JOURNAL OF VIROLOGY, Sept. 1990, p. 4137–4145
0022-538X/90/094137-09$02.00/0

Vol. 64, No. 9

RNA Editing by G-Nucleotide Insertion in Mumps Virus P-Gene mRNA Transcripts

REAY G. PATERSON* AND ROBERT A. LAMB

Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208-3500

Received 21 March 1990/Accepted 23 May 1990

A guanine nucleotide insertion event has been shown to occur at a specific site within mumps virus P-gene mRNA transcripts. The region of the mRNA containing the site expected to be used for RNA editing and the complementary portion of the genomic RNA were cloned, and their nucleotide sequences were obtained. The genomic RNA was found to possess six C residues at the insertion site, whereas 63% of the P-gene-specific mRNA transcripts were found to have from two to five G residues inserted at this position in the RNA. An unedited mRNA was shown to encode the mumps virus cysteine-rich protein V, and mRNA transcripts containing two and four inserted G residues were translated to yield the mumps virus P and I proteins, respectively.

We and others have recently demonstrated that an RNA-editing event occurs to RNA transcripts derived from the P genes of the paramyxoviruses simian virus 5 (SV5), measles virus, and Sendai virus (3, 35, 36), resulting in the synthesis of more than one mRNA species that differ from each other by the number of G residues at a specific insertion point in the RNA. In each case, an mRNA is transcribed that is a faithful copy of the P gene and which can be translated in the case of SV5 to produce the cysteine-rich polypeptide V (35) and in the case of measles virus and Sendai virus to produce the P protein (3, 36). In addition, mRNA species that contain G residues not templated by the genomic RNA (vRNA) are transcribed that are translated to yield the P protein in the case of SV5 and the V polypeptide in the case of measles virus and Sendai virus.

Mumps virus is a member of the paramyxovirus family of RNA viruses. It has a single-stranded negative-sense RNA genome that is approximately 15,000 nucleotides in length. The order on the genomic RNA of the genes encoding the mumps virus proteins has been determined, and the results show that mumps virus is similar to SV5 in having an additional gene (SH) in the genome between those encoding the fusion protein and the hemagglutinin-neuraminidase (7–9, 15, 25) that is lacking from other paramyxoviruses so far examined. The nucleotide sequences for several mumps virus genes have been determined, and a comparison of the predicted amino acid sequences of the hemagglutinin-neuraminidase, fusion, and matrix proteins with the predicted amino acid sequences of these proteins from other paramyxoviruses indicates that mumps virus is most closely related to SV5 (9, 32, 37). There are conflicting data in the literature with respect to the sequence of the mumps virus P gene which has recently been published (6, 34). A significant variation was detected in a G-rich region (mRNA sense) that results in the addition of one amino acid in one of the published sequences (34), although both sequences were predicted to encode the mumps virus P protein. As a result of comparing the published mumps virus P-gene sequences (6, 34) with the sequence in the vicinity of the known insertion site in the SV5 P gene (35), we predicted that if the insertion of nontemplated G residues was occurring to a mumps virus P-gene transcript, it would occur in the G-rich region where the observed variation in the two mumps virus P gene sequences exists (26).

When the polypeptides synthesized in mumps virus-infected cells were analyzed, two low-molecular-weight polypeptides (M_rs of ~19,000 and ~28,000) were detected that were not present in purified virus (14). Peptide mapping experiments suggested that the two polypeptides had sequences in common with each other and also that they were related to the mumps virus P protein (14). In addition, both pulse-chase protocols and experiments in which the mumps virus-infected cell proteins were labeled in the presence of *N*-tosyl-L-phenylalanine chloromethyl ketone suggested that the polypeptides were primary translation products and not the result of proteolytic cleavage (28). It therefore became compelling, in view of the similarity between mumps virus and SV5 in other respects, to determine both whether an insertion event is occurring in the mumps virus P gene and, if so, which protein (P or a P-related protein) is encoded by an mRNA that is a faithful transcript of the P gene.

Here we report that in addition to P, two other proteins are encoded by the mumps virus P gene. The three proteins are amino coterminal and are translated from three mRNAs that differ from each other in the number of G residues inserted at the predicted site. One of the low-molecular-weight P-gene-specific proteins is a cysteine-rich protein analogous to protein V from SV5.

MATERIALS AND METHODS

Cells. Monolayer cultures of the TC7 clone of CV-1 cells were grown in Dulbecco modified Eagle medium (DMEM) supplemented with 10% NU-serum IV (Collaborative Research, Bedford, Mass.).

Virus infections, metabolic labeling of cells, and SDS-PAGE. The RW strain of mumps virus was grown in monolayer cultures of CV-1 cells in DMEM supplemented with 2% NU-serum IV and harvested at 4 days postinfection. The W3 strain of SV5 (4) was grown as described previously (27). Metabolic labeling of infected cell proteins was carried out as described previously (25), using either Tran[^{35}S]-label (ICN Radiochemicals, Irvine, Calif.), [^{35}S]cysteine, or

* Corresponding author.

[^{35}S]methionine (Amersham Corp., Arlington Heights, Ill.) in methionine- and cysteine-free DMEM. Protein samples were prepared for electrophoresis and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on 15% polyacrylamide gels as previously described (21).

Isolation of poly(A)-containing mRNAs. mRNAs were isolated as described previously (25).

Molecular cloning and nucleotide sequencing. Nucleotides 381 to 884 of mumps virus P-gene-specific mRNA species and also the complementary region of the mumps vRNA were cloned as follows. Briefly, first-strand synthesis was primed by using either an oligonucleotide complementary to nucleotides 866 to 884 in the mRNA or an oligonucleotide complementary to nucleotides 381 to 398 in the P gene. The single-stranded cDNA was then used directly as the template in a polymerase chain reaction (PCR) as described by Doherty and co-workers (5), using an excess of both the oligonucleotide primers mentioned above and *Taq* DNA polymerase (Amplitaq; Perkin-Elmer Cetus Corp., Norwalk, Conn.). Upon completion of 35 cycles of PCR, the resulting double-stranded cDNA was digested with the restriction endonucleases *Bam*HI and *Pvu*II and ligated into *Bam*HI- and *Pvu*II-digested pGEM-3 (Promega Corp., Madison, Wis.). Direct sequencing of the double-stranded plasmid DNA was carried out by the dideoxy-chain termination method, using the Klenow fragment of *Escherichia coli* DNA polymerase (30). Restriction endonucleases, T4 DNA ligase, and Klenow fragment were obtained from Bethesda Research Laboratories, Inc., Gaithersburg, Md. Avian myeloblastosis virus reverse transcriptase was obtained from Molecular Genetics Resources, Tampa, Fla. Oligonucleotides were synthesized by the Northwestern University Biotechnology Facility on a DNA synthesizer (model 380B; Applied Biosystems Inc., Foster City, Calif.).

Site-specific mutagenesis and plasmid construction. Plasmid pMP1 has been described previously (6) and was kindly provided by N. Elango. To construct pG3MP1, pMP1 was digested with *Nco*I, which cuts uniquely at a site one nucleotide upstream of the initiation codon for the mumps virus P protein. The *Nco*I sticky ends were filled in with T4 DNA polymerase (Bethesda Research Laboratories), and 8-mer *Xba*I linkers (New England BioLabs, Inc., Beverly, Mass.) were added. After linker ligation, pMP1 was digested with *Xba*I and *Pst*I, and the *Xba*I-*Pst*I insert was gel purified and ligated into *Xba*I- and *Pst*I-digested pGEM-3 (Promega Corp.). To construct the DNA templates for transcription of synthetic mRNAs encoding the mumps virus P and V proteins, the *Xba*I-*Pst*I insert from pG3MP1 was ligated into *Xba*I- and *Pst*I-digested pGEM-3Zf(+) (Promega Corp.). The resulting recombinant plasmid DNA was used as the template for oligonucleotide-directed mutagenesis as instructed by the manufacturer and as described previously (38), using R408 (29) as the helper bacteriophage. The mutations were verified by direct plasmid DNA sequencing as described above.

In vitro RNA synthesis. For transcription of the entire coding region, plasmid DNAs were linearized downstream of the T7 promoter and the cDNA insert, using *Hin*dIII. For synthesis of truncated forms of the mRNA, the DNA template was linearized by using either *Cla*I, *Pvu*II, or *Sau*I, which recognize sites within the coding region of the cDNA. In vitro synthesis of mRNA was carried out as previously described (16), and 1 μg of RNA was translated in vitro by using a rabbit reticulocyte lysate as described below. T7 DNA-dependent RNA polymerase was obtained from Bethesda Research Laboratories, RNasin and RQ DNase were obtained from Promega Corp., and ^{7m}G (5′)ppp(5′) G (sodium salt) was obtained from Pharmacia-LKB Biotechnology Inc., Piscataway, N.J.

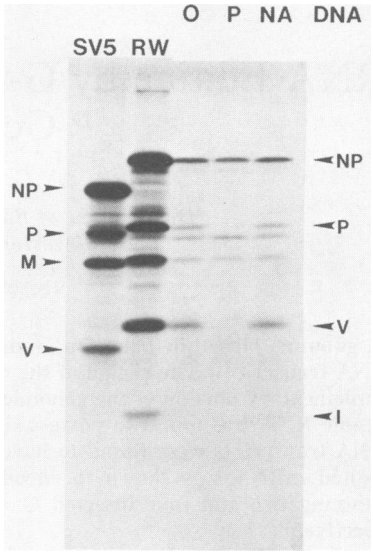

FIG. 1. Hybrid-arrested translation of mumps virus mRNAs, using pMP1 cDNA. Hybrid-arrested translation and analysis of the in vitro-translated products on 15% SDS-PAGE was done as described in Materials and Methods. Lanes: SV5, in vitro translation of SV5 poly(A)-containing mRNA; RW, in vitro translation of mumps virus mRNA; 0, control with no added DNA in the hybridization with mumps mRNA; P, in vitro translation of mumps virus mRNA hybridized to pMP1 insert DNA; NA, in vitro translation of mumps virus mRNA after hybridization to an irrelevant control DNA. The designations for the SV5 polypeptides are given on the left, and the designations for the mumps virus polypeptides are given on the right.

In vitro translation and hybrid-arrested translation of mRNAs. mRNAs were translated in vitro by using a micrococcal nuclease-treated rabbit reticulocyte lysate (Promega Corp.) as instructed by the manufacturer. The in vitro-synthesized products were labeled with [^{35}S]methionine or [^{35}S]cysteine. One-fifth volume of each translation reaction was analyzed by SDS-PAGE as described above. Hybrid-arrested translation was carried out as described previously (19, 20), using approximately 2 μg of double-stranded cDNA and the poly(A)$^+$ mRNA (0.5 μg) obtained from a 10-cm plate of mumps virus-infected CV-1 cells.

RESULTS

Identification of mumps virus P-gene-specific protein products. To identify mumps virus P-gene-encoded proteins, insert DNA from the plasmid pMP1 (6) was hybridized to poly(A)-containing RNA isolated from mumps virus-infected CV-1 cells. The RNAs were then used to program a rabbit reticulocyte lysate cell-free translation system as described in Materials and Methods. pMP1 insert DNA specifically prevented translation of the P polypeptide as well as two other lower-molecular-weight polypeptides (M_rs of ~28,000 and ~19,000), which we have designated V and I, respectively, indicating that polypeptides P, V, and I are encoded by the P gene (Fig. 1, lane P). The products of in vitro translation of SV5 mRNAs are shown for comparison; the mumps virus protein designated V had a slower electropho-

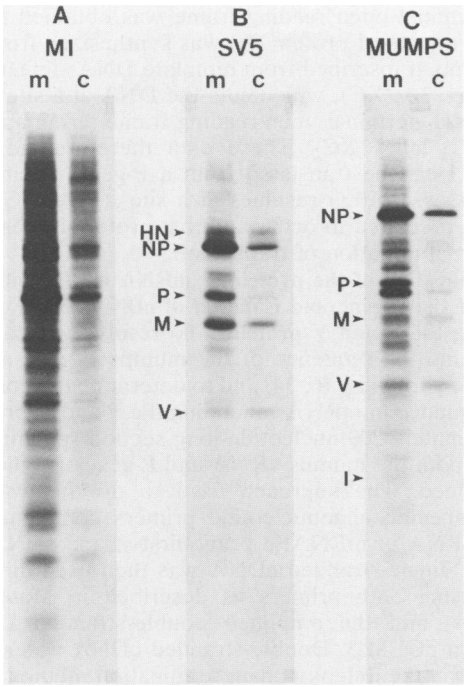

FIG. 2. Selective radiolabeling of mumps virus-infected cell proteins, using [^{35}S]methionine or [^{35}S]cysteine. Mumps virus-infected, SV5-infected, and mock-infected (MI) CV-1 cells were labeled at 18 h postinfection for 1.5 h with either 125 μCi of [^{35}S]methionine per ml or 250 μCi of [^{35}S]cysteine per ml, the cells were lysed, and the labeled proteins were analyzed by SDS-PAGE on 15% polyacylamide gels as described in Materials and Methods. (A) Mock-infected cell lysate; (B) SV5-infected cell lysate; (C) mumps virus-infected cell lysate. Lanes: m, proteins labeled with [^{35}S]methionine; c, proteins labeled with [^{35}S]cysteine. HN, Hemagglutinin-neuraminidase; NP, nucleoprotein; P, phosphoprotein; M, matrix protein; V, protein V; I, protein I.

retic mobility than did protein V from SV5 (Fig. 1, lanes SV5 and RW).

Comparison of mumps virus-infected cell polypeptides metabolically labeled with [^{35}S]methionine and [^{35}S]cysteine. To examine the relative ability with which mumps virus P-gene-encoded proteins could be selectively radiolabeled, mumps virus-infected cells were labeled with either [^{35}S]methionine or [^{35}S]cysteine, and the labeled proteins were analyzed by SDS-PAGE. For comparison, a parallel experiment was carried out with SV5-infected cells, since the SV5 P-gene-encoded protein V is known to be a cysteine-rich protein (35). Although the mumps virus P, V, and I proteins were all readily labeled with methionine (Fig. 2C, lane m), of the P-gene-encoded proteins, only V was easily detected when labeled with [^{35}S]cysteine (lane c). This observation is in agreement with results obtained with SV5 (Fig. 2B), showing that protein V was easily detected when labeled with methionine or cysteine and the P protein, although labeled well with methionine, was poorly labeled with cysteine (35). These data therefore suggest that the mumps virus polypeptide designated as V contains the cysteine-rich open reading frame previously identified in the mumps virus P-gene nucleotide sequence (35) that is homologous to the COOH-terminal region of the SV5 protein V.

Mapping of the coding region for protein I. Plasmid pMP1 was derived from a cDNA library constructed by using mRNA isolated from mumps virus-infected Vero cells (8) and was shown to hybrid select an mRNA that could be translated in vitro to yield the P protein (6). The nucleotide sequence of pMP1 (6) differs from both of the published sequences of the mumps virus P gene (6, 34) over the G-rich region where a G insertion event was predicted to occur (26). When the nucleotide sequence of pMP1 is translated, it is predicted to contain two open reading frames, each having the capacity to encode a protein of 170 amino acids. Neither of these open reading frames contains the cysteine-rich region, indicating that they do not encode the mumps virus protein V, nor are they large enough to encode the P protein. However, a reading frame of 170 amino acids is sufficient to encode protein I, which has an apparent molecular weight on SDS-PAGE of approximately 19,000.

To determine which of the two reading frames gives rise to protein I, the insert from pMP1 was adapted for expression by removing the G/C tail from the 5′ end of the DNA and subcloning the modified insert into pGEM-3 such that synthetic mRNA could be transcribed by using T7 RNA polymerase. Both full-length and truncated transcripts were prepared as described in Materials and Methods and are shown in schematic form in Fig. 3A. The in vitro-transcribed mRNAs were translated in vitro by using a rabbit reticulocyte lysate system, and the resulting proteins were analyzed by SDS-PAGE. The [^{35}S]methionine- or [^{35}S]cysteine-labeled proteins synthesized in vitro from full-length pMP1 RNA transcripts are shown in Fig. 3B. When the in vitro-synthesized proteins were labeled with [^{35}S]methionine, the major translation product (met panel, lane XH) had an electrophoretic mobility indistinguishable from that of protein I translated in vitro by using poly(A)-containing RNA from mumps virus-infected cells (met panel, lane A^+) or synthesized in mumps virus-infected cells (in vivo panel, lane met). The other, less abundant polypeptides species observed when the full-length pMP1 RNA transcripts were translated in vitro by using methionine as the radiolabel are internal initiation products (the polypeptide with an electrophoretic mobility intermediate between that of proteins V and I is thought to have initiated at the first AUG codon in the COOH-terminal open reading frame) (see below). As predicted from the sequence of pMP1, translation of pMP1-derived in vitro transcripts did not give rise to in vitro translation products that could be labeled with [^{35}S]cysteine (cys panel, lane XH), although the mumps virus-encoded NP, M, and V proteins were readily labeled with [^{35}S]cysteine when poly(A)-containing RNA was translated in vitro (cys panel, lane A^+). Further evidence that protein I is encoded by the NH_2-terminal open reading frame was obtained from in vitro translation of truncated pMP1-derived RNA transcripts (Fig. 3C). When the DNA template for transcription was linearized with *Pvu*II or *Sau*I (Fig. 3A, XPv and XS, respectively), which cut the insert at unique sites downstream of the predicted termination codon in the NH_2-terminal reading frame, translation of protein I was unaffected (Fig. 3C, lanes XPv and XS). However, translation of the minor higher-molecular-weight band (Fig. 3C, lane XH, *) was eliminated when the template was linearized with either *Pvu*II or *Sau*I, as was the smallest of the three minor polypeptide species (compare lanes XH, XPv, and XS in Fig. 3C), which suggests that they arose by internal initiation of translation in the COOH-terminal reading frame. The minor polypeptide that migrated just below protein I on SDS-PAGE was translated from the truncated transcripts (Fig. 3C, lanes XH, XPv, and XS), indicating that it was an internal initiation product from the NH_2-terminal reading frame. Further evidence that protein I is derived from the

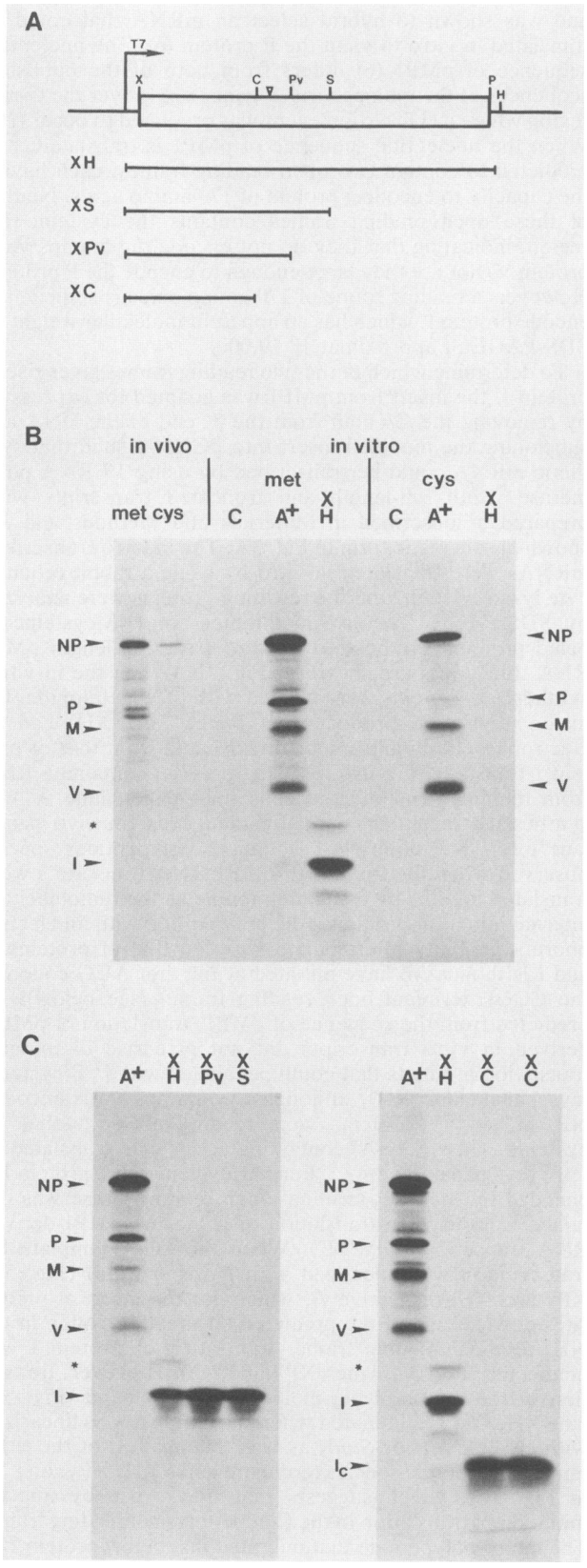

FIG. 3. Mapping the region of the mumps virus P gene encoding protein I, using bacteriophage T7 DNA-dependent RNA polymerase runoff transcripts and in vitro translation. The pMP1 insert DNA modified as described in Materials and Methods was placed under the control of the T7 promoter in the plasmid pGEM-3. The template DNA was linearized downstream of the T7 promoter and insert DNA by endonuclease digestion with *Hin*dIII; for the synthesis of truncated transcripts, the DNA was digested with either *Sau*I, *Pvu*II, or *Cla*I, which recognize sites within the protein-coding region. The in vitro-synthesized runoff transcripts were translated in vitro, using a rabbit reticulocyte lysate and either [^{35}S]methionine or [^{35}S]cysteine as the radioactive precursor. (A) Schematic representation of the MP1 insert showing the positions of the restriction endonuclease recognition sites used to generate the templates for the synthesis of runoff transcripts, with the transcripts shown in diagrammatic form below. Restriction sites: X, *Xba*I; C, *Cla*I; Pv, *Pvu*II; S, *Sau*I; P, *Pst*I; H, *Hin*dIII. T7, Site of initiation of transcription by T7 RNA polymerase; ∇, site of G insertion. The MP1 insert is represented by the open box. In vitro-synthesized transcripts are represented by the horizontal lines and are named according to the endonuclease used to linearize the template DNA; e.g., XS, represents transcript produced from a DNA template linearized by using *Sau*I. (B) Proteins synthesized in vitro in rabbit reticulocyte lysates programmed with full-length MP1 transcripts. The proteins were labeled with either [^{35}S]methionine (met) or [^{35}S]cysteine (cys) and analyzed on a 15% SDS-polyacrylamide gel. in vivo, Metabolically labeled proteins synthesized in mumps virus-infected CV-1 cells; in vitro, proteins synthesized in rabbit reticulocyle lysates. Lanes: C, no RNA control translation; A$^+$, translation of poly(A)-containing mRNA from mumps virus-infected CV-1 cells; XH, translation of full-length MP1 T7 RNA polymerase transcript. NP, Nucleoprotein; P, phosphoprotein; M, matrix protein; V, protein V; *, internal initiation product derived from the C-terminal open reading frame in MP1; I, protein I. (C) [^{35}S] methionine-labeled proteins synthesized in vitro in rabbit reticulocyte lysates programmed with truncated MP1 transcripts. Lanes: A$^+$, translation of poly(A)-containing mRNA from mumps virus-infected CV-1 cells; XH, translation of the full-length MP1 T7 transcript; XPv, translation of RNA transcribed from templates linearized with *Pvu*II; XS, translation of RNA transcribed from templates linearized with *Sau*I; XC, translation of RNA transcribed from templates linearized with *Cla*I. I_c truncated form of protein I. Other designations as given above.

NH_2-terminal open reading frame was obtained by finding that a truncated I protein (I_c) was synthesized from in vitro transcripts transcribed from template DNAs linearized with *Cla*I (Fig. 3A, XC), which cuts the DNA at a site upstream of the NH_2-terminal open reading frame termination codon (Fig. 3C, lanes XC). These data therefore indicate that protein I can be translated from a P-gene-specific mRNA that possesses 10 G residues at a site where a G insertion event is predicted to occur and that protein I does not arise by internal initiation of translation.

Identification of the protein V mRNA as the faithful transcript of the P gene and isolation of cDNA copies of mRNAs encoding the P and V proteins. To resolve the controversy concerning the sequence of the mumps P gene across the region of variability (6, 34) and to determine the complementary sequence in mRNAs encoding the P and V proteins, an approximately 200-nucleotide-long section spanning this region of both the mumps vRNA and P-gene-specific mRNAs was cloned. The approach used to do this was to use P-gene-specific oligonucleotide primers complementary to either vRNA or mRNA to prime first-strand cDNA synthesis. The single-stranded cDNA was then used directly in a PCR, using both primers as described in Materials and Methods, and the amplified double-stranded DNA was cloned in pGEM-3. Double-stranded cDNA was sequenced directly by the dideoxy-chain termination method (30). The nucleotide sequence over the region of interest was obtained for 70 vRNA-specific clones and 54 mRNA-specific clones.

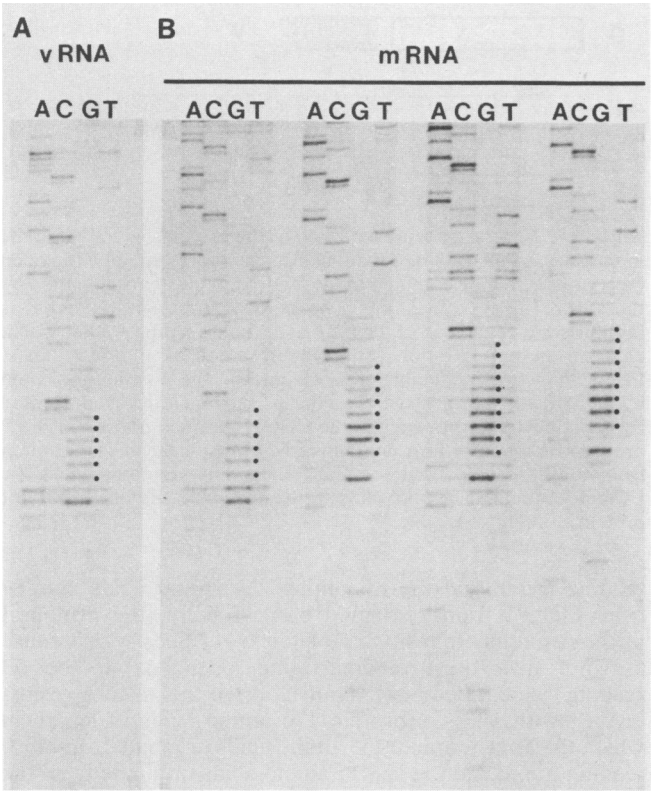

FIG. 4. Sequences of four mRNA cDNA clones in comparison with the sequence of a vRNA cDNA clone. The nucleotide sequences of a vRNA cDNA clone and four mRNA cDNA clones in the region of nucleotides 512 to 568 (numbered according to reference 6) are shown to illustrate the different number of G residues at the predicted insertion site. Sequencing of double-stranded plasmid DNA was carried out by the dideoxy-chain termination method (30). (A) Sequence of the vRNA cDNA clone. (B) Sequences of (from left to right) a V cDNA (6 G residues), a P cDNA (8 Gs), an I cDNA (10 G residues), and a cDNA with 9 G residues at predicted site. The G residues in question are indicated by dots.

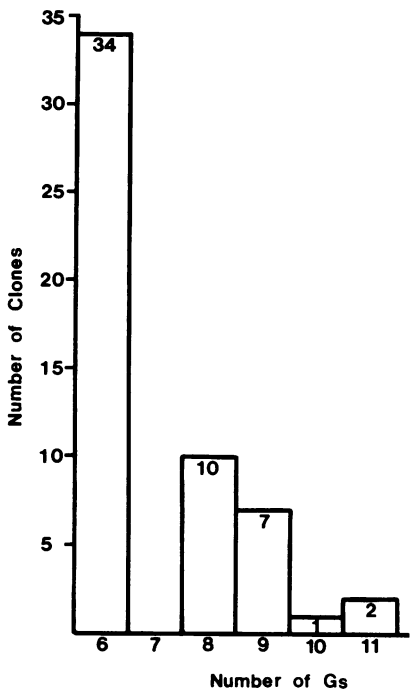

FIG. 5. Distribution of the mRNA cDNA clones according to the number of G residues at the predicted insertion site. A total of 54 mRNA cDNA clones were sequenced over the region of interest, and the distribution of the clones with respect to the number of G residues at the insertion site is shown in the form of a histogram. In this figure, 6 G residues corresponds to no G insertion. The number of clones in each group is indicated by the number inside the relevant column. No clones with seven G residues (i.e., one inserted G) were detected.

The sequences of the vRNA- and mRNA-derived clones differed from each other only by the number of G residues present between nucleotides 530 and 537 (numbering system as in reference 6) (Fig. 4). In 69 of the 70 vRNA-derived cDNAs there were six G residues (mRNA sense) in this region, whereas in one clone there were seven G residues. Thus, from these results we can conclude that the mumps P gene contains six C residues in the vRNA at this site. The most likely explanation for the existence of the single clone that was found to contain seven G residues is that it arose as a result of either an error introduced by reverse transcriptase during synthesis of the first strand of cDNA or by *Taq* polymerase during the PCR amplification. The majority of the mRNA-derived cDNAs (34 of 54) contained six G residues and therefore were direct copies of the vRNA template; of the remaining 20 clones, 10 possessed 8 G residues, 7 had 9, 1 had 10, and 2 had 11 (Fig. 5). These data indicate that several P-gene-specific mRNAs exist in mumps virus-infected cells that differ from each other by the presence of different numbers of nontemplated G residues at a specific point in the mRNA. Examination of the nucleotide sequence of the mRNA species that contains six G residues shows that it has the potential to encode a protein of 224 amino acids which would possess at its COOH terminus the cysteine-rich domain homologous to that found in protein V of SV5 (35). Therefore, as in the case of SV5, the mRNA species that is a faithful copy of the P gene encodes the mumps virus protein V. The predicted amino acid sequence of the second most common class of mRNA, those with two nontemplated residues, contains 391 amino acids, and it therefore has the potential to encode the mumps virus P protein, whereas an mRNA containing four nontemplated G residues has the ability to encode protein I. The cDNA clones that have nine G residues at the insertion site would be predicted to encode a protein V containing an additional glycine residue. Whether such a protein is biologically active is not known.

Expression of the mumps P and V proteins from in vitro-transcribed mRNA. To provide further evidence that the P, V, and I proteins are translated from mRNAs that differ from each other only by the number of G residues at the specific insertion point in the RNA, the P and V proteins were translated in vitro by using synthetic mRNAs. DNA templates containing either six or eight G residues in the region of interest were prepared by oligonucleotide-directed mutagenesis of pMP1 DNA as described in Materials and Methods. Full-length RNA transcripts were prepared by using DNA templates linearized with *Hin*dIII (Fig. 3A) and were translated in vitro by using a rabbit reticulocyte lysate to yield either [^{35}S]methionine- or [^{35}S]cysteine-labeled proteins (Fig. 6). When a transcript containing two nontemplated residues (eight G's) was translated in vitro in the presence of either [^{35}S]methionine or [^{35}S]cysteine (Fig. 6, lanes P), the major protein synthesized had an electropho-

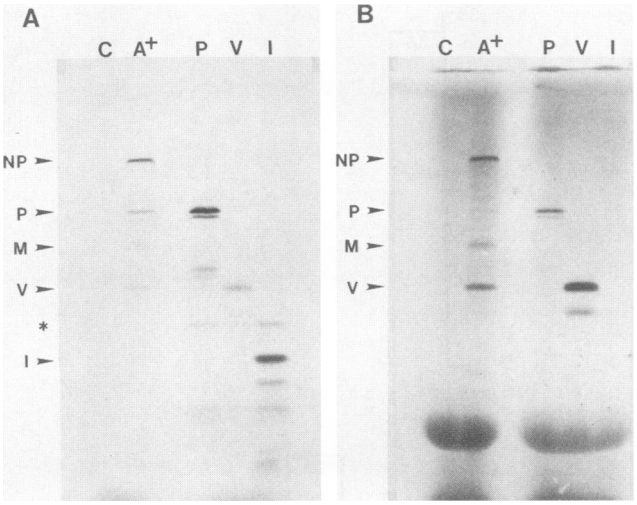

FIG. 6. Expression of the P and V proteins from in vitro-transcribed mRNAs. P and V cDNA clones were constructed by subcloning the I cDNA clone MP1 (containing 10 G residues at the predicted insertion site) into pGEM-3Zf(+), followed by site-specific mutagenesis to delete either 2 or 4 G residues from the G insertion site; RNA was transcribed with T7 RNA polymerase from the P and V cDNA clones (8 and 6 G residues, respectively) and translated in vitro, using a rabbit reticulocyte lysate. (A) In vitro translation products labeled with [^{35}S]methionine. (B) [^{35}S]cysteine-labeled in vitro-synthesized proteins. In vitro-translated RNAs were as follows: lane C, no RNA control; lane A$^+$, poly(A)-containing mRNAs from mumps virus-infected CV-1 cells; lane P, P mRNA; lane V, V mRNA; lane I, I mRNA. NP, Nucleoprotein; P, phosphoprotein; M, matrix protein; V, protein V; *, internal initiation product from the COOH-terminal open reading frame that is translated to yield the COOH-terminal portion of the P protein; I, I protein.

retic mobility indistinguishable from that of the P protein translated by using poly(A)-containing mRNAs from mumps virus-infected cells (Fig. 6, lanes A$^+$). Further evidence in support of the identification of this species as the mumps virus P protein is that it can be immunoprecipitated by a monoclonal antibody that recognizes the mumps virus P protein synthesized in infected cells (data not shown). As expected, translation of an mRNA with six G residues yielded protein V (Fig. 6, lanes V). Several minor bands were observed as seen previously (and shown here for comparison in lanes I of Fig. 6) when protein I was synthesized in vitro, and they most likely arose from initiation at internal methionine codons, a common finding when synthetic RNAs are translated in vitro (18, 35).

The accumulated data suggest a coding strategy for the mumps virus P gene as depicted schematically in Fig. 7. The P, V, and I proteins are NH_2 coterminal and have 155 amino acids in common up until the G insertion site (depicted in Fig. 7 by the NH_2-terminal open box, with the G's indicated by the vertical line). Protein V is translated in the 0 reading frame (open box) from a mRNA that is an exact copy of the P gene; it is 224 amino acids long and contains the conserved COOH-terminal cysteine-rich domain (indicated by the wavy lines). The mumps virus P protein is translated from an mRNA that contains two nontemplated nucleotides that cause translation to switch from the 0 reading frame to the +1 reading frame at the insertion site (open arrowhead). Thus, the P protein is predicted to contain 391 amino acids, the NH_2-terminal 155 amino acids originating from the 0

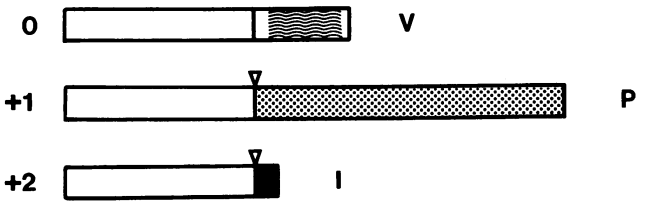

FIG. 7. Schematic representation of the proteins encoded by the mumps virus P gene. The 0 reading frame is represented by the open box, with the region common to the P, V, and I proteins delimited by the vertical line. Translation of protein V from a mRNA that is faithful copy of the vRNA continues in the 0 reading frame, which encodes the cysteine-rich domain represented by the wavy lines. Insertion of two nontemplated G residues at the position indicated by the arrow results in a switch to the +1 reading frame (represented by the stippled box) to encode the COOH-terminal portion of the P protein. Insertion of four nontemplated G residues causes a switch from the 0 reading frame to the +2 reading frame to yield the COOH-terminal 15 amino acids (represented by the black box) of protein I.

reading frame and the remaining 236 amino acids coming from the +1 frame (stippled box). Finally, the protein I mRNA contains four nontemplated G residues which cause a switch from the 0 reading frame (open box) to the +2 reading frame (filled box). Translation of this mRNA results in the synthesis of a protein 170 amino acids in length in which the NH_2-terminal 155 amino acids originate from the 0 reading frame and the last 15 amino acids originate from the +2 reading frame.

DISCUSSION

The RNA-editing phenomenon involving the addition of nontemplated G residues at a specific site in a paramyxovirus mRNA was first identified as the mechanism by which the P gene of SV5 encodes the amino-coterminal P and V proteins (35). In the case of SV5, protein V is translated from an unedited mRNA, whereas the P protein is translated from an mRNA that possesses two G residues that are not templated by the virion RNA. Upon examination of the available nucleotide sequences of the P genes of other paramyxoviruses, including Newcastle disease virus, Sendai virus, parainfluenza virus type 3, measles virus, and canine distemper virus (1, 10, 11, 22, 23, 31, 33), a sequence was identified (in the mRNA sense) consisting of a run of G residues preceded by a sequence bearing some resemblance to a polyadenylation signal (3, 26). Insertion of G residues at this sequence would permit access to a reading frame encoding a cysteine-rich domain in each case homologous to the COOH-terminal domain in protein V of SV5 (35). These observations led to the prediction (35) that a protein homologous to protein V would be encoded by all paramyxoviruses and also that it would be translated from an mRNA containing nontemplated G residues inserted at the predicted insertion site (3, 26). The observations with SV5 have been extended, since RNA editing of RNA transcripts has been found to occur in the case of measles virus and Sendai virus P-gene transcripts in which one G residue is inserted at the predicted site to yield the protein V mRNA (3, 36).

The paramyxovirus P protein together with the L protein is thought to form the paramyxovirus transcriptase (2, 12) and is therefore probably an essential protein. Interestingly, the cysteine-rich domain of protein V is more highly conserved between different paramyxoviruses than is the amino acid sequence of the P protein encoded in a different reading

frame by the same nucleotides, which suggests that protein V has a necessary function (35). An attractive role for protein V is as a factor involved in transcription or replication (or both) of the vRNA, since the cysteine-rich region is reminiscent of a zinc finger domain, a motif that has been identified in nucleic acid binding regulatory proteins (for reviews, see references 17 and 24).

Prompted by the conflicting data reported in the literature concerning the sequence of the mumps virus P gene (6, 34), where the observed variation occurred in the region of the predicted insertion site (3, 26), and by the fact that two proteins had been detected in mumps virus-infected cells that appeared to be related to the P protein (14, 28), we decided to examine the possibility that RNA editing existed in the case of mumps virus. The relevant region of the mumps virus genomic RNA and P-gene-specific mRNA species was cloned. The results obtained upon sequencing 70 vRNA-specific clones indicated that the vRNA contains six G residues (mRNA sense) as reported by Elango and co-workers (6) and not eight G residues (34), but it also contains a nucleotide (G in the mRNA sense, nine nucleotides after the insertion site) which is present in the sequence reported by Takeuchi and co-workers (34). Thus, the corrected sequence of the P gene across this region (nucleotides 527 to 556, numbered according to reference 6) is 5′-AAGA **GGGGGG**CCGGGAGCGCGTGCTCAAGG-3′. The net result of this sequence in and surrounding the insertion site of mumps virus is that an exact transcript of the P gene encodes the mumps protein V and not the P protein as reported previously (6, 34). The majority of the mRNA-specific cDNA clones sequenced (34 of 54; Fig. 5) possessed the same sequence as that obtained for the vRNA-specific cDNA clones (six G's). The remaining 20 mRNA-specific clones differed from the vRNA sequence only by the number of G residues at the predicted insertion site (Fig. 4 and 5). The number of inserted G residues in the mRNA-specific cDNA clones sequenced varied from two to five, the most common class containing two inserted G's. If the nucleotide sequence of clones containing two nontemplated G residues is translated, it is found to yield a reading frame large enough to encode the mumps virus P protein. Further evidence that the mumps protein V is translated from an mRNA that is an exact copy of the vRNA, whereas the P protein is derived from an edited mRNA containing two nontemplated nucleotides and a third P-gene-specified protein (I) is translated from a mRNA having four nontemplated nucleotides inserted at the predicted site, was obtained by in vitro translating synthetic RNA transcripts containing either 6, 8, or 10 G residues at the specific position in the RNA. The relative abundance of [^{35}S]methionine- or [^{35}S]cysteine-labeled P, V, and I proteins seen in mumps virus-infected cells approximately correlates with the ratio of mRNA species observed here, taking into account the number of methionine and cysteine residues in each protein (3 methionines and 8 cysteines in V; 12 methionines and 1 cysteine in P; 3 methionines and 0 cysteine in I).

Of the 70 vRNA cDNA clones sequenced, one possessed seven G residues rather than six. The simplest explanation for this finding is that the additional G residue is the result of an error introduced during the synthesis of the first strand of cDNA by reverse transcriptase or during amplification of the cDNA using *Taq* DNA polymerase; there is no biological evidence for the involvement of more than one form of virus genome in the paramyxovirus infectious cycle. One interesting point concerning the vRNA clone containing seven G residues is that this extra nucleotide is the only difference that was detected among the 100 nucleotides sequenced in each of 70 vRNA-specific clones and in the mRNA cDNAs. Apart from the variation in the number of G residues present in the insertion site, there were no other nucleotide differences detected in the 5,400 nucleotides sequenced. This finding suggests that there is some unusual feature in this particular region of the vRNA. The introduction of error during the cloning and amplification of the mRNA-specific cDNA clones is unlikely to be a valid explanation for the variability observed in the run of G's because of the high frequency of clones possessing more than six G residues (37%) and the demonstration that the mumps virus P, V, and I proteins are translated from unique mRNAs differing only in the number of G's at the insertion site.

The sizes of the open reading frames that have been shown here to give rise to the V and I proteins and the observed relative electrophoretic mobility on SDS-PAGE are in good agreement with the observed sizes of the low-molecular-weight nonstructural polypeptides, M_rs of 23,000 to 28,000 and 17,000 to 19,000, detected previously in mumps virus-infected cells (14, 28). In addition, the finding that the P, V, and I polypeptides have 155 amino acids in common is in good agreement with previous results of peptide mapping experiments which indicated that the three polypeptides shared amino acids (14).

The mechanism responsible for the insertion of the nontemplated G residues in mRNAs derived from paramyxovirus P genes remains elusive. It was suggested that it could be a function of the virus-encoded transcriptase complex "stuttering" while copying this region of the genomic RNA in a process analogous to that by which the virus transcriptase polyadenylates the virus-specific mRNAs (35). This model was attractive because just upstream of the known or predicted insertion sites a sequence could be identified that resembled a putative polyadenylation signal (26). Support for this hypothetical mechanism has come recently from work carried out with Sendai virus, in which it has been possible to show that in addition to occurring in vivo in infected cells, the G insertion event also takes place in vitro, using purified virions as the source of polymerase and template (36). Evidence was also provided to suggest that the G insertion was occurring cotranscriptionally by coexpressing Sendai virus-transcribed P-gene-specific mRNAs and P-gene-specific mRNAs expressed by using recombinant vaccinia virus. Only mRNAs transcribed by Sendai virus were found in the edited form; those mRNAs derived from the recombinant vaccinia virus were unmodified (36). Taken together, these results suggest that Sendai virus proteins mediate the RNA-editing process and that they cannot act posttranscriptionally in *trans* on P-gene-specific mRNAs expressed by vaccinia virus. The mechanism of G insertion is relatively specific in that the number of residues inserted is quite small. In SV5, only edited RNAs with two inserted G residues were detected, although the 22 clones examined is perhaps too small a number in which to have detected less frequently occurring insertion events (35), whereas of the 19 measles virus P-specific cDNA clones examined, 7 were derived from unedited mRNAs, 10 had one G inserted, and the remaining 2 clones had three G residues inserted (3). Although more variable in the cases of Sendai virus and mumps virus, where from one to eight nontemplated G residues have been detected in P-gene-specific mRNA cDNA clones (36; this report), this number is still much smaller than the number of A residues added to negative-strand virus mRNAs as a result of polyadenylation (13). The specificity of the G insertion event is therefore thought to be

dependent on more than the simple nucleotide sequence surrounding the insertion site. It has been suggested that the secondary structure in this region of the vRNA might be involved, since a stable stem-loop structure can be predicted in this area for the SV5 genomic RNA (35). The evidence is less compelling for the existence of such a structure in the vRNA of Sendai virus and measles virus (36), although the observation of dideoxy-chain termination sequencing artifacts in this region of P-gene-specific cDNA clones provides some indication that at least in the DNA there is some secondary structure (35, 36).

Thus, although there is increasing evidence to suggest that the G insertion event is a virus-encoded cotranscriptional mechanism (36) and that this form of RNA editing may be a common feature in paramyxovirus P-gene-derived mRNAs, the factors responsible for determining the specificity remain to be defined. In addition, the function of the V protein of paramyxoviruses has yet to be elucidated despite the fact that its conservation among at least the four viruses examined to date suggests that it plays a biologically important role in the paramyxovirus life cycle.

ACKNOWLEDGMENTS

We thank Margaret Shaughnessy for excellent technical assistance and Dennis Wong for performing some of the nucleotide sequencing as part of a C99 project for undergraduates at Northwestern University. We are very grateful to N. Elango, currently at the National Research Council of Canada Biotechnology Center, Montreal, Canada, for making available the mumps virus cDNA clone pMP1 and N. Waxham for the RW strain of mumps virus.

This research was supported by Public Health Service research grant AI-23173 from the National Institute of Allergy and Infectious Diseases.

LITERATURE CITED

1. **Bellini, W. J., G. Englund, S. Rozenblatt, H. Arnheiter, and C. D. Richardson.** 1985. Measles virus P gene codes for two proteins. J. Virol. **53**:908–919.
2. **Buetti, E., and P. W. Choppin.** 1977. The transcriptase complex of the paramyxovirus SV5. Virology **82**:493–508.
3. **Cattaneo, R., K. Kaelin, K. Baczko, and M. A. Billeter.** 1989. Measles virus editing provides an additional cysteine-rich protein. Cell **56**:759–764.
4. **Choppin, P. W.** 1964. Multiplication of a myxovirus (SV5) with minimal cytopathic effects and without interference. Virology **23**:224–233.
5. **Doherty, P. J., M. Huesca-Contreras, H. M. Dosch, and S. Pan.** 1989. Rapid amplification of complementary DNA from small amounts of unfractionated RNA. Anal. Biochem. **177**:7–10.
6. **Elango, N., J. Kovamees, and E. Norrby.** 1989. Sequence analysis of the mumps virus mRNA encoding the P protein. Virology **169**:62–67.
7. **Elango, N., J. Kovamees, T. M. Varsanyi, and E. Norrby.** 1989. mRNA sequence and deduced amino acid sequence of the mumps virus small hydrophobic protein gene. J. Virol. **63**: 1413–1415.
8. **Elango, N., T. M. Varsanyi, J. Kovamees, and E. Norrby.** 1988. Molecular cloning and characterisation of six genes, determination of gene order and intergenic sequences and leader sequence of mumps virus. J. Gen. Virol. **69**:2893–2900.
9. **Elliott, G. D., M. A. Afzal, S. J. Martin, and B. K. Rima.** 1989. Nucleotide sequence of the matrix, fusion and putative SH protein genes of mumps virus and their deduced amino acid sequences. Virus Res. **12**:61–76.
10. **Galinski, M. S., M. A. Mink, D. M. Lambert, S. L. Wechsler, and M. W. Pons.** 1986. Molecular cloning and sequence analysis of the human parainfluenza 3 virus mRNA encoding the P and C proteins. Virology **155**:46–60.
11. **Giorgi, C., B. M. Blumberg, and D. Kilakofsky.** 1983. Sendai virus contains overlapping genes expressed from a single mRNA. Cell **35**:829–836.
12. **Hamaguchi, M., T. Yoshida, K. Nishikawa, H. Naruse, and Y. Nagai.** 1983. Transcriptive complex of Newcastle disease virus. I. Both L and P proteins are required to constitute an active complex. Virology **128**:105–117.
13. **Herman, R., S. Adler, R. A. Lazzarini, R. J. Colonno, A. K. Banerjee, and H. Westphal.** 1978. Mapping of vesicular stomatitis virus messenger RNA and intervening polyadenylate sequences by electron microscopy. Cell **15**:587–596.
14. **Herrler, G., and R. W. Compans.** 1982. Synthesis of mumps virus polypeptides in infected Vero cells. Virology **119**:430–438.
15. **Hiebert, S. W., R. G. Paterson, and R. A. Lamb.** 1985. Identification and predicted sequence of a previously unrecognized small hydrophobic protein, SH, of the paramyxovirus simian virus 5. J. Virol. **55**:744–751.
16. **Hull, J. D., R. Gilmore, and R. A. Lamb.** 1988. Integration of a small integral membrane protein, M_2, of influenza virus into the endoplasmic reticulum: analysis of the internal signal-anchor domain of a protein with an ectoplasmic NH_2 terminus. J. Cell Biol. **106**:1489–1498.
17. **Klug, A., and D. Rhodes.** 1987. 'Zinc fingers': a novel protein motif for nucleic acid recognition. Trends Biochem. Sci. **12**: 464–469.
18. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell Biol. **108**:229–241.
19. **Lamb, R. A., and P. W. Choppin.** 1981. Identification of a second protein (M_2) encoded by RNA segment 7 of influenza virus. Virology **112**:729–737.
20. **Lamb, R. A., P. W. Choppin, R. M. Chanock, and C.-J. Lai.** 1980. Mapping of the two overlapping genes for polypeptides NS_1 and NS_2 on RNA segment 8 of influenza virus genome. Proc. Natl. Acad. Sci. USA **77**:1857–1861.
21. **Lamb, R. A., P. R. Etkind, and P. W. Choppin.** 1978. Evidence for a ninth influenza viral polypeptide. Virology **91**:60–78.
22. **Luk, D., A. Sanchez, and A. K. Banerjee.** 1986. Messenger RNA encoding the phosphoprotein (P) gene of human parainfluenza virus 3 is bicistronic. Virology **153**:318–325.
23. **McGinnes, L., C. McQuain, and T. Morrison.** 1988. The P protein and the nonstructural 38K and 29K proteins of Newcastle disease virus are derived from the same open reading frame. Virology **164**:256–264.
24. **Moran, E., and M. B. Mathews.** 1987. Multiple functional domains in the adenovirus E1A gene. Cell **48**:177–178.
25. **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.
26. **Paterson, R. G., S. M. Thomas, and R. A. Lamb.** 1989. Specific non-templated nucleotide addition to a simian virus 5 mRNA: prediction of a common mechanism by which unrecognized hybrid P-cysteine-rich proteins are encoded by paramyxovirus "P" genes, p. 232–245. *In* D. Kolakofsky and B. W. J. Mahy (ed.), Genetics and pathogenicity of negative stranded viruses. Elsevier/North-Holland Publishing Co., Amsterdam.
27. **Peluso, R. W., R. A. Lamb, and P. W. Choppin.** 1977. Polypeptide synthesis in simian virus 5-infected cells. J. Virol. **23**: 177–187.
28. **Rima, B. K., M. W. Roberts, W. D. McAdam, and S. J. Martin.** 1980. Polypeptide synthesis in mumps virus-infected cells. J. Gen. Virol. **46**:501–505.
29. **Russel, M., S. Kidd, and M. R. Kelley.** 1986. An improved filamentous helper phage for generating single-stranded plasmid DNA. Gene **45**:333–338.
30. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
31. **Sato, H., M. Oh-Hira, N. Ishida, Y. Imamura, S. Hattori, and M. Kawakita.** 1987. Molecular cloning and nucleotide sequence of P, M and F genes of Newcastle disease virus avirulent strain D26. Virus Res. **7**:241–255.
32. **Sheshberadaran, H., and R. A. Lamb.** 1990. Sequence characterization of the membrane protein gene of paramyxovirus simian virus 5. Virology **176**:234–243.
33. **Shioda, T., Y. Hidaka, T. Kanda, H. Shibuta, A. Nomoto, and K.**

Iwasaki. 1983. Sequence of 3,687 nucleotides from the 3′ end of Sendai virus genome RNA and the predicted amino acid sequences of viral NP, P and C proteins. Nucleic Acids Res. **11**:7317–7330.

34. **Takeuchi, K., M. Hishiyama, A. Yamada, and A. Sugiura.** 1988. Molecular cloning and sequence analysis of the mumps virus gene encoding the P protein: mumps virus P gene is monocistronic. J. Gen. Virol. **69**:2043–2049.
35. **Thomas, S. M., R. A. Lamb, and R. G. Paterson.** 1988. Two mRNAs that differ by two nontemplated nucleotides encode the amino coterminal proteins P and V of the paramyxovirus SV5. Cell **54**:891–902.
36. **Vidal, S., J. Curran, and D. Kolakofsky.** 1990. Editing of the Sendai virus P/C mRNA by G insertion occurs during mRNA synthesis via a virus-encoding activity. J. Virol. **64**:239–246.
37. **Waxham, M. N., J. Aronowski, A. C. Server, J. S. Wolinsky, J. A. Smith, and H. M. Goodman.** 1988. Sequence determination of the mumps virus HN gene. Virology **164**:318–325.
38. **Zoller, M. J., and M. Smith.** 1982. Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment. Nucleic Acids Res. **10**:6487–6500.