetic elements.

The genomes of prokaryocytes such as bacteriophage λ contain certain regions termed "hot spots," which promote re-

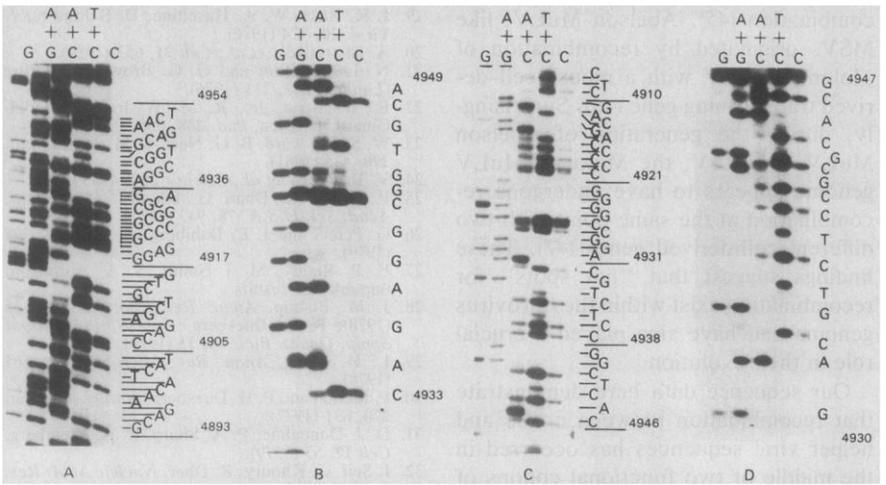



Fig. 4. DNA-sequencing gels in the region between positions 4900 and 4950. The DNA fragments were end labeled as described, subjected to base-specific cleavage reactions, and analyzed on 0.3-mm thick 10 or 20 percent polyacrylamide gels. (A) Sequence analysis of the plus strand (5′ end labeled) from Rsa I site at position 4806 to 4809. (B) Sequence analysis of plus strand (5′ end labeled) from Tha I site at position 4924 to 4927. (C) Sequence analysis of the minus strand (5′ end labeled) from Hinf I site at position 4995 to 4999. (D) Sequence analysis of the minus strand (3′ end labeled) from Sst II site at position 4923 to 4928. Restriction enzyme analysis of this region shows the presence of an Sst II site at position 4923 to 4928 and a Hae III site at position 4937 to 4940. No Hha I or Msp I sites could be detected in this region. Interpretation of the sequence data is given at the right.

```
                    4910      4920      4930      4940
                     .         .         .         .
A) (+ strand)  GCTGGGAGGCCCGCGGC--GCAGAGGCCG--GTGCAG
B) (+ strand)                    CAGAGG-CG--GTGCAG
C) (- strand)  CGACCC-CCGGG-GCCGGACGTCTCCGGCT-CACGTC
D) (- strand)                    GACGTCTCC-GC-GCACG

Consensus      GCTGGGAGGCCCGCGGCCTGCAGAGGCCGACGTGCAG
Sequence       CGACCCTCCGGGCGCCGGACGTCTCCGGCTGCACGTC
```

combination (*45*). Abelson MuLV, like MSV, originated by recombination of Moloney MuLV with a mouse cell-derived transforming gene (*46*). Surprisingly, during the generation of Abelson MuLV and MSV, the Moloney MuLV genome appears to have undergone recombination at the same point with two different cell-derived genes (*47*). These findings suggest that "hot spots" for recombination exist within the retrovirus genome and have also played a crucial role in their evolution.

Our sequence data here demonstrate that recombination between c-*mos* and helper viral sequences has occurred in the middle of two functional codons of the c-*mos* gene. Hence v-*mos* lacks regulatory signals for its transcription and translation. To render such an incomplete gene biologically active, the helper virus has provided this gene with transcriptional promoter and terminator signals as well as the initiating and terminating codons for translation. Recent findings have shown that molecularly cloned c-*mos* can be rendered biologically active as a transforming gene by the addition of the helper virus LTR (*3*). Detailed structural comparisons of v-*mos* and c-*mos*, as well as analysis of c-*mos* flanking sequences, may provide insights as to how c-*mos* might be transcriptionally activated in naturally occurring tumors.

E. PREMKUMAR REDDY
MARY JANE SMITH
STUART A. AARONSON
*Laboratory of Cellular and Molecular Biology, National Cancer Institute, Bethesda, Maryland 20205*

### References and Notes

1. A. E. Frankel and P. J. Fischinger, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 3705 (1976).
2. S. R. Tronick, K. C. Robbins, E. Canaani, S. G. Devare, S. A. Aaronson, *ibid.* **76**, 6314 (1979).
3. M. Oskarsson, W. L. McClements, D. G. Blair, J. V. Maizel, G. F. Vande Woude, *Science* **207**, 1222 (1980).
4. P. Andersson, M. P. Goldfarb, R. A. Weinberg, *Cell* **16**, 63 (1979).
5. E. Canaani, K. C. Robbins, S. A. Aaronson, *Nature (London)* **282**, 378 (1979).
6. J. M. Coffin *et al.*, *J. Virol.*, in press.
7. A. M. Maxam and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 560 (1977).
8. A. J. Bukhari, J. A. Shapiro, S. L. Adhya, Eds., *DNA Insertion Elements, Plasmids and Episomes* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1977).
9. H. Ohtsubo, H. R. Ohmori, E. Ohtsnboi, *Cold Spring Harbor Symp. Quant. Biol.* **53**, 1269 (1978).
10. K. Shimotohno, S. Mizutani, H. M. Temin, *Nature (London)* **285**, 550 (1980).
11. R. Dhar, W. McClements, L. W. Enquist, G. W. Vande Woude, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3937 (1980).
12. J. G. Sutcliffe, T. M. Shinnick, I. M. Verma, R. A. Lerner, *ibid.* p. 3302.
13. E. P. Reddy *et al.*, *ibid.*, p. 5234.
14. G. Ju and A. M. Skalka, *Cell* **22**, 379 (1980).
15. J. E. Majors and H. E. Varmus, *Nature (London)* **289**, 253 (1981).
16. J. M. Coffin, T. C. Hageman, A. M. Maxam, W. A. Haseltine, *Cell* **13**, 761 (1978).
17. D. Pribnow, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 784 (1975).
18. M. Rosenberg and D. Court, *Annu. Rev. Genet.* **13**, 319 (1979).
19. J. K. Rose, W. A. Haseltine, D. Baltimore, *J. Virol.* **20**, 324 (1976).
20. A. Efstratiadis *et al.*, *Cell* **21**, 653 (1980).
21. N. J. Proudfoot and G. G. Brownlee, *Nature (London)* **263**, 211 (1976).
22. E. W. Benz, Jr., R. M. Wydro, B. Nadal-Ginard, D. Dina, *ibid.* **288**, 665 (1980).
23. W. S. Hayward, B. G. Neel, S. M. Astrin, *ibid.* **290**, 475 (1981).
24. V. B. Reddy *et al.*, *Science* **200**, 494 (1978).
25. P. Gruss, R. Dhar, G. Khoury, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 943 (1981).
26. G. Peters and J. E. Dahlberg, *J. Virol.* **31**, 398 (1979).
27. E. P. Reddy, M. J. Smith, S. A. Aaronson, unpublished results.
28. J. M. Bishop, *Annu. Rev. Biochem.* **47**, 35 (1978); P. H. Duesberg, *Cold Spring Harbor Symp. Quant. Biol.* **44**, 13 (1979).
29. L. H. Wang, *Annu. Rev. Microbiol.* **32**, 561 (1978).
30. P. Mellon and P. H. Duesberg, *Nature (London)* **270**, 631 (1977).
31. D. J. Donoghue, P. A. Sharp, R. A. Weinberg, *Cell* **17**, 53 (1979).
32. I. Seif, G. Khoury, R. Dhar, *Nucleic Acids Res.* **6**, 3387 (1979).
33. M. Kozak, *Cell* **15**, 1109 (1978).
34. P. Leder, J. N. Hansen, D. Konkel, A. Leder, Y. Nishioka, C. Talkington, *Science* **209**, 1336 (1980).
35. S. Oroszlan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1404 (1978); S. Oroszlan and R. V. Gilden, in *Molecular Biology of RNA Tumor Viruses*, J. R. Stephenson, Ed. (Academic Press, New York, 1980), p. 299.
36. S. Oroszlan, personal communication.
37. M. Barbacid, J. R. Stephenson, S. A. Aaronson, *Nature (London)* **262**, 554 (1976).
38. R. B. Naso, L. J. Arcement, R. B. Arlinghaus, *Cell* **4**, 31 (1975).
39. J. J. Kopchick, G. A. Jamjoom, K. F. Watson, R. B. Arlinghaus, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2016 (1978).
40. E. C. Murphy, Jr., J. J. Kopchick, K. F. Watson, R. B. Arlinghaus, *Cell* **13**, 359 (1978).
41. L. Philipson, P. Andersson, V. Olshevsky, R. Weinberg, D. Baltimore, R. Gesteland, *ibid.*, p. 189.
42. S. Hu, N. Davidson, I. M. Verma, *ibid.* **10**, 469 (1977).
43. C. Van Beveren, J. A. Galleshaw, V. Jonas, A. J. M. Berns, R. F. Doolittle, D. J. Donoghue, I. M. Verma, *Nature (London)* **289**, 258 (1981).
44. K. Cremer, E. P. Reddy, S. A. Aaronson, *J. Virol.* **38**, 704 (1981).
45. K. D. McMilin, M. M. Stahl, F. W. Stahl, *Genetics* **77**, 409 (1974).
46. S. P. Goff, E. Gilboa, O. N. Witte, D. Baltimore, *Cell* **22**, 777 (1980).
47. E. P. Reddy, unpublished data.
48. C. P. D. Tu and S. N. Cohen, *Gene* **10**, 177 (1980).

24 June 1981; revised 10 September 1981

## An Efference Copy Which is Modified by Reafferent Input

Abstract. *In electric fish of the mormyrid family, an efference copy is present in the brain region that receives afferent input from ampullary electroreceptors. The efference copy is elicited by the motor command to fire the electric organ. Its effect is always opposite that of ampullary afferents responding to the electric organ discharge, and it changes to match variations in this afferent input. It probably reduces the central effects of activity in ampullary receptors evoked by the electric organ discharge.*

The motor behavior of an animal will normally elicit activity in its own receptors and sensory afferents. This self-induced sensory input was termed reafference by von Holst and Mittelstaedt (*1*). An animal must always distinguish between such reafferent input and sensory input from external sources. Behavioral experiments of von Holst and Mittelstaedt (*1*) and Sperry (*2*) suggested that the problem is solved by signals from motor centers to sensory receiving areas; these signals prepare such areas for the expected reafference. Such signals were termed "efference copies" by von Holst and Mittelstaedt and "corollary discharges" by Sperry. Effects of motor commands on sensory centers have since been seen physiologically in a variety of preparations (*3–6*).

In many sensory-motor systems, reafferent input must be nullified to prevent inappropriate reflexes or interference with detection of external sources of stimulation. In the lateral line system of fish and amphibia (*4*), the crayfish escape response (*3*), or the knollenorgan electroreceptor afferents in mormyrids (*5*), the motor command briefly inhibits the expected reafference. Such a simple inhibition does not seem functionally useful, however, when the effects of the motor act are complex and of long duration. Von Holst and Mittelstaedt (*1*) and Sperry (*2*) argued that the inhibition of the reafference during voluntary movement could not explain their results. They inferred instead that a kind of negative image of the expected reafference is conveyed to the sensory centers. Such an image could be excitatory, inhibitory, or both. When summed with the actual sensory input, the result is a nulling or reduction of the effect of the reafference. This report describes an efference copy of the latter type in electric fish.

There are three distinct types of electroreceptors in mormyrids: mormyromasts, knollenorgans, and ampullary receptors (*7*). All three types respond, with different time courses, to the electric organ discharge (EOD). However, only the responses of mormyromasts seem to be involved in measuring object-induced distortions in the electric field created by the EOD, that is, in active electrolocation (*7–9*). Knollenorgans probably assist in detecting the EOD's of other fish, that is, in communication. Ampullary receptors in mormyrids, like similar receptors in catfish or sharks, measure the low-frequency external electric fields generated by other aquatic animals (*10*). Afferents from the three types of electrore-


0036-8075/81/1023-0450$01.00/0 