JOURNAL OF VIROLOGY, Feb. 1982, p. 527–534
0022-538X/82/020527-08$02.00/0

Vol. 41, No. 2

The Leader Sequence of the Subgenomic mRNA's of Rous Sarcoma Virus Is Approximately 390 Nucleotides

PERRY B. HACKETT,[1]* RONALD SWANSTROM,[2] HAROLD E. VARMUS,[2] AND J. MICHAEL BISHOP[2]

Department of Genetics and Cell Biology, University of Minnesota, St. Paul, Minnesota 55108,[1] and Department of Microbiology and Immunology, University of California, San Francisco, California 94143[2]

Received 3 August 1981/Accepted 28 September 1981

The subgenomic mRNA's of Rous sarcoma virus share a common 5′ leader sequence spliced from genomic RNA. We have examined the 5′ terminal sequences of four Rous sarcoma virus RNAs: virion RNA and three species of intracellular mRNA which direct the synthesis of the RSV gene products. The lengths of the leaders on the RNAs were determined by the extent that they could protect cloned Rous sarcoma virus DNA fragments from S1 nuclease digestion after RNA-DNA hybridization. We found that the subgenomic mRNA's that direct the synthesis of the *env* and *src* gene products have uninterrupted spliced leader sequences of approximately 390 nucleotides, whereas virion RNA and full-length intracellular viral RNA have 5′ termini homologous to the cloned viral DNA probe over at least the first 735 bases. In the accompanying manuscript we have determined the nucleotide sequence of the 5′ end of the Rous sarcoma virus genome, including the candidate splice donor site identified here (Swanstrom et al., J. Virol. 41:535–541, 1982).

The four genes of Rous sarcoma virus (RSV) are expressed via at least four distinct mRNA species. The largest is mRNA^{gag}, approximately 9.5 kilobases (15, 46), which encodes the entire viral genome and indeed may be identical to both the viral genome and the primary transcript of the provirus (13a). mRNA^{gag} is translated in vivo and in vitro to produce a polyprotein precursor, Pr76^{gag}, which contains the viral core proteins encoded in the first (or leftmost) gene, *gag* (27, 28, 43). An as yet unidentified RNA species encodes the *pol* precursor, Pr180$^{gag/pol}$, which is a fusion protein containing the *gag* proteins and the *pol* gene product, RNA-dependent DNA polymerase (24, 26, 31, 45). Nucleotide sequence analysis indicates that mRNA$^{gag/pol}$ probably contains a splice in the RNA at the *gag/pol* junction since these coding regions are in different reading frames in genomic RNA (D. Schwartz, personal communication). Although the third and fourth genes, *env* and *src*, are encoded in mRNA^{gag}, they are apparently not expressed by this mRNA. Instead, two other subgenomic mRNA's are produced in RSV infected cells: mRNA^{env} and mRNA^{src}. mRNA^{env}, approximately 5.1 kilobases, encodes both *env* and *src*, but is translated in vivo and in vitro only into a protein percursor which is subsequently processed into the envelope glycoprotein (15, 27, 31, 36, 46). mRNA^{src}, approximately 3.1 kilobases, encodes only the 3′ proximal gene, *src* (15, 46), whose product (pp60^{src}) is responsible for neoplastic transformation of infected cells (5, 30).

The mRNA's of all three size classes are capped (7, 23) and polyadenylated (15, 46) and contain a common leader sequence at their 5′ ends; this sequence apparently is spliced from a primary transcript of the RSV genome to the subgenomic mRNA's *env* and *src* (7, 19, 23, 37, 46). Previous studies of the mRNA's showed that the leader was composed of at least 104 nucleotides (7). The 5′ ends of the RSV mRNA's contain sites for ribosome binding, initiation of protein synthesis, and initiation of DNA synthesis by reverse transcriptase in addition to a splice donor site (3). A more detailed characterization of this region of RSV RNA would help to provide an insight into how these control sites are organized relative to one another.

The purpose of the studies described in this report was to determine the total lengths of the spliced leader sequences on the two subgenomic RSV mRNA's. Our results indicate that the 5′ leader sequences for mRNA^{env} and mRNA^{src} are identical and uninterrupted units of about 390 bases that are spliced from the primary RSV transcript to the subgenomic mRNA's.

MATERIALS AND METHODS

Oligodeoxythymidylic acid-cellulose, T3 grade, was purchased from Collaborative Research, Inc.; S1 nuclease, type III from *Aspergillus oryzae*, was from

Sigma Chemical Co.; all restriction endonucleases were from New England Biolabs. avian myeloblastosis virus polymerase was obtained from Life Sciences Inc.

Cells and viruses. Primary chicken embryo fibroblasts infected with RSV were propagated as described previously (4). Chicken cells were infected with a strain of RSV, Schmidt-Ruppin A (SR-A), derived from quail cells that had been transfected by a clone of viral DNA (SRA-2) (10). The SRA-2 clone of viral DNA was isolated by insertion into the procaryotic vector λgtWES λB (10).

Isolation of virion RNA. Every 3 to 4 h, 35 ml of standard medium was collected from each of several roller bottles that contained SR-A-infected chicken embryo fibroblasts. The media were pooled and clarified by centrifugation at 5,000 × *g* for 10 min at 4°C. Virus was recovered from 200 ml of clarified medium by sedimentation at 25,000 rpm for 60 min at 4°C in a Beckman SW27 rotor and suspended in 2 ml of 100 mM NaCl–10 mM Tris (pH 7.5)–1 mM EDTA. The RSV virion RNA was isolated by adding 400 μg of proteinase K (self-digested for 20 min at 37°C) and sodium dodecyl sulfate to 0.2%, incubating for 20 min at 37°C, and deproteinizing twice with phenol-chloroform-isoamyl alcohol (1:1:0.01).

Isolation and fractionation of cellular RNA by rate-zonal centrifugation. Polyadenylated RNA from the total cellular RNA of infected cells was isolated as previously described (42a, 45). After ethanol precipitation, the mRNA's were resuspended in 2 mM EDTA (pH 7.0)–0.5% sodium dodecyl sulfate, heated to 60°C for 2 min, and fractionated by sedimentation through 36-ml gradients of 15 to 30% sucrose in 5 mM EDTA (pH 6.8)–0.5% sodium dodecyl sulfate at 20,000 rpm at 18°C in a Beckman SW27 rotor for 17 h. Pools of RNA sedimenting between 19 and 22S, 25 and 30S, and 34 and 40S, as determined from 18 and 28S rRNA markers, were collected from the gradients, ethanol precipitated, individually suspended in 5 mM EDTA (pH 7.0)–0.5% sodium dodecyl sulfate, and sedimented through a second set of sucrose gradients as described above. Again, size fractions of 19 to 22S, 25 to 30S, and 34 to 40S were collected. The individual pools of RNA were passed over oligodeoxythymidylic acid-cellulose to remove the 5′ ends of degradation products from 38S RSV RNA. The size-fractionated pools of RNA were analyzed by agarose gel electrophoresis under denaturing conditions, followed by analysis of RSV-specific sequences with radioactive viral DNA as a hybridization probe (20). In this analysis there was no significant cross-contamination of either larger or smaller RNA species in the size-fractionated RNA species. The RNA pools are henceforth designated by their predominant RSV mRNA species: $mRNA^{src}$, $mRNA^{env}$, and $mRNA^{gag}$.

Preparation of specific fragments of DNA labeled at their 3′ ends with ^{32}P. A fragment of the SRA-2 clone of RSV DNA (10) containing about 1,700 base pairs of DNA, including sequences homologous to the 5′ end of the virion RNA, was subcloned into a plasmid designated pPvu DG (41). The RSV-specific DNA insert was isolated and further cleaved with restriction endonuclease *Hinf* or *Hpa*II, leaving 5′ overhanging ends (Fig. 1). The 5′ overhanging ends were "filled in" in a repair reaction by using avian myeloblastosis virus polymerase in a 30-μl reaction mixture, containing 50 μCi of either [α-^{32}P]dATP or [α-^{32}P]dGTP, 50 μM each of the remaining deoxynucleoside triphosphates, 60 mM NaCl, 50 mM Tris (pH 8.1), 7 mM $MgCl_2$, 10 mM dithiothreitol, which was incubated at 37°C for 1 h. After removal of unincorporated ^{32}P-deoxynucleoside triphosphate by chromatography over Sephadex G-50, the DNA fragments were separated by electrophoresis in a 5% acrylamide gel, and the location of each fragment was determined by autoradiography. The appropriate DNA bands were eluted from the gel by the procedure of Maxam and Gilbert (21) and precipitated with 2 volumes of 95% ethanol. DNA fragments of simian virus 40 cleaved with *Hin*dIII, pBR322 cleaved with *Hinf*, or various restriction fragments of viral DNA were labeled at their ends as described above and used as reference markers in our experiments.

Hybridization of RSV RNA and RSV DNA. We used the protocol of Casey and Davidson (6) for hybridization of the RSV mRNA fractions and virion RNA to the DNA probe. Each reaction mixture contained approximately 10,000 cpm of an isolated RSV DNA fragment (the concentration of DNA was not determined), 5 μg of yeast RNA, and either 2 μg of cellular polyadenylated RNA or 0.02 μg of virion RNA; the nucleic acids were mixed, precipitated in ethanol, and suspended by rapid pipetting in 16 μl of a buffer containing 80% formamide, 400 mM NaCl, 40 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid] buffer, and 1 mM EDTA, (pH 6.4). The hybridization mixtures were immersed in boiling water for 10 min to denature the DNA duplexes and then incubated at 53°C for 5 h. The optimal amounts of RNA and DNA required in the reactions were empirically determined. The DNA probe was estimated to be in 3- to 10-fold molar excess over RSV mRNA, as determined by the amount of DNA protected by the RSV mRNA after S1 nuclease hydrolysis; the maximal level of hybridization was reached by 5 h (data not shown).

S1 nuclease treatment of RSV RNA-DNA hybrids. After hybridization, a 4-μl sample of the reaction mixture was diluted into 120 μl of ice-cold S1 nuclease buffer (280 mM NaCl, 30 mM sodium acetate [pH 4.5], 3 mM $ZnSO_4$, 20 μg of denatured salmon sperm DNA per ml) (2) that lacked S1 enzyme; 10 μg of yeast RNA, 12 μl of 300 mM EDTA (pH 7.0), and 300 μl of 95% ethanol were added, and the nucleic acids were precipitated. To the remaining 12 μl of the hybridization mixture, 120 μl of S1 nuclease buffer containing 100 U of S1 nuclease was added, followed by incubation at 37°C for 10 min; the hydrolysis reaction was terminated by the addition of EDTA, yeast RNA, and 95% ethanol as described above.

Analysis of DNA fragments by electrophoresis. S1 nuclease-treated and untreated nucleic acids from the hybridization reactions with the *Hinf*-A probe were analyzed before and after denaturation with glyoxal on 4% polyacrylamide gels in 10 mM phosphate buffer, pH 7.4 (22). Experiments with the *Hpa*II-C probe were examined on 10% acrylamide gels in 10 mM phosphate buffer (pH 7.4). Glyoxal and formamide were deionized by using Bio-Rad mixed bed resin AG 501—X8(D) before each experiment. Samples were electrophoresed on 3-mm-thick gels at 7 V/cm until a bromophenol blue dye marker had migrated 15 cm. Some of the samples were analyzed by electrophoresis in a 40-cm-long, 10% acrylamide gel containing 8 M urea (34). We

also used ^{32}P-labeled, dideoxynucleotide-terminated fragments from M13mp2 for size determinations on the 40-cm-long acrylamide gels (16). All gels were dried onto Whatman no. 1 paper for autoradiographic analysis as described previously (40).

RESULTS

Experimental strategy. Our strategy for determining the lengths of the leader sequences of the RSV mRNA's was to hybridize end-labeled restriction fragments of cloned RSV DNA to either whole cell RNA from productively infected cells or to size fractions of infected cell RNA enriched for each of the viral mRNA species. After annealing, the RNA-DNA hybrids were treated with S1 nuclease, and the size of the probe was determined. This methodology, first developed by Berk and Sharp (2) and later modified by Weaver and Weissmann (44), can reveal the length of a common leader and whether the leader is contiguous or broken up into distinct units.

The virus used to establish the productive infection was obtained after transfection of cloned viral DNA into permissive cells (10). This insured complete homology between the DNA probe and viral RNA. Between 20 and 50% of the viral RNA in these cells was the size of RNA of transformation-defective (*td*) deletion mutants of RSV, and we estimate that this percentage of the mRNA^{src} fraction was td mRNA^{env}. In experiments with size-fractionated RNA the smaller size fractions contained fragments of genome RNA, but since these fragments were polarized to their 3′ ends by selection on oligodeoxythymidylic acid-cellulose, these contaminants could not react with the 5′ sequence-specific probes. The mRNA^{gag} fraction contained intracellular genome RNA, mRNA^{gag}, and mRNA$^{gag/pol}$ We do not fully understand the relationship of these different RNA species to one another, although we assume that genome RNA serves as the mRNA for *gag* (27, 43) and that mRNA$^{gag/pol}$ is a minor species (26, 31, 45).

An abbreviated physical map of the sequences at the 5′ end of the viral genome is presented in Fig. 1. Below the map are shown the positions of the DNA probes used in S1 mapping. The *Hinf*-A probe was a 680-base-pair long *Hinf* fragment that had been trimmed to 470 base pairs with *Bam*HI. In some experiments the entire *Hinf* fragment was used. The probe covered sequences starting 55 nucleotides from the position of the cap to within the coding region for pl9 (41). A *Hpa*II fragment (termed *Hpa*II-C), 245 base pairs long and containing only 105 base pairs of the 5′ noncoding region, was also used to permit greater resolution of the splice point.

S1 mapping with the *Hinf* A probe. The *Hinf*-A probe was annealed to whole cell polyadenylated RNA from SR-A-infected chicken cells. After annealing, the hybrids were treated with S1 and then denatured with glyoxal, and the size of the probe was determined by migration in a polyacrylamide gel relative to known size standards. The results of such an analysis are shown in Fig. 2, lane f. About one-half of the probe was protected from S1 digestion, migrating at 470 bases. The other half was trimmed to about 340 bases. As controls the probe was analyzed after annealing but without S1 treatment (lanes c and e) or after annealing to yeast RNA followed by S1 treatment (lane d); in the latter case the probe was completely digested. There was no evidence for any spliced RNA species in which the splice point was closer to the 5′ end of the RNA. Our analysis was designed to detect splice sites as close as 100 nucleotides to the 5′ end of the RNA. From this experiment it appears that the spliced subgenomic RNAs share a common leader sequence of between 380 to 400 nucleotides. This conclusion was confirmed by examining the size of the leader of the fractionated RNAs.

When the *Hinf*-A probe was annealed to the mRNA^{gag} fraction, most of the probe was completely protected from digestion (lane h). A small but detectable amount of the probe was shortened to about 340 bases. The fraction of the probe that was shortened can probably be accounted for by contamination with the subgenomic mRNAs in the mRNA^{gag} fraction. We infer that the vast majority of intracellular genome-size RNA is contiguous with the probe over its entire length. S1 mapping of virion RNA with the *Hinf*-A probe showed no evidence of a spliced species (data not shown).

Both the mRNA^{env} fraction (lane j) and the mRNA^{src} fraction (lane l) showed a unique splice point with the probe being trimmed to about 340 bases. In each case a small percentage of the probe was fully protected. We assume this was due to a trace amount of genome-size RNA contaminating the smaller fractions of RNA. The subgenomic RNAs appear to share a common splice site 380 to 400 nucleotides from the 5′ end of genome RNA.

We also examined the hybrids between the DNA probe and viral RNA after S1 digestion, but without denaturation before electrophoresis. In this experiment the entire *Hinf* fragment (680 nucleotides) was used as probe. The results (not shown) were essentially the same as in Fig. 2, indicating that the 5′ ends of virion RNA and mRNA^{gag} are coextensive with the cloned viral DNA over at least 735 nucleotides and that the leader sequence of the subgenomic mRNA's is 380 to 400 nucleotides long and does not contain any other sequences through position 735. Based on this initial mapping, we designed a

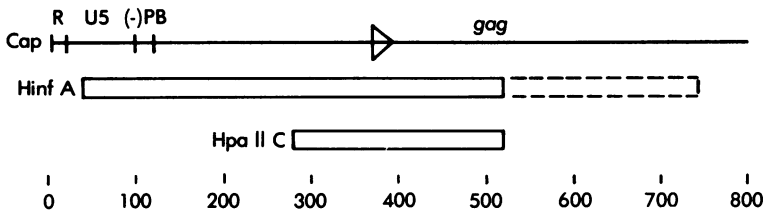

FIG. 1. DNA probes used in S1 mapping. The solid line represents viral RNA. Regions labeled R, U5, and (−)PB represent functional units in the 5′ noncoding domain of the viral genome. R denotes a 21-nucleotide-long sequence present at each end of viral RNA; (−)PB is the binding site for $tRNA^{Trp}$, the primer for the first strand of viral DNA; U5 is the sequence at the 5′ end of viral RNA that is duplicated during the synthesis of viral DNA [U5 includes the entire sequence between R and (−)PB] (H. E. Varmus and R. Swanstrom, *in* R. Weiss, N. Teich, H. E. Varmus, and J. Coffin, (ed.), *Molecular Biology of Tumor Viruses, Part III, RNA tumor viruses*, in press). The arrowhead represents the start of the *gag* coding region, 372 nucleotides from the 5′ end of viral RNA (41). The numbering at the bottom is from the cap site in viral RNA, going in the 3′ direction. DNA probes used in S1 mapping are shown as open boxes. The *Hinf*-A probe extends from a *Hinf* site 55 bases from the RNA capping site to a *Bam*HI site at position 525. In some cases the *Hinf* fragment was used without *Bam*HI cleavage; that probe extended to position 735 (dashed line). The *Hpa*II-C probe extends from *Hpa*II sites at 270 and 515.

shorter probe to map the splice donor site with greater precision.

S1 mapping with the *Hpa*II-C probe. The *Hpa*II-C probe spans sequences starting 105 nucleotides upstream from the start of *gag* and includes about 140 nucleotides of the *gag*-coding region (Fig. 1). This probe was annealed to whole cell RNA followed by digestion with S1 nuclease. After digestion the hybrids were denatured in the presence of glyoxal, and the size of the probe was determined by gel electrophoresis (Fig. 3). As was the case with the *Hinf*-A probe, about 50% of the *Hpa*II-C probe was completely protected from S1 (lane a, 245 bases). The remainder of the probe was trimmed to between 120 and 125 nucleotides in length (lane a). When the probe was annealed, but not treated with S1, it migrated as a unique band of about 245 bases (lanes b and d). When the probe was annealed to yeast RNA and then treated with S1, it was completely digested (lane c).

The *Hpa*II-C probe was also annealed to the size-fractionated intracellular RNAs. After S1 digestion the protected hybrids were denatured, and the probe was sized by electrophoresis in a 40-cm-long, 10% acrylamide gel containing 8 M urea (Fig. 4). M13mp2 bacteriophage DNA markers prepared for nucleotide sequencing were separated in parallel lanes (a through d) to allow us to accurately estimate the sizes of the S1 cleavage fragments from hybridizations of *Hpa*II-C to the $mRNA^{src}$, $mRNA^{env}$, and $mRNA^{gag}$ fractions (lanes e through g). For reference, viral DNA fragments of 98, 110, and 145 bases were electrophoresed in a parallel lane (h). In some cases the marker DNAs and the *Hpa*II-C S1 fragments were resolved into doublets on these gels. The doublet bands in the marker lanes are a consequence of unequal migration rates of the complementary strands. The presence of doublets in some of the S1-treated samples was probably the consequence of S1 nuclease producing frayed ends of the hybrids; such multiple bands have been previously observed in other studies (11). The $mRNA^{src}$ fraction and the $mRNA^{env}$ fraction each protected between 123 and 125 nucleotides of the probe (lanes e and f). The $mRNA^{gag}$ fraction protected the entire length of the probe (lane g). Since the probe starts 105 nucleotides upstream from the beginning of *gag*, the splice donor site must lie about 20 nucleotides within the coding region of *gag*, 390 nucleotides from the 5′ end of the RNA. The resolution of the gel system, the variety of markers, and the short length of the probe argue that the length assignment is fairly precise, probably with less than a 5-base error. An examination of the nucleotide sequence in this region supports the assignment of the splice donor site to this position (41).

DISCUSSION

The results of our examination of the RSV mRNA's for *env* and *src* indicate that these mRNA's contain leader sequences of ca. 390 bases of uninterrupted homology with the 5′ terminus of RSV virion RNA. In contrast, virion RNA and full-length RSV $mRNA^{gag}$ have 5′ termini homologous to the cloned viral DNA probe for at least the first 735 bases. These results are based on the degree of protection against S1 nuclease digestion which the various RSV RNAs bestowed on single stranded RSV DNA. Owing to the inexactness in the measurement of the fragment sizes by electrophoresis and in the cleavage around the single strand-double strand junction (11, 12), our resolution

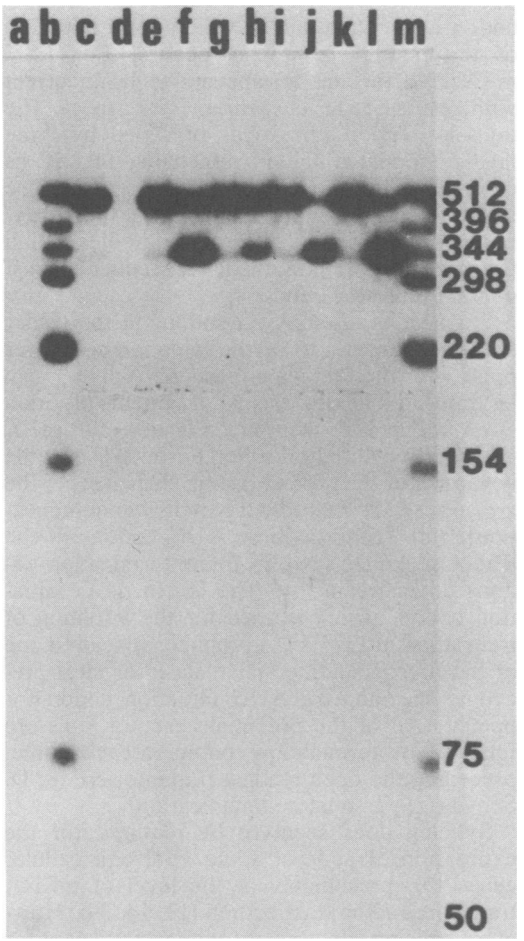

FIG. 2. S1 mapping with the *Hinf*-A probe. Viral RNAs were annealed to end-labeled *Hinf*-A probe. The resulting RNA-DNA hybrids were treated with S1 nuclease. After S1 digestion the hybrids were denatured in the presence of glyoxal, and the probe was sized by electrophoresis in a 4% polyacrylamide gel. Lanes c, e, g, i, and k contain samples not treated with S1 nuclease. Lanes d, f, h, j, and l contain samples that had been treated with S1. RNAs used in the annealings: c and d, yeast RNA: e and f, whole cell polyadenylated RNA; g and h, mRNA^{gag}; i and j, mRNA^{env}; k and l, mRNA^{src}. Lanes b and m are end-labeled *Hinf* fragments from pBR322. Lane a is a 50-base fragment of the *Hinf*-A probe generated by cleavage with *Bst*EII. Marker DNAs were denatured in the presence of glyoxal before electrophoresis.

may not be precise, but is probably accurate to within 5 nucleotides. The nucleotide sequence of this region of the RSV genome reveals a consensus splice donor site at position 389 (41).

If mRNA^{gag} contains a splice further downstream, or if mRNA's *env* and *src* contain further leader elements, they must exist beyond the 735 nucleotide limit of these studies. We cannot rigorously exclude the possibility that further sequences are spliced to mRNA^{src} and mRNA^{env}. Also, mRNA^{gag} could conceivably be spliced within the first 735 nucleotides, but represent a minor species distinct from genome RNA. If there is a distinct RNA species serving as mRNA^{gag} it probably is not spliced at position 390 since this region is known to code for part of Pr76^{gag} (25, 41).

The observation of leader sequences of 390 nucleotides on mRNA^{env} and mRNA^{src} raises the question of their purpose(s) in a genome which is remarkably compact relative to most eucaryotic genes. It appears that the 390-nucleotide leader is not a determinant in the affiliation of mRNA's with membrane-bound or unbound polysomes since mRNA^{env} is translated on membrane-bound polysomes, whereas mRNA-^{gag} and mRNA^{src} are translated on free polysomes (20, 32). Two loci expected to reside near the 5′ termini of viral mRNA's are ribosome binding sites and protein initiation codons

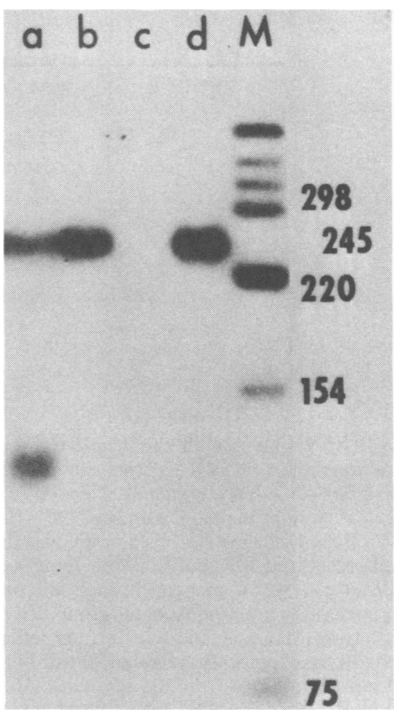

FIG. 3. S1 mapping with the *Hpa*II-C probe and whole cell polyadenylated RNA. RNAs and end-labeled *Hpa*II-C DNA probe were annealed, and the DNA from the resulting hybrids was analyzed by electrophoresis in a 10% polyacrylamide gel, as described in Fig. 2, before (lanes b and d) and after (lanes a and c) treatment with S1. RNAs used in the annealings: a and b, whole cell polyadenylated RNA; c and d, yeast RNA. Lane M contains end-labeled *Hinf* fragments of pBR322 as size markers.

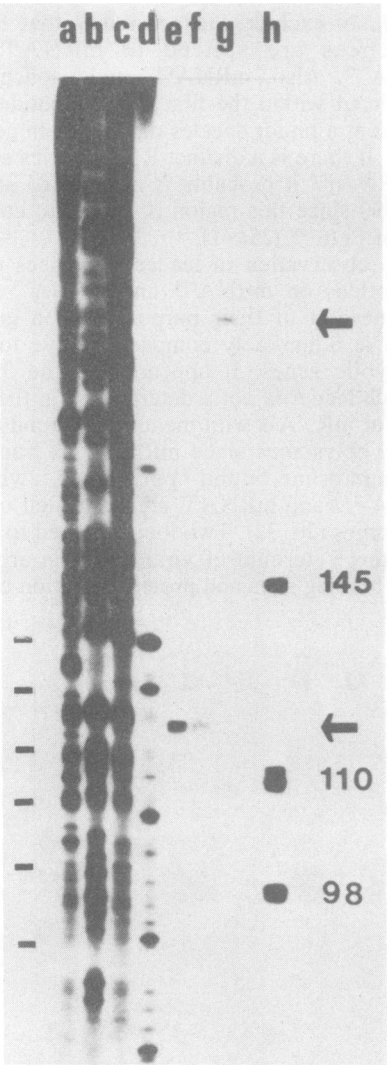

FIG. 4. S1 mapping with the *Hpa*II-C probe and size fractions of viral RNA. Analysis of ^{32}P end-labeled *Hpa*II-C DNA fragments cleaved with S1 nuclease, after hybridization with mRNA^{src} (lane e), mRNA^{env} (lane f), or mRNA^{gag} (lane g), was done by electrophoresis in a 10% polyacrylamide gel containing 8 M urea (34). Lanes a through d contain a dideoxynucleotide monophosphate-terminated marker sequence from M13mp2; the lines on the left margin represent 10 base intervals, between 90 and 140 nucleotides long, determined from the sequence (16). Lane h contains the RSV *Hpa*II-C probe fragment cleaved with *Sau*3A and end-labeled. The sizes of the *Sau*3A markers in nucleotides are shown to the right. The arrows indicate the positions of the labeled *Hpa*II-C fragments after S1 nuclease treatment.

(3). Sherman and Stewart (35) and Kozak (18) have proposed that 40S ribosomal subunits bind to the 5′ cap sites on eucaryotic mRNA's and then drift down the mRNA until the first AUG codon is encountered, at which point initiation of polypeptide synthesis occurs. For RSV mRNA^{gag}, this model appears to be incorrect with respect to at least the second aspect. The initiation codon for *gag* is preceded by three AUG codons in the 5′ noncoding region, as determined in the accompanying paper (41). The situation is even more complex for the subgenomic mRNA's; since the splice site lies within the *gag* coding region, the leader sequence must contain four AUG codons.

None of the four AUG codons in the leader sequence appears to be the initiation codon for pp60^{src} for the following reasons. (i) pp60^{src} can be translated in vitro from 3′ fragments of virion RNA which lack the spliced leader (30). (ii) A presumptive initiation codon for pp60^{src} is at the beginning of the protein coding sequences of the *src* gene (8; D. Schwartz, personal communication). (iii) The first three AUG codons in the leader sequence have in frame termination codons downstream. (iv) The fourth AUG initiation codon, which is used for the initiation of translation of Pr76^{gag}, is probably spliced to one of the four candidate splice acceptor sites preceding the putative AUG initiation codon for pp60^{src}. All of the potential acceptor sites are followed by termination codons in each frame, preceding the open reading frame for *src* (8; D. Schwartz, personal communication).

Splicing does seem to be required for the expression of at least some viral and cellular genes (9), presumably at the level of mRNA transport and/or stabilization (13, 14, 39). However, splicing is not an absolute requirement for stabilization of all mRNA's since neither histone mRNA's (17) nor the message for adenovirus protein IX (1) is processed in this fashion. Polyadenylated messages from yeast and *Neurospora* exist which are also unspliced transcripts of their respective genes (33, 38, 42). Our studies indicate that there is probably no splice preceding the initiation codon for Pr76^{gag}, but we cannot say whether splicing occurs within or following the *gag* gene in mRNA^{gag}.

ACKNOWLEDGMENTS

We thank Bill DeLorbe and Richard Parker for making the cloned RSV DNAs available, Paul Luciw for virus, Jeff Vieira for running the DNA sequencing gel (Fig. 4), and Janet McCallister and Bertha Cook for help in preparing the manuscript. We also thank Dennis Schwartz for communication of results before publication.

P.B.H. was supported by an Institutional Research Service Award from the National Institutes of Health and a grant from the National Science Foundation. R.S. was supported by Public Health Service training grant 1T32 CA 09043 from the National Institutes of Health. The research was supported by Public Health Service grants CA 12705 and CA 19287 from the National Institutes and the American Cancer Society grant MV48G to J.M.B. and H.E.V.

LITERATURE CITED

1. **Alestrom, P., G. Akusjarvi, M. Perricaudet, M. B. Mathews, D. F. Klessig, and U. Pettersson.** 1980. The gene for polypeptide *IX* of adenovirus type 2 and its unspliced messenger RNA. Cell **19**:671–681.
2. **Berk, A. J., and P. A. Sharp.** 1978. Spliced early mRNAs of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **75**:1274–1278.
3. **Bishop, J. M.** 1978. Retroviruses. Annu. Rev. Biochem. **47**:35–88.
4. **Bishop, J. M., W. E. Levinson, N. Quintrell, D. Sullivan, L. Fanshier, and J. Jackson.** 1970. The low molecular weight RNAs of Rous sarcoma virus. I. The 4S RNA. Virology **42**:182–195.
5. **Brugge, J., E. Erikson, M. S. Collett, and R. L. Erikson.** 1977. Peptide analysis of the transformation-specific antigen from avian sarcoma virus-transformed cells. J. Virol. **26**:773–782.
6. **Casey, J., and N. Davidson.** 1977. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. Nucleic Acids Res. **4**:1539–1552.
7. **Cordell, B., S. R. Weiss, H. E. Varmus, and J. M. Bishop.** 1978. At least 104 nucleotides are transposed from the 5′ terminus of the avian sarcoma virus genome to the 5′ termini of smaller viral mRNAs. Cell **15**:79–91.
8. **Czernilofsky, A. P., A. D. Levinson, H. E. Varmus, J. M. Bishop, E. Tisher, and H. M. Goodman.** 1980. Nucleotide sequence of an avian sarcoma virus oncogene (*src*) and proposed amino acid sequence for gene product. Nature (London) **287**:198–203.
9. **Darnell, J. E.** 1978. Implications of RNA-RNA splicing in evolution of eukaryotic cells. Science **202**:1257–1260.
10. **DeLorbe, W. J., P. A. Luciw, H. M. Goodman, H. E. Varmus, and J. M. Bishop.** 1980. Molecular cloning and characterization of avian sarcoma virus circular DNA molecules. J. Virol. **36**:50–61.
11. **Green, M. R., and R. G. Roeder.** 1980. Definition of a novel promoter for the major adenovirus-associated virus mRNA. Cell **22**:231–242.
12. **Grosschedl, R., and M. L. Birnstiel.** 1980. Identification of regulatory sequences in the prelude sequences of an H2A histone gene by the study of specific deletion mutants in vivo. Proc. Natl. Acad. Sci. U.S.A. **77**:1432–1436.
13. **Gruss, P., C.-J. Lai, R. Dhar, and G. Khoury.** 1979. Splicing as a requirement for biogenesis of functional 16S mRNA of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **76**:4317–4321.

13a. **Hackett, P. B., H. E. Varmus, and J. M. Bishop.** 1981. The genesis of Rous sarcoma virus messenger RNAs. Virology **112**:714–728.

14. **Hamer, D. H., and P. Leder.** 1979. Splicing and the formation of stable RNA. Cell **18**:1299–1302.
15. **Hayward, W. S.** 1977. Size and genetic content of viral RNAs in avian oncovirus-infected cells. J. Virol. **24**:47–63.
16. **Heidecker, G., J. Messing, and B. Cronenborn.** 1980. A versatile primer for DNA sequencing in the M13mp2 cloning system. Gene **10**:68–73.
17. **Kedes, L. H.** 1979. Histone genes and histone messengers. Annu. Rev. Biochem. **48**:837–870.
18. **Kozak, M.** 1978. How do eucaryotic ribosomes select initiation regions in messenger RNA? Cell **15**:1109–1123.
19. **Krzyzek, R. A., M. S. Collett, A. F. Lau, M. L. Perdue, J. P. Leis, and A. J. Faras.** 1978. Evidence for splicing of avian sarcoma virus 5′-terminal genomic sequences onto viral-specific RNA in infected cells. Proc. Natl. Acad. Sci. U.S.A. **75**:1284–1288.
20. **Lee, J. S., H. E. Varmus, and J. M. Bishop.** 1979. Virus-specific messenger RNAs in permissive cells infected by avian sarcoma virus. J. Biol. Chem. **254**:8015–8022.
21. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
22. **McMaster, G. K., and G. G. Carmichael.** 1977. Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. Proc. Natl. Acad. Sci. U.S.A. **74**:4835–4838.
23. **Mellon, P., and P. H. Duesberg.** 1977. Subgenomic, cellular Rous sarcoma virus RNAs contain oligonucleotides from the 3′ half and the 5′ terminus of virion RNA. Nature (London) **270**:631–634.
24. **Oppermann, H., J. M. Bishop, H. E. Varmus, and L. Levintow.** 1977. A joint product of the genes *gag* and *pol* of avian sarcoma virus: a possible precursor of reverse transcriptase. Cell **12**:993–1005.
25. **Palmiter, R. D., J. Gagnon, V. M. Vogt, S. Ripley, and R. N. Eisenman.** 1978. The NH_2-terminal sequence of the avian oncovirus *gag* precursor polyprotein ($Pr76^{gag}$). Virology **91**:423–433.
26. **Paterson, B. M., D. J. Marciani, and T. S. Papas.** 1977. Cell-free synthesis of the precursor polypeptide for avian myeloblastosis virus DNA polymerase. Proc. Natl. Acad. Sci. U.S.A. **74**:4951–4954.
27. **Pawson, T., R. Harvey, and A. E. Smith.** 1977. The size of Rous sarcoma virus mRNAs active in cell-free translation. Nature (London) **268**:416–420.
28. **Pawson, T., G. S. Martin, and A. E. Smith.** 1976. Cell-free translation of virion RNA from nondefective and transformation-defective Rous sarcoma viruses. J. Virol. **19**:950–967.
29. **Pawson, T., P. Mellon, P. H. Duesberg, and G. S. Martin.** 1980. *env* gene of Rous sarcoma virus: identification of the gene product by cell-free translation. J. Virol. **33**:993–1003.
30. **Purchio, A. F., E. Erikson, J. S. Brugge, and R. L. Erikson.** 1978. Identification of a polypeptide encoded by the avian sarcoma virus *src* gene. Proc. Natl. Acad. Sci. U.S.A. **75**:1567–1571.
31. **Purchio, A. F., E. Erikson, and R. L. Erikson.** 1977. Translation of 35S and of subgenomic regions of avian sarcoma virus RNA. Proc. Natl. Acad. Sci. U.S.A. **74**:4661–4665.
32. **Purchio, A. F., S. Jovanovich, and R. L. Erikson.** 1980. Sites of synthesis of viral proteins in avian sarcoma virus-infected chicken cells. J. Virol. **35**:629–636.
33. **Ratzkin, B., and J. Carbon.** 1977. Functional expression of cloned yeast DNA in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **74**:487–491.
34. **Sanger, F., and A. R. Coulson.** 1978. The use of thin acrylamide gels for DNA sequencing. FEBS Lett. **87**:107–110.
35. **Sherman, F., and J. W. Stewart.** 1975. The use of iso-1-cytochrome c mutants of yeast for elucidating the nucleotide sequences that govern initiation of translation, p. 175–191. *In* G. Bernardi and F. Gros (ed.), Organization and expression of the eukaryotic genome: biochemical mechanisms of differentiation in prokaryotes and eukaryotes, Proceedings of the 10th FEBS Meeting. American Elsevier, New York.
36. **Stacey, D. W., V. G. Allfrey, and H. Hanafusa.** 1977. Microinjection analysis of envelope-glycoprotein messenger activities of avian leukosis virus RNAs. Proc. Natl. Acad. Sci. U.S.A. **74**:1614–1618.
37. **Stolzfus, C. M., and L. K. Kuhnert.** 1979. Evidence for the identity of shared 5′-terminal sequence between genome RNA and subgenomic mRNA's of B77 avian sarcoma virus. J. Virol. **32**:536–545.
38. **Struhl, K., and R. W. Davis.** 1977. Production of a functional eukaryotic enzyme in *Escherichia coli*: cloning and expression of the yeast structural gene for imidazoleglycerolphosphate dehydratase (his3). Proc. Natl. Acad. Sci. U.S.A. **74**:5255–5259.
39. **Subramanian, K. N.** 1979. Segments of simian virus 40 DNA spanning most of the leader sequence of the major late viral messenger RNA are dispensable. Proc. Natl. Acad. Sci. U.S.A. **76**:2556–2560.
40. **Swanstrom, R., and P. R. Shank.** 1978. X-ray intensifying screens greatly enhance the detection by autoradiography

of the radioactive isotopes ^{32}P and ^{125}I. Anal. Biochem. 86:184–192.

41. **Swanstrom, R., H. E. Varmus, and J. M. Bishop.** 1982. Nucleotide sequence of the 5′ noncoding region and part of the *gag* gene of Rous sarcoma virus. J. Virol. 41:535–541.
42. **Vapnek, D., J. A. Hautala, J. W. Jacobson, N. H. Giles, and S. R. Kushner.** 1977. Expression in *Escherichia coli* K-12 of the structural gene for catabolic dehydroquinase of *Neurospora crassa*. Proc. Natl. Acad. Sci. U.S.A. 74:3508–3512.

42a. **Varmus, H. E., N. Quintrell, and S. Ortiz.** 1981. Retroviruses as mutagens: insertion and subsequent excision of a nontransforming provirus alter expression of a resident transforming provirus. Cell 25:23–36.

43. **von der Helm, K., and P. H. Duesberg.** 1975. Translation of Rous sarcoma virus RNA in a cell-free system from Ascites Krebs II cells. Proc. Natl. Acad. Sci. U.S.A. 72:614–618.
44. **Weaver, R. F., and C. Weissmann.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15S β-globin mRNA precursor and mature 10S β-globin mRNA have identical map coordinates. Nucleic Acids Res. 7:1175–1193.
45. **Weiss, S. R., P. B. Hackett, H. Oppermann, A. Ullrich, L. Levintow, and J. M. Bishop.** 1978. Cell-free translation of avian sarcoma virus RNA: suppression of the *gag* termination codon does not augment synthesis of the joint *gag/pol* product. Cell 15:607–614.
46. **Weiss, S. R., H. E. Varmus, and J. M. Bishop.** 1977. The size and genetic composition of virus-specific RNAs in the cytoplasm of cells producing avian sarcoma-leukosis virus. Cell 12:983–992.