JOURNAL OF VIROLOGY, Sept. 1985, p. 744–751
0022-538X/85/090744-08$02.00/0

Vol. 55, No. 3

Identification and Predicted Sequence of a Previously Unrecognized Small Hydrophobic Protein, SH, of the Paramyxovirus Simian Virus 5

SCOTT W. HIEBERT, REAY G. PATERSON, AND ROBERT A. LAMB*

Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60201

Received 18 April 1985/Accepted 28 May 1985

A previously unrecognized gene (SH) has been identified on the virion RNA of the paramyxovirus simian virus 5 between the genes for the fusion protein and the hemagglutinin-neuraminidase. An SH mRNA of 292 nucleotides (plus polyadenylate residues), transcribed from the SH gene, has been identified. The SH mRNA contains a single open reading frame which encodes a polypeptide of 44 amino acids with a molecular weight of 5,012. The SH polypeptide is predicted to contain an extensive hydrophobic region. This protein has been identified in simian virus 5-infected cells, and it has been shown to be encoded by the SH mRNA by in vitro translation of size-fractionated mRNAs, hybrid-arrest translation, and hybrid-selection translation.

Simian virus 5 (SV5) is a prototype of the paramyxovirus family. The SV5 virion contains an envelope consisting of a lipid bilayer with a non-glycosylated protein (M) associated with its inner surface and two integral membrane glycoproteins (HN and F) that form spike-like projections on the outer surface (8). Extensive studies involving the isolation and purification of biologically active forms of the glycoproteins HN and F have demonstrated that HN has both receptor-binding (hemagglutinating) and neuraminidase activities, while the F protein causes membrane fusion (41, 42; for a review, see reference 9). Inside the envelope is the ribonucleoprotein core, which consists of a single ca. 15,000-nucleotide chain of negative-sense RNA, and the major structural component, the nucleocapsid protein (NP). Two other polypeptides, P and L, are associated with the nucleocapsid, and together they are thought to have RNA transcriptase activity (6). Once the virus has entered a cell, the virion polymerase transcribes the 50S RNA into mRNA species encoding the six major structural proteins (NP, P, M, F, HN and L) plus a nonstructural protein designated V (16, 37).

To better understand the primary structure-function relationship of the SV5 polypeptides, a library of cDNA clones to mRNAs from SV5-infected CV1 cells was constructed (35). Clones containing sequences specific for the mRNA species coding for NP, P/V, M, F, and HN proteins have been isolated from the library. Some of these cloned cDNAs are thought to have been derived from polycistronic mRNA transcripts as judged by their ability to prevent translation in vitro of more than one polypeptide in hybrid-arrest of translation experiments and by their hybridization pattern to two monocistronic mRNA species and larger RNA species in blot-hybridization experiments (35).

The nucleotide sequence of two clones, Fc (36) and HN177 (20), has been determined. The clone Fc has been shown to contain not only a full-length copy of the F mRNA but also 106 nucleotides of the M mRNA (36). Clone HN177 contains a complete copy of the HN mRNA plus 253 nucleotides before the position to which the 5′ end of the HN mRNA was mapped (20). These 253 nucleotides are not derived from the 3′ end of the F mRNA, and therefore it seems possible that on the SV5 genome there is a previously undetected gene between F and HN, of which we have obtained part of the sequence.

In the present study, sequences derived from the region 3′ to the HN coding sequences on the SV5 genome have been analyzed further. The complete nucleotide sequence of the region between F and HN on the 50S RNA was determined by both direct and indirect means, and a small mRNA transcribed from this region was identified. There is a single open reading frame on the mRNA, and use was made of the deduced amino acid sequence to identify the small and highly hydrophobic polypeptide, designated SH, encoded by the mRNA.

MATERIALS AND METHODS

Cells. Monolayer cultures of the TC7 clone of CV1 cells and a variant of the MDBK line of bovine kidney cells were grown in Dulbecco modified Eagle medium with 10% fetal calf serum.

Virus infection and labeling of cells. The W3 strain of SV5 was grown in MDBK cells (7, 37). For biochemical experiments, CV1 cells were infected as previously described (37) with actinomycin D (5 μg/ml) added immediately after infection (35). For protein labeling, infected cultures at 10 h postinfection were washed in minimal essential medium deficient in phenylalanine and isoleucine, labeled with [^{3}H]phenylalanine plus [^{3}H]isoleucine, and harvested in lysis buffer (29).

Primer extension sequence analysis. Polyadenylic acid-containing [poly(A)$^+$] mRNAs from SV5-infected CV1 cells were isolated as previously described (35). Primer extension sequence analysis of the 5′ end of the SV5 SH mRNA was done as previously described (28) with a *Taq*I-to-*Hpa*II DNA primer fragment of clone HN177 (20) (nucleotides −246 to −203) which was 5′ uniquely labeled at nucleotide −203. The extended products were separated by gel electrophoresis, and the nucleotide sequence of the DNA was determined (28, 34).

Primer extension sequence analysis on the SV5 50S RNA. Two synthetic oligonucleotides, 5′d(GAACAAGATAAGAC AGTCAT)3′ and 5′d(TGCCATACTTCCTACTCAC)3′, were synthesized by Patrick Groody with a Vega Coder 300

* Corresponding author.

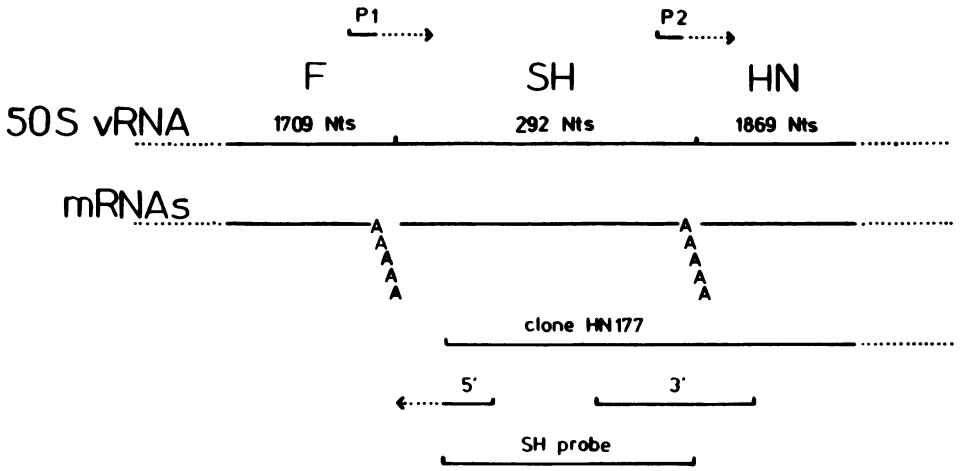

FIG. 1. Schematic diagram showing the region of 50S SV5 RNA described in the experiments and the mRNAs transcribed from this region. The DNA fragments from clone HN177 (20) and the synthetic oligonucleotides used to establish the SH region are indicated, and a detailed description of the experiments is described in the text and subsequent figures. P1, Synthetic oligonucleotide 5'd(GAACAAGATAAGACAGTCAT)3' corresponding to the 3' end of the F mRNA (nucleotides 1675 to 1694 [36]). P2, Synthetic oligonucleotide 5'd(TGCCATACTTCCTACTCAC)3' corresponding to the region upstream of the 5' end of the HN mRNA (clone HN177 sequences −45 to −27 [20]). 5', *Taq*I-*Hpa*II DNA primer from clone HN177 nucleotides −246 to −203; 3', *Dde*I-*Hpa*II DNA fragment from clone HN177 nucleotides −100 to +20; SH probe, *Pst*I-*Hae*III DNA fragment from clone HN177 nucleotides −253 to +2.

oligonucleotide synthesizer and purified by polyacrylamide gel electrophoresis. The synthetic primers were hybridized to purified 50S virion RNA and extended with reverse transcriptase in the presence of deoxynucleotides and limiting concentrations of di-deoxynucleotides as previously described (2). Samples were analyzed on DNA sequencing gels (33).

Nuclease S1 analysis. To determine the position of the 3' end of the SH mRNA, nuclease S1 analysis was done as previously described (36) with a ^{32}P-3'-end-labeled *Dde*I-to-*Hpa*II DNA fragment of clone HN177 (20). Specific labeling at the *Dde*I site was done with the Klenow fragment of *Escherichia coli* DNA polymerase I and [α-^{32}P]dATP in the presence of unlabeled dTTP (33).

RNA blot analysis. SV5-specific mRNAs were separated on a 1.5% agarose gel containing 2.2 M formaldehyde, transferred to nitrocellulose, and hybridized with ^{32}P-labeled cloned DNA probes (39) specific for F, SH, and HN regions, as previously described (28).

Hybrid-arrested translation and hybrid selection of mRNAs. Hybridization of SV5-specific mRNA with a 255-nucleotide *Pst*I-to-*Hae*III double-stranded DNA fragment from clone HN177 (20), which includes the entire coding region of the SH mRNA, was done as previously described (26). Approximately 3 μg of double-stranded cDNA was hybridized to poly(A)$^+$ mRNAs (ca. 0.5 μg) obtained from a 10-cm plate of SV5-infected CV1 cells treated with actinomycin D. Hybrid selection of SV5-specific mRNA was done with the same *Pst*I-to-*Hae*III DNA fragment from clone HN177 bound to nitrocellulose by using previously described methods (38). mRNAs that hybridized to the immobilized DNA were eluted and translated in wheat germ extracts as previously described (27).

Methyl mercury gel electrophoresis. A vertical, 1.4% agarose (with low gelling temperature) gel containing 10 mM methyl mercury hydroxide was prepared as previously described (33). Poly(A)$^+$ RNA (10 μg) was denatured in 5 mM methyl mercury hydroxide and subjected to electrophoresis for 16 h at 40 V as previously described (35). In vitro translation of the fractionated mRNA species was done essentially as previously described (32), except that wheat germ extracts were used.

Polyacrylamide gel electrophoresis, autoradiography, and fluorography. Gels contained 20% acrylamide and an acrylamide–bis-acrylamide ratio of 300:1 (25, 29) or 9% acrylamide and 8 M urea with an acrylamide–bis-acrylamide

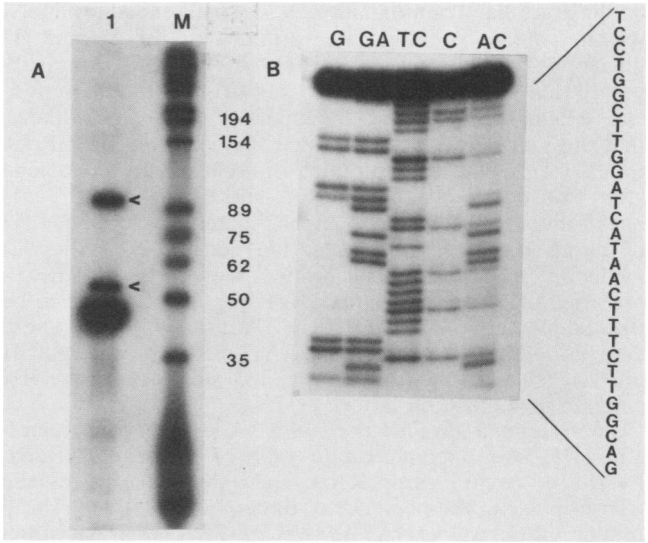

FIG. 2. Primer extension and sequence analysis of SV5 mRNA with a DNA primer from the SH region. (A) The 43-nucleotide *Taq*I-*Hpa*II DNA fragment (HN177 sequences −246 to −203) was 5' uniquely end labeled at the *Hpa*II site and hybridized to poly(A)$^+$ mRNAs isolated from SV5-infected CV1 cells. The primer was extended with reverse transcriptase as described in the text. Lane 1, cDNAs were analyzed on an 8% polyacrylamide gel containing 9 M urea; lane M, ^{32}P-labeled size markers from *Hin*fI-digested pFVM/45 DNA (27). (B) Nucleotide sequence of the larger 92-nucleotide extended product from Fig. 2A (lane 1) showing the 5'-terminal region. The base-specific chemical cleavage is indicated above each lane of the gel. The sequence is the complement of the mRNA sense shown in Fig. 6.

ratio of 10:1 (45). Fluorography and autoradiography were done as previously described (37).

Computer-assisted analysis. The DNA and amino acid sequence analysis programs of Lagrimini et al. (24) and Devereux et al. (14) were used. The relative hydrophobicity was calculated by the program of Kyte and Doolittle (23).

Special materials. Actinomycin D was from Merck & Co., Inc., Rahway, N.J. Klenow fragment of *E. coli* DNA polymerase I and nuclease S1 were from Boehringer-Mannheim Biochemicals, Indianapolis, Ind. Methyl mercury hydroxide was from Thiokol, Danvers, Mass. Avian myeloblastosis virus reverse transcriptase was from Life Sciences, Inc., St. Petersburg, Fla.

RESULTS

Identification of a previously unrecognized region (SH) on the 50S genome RNA. The nucleotide sequence of the cDNA clones HN177 and HN36 indicate that they contain 253 and 247 nucleotides, respectively, before the site to which the 5′ end of the HN mRNA was mapped (20). We have designated this region SH. As both clones were of almost exactly the same length, it is suggested that the 253 or 247 nucleotides of the SH region were derived by reverse transcriptase terminating prematurely on a polycistronic mRNA, or that the clones mapped to the 5′ end of a polycistronic mRNA (20). To investigate these possibilities, primer-extension sequencing analysis of SV5-specific mRNAs was done.

A DNA primer derived from the 5′-proximal region of clone HN177 was isolated (*Taq*I-*Hpa*II; HN177 nucleotides −246 to −203) and 5′ uniquely labeled at nucleotide −203 (Fig. 1). The primer was hybridized to poly(A)$^+$ mRNAs isolated from SV5-infected cells, the primer was extended with reverse transcriptase, and the products were analyzed on denaturing polyacrylamide gels. Two extended products, which were larger than the primer and of 51 and 92 nucleotides in length, were observed (Fig. 2A). This suggested that either two poly(A)$^+$ transcripts with different 5′ ends originated from the 50S genome RNA, or, more likely, that the longer product contained the 5′ end of an mRNA transcript; the shorter product was caused by a "strong stop" of reverse transcriptase. The 5′ end of the 51-nucleotide product maps to approximately nucleotide −253 of clone HN177. The nucleotide sequence of the 92-nucleotide cDNA extended product was obtained. The 5′-proximal region of the sequence (Fig. 2B) was found to be the complement of 5′-AGGACCGAACCTAG . . . 3′. These data indirectly suggest that a poly(A)$^+$ mRNA transcript (SH mRNA) is transcribed from a region 3′ upstream of the HN gene on the 50S RNA.

To investigate directly the 50S RNA sequence between F and HN, two oligonucleotide primers were synthesized, hybridized to 50S virion RNA, and extended with reverse transcriptase in the presence of di-deoxynucleotides. The F primer 5′d(GAACAAGATAAGACAGTCAT)3′ was synthesized and represents the 3′ end of the F mRNA (nucleotides 1675 to 1694) (36); the second synthetic primer, 5′d(TGCCATACTTCCTACTCAC)3′, corresponds to a region upstream of the 5′ end of the HN mRNA (clone HN177 sequences −45 to −27 [20]) (see Fig. 1 for location of primers). The F primer yielded the sequence, as shown below and in Fig. 3,

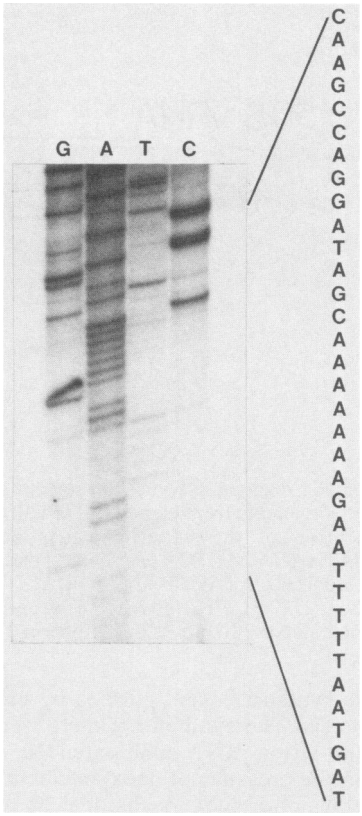

FIG. 3. Primer extension sequence analysis of the SV5 50S RNA between the F and SH genes. The primer 5′d(GAACAAGATAA GACAGTCAT)3′, which represents nucleotides 1675 to 1694 of the F mRNA (36), was synthesized. The primer was hybridized to SV5 50S virion RNA and extended with reverse transcriptase in the presence of deoxynucleotides and di-deoxynucleotides (2). Samples were analyzed on a 20% polyacrylamide gel containing 9 M urea (34). The base-specific reaction is indicated at the top of each lane of the gel, and the sequence is the mRNA sense. The region of the gel showing the sequence between the F and SH genes is shown. The diagonal line in the G lane is an artifact introduced in the autoradiography of the sequencing gel.

which is contiguous with that found for the 3′ end of the F mRNA, except for the G residue marked with a star (36) and then, after 4 nucleotides, matched the sequence of the SH region obtained by primer extension sequence analysis of mRNAs (Fig. 2B). The second primer yielded the identical sequence to that found for clone HN177 (20). Thus, these data indicate there is a region (SH) of 298 nucleotides between the F and HN genes on the SV5 50S RNA (see Fig. 6).

Identification of the SH mRNA. To obtain direct evidence that an mRNA species was transcribed from the 50S virion RNA, and if so, to examine the relationship of possible polycistronic transcripts, RNA blot hybridization was performed. Poly(A)$^+$ RNAs from SV5-infected CV1 cells were separated by electrophoresis on formaldehyde-agarose gels and transferred to nitrocellulose paper. Strips of paper corresponding to individual lanes of the gel were hybridized

CCATTAGTAATTTTTAAG*AAAAAAACGATAGGACCGAACC
F SH

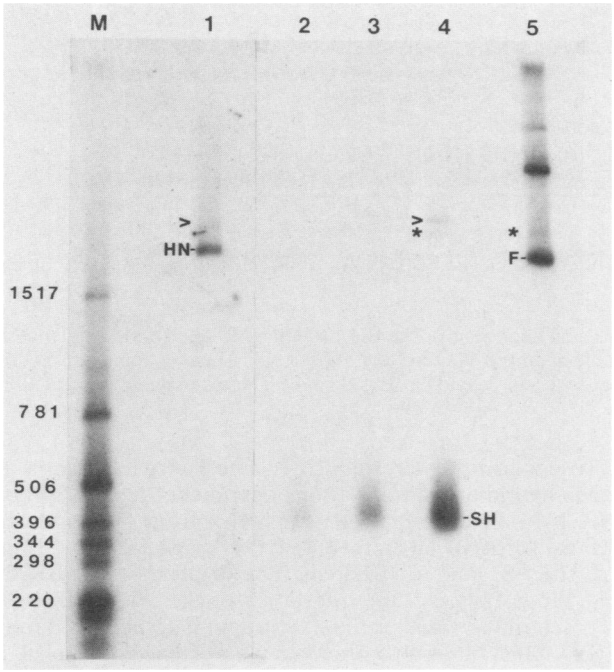

FIG. 4. RNA blot-hybridization analysis of SV5-specific mRNAs probed with F-, SH-, and HN-specific DNAs. Poly(A)$^+$ mRNAs from SV5-infected CV1 cells were separated on a formaldehyde-agarose gel and transferred to nitrocellulose, and the individual lanes hybridized to ^{32}P-labeled DNA probes for F, SH, and HN. Lane M, ^{32}P-labeled marker DNA fragments from pFVM/45 (27); lane 1, hybridization with a ^{32}P-labeled, HN-specific probe; lanes 2 to 4, increasing concentrations of mRNA in the ratio 1:2:4 hybridized to the SH-specific ^{32}P-labeled probe (clone HN177 nucleotides −253 to +2); and lane 5, hybridization with a ^{32}P-labeled, F-specific probe. Symbols: (*) a transcript shared between F and SH and (>) a transcript shared between SH and HN.

to ^{32}P-labeled DNA probes specific for the F, SH, and HN genes. The SH region probe (clone HN177 nucleotides −253 to +2) hybridized to a major species of ca. 400 nucleotides, two minor species of ca. 2,100 and ca. 2,200 nucleotides, and a very minor species of ca. 3,600 nucleotides (Fig. 4, lanes 2 to 4). The major ca. 400-nucleotide mRNA species that specifically hybridized to the SH probe is of the size expected for an mRNA transcribed from the SH region of the 50S RNA [292 nucleotides plus ca. 100 poly(A) residues]. The F probe hybridized to two major species, one of ca. 1,800 nucleotides and the other of ca. 3,200 nucleotides, and, to a lesser extent, to a species of ca. 4,500 nucleotides (Fig. 4, lane 5). These three species are thought to represent monocistronic F mRNA and polycistronic 5′-M-F-3′ and 5′-P/V-M-F-3′ mRNAs, respectively (35). The F probe also hybridized to a minor species of ca. 2,100 nucleotides that is shared between F and SH. The HN-specific probe hybridized to a single major species of 1,900 nucleotides corresponding to monocistronic HN mRNA (20, 35) and a minor species of ca. 2200 nucleotides that is shared between SH and HN. The ca. 2,100- and ca. 2,200-nucleotide species of RNA were also seen previously (35), but their significance was not appreciated. Thus, these data indicate that an mRNA is transcribed from the SH region, and the ca. 2,100- and ca. 2,200-nucleotide species probably represent polycistronic transcripts of F + SH and SH + HN, respectively.

Determination of the 3′ end of monocistronic SH mRNAs. The 50S virion RNA and clones HN177 and HN36 contain the sequence 3′-UUCUUUUUU-5′ (vRNA sense) at position −11 to −3 of clone HN177 just before the location of the 5′ end of the HN mRNA (see above and reference 20). This sequence is very similar to the consensus sequence for poly(A) addition in Sendai virus (19). To determine the 3′ end of monocistronic SH mRNAs, a 120-nucleotide *Dde*I-*Hpa*II DNA fragment (HN177 nucleotides −100 to +20 [Fig. 1]) was 3′-^{32}P labeled at the *Dde*I site. This DNA fragment was hybridized to poly(A)$^+$ mRNAs from SV5-infected CV1 cells and digested with nuclease S1; the protected DNA fragments were analyzed on denaturing polyacrylamide gels. A major protected fragment of ca. 100 nucleotides was observed (Fig. 5), which suggests that poly(A) is added to the SH mRNA in the region of nucleotides −8 to −3 (clone HN177 numbering) as described above.

Nucleotide sequence of the SH mRNA and the predicted amino acid sequence. The direct sequencing of the SH region of the 50S virion RNA, the primer extension sequencing to map the 5′ end of the SH mRNA, the 3′-end mapping of the SH mRNA, and the nucleotide sequences of clones HN177 and HN36 (20) make it possible to predict the entire sequence of the SH mRNA. The sequence (mRNA sense [Fig. 6]) contains ca. 292 nucleotides; however, the exact position within the 3′ six adenosine residues in which transcription termination occurs and poly(A) addition begins cannot be determined. The nucleotide sequence contains a single open reading frame beginning at the ATG codon at nucleotides 80 to 82 and terminating at TAA (nucleotides 212 to 214). The open reading frame could encode a protein of 44 amino acids with a molecular weight of 5,012. The predicted SH protein does not contain potential glycosylation sites for the N-glycosidic linkage of oligosaccharides (i.e., Asn-X-Ser or Thr). However, the predicted SH protein is highly hydrophobic in nature, as a large percentage of its residues are hydrophobic and it contains a region (residues 19 to 38)

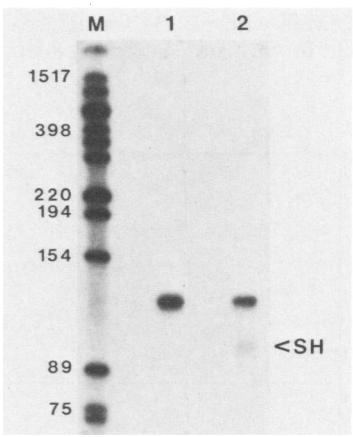

FIG. 5. Nuclease S1 analysis to determine the 3′ end of the SH mRNA. A 120-nucleotide *Dde*I-to-*Hpa*II DNA fragment (HN177 nucleotides −100 to +20, [20]) was 3′ end labeled and hybridized to poly(A)$^+$ RNA from SV5-infected CV1 cells and digested with nuclease S1, and the protected DNA fragments were analyzed on 6% denaturing polyacrylamide gels (28). Lane M, ^{32}P-labeled marker DNA fragments from pFVM/45 (27); lane 1, hybridization in the presence of no added mRNA as a control for reassociation of the DNA; and lane 2, hybridization with poly(A)$^+$ RNA from SV5-infected CV1 cells. The protected fragment of ca. 100 nucleotides is indicated as SH.

```
.1          .20          .40          .60          .80          .100
AGGACCGAACCTAGTATTGAAAGAACCGTCTCGGTCAATCTAGGTAATCGAGCTGATACCGTCTCGGAAAGCTCAAATC   ATG CTG CCT GAT CCG GAA GAT CCG
                                                                                   Met-Leu-Pro-Asp-Pro-Glu-Asp-Pro-(8)

          .120         .140         .160         .180
GAA AGC AAG AAA GCT ACA AGG AGA GCA GGA AAC CTA ATT ATC TGC TTC CTA TTC ATC TTC TTT CTG TTT GTA ACC TTC ATT GTT CCA
Glu-Ser-Lys-Lys-Ala-Thr-Arg-Arg-Ala-Gly-Asn-Leu-Ile-Ile-Cys-Phe-Leu-Phe-Ile-Phe-Phe-Leu-Phe-Val-Thr-Phe-Ile-Val-Pro-(37)

          .200         .220         .240         .260         .280
ACT CTA AGA CAC TTG CTG TCC TAA    CACCTGCTATAGGCTATCCACTGCATCATCTCTCCTGCCATACTTCCTACTCACATCATATCTATTTTAAAGAAAAAA
Thr-Leu-Arg-His-Leu-Leu-Ser (44)
```

FIG. 6. Nucleotide sequence of the SH mRNA and the predicted amino acid sequence of the SH protein. The nucleotide sequence is derived from direct sequencing of the SV5 50S RNA, primer extension sequencing of the 5′ end of the SH mRNA, 3′-end mapping of the SH mRNA, and the nucleotide sequences of clones HN177 and HN36 (20). (See text and legends to Fig. 1 to 4.) The extensively hydrophobic region in the predicted SH protein is underlined.

of sufficient length and hydrophobicity to anchor the protein in membranes. The hydropathy plot (Fig. 7) demonstrates that residues 19 to 38 have an average hydropathic index (23) of >2, a value normally found only in proteins that are embedded in membranes (23).

Identification of the SH protein synthesized in vitro and in vivo. Examination of the predicted amino acid sequence of the SH protein indicates that it is rich in phenylalanine (six residues) and isoleucine (four residues) but only contains one cysteine residue and an initiator methionine residue that may be cleaved from the mature protein. Poly(A)$^+$ mRNAs isolated from SV5-infected CV1 cells were translated in wheat germ extracts with [^{3}H]phenylalanine and [^{3}H]isoleucine as radioactive precursors, and the translation products were analyzed on highly cross-linked polyacrylamide gels (45). A previously unrecognized polypeptide, SH (Fig. 8, lanes 1 and 4), of $M_r \sim 5{,}000$ was readily detected and shown to be the predicted SH polypeptide by hybrid-selection and hybrid-arrest translation experiments. The SH probe DNA (*Pst*I-*Hae*III DNA [Fig. 1]) selected mRNAs from SV5-infected CV1 cells that on translation in vitro yielded a polypeptide (SH) of $M_r \sim 5{,}000$ (Fig. 8, lane 3). A control hybridization with pBR322 sequences did not yield any recognizable in vitro translation products (Fig. 8, lane 2). To confirm these results, the same DNA fragment was used in hybrid-arrest translation experiments. The SH probe DNA hybridized to poly(A)$^+$ mRNAs from SV5-infected CV1 cells specifically prevented the translation of SH in vitro (Fig. 8, lane 5), whereas hybridization with pBR322 DNA did not prevent the translation of SH (Fig. 8, lane 6).

To determine whether the SH protein is translated during an SV5 infection in vivo, infected CV1 cells were labeled at 10 h postinfection with [^{3}H]phenylalanine and [^{3}H]isoleucine. A protein (Fig. 8, lane 9) of the same mobility as SH synthesized in vitro (Fig. 8, lane 8) could be readily detected in SV5-infected cells but not in uninfected cells (Fig. 8, lane 10).

SH is translated from a small mRNA of ca. 400-nucleotide chain length. To correlate directly the size of the SH mRNA described above with the translation of the SH polypeptide, poly(A)$^+$ mRNAs isolated from SV5-infected CV1 cells were denatured in methyl mercury hydroxide and separated on an agarose gel (with low gelling temperature) containing methyl mercury hydroxide. After electrophoresis, the gel was fractionated into 2-mm slices. A portion of each slice was melted, diluted, and translated directly in wheat germ extracts. The following fractions containing the SV5-specific polypeptides described previously (Fig. 9 [35]) were ob-

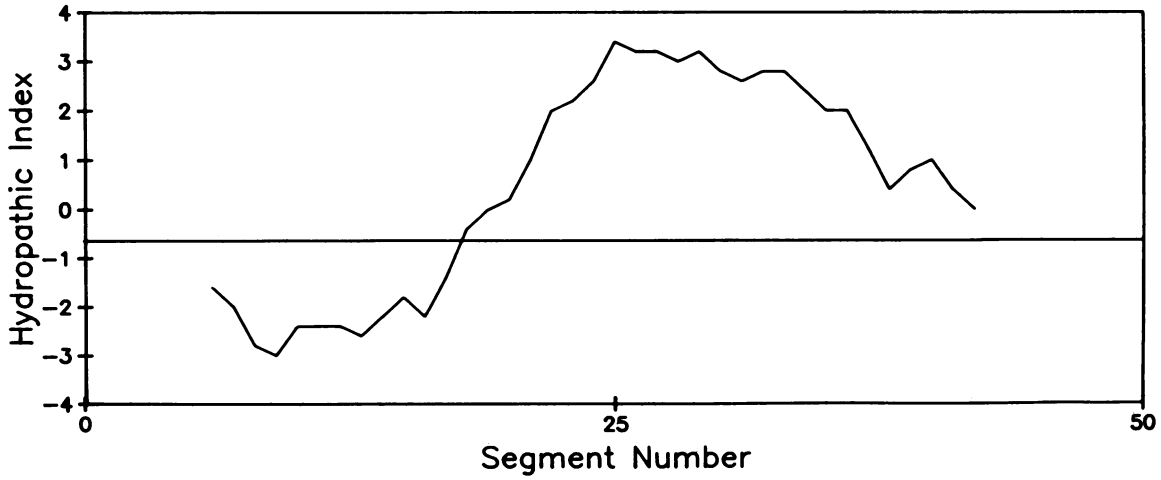

FIG. 7. Hydropathy plot of the predicted SH protein. The relative hydrophobicity and hydrophilicity of the protein along its amino acid sequence was calculated as previously described (23) with a segment length of 7 amino acids. The consecutive scores are plotted from the N to C terminus of the protein. The midpoint line corresponds to the grand average of the hydropathy of the amino acid compositions found in most sequenced proteins (23).

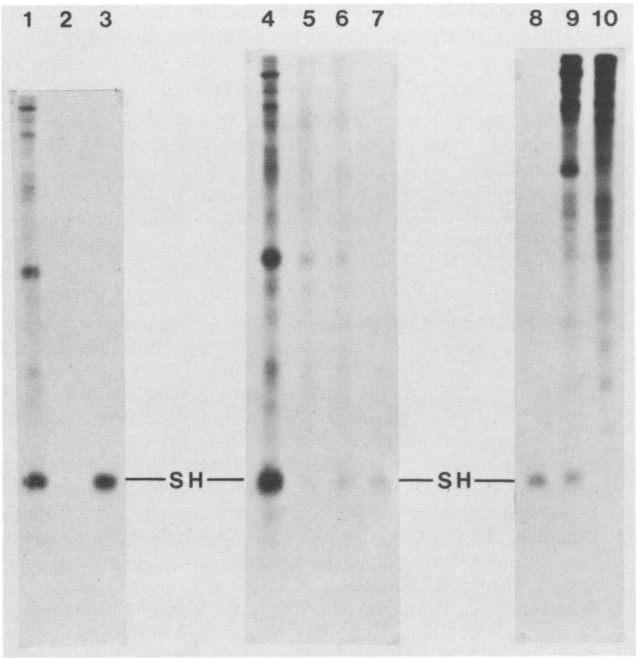

FIG. 8. Identification of the SH polypeptide synthesized in vitro and in vivo was as follows: hybrid-selection translation, hybrid-arrest of translation, and in vivo metabolic labeling. [^{3}H]phenylalanine and [^{3}H]isoleucine were used as radioactive precursors to label the SH protein. Lane 1, In vitro translation of total SV5-infected cell mRNA; lane 2, in vitro translation of mRNAs selected by pBR322 DNA; lane 3, in vitro translation of mRNAs selected by SH-specific DNA (clone HN177 nucleotides −253 to +2); lane 4, as in lane 1; lane 5, hybrid-arrest of translation of SH after hybridization of SV5-infected cell mRNA with SH-specific DNA; lane 6, Hybrid-arrest of translation of SV5-infected cell mRNA hybridized with pBR322 DNA as a control; lane 7, as in lane 3; lane 8, as in lane 3; lane 9, [^{3}H]phenylalanine- and [^{3}H]isoleucine-labeled, SV5-infected cells; and lane 10, [^{3}H]phenylalanine- and [^{3}H]isoleucine-labeled, uninfected cells.

tained: fraction 23 (HN), fractions 24 and 25 (NP and F), fractions 28 and 29 (M) and fractions 28 to 30 (P and V). Fractions 59 to 61 yielded the SH polypeptide, and the mRNA in these fractions corresponds to a size of ca. 400 nucleotides as compared with the mobility of single-stranded DNA standards. To separate all of the SV5-specific polypeptides on one polyacrylamide gel, a 20% gel with low cross-linking had to be used, and it was found that SH migrates as a diffuse band. This is probably attributable to the hydrophobic nature of SH. Thus, these data indicate that the SH polypeptide is translated from the SH mRNA.

DISCUSSION

We have shown that there is a region (SH) on the SV5 50S RNA located between the genes for F and HN that is transcribed to yield a small poly(A)$^+$ mRNA. The nucleotide sequence of the SH mRNA was obtained from the sequence of this region of the 50S virion RNA, the sequences of clone HN177 and HN36 (20), the 5′ primer extension sequencing of the SH mRNA, and mapping of the 3′ end of the SH mRNA. The SH mRNA contains a single open reading frame encoding 44 amino acids, beginning at the ATG codon at nucleotides 80 to 82 and terminating at TAA (nucleotides 212 to 214). By the choosing of radioactive precursors ([^{3}H]phenylalanine and [^{3}H]isoleucine) predicted to be contained in the protein sequence, we have identified a previously undetected polypeptide (SH) (M_r = 5,012) which is synthesized both in vitro and in vivo. The SH polypeptide was shown to be encoded by the SH region of the 50S genome RNA and the SH mRNA by hybrid-arrest translation and hybrid-selection translation experiments, and translation in vitro of mRNAs fractionated on methyl mercury hydroxide-containing agarose gels.

The SH (small hydrophobic) polypeptide contains a large percentage of hydrophobic residues and possesses a region (residues 19 to 38) which is sufficiently hydrophobic and of appropriate length to anchor the polypeptide in cell membranes (23). We have not been able to identify SH in purified virions, although it can be readily detected in infected cells, and thus it may be a viral nonstructural protein. An example of nonstructural integral membrane proteins exists in influenza A viruses, in which it has been found that the nonglycosylated M_2 protein is expressed at the infected cell surface (31). In influenza B viruses, a nonstructural glycoprotein (NB) has also been identified in infected cells (30, 43, 44).

The SH mRNA was readily detected in RNA blotting experiments, but its relative rate of transcription compared

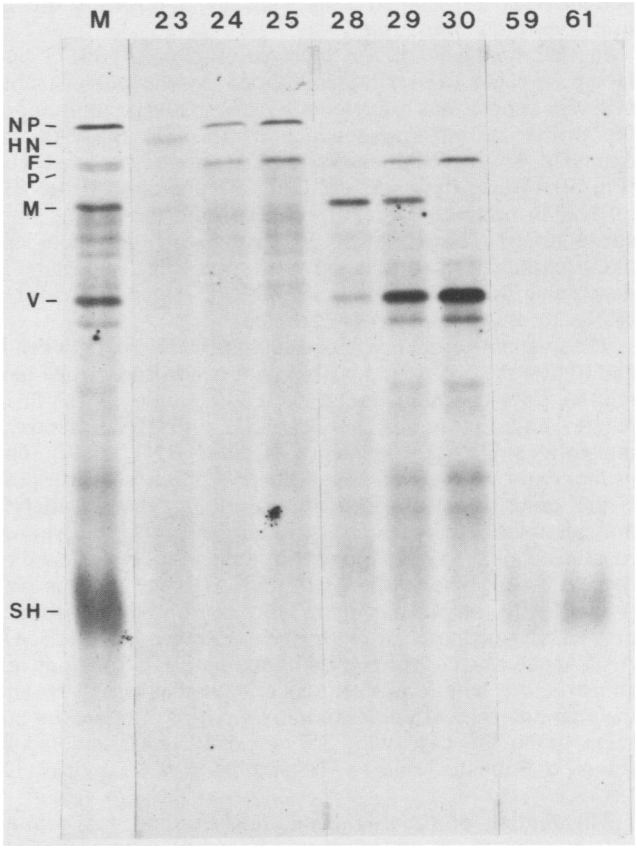

FIG. 9. In vitro translation of SV5 mRNAs fractionated on a methyl mercury agarose gel (low gelling temperature). The gel electrophoresis, fractionation, and in vitro translation, with [^{3}H]phenylalanine and [^{3}H]isoleucine as labeled precursors, were done as described in the text. M, In vitro translation of total mRNA from SV5-infected CV1 cells. The lane numbers represent fractions taken from the region of the gel corresponding to the migration of markers of 2,300 to 200 nucleotides. Representative fractions only are shown.

with that of the F and HN mRNAs and its level of accumulation cannot be determined from these experiments. Polycistronic transcripts shared between NP + P/V, P/V + M, and M + F and larger transcripts between P/V + M + F could be readily detected (35). However, much lesser amounts of polycistronic transcripts between F + SH and SH + HN can be detected. It has been indicated, from UV transcriptional mapping analysis, that for the rhabdoviruses (e.g., vesicular stomatitis virus) and paramyxoviruses (e.g., Sendai virus, Newcastle disease virus, measles, and respiratory syncytial virus), there is a single promoter at the 3′ end of the genome RNA, and transcription of each mRNA has been shown to be sequential to that of the 3′-proximal gene (1, 3, 10, 15, 17, 18). With vesicular stomatis virus it has been found that transcription pauses between genes (21, 22). It would seem likely that the relative lack of polycistronic transcripts between F and SH and SH and HN in SV5 is due to transcription pausing at these junctions to a much greater extent than at the NP, P/V, M, or F gene junctions. A second less likely possibility is that the SV5 genome contains more than one transcriptase entry site that is usually activated after transcription of the SH 3′-proximal genes. With respiratory syncytial virus, it has also been observed that there is a lack of polycistronic transcripts between the genes for the 9,500-molecular weight (9.5K) polypeptide and the G polypeptide (12, 15), which also suggests a possible site of transcriptional attenuation.

In the synthesis of the cDNAs of clones HN177 and HN36, reverse transcriptase stopped on the polycistronic mRNAs at positions 4 nucleotides apart (20). In addition, in the primer extension sequence of the SH mRNAs, an extended product that maps to this region was observed (Fig. 2). These data suggest that this position on the SH mRNA in monocistronic or polycistronic transcripts may cause reverse transcriptase to strong-stop. A much less likely explanation is that the SH mRNA has alternate 5′ ends, and this possibility can only be eliminated once a cDNA to the SH mRNA is obtained.

The available evidence suggests that the region between F and SH on the 50S genome RNA that is not transcribed into mRNA (intercistronic region) is 4 nucleotides (CGAT) in length, and the region between SH and HN extends 2 nucleotides (TA) (see sequence of clone HN177 [20]). The primer extension sequencing of the SV5 50S RNA across the F-SH gene junction indicates that, at the F mRNA polyadenylation site, a G is found before the seven A residues which was not found in the nucleotide sequence of clone Fc (36). The synthesis of cDNAs from mRNAs was primed with oligodeoxythymidylic acid, and the primer may have mismatched with the sequence 5′-AAGAAA AAAA(poly A)-3′. Although the intercistronic junctions are conserved in length and sequence in Sendai virus (19) and the rhabdovirus vesicular stomatis virus (40), this does not seem to be the case with SV5 (see above; Paterson and Lamb, unpublished data) or respiratory syncytial virus (11, 13).

The finding of the SH gene, mRNA, and polypeptide indicates that the paramyxovirus SV5 contains at least seven known genes (NP, P/V, M, F, SH, HN, and L). The available nucleotide sequence of the Sendai virus 50S RNA suggests that a counterpart to SH is not found between the F and HN genes (4, 5). However, respiratory syncytial virus, which is a member of the pneumovirus subgroup of paramyxoviruses, is known to contain a minimum of 10 genes (11, 15). Thus, the paramyxoviruses may contain more variation in gene structure than previously thought. The extra gene products may be important for the pathogenicity and virulence of the viruses in their natural animal hosts.

ACKNOWLEDGMENTS

We thank Diane Braun for excellent technical assistance and Patrick Groody and Robert Letsinger for synthesizing the oligonucleotides.

This research was supported by a National Institutes of Health research grant (AI-20201). R.A.L. is an Established Investigator of the American Heart Association.

LITERATURE CITED

1. **Abraham, G., and A. K. Banerjee.** 1976. Sequential transcription of the genes of vesicular stomatitis virus. Proc. Natl. Acad. Sci. U.S.A. **73**:1504–1508.
2. **Air, G. M.** 1979. Nucleotide sequence coding for the 'signal peptide' and N terminus of the hemagglutinin from an Asian (H2N2) strain of influenza virus. Virology **97**:468–472.
3. **Ball, L. A., and C. N. White.** 1976. Order of transcription of genes of vesicular stomatitis virus. Proc. Natl. Acad. Sci. U.S.A. **73**:442–446.
4. **Blumberg, B. M., C. Giorgi, K. Rose, and D. Kolakofsky.** 1985. Sequence determination of the Sendai virus fusion protein gene. J. Gen. Virol. **66**:317–331.
5. **Blumberg, B., C. Giorgi, L. Roux, R. Raju, P. Dowling, A. Chollet, and D. Kolakofsky.** 1985. Sequence determination of the Sendai virus HN gene and its comparison to the influenza virus glycoproteins. Cell **41**:269–278.
6. **Buetti, E., and P. W. Choppin.** 1977. The transcriptase complex of the paramyxovirus SV5. Virology **82**:493–508.
7. **Choppin, P. W.** 1964. Multiplication of a myxovirus (SV5) with minimal cytopathic effects and without interference. Virology **23**:224–233.
8. **Choppin, P. W., and R. W. Compans.** 1975. Reproduction of paramyxoviruses, p. 95–178. *In* H. Fraenkel-Conrat and R. R. Wagner (ed.), Comprehensive virology. Plenum Publishing Corp., New York.
9. **Choppin, P. W., and A. Scheid.** 1980. The role of viral glycoproteins in adsorption, penetration and pathogenicity of viruses. Rev. Infect. Dis. **2**:40–61.
10. **Collins, P. L., L. E. Hightower, and L. A. Ball.** 1980. Transcriptional map for Newcastle disease virus. J. Virol. **35**:682–693.
11. **Collins, P. L., Y. T. Huang, and G. W. Wertz.** 1984. Identification of a tenth mRNA of respiratory syncytial virus and assignment of polypeptides to the 10 viral genes. J. Virol. **49**:572–578.
12. **Collins, P. L., and G. W. Wertz.** 1983. cDNA cloning and transcriptional mapping of nine polyadenylated RNAs encoded by the genome of human respiratory syncytial virus. Proc. Natl. Acad. Sci. U.S.A. **80**:3208–3212.
13. **Collins, P. L., and G. W. Wertz.** 1985. The envelope-associated 22K protein of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polytranscript. J. Virol. **54**:65–71.
14. **Devereux, J., P. Haeberli, and O. Smithies.** 1982. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
15. **Dickens, L. E., P. L. Collins, and G. W. Wertz.** 1984. Transcriptional mapping of human respiratory syncytial virus. J. Virol. **52**:364–369.
16. **Etkind, P. R., R. K. Cross, R. A. Lamb, D. C. Merz, and P. W. Choppin.** 1980. In vitro synthesis of structural and nonstructural proteins of Sendai and SV5 viruses. Virology **100**:22–33.
17. **Flamand, A., and J. F. Delagneau.** 1978. Transcriptional mapping of rabies virus in vivo. J. Virol. **28**:518–523.
18. **Glazier, K., R. Raghow, and D. W. Kingsbury.** 1977. Regulation of Sendai virus transcription: evidence for a single promoter in vivo. J. Virol. **21**:863–871.
19. **Gupta, K. C., and D. W. Kingsbury.** 1984. Complete sequence of the intergenic and mRNA start signals in the Sendai virus genome: homologies with the genome of vesicular stomatitis virus. Nucleic Acids Res. **12**:3829–3841.

20. **Hiebert, S. W., R. G. Paterson, and R. A. Lamb.** 1985. Hemagglutinin-neuraminidase protein of the paramyxovirus simian virus 5: nucleotide sequence of the mRNA predicts an N-terminal membrane anchor. J. Virol. **54**:1–6.
21. **Iverson, L. E., and J. K. Rose.** 1981. Localized attenuation and discontinuous synthesis during vesicular stomatitis virus transcription. Cell **23**:477–484.
22. **Iverson, L. E., and J. K. Rose.** 1982. Sequential synthesis of 5′-proximal vesicular stomatitis virus mRNA sequences. J. Virol. **44**:356–365.
23. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
24. **Lagrimini, L. M., S. T. Brentano, and J. E. Donelson.** 1984. A DNA sequence analysis package for the IBM personal computer. Nucleic Acids Res. **12**:605–614.
25. **Lamb, R. A., and P. W. Choppin.** 1977. The synthesis of Sendai virus polypeptides in infected cells. III. Phosphorylation of polypeptides. Virology **81**:382–397.
26. **Lamb, R. A., and P. W. Choppin.** 1981. Identification of a second protein (M2) encoded by RNA segment 7 of influenza virus. Virology **112**:729–737.
27. **Lamb, R. A., P. R. Etkind, and P. W. Choppin.** 1978. Evidence for a ninth influenza viral polypeptide. Virology **91**:60–78.
28. **Lamb, R. A., and C.-J. Lai.** 1982. Spliced and unspliced messenger RNAs synthesized from cloned influenza virus M DNA in an SV40 vector: expression of the influenza virus membrane protein (M1). Virology **123**:237–256.
29. **Lamb, R. A., B. W. J. Mahy, and P. W. Choppin.** 1976. The synthesis of Sendai virus polypeptides in infected cells. Virology **69**:116–131.
30. **Lamb, R. A., M. W. Shaw, D. J. Briedis, and P. W. Choppin.** 1983. The nucleotide sequence of the neuraminidase gene of influenza B virus reveals two overlapping reading frames, p. 77–86. *In* W. G. Laver (ed.), The origin of pandemic influenza viruses. Elsevier/North-Holland Publishing Co., New York.
31. **Lamb, R. A., S. L. Zebedee, and C. D. Richardson.** 1985. Influenza virus M2 protein is an integral membrane protein expressed on the infected-cell surface. Cell **40**:627–633.
32. **Lonberg, N., and W. Gilbert.** 1983. Primary structure of chicken muscle pyruvate kinase mRNA. Proc. Natl. Acad. Sci. U.S.A. **80**:3661–3665.
33. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
35. **Paterson, R. G., T. J. R. Harris, and R. A. Lamb.** 1984. Analysis and gene assignment of mRNAs of a paramyxovirus, simian virus 5. Virology **138**:310–323.
36. **Paterson, R. G., T. J. R. Harris, and R. A. Lamb.** 1984. Fusion protein of the paramyxovirus simian virus 5: nucleotide sequence of mRNA predicts a highly hydrophobic glycoprotein. Proc. Natl. Acad. Sci. U.S.A. **81**:6706–6710.
37. **Peluso, R. W., R. A. Lamb, and P. W. Choppin.** 1977. Polypeptide synthesis in simian virus 5-infected cells. J. Virol. **23**:177–187.
38. **Ricciardi, R. P., J. S. Miller, and B. E. Roberts.** 1979. Purification and mapping of specific mRNAs by hybridization-selection and cell-free translation. Proc. Natl. Acad. Sci. U.S.A. **76**:4927–4931.
39. **Rigby, P. W. J., M. Dreckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
40. **Rose, J. K.** 1980. Complete intergenic and flanking gene sequences from the genome of vesicular stomatitis virus. Cell **19**:415–421.
41. **Scheid, A., L. A. Caliguiri, R. W. Compans, and P. W. Choppin.** 1972. Isolation of paramyxovirus glycoproteins: association of both hemagglutinating and neuraminidase activities with the larger SV5 glycoprotein. Virology **50**:640–652.
42. **Scheid, A., and P. W. Choppin.** 1974. Identification of the biological activities of paramyxovirus glycoproteins. Activation of cell fusion, hemolysis, and infectivity by proteolytic cleavage of an inactive precursor protein of Sendai virus. Virology **57**:475–490.
43. **Shaw, M. W., and P. W. Choppin.** 1984. Studies on the synthesis of the influenza B virus NB glycoprotein. Virology **139**:178–184.
44. **Shaw, M. W., P. W. Choppin, and R. A. Lamb.** 1983. A previously unrecognized influenza B virus glycoprotein from a bicistronic mRNA that also encodes the viral neuraminidase. Proc. Natl. Acad. Sci. U.S.A. **80**:4879–4883.
45. **Swank, R. T., and K. D. Munkries.** 1971. Molecular weight analysis of oligopeptides by electrophoresis in polyacrylamide gel with sodium dodecyl sulfate. Anal. Biochem. **39**:462–477.