JOURNAL OF VIROLOGY, Feb. 1984, p. 572–578
0022-538X/84/020572-07$02.00/0

Vol. 49, No. 2

Identification of a Tenth mRNA of Respiratory Syncytial Virus and Assignment of Polypeptides to the 10 Viral Genes

PETER L. COLLINS,* YUNG T. HUANG, AND GAIL W. WERTZ

Department of Microbiology and Immunology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27514

Received 1 June 1983/Accepted 28 October 1983

Nine mRNAs, their cDNA clones, and a genome transcriptional map have been reported previously for respiratory syncytial virus (P. L. Collins and G. W. Wertz, Proc. Natl. Acad. Sci. U.S.A. **80**:3208–3212, 1983). We report here the identification of a 10th viral mRNA, designated mRNA 2b (molecular weight [MW] ca. 0.39×10^6), that was detected by RNA (Northern) blot hybridization with cDNA clones. Analysis of a polycistronic readthrough transcript was used to deduce the position in the viral transcriptional map of the gene encoding the newly identified mRNA. The polypeptide coding assignments of 9 of the 10 respiratory syncytial virus mRNAs were determined. Individual viral mRNAs were purified by hybridization selection with nine unique, nonoverlapping cDNA clones and analyzed by translation in vitro. Each of the nine mRNAs encoded a single polypeptide chain. The coding assignments were as follows: RNA 1a (MW ca. 0.24×10^6), a 9,500-dalton (9.5K) protein; RNA 1b (MW 0.26×10^6), an 11K protein; RNA 1c (MW 0.26×10^6), a 14K protein; RNA 2a (MW 0.38×10^6), the 34K phosphorylated (P) protein; RNA 2b (MW 0.39×10^6), a 36K protein; RNA 3a (MW 0.40×10^6), the 26K matrix (M) protein; RNA 3b (MW 0.40×10^6), a 24K protein; RNA 4 (MW 0.47×10^6), the 42K major nucleocapsid (N) protein; and RNA 5 (MW 0.74×10^6), a 59K protein. The cDNA clones used for the hybridization selections were respiratory syncytial virus specific and did not hybridize with uninfected-cell mRNA; therefore the proteins synthesized with the selected mRNAs were virus specific. The 9.5K, 11K, 14K, 24K, M, P, 36K, N, and 59K proteins were encoded by different mRNAs; therefore these nine proteins are all unique. The 9.5K, 11K, 14K, 24K, M, P, and N proteins synthesized in vitro with hybrid-selected mRNAs each had counterparts with the same electrophoretic mobilities in extracts of virus-infected cells. The in vitro polypeptides and their authentic counterparts were shown to be closely related by limited digest peptide mapping. The 36K and 59K polypeptides lacked counterparts with the same electrophoretic mobilities in infected cells and therefore are candidates for the unprocessed precursors of the viral F and G glycoproteins. The 10th viral mRNA, the 2,500K RNA 7, was not tested directly but is the only known mRNA of the appropriate size to encode the 200K large (L) protein of the viral nucleocapsid. These assignments account for all 10 of the reported viral mRNAs and bring to 10 the number of known unique viral proteins.

Respiratory syncytial (RS) virus, a paramyxovirus (5, 23), is an important agent of human respiratory tract disease. The RS virus genome is a single negative strand of RNA, with a molecular weight (MW) greater than 5.0×10^6, that is transcribed in vivo to generate subgenomic polyadenylated [poly(A)$^+$] transcripts (20, 21, 27).

We recently prepared cDNA clones of nine viral transcripts with intracellular poly(A)$^+$ RNA used as template (8; P. L. Collins and G. W. Wertz, Abstr. Annu. Meet. Am. Soc. Microbiol. 1983, T23, p. 299). Independently, Venkatesan and co-workers have reported the construction and characterization of cDNA clones of four viral messengers (34). In our previous report (8), the sizes of nine unique viral poly(A)$^+$ transcripts, MWs ca. 0.24×10^6 to 2.50×10^6, were demonstrated by hybridizing the cDNA clones with RNA (Northern) blots prepared with infected-cell RNA. The RNA blot analyses also identified a series of polycistronic poly(A)$^+$ RNAs that, by analogy to the vesicular stomatitis virus (VSV) system (16, 17), were concluded to be derived from transcriptional readthrough of adjacent genes. The polycistronic RNAs were characterized by size and sequence homologies and used to prepare a transcriptional map of the RS viral genome (8).

There has been no consensus about the number and identities of unique RS virus proteins. The structural proteins reported to date for the RS virion include a major nucleocapsid protein (N; MW ca. 42,000), a nucleocapsid phosphoprotein (P; MW ca. 34,000), a large nucleocapsid protein (L; MW ca. 200,000), an envelope matrix protein (M; MW ca. 26,000), and two envelope glycoproteins, the fusion glycoprotein (F; MW ca. 68,000 to 70,000) and a second, methionine-poor glycoprotein (G; MW ca. 84,000 to 90,000) whose function has not been determined (2–4, 11, 13, 21, 28–30, 33–35, 37). Other polypeptides also have been suggested to be virus specific, including a 24,000 to 25,000-dalton virion polypeptide and additional, smaller polypeptides from infected cells that will be described below.

It is important to establish which proteins among those observed in virion preparations and in extracts of infected cells are both virus specific and unique. The two glycoproteins and the N, M, and P proteins appear to be virus specific because they are consistently precipitated with antiviral antibody; precipitation of the other proteins cited above has been inconsistent, and their status as virus specific determined on this basis is tentative (2, 11, 13, 30, 35). Several of the proteins have been shown to be unique and virus specific by in vitro translation of purified individual viral mRNAs (21, 34). Specifically, in vitro translation of viral mRNAs separated by gel electrophoresis showed that mRNAs from separate gel bands encoded the N, P, and M proteins and a 59K polypeptide thought to be the unprocessed precursor to

* Corresponding author.

one of the glycoproteins (21). Additionally, a separate gel band contained at least two RNAs and encoded three polypeptides, 9.5K, 11K, and 14K (21). In other work, the N, P, and M proteins and a possible nonstructural protein have been synthesized in vitro by using mRNA selected by hybridization to cDNA clones (34). Peptide mapping confirmed that the N, P, and M proteins are unrelated (21, 34). Thus, several proteins have been identified as the products of unique RS virus genes.

In the work reported here, the existence of a 10th unique poly(A)$^+$ transcript was demonstrated by cDNA cloning and RNA blot analysis. In addition, clones representing the RS virus mRNAs were used to purify their homologous mRNAs by hybrid selection for translation in vitro. This determined the polypeptide coding assignments for the RS virus mRNAs. These results showed that the 10 viral transcripts are all mRNAs, confirmed that they are unrelated, and catalogued the proteins encoded by the 10 viral genes.

MATERIALS AND METHODS

Viruses and cells. RS virus strain A_2 was propagated in monolayer cultures of HEp-2 cells as described previously (20).

mRNA purification. RNA was purified from cytoplasmic extracts of infected, antinomycin D-treated cultures (20) by sedimentation through cesium chloride (10), and poly(A)$^+$ RNA was selected on columns of oligodeoxythymidylate-cellulose (10).

cDNA clones. cDNA clones of viral poly(A)$^+$ transcripts were prepared and cloned in *Escherichia coli* cells by using deoxycytidylate-deoxyguanylate homopolymeric tailing and other conventional methods as described elsewhere (8). The specific cDNA clones used for the experiments shown here were: for RNA 1a, cDNA clone 3/39; for RNA 1b, clone 6/9; for RNA 1c, clone 3/55; for RNA 2a, clone 6/6; for RNA 2b, clone 6/58; for RNA 3a, clone 3/26; for RNA 3b, clone 3/51; for RNA 4, clone 6/63; and for RNA 5, clones 4/41 and 1/63. The characterization of most of these clones has been reported elsewhere (8).

RNA (Northern) blots. Intracellular viral RNAs, labeled metabolically with [^{3}H]uridine, were separated by electrophoresis in agarose-urea gels, transferred to diazobenzyloxymethyl paper (Schleicher & Schuell, Inc.), and analyzed by hybridization to nick-translated (32) cDNA clones. These procedures have been described previously (8).

DNA dot blots. Intact plasmids bearing cDNA inserts were denatured and fixed to nitrocellulose as previously described (24). For screening the cDNA library, the plasmids were prepared by a rapid small-scale method (19). These DNA dot blots were analyzed by hybridization to cDNA inserts that had been cut from the plasmid with endonuclease *Pst*I, purified by gel electrophoresis, and radiolabeled by nick-translation (32).

Hybrid selection. Hybrid selection of mRNAs was performed by the procedure of Ricciardi et al. (31), with the following additional details and modifications. The plasmids bearing cDNA inserts were linearized by digestion with nuclease *Pst*I or *Eco*RI and purified by extraction with phenol-chloroform and precipitation with ethanol before denaturation and filter binding. A total of 15 μg of recombinant plasmid was bound to each 35-mm nitrocellulose filter, and each filter was hybridized with an amount of poly(A)$^+$ RNA representing about one 15-cm plate of RS virus-infected cells. Hybridizations were done for 3 to 8 h at 50°C in the presence of 65% formamide, as recommended (31), except for cDNA clones of RNAs 1a, 1b, and 3a, which were hybridized at 42°C, and cDNA clones of RNA 2b, which were hybridized in the presence of 75% formamide at 50°C. Washes were done at 65°C instead of the recommended 60°C, except that the washes for the cDNA clones of RNAs 1a, 1b, and 3a were done at 50°C. The final washes were buffered with 10 mM Tris-hydrochloride, pH 7.5. Typically, one-fourth of the hybridized RNA subsequently eluted from each filter was used to program each 25- to 30-μl cell-free translation reaction mixture, and 1 to 2 μl (reticulocyte lysates) or 2 to 10 μl (wheat germ extracts) of the reaction mixtures were analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE).

Cell-free translation. In vitro translations were performed with nuclease-treated reticulocyte lysates (Green Hectares) as described previously (10) or with wheat germ extracts (not nuclease treated) (Bethesda Research Laboratories) under the conditions recommended by the supplier.

Preparation of [^{35}S]methionine-labeled polypeptides from infected cells. Monolayer cultures of HEp-2 cells were infected with ca. 1 PFU of RS virus per cell and incubated at 35°C. At 18 h after virus addition, 5 μg of actinomycin D per ml was added, and at 22 h the cells were incubated with 50 μCi of [^{35}S]methionine per ml for 0.5 h. The cells were washed, solubilized in SDS-PAGE sample buffer (25), and immediately boiled for 5 min.

SDS-PAGE. SDS-PAGE was performed on gels (16 cm long by 0.15 cm thick) with the buffer system of Laemmli (9, 10, 25). The stacking (upper) gel contained 4.5% acrylamide–0.12% *N,N'*-methylenebisacrylamide (bisacrylamide). The separating (lower) gels contained 11.5, 13, or 15% acrylamide, with acrylamide-bisacrylamide ratios of 38:1 (Fig. 3, lanes a through h; Fig. 5) or 175:1 (Fig. 3, lanes i through r; Fig. 4). The gels were fixed and analyzed by fluorography (9, 10).

Peptide mapping. [^{35}S]methionine-containing polypeptides, detected in dried gels by fluorography, were excised, rehydrated, and analyzed by limited protease digestion and SDS-PAGE by the method of Cleveland et al. (6, 7, 21). The proteases and protease concentrations used are indicated in the figure legends. The digestion fragments were separated in gels containing 17% acrylamide–0.32% bisacrylamide and fixed and analyzed by fluorography.

RESULTS

Identification of a 10th unique viral poly(A)$^+$ transcript. One hundred seventy-six cDNA clones, prepared from infected-cell poly(A)$^+$ RNA and characterized as virus specific in previous work (8; unpublished data), were screened both by dot blot hybridization with cDNA clones of the nine previously identified mRNAs and by hybridization with RNA blots of intracellular viral RNAs. Fourteen clones were identified that (i) did not hybridize detectably with the cloned cDNAs of the nine known viral mRNAs (not shown), (ii) hybridized with each other, and (iii) yielded a single, new pattern of hybridization to RNA blots, shown for representative clone 6/58 in Fig. 1 (lanes b, h, and n). Clone 6/58 hybridized strongly with RNA 8, previously identified as intracellular viral genomic RNA (8, 20), but did not hybridize detectably with RNA from uninfected cells (lane f). This was proof of virus specificity. cDNA clone 6/58 also hybridized strongly at the position of viral RNA gel band 2 and, as shown by longer film exposures (lanes h and n), at the position of viral RNA G. This hybridization pattern differed from the patterns obtained with cDNA clones of the previously identified mRNAs 2, 3a, 3b (lanes c, d, and e, respectively), and the remaining known mRNAs (8; unpub-

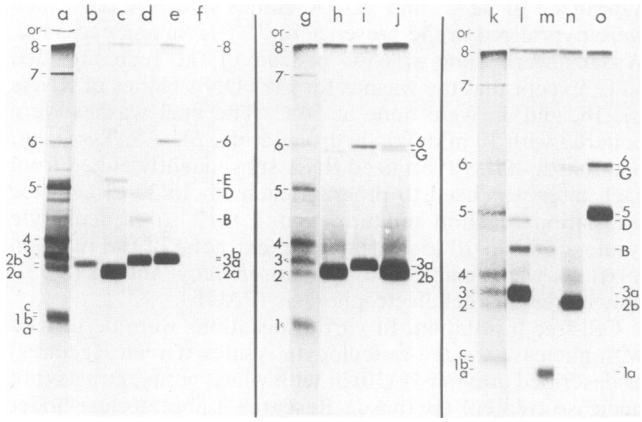

FIG. 1. Hybridization of RS virus cDNA clones with blots of intracellular viral RNAs. RNA was labeled with [^{3}H]uridine in the presence of antinomycin D, extracted from uninfected (lane f) or RS virus-infected (lanes a through e and g through o) HEp-2 cells, and separated by electrophoresis in agarose-urea gels (three different gels are represented). Lane a was analyzed by fluorography, and the remaining lanes were transferred to diazobenzyloxymethyl paper. After the transfer, lanes g and k were analyzed by fluorography. The remaining strips were subjected to hybridization with the following ^{32}P-labeled cDNA clones: the mRNA 2b clone 6/58 (lanes b, f, h, and n), the mRNA 2a clone 6/6 (lane c), the mRNA 3a clone 3/26 (lanes d and l), the mRNA 3b clone 3/51 (lanes e and i), the mRNA 1a clone 3/39 (lane m), the mRNA 5 clone 4/41 (lane o), and a mixture of the mRNA 2b clone 6/58 and the mRNA 5 clone 3/51 (lane j). The strips were analyzed by autoradiography with intensifying screens at −70°C. The positions of the unique mRNAs (mRNAs 1a, 1b, 1c, 2a, 2b, 3a, 3b, 4, 5, and 7) and some of the polycistronic RNAs (B, D, E, G, and 6) are shown. RNA 8 consists of genomic-length RNAs of plus and minus polarities (8, 20). The molecular sizes (in nucleotides), calculated from the MW estimates (8, 20) and including polyadenylate, for the unique mRNAs are: 1a, 720; 1b and 1c, 770; 2a, 1,170; 2b, 1,185; 3, 1,240; 4, 1,420; 5, 2,240; and 7, 7,500.

lished data). These results showed that RNA gel band 2 contained two mRNAs: the previously reported mRNA 2 (8), which is now designated mRNA 2a, and the newly identified, slightly larger mRNA that is reported here and is designated mRNA 2b. As further proof of this interpretation, the hybrid selection experiments described below showed that mRNAs 2a and 2b encode two different viral proteins.

The size of cDNA clone 6/58 was estimated by gel electrophoresis to be approximately 740 base pairs; related clones of up to 950 base pairs have also been isolated (not shown). The size of mRNA 2b, estimated by gel electrophoresis in parallel with mRNAs from VSV-infected cells (not shown), is approximately 1,185 nucleotides, including polyadenylate.

Transcriptional map position for the 10th viral gene. As described above, cloned cDNAs of mRNA 2b hybridized to genomic RNA, mRNA 2b, and RNA G. RNA G also hybridized with cDNA clones of RNA 5 (8) (Fig. 1, lane o), suggesting that RNA G might be a polycistronic RNA and might be useful in deducing the map position of gene 2b. However, because RNA G was similar in size to RNA 6, it was necessary to establish that these RNAs were distinct species. The following mixing experiment was performed: ^{32}P-labeled clone 6/58, which hybridized to RNA G (lane h), was mixed with a two- to threefold lesser amount of ^{32}P-labeled mRNA 3b clone 3/51, which hybridized with RNA 6 (lane i). This mixed probe hybridized to a doublet in the region of RNAs G and 6 (lane j), showing that these RNAs were distinct. The MW of RNA G was approximately the sum of mRNAs 2b and 5 (Fig. 2). Thus, RNA G met the criteria for a dicistronic transcript of the genes encoding mRNAs 2b and 5, indicating that these genes are adjacent in the transcriptional map.

Accordingly, the transcriptional map determined previously by analysis of polycistronic RNAs (8) has been expanded, based on the analysis of RNAs 2b and G. The genes encoding mRNAs 1c, 1b, 4, 2a, 3a, and 1a constitute one group of contiguous genes, and the genes encoding mRNAs 2b, 5, 3b, and 7 constitute a second group (Fig. 2). Because this method of mapping did not indicate the direction of transcription, the arrangement of the two blocks of genes shown is one of eight possible permutations. The arrangement shown is favored because it is consistent with preliminary results from UV inactivation studies (L. E. Dickens, P. L. Collins, and G. W. Wertz, Abstr. Annu. Meet. Am. Soc. Microbiol. 1983, T24, p. 300; manuscript in preparation) and places the genes in the approximate order (left to right) of decreasing intracellular mRNA abundance (estimated by comparing the abundance of individual transcripts with that of genomic RNA in RNA blots [Fig. 1] [8]).

None of the polycistronic RNAs identified to date links the two groups of genes in the deduced transcriptional map (Fig. 2). To illustrate that point, RNA blot hybridization patterns are shown (Fig. 1) for cDNAs representing the genes surrounding the possible "gap" in the gene order: mRNA 3a cDNA (lane l), mRNA 1a cDNA (lane m), mRNA 2b cDNA (lane n), and mRNA 5 cDNA (lane o). A dicistronic transcript of genes 1a and 2b would be the same size or slightly smaller than RNA B and should have been detectable in lanes m and n. Neither this RNA nor any other transcript derived from transcription across the gap has been detected.

Cell-free translation of total and hybrid-selected mRNAs. Cloned cDNAs of nine unique RS virus poly(A)$^+$ RNAs (RNAs 1a, 1b, 1c, 2a, 2b, 3a, 3b, 4, and 5) were used to purify individual mRNAs by hybridization selection. The products synthesized in vitro in response to each of the hybrid-selected mRNAs were analyzed by SDS-PAGE in parallel with the products obtained by translation of total infected or uninfected cell mRNAs and with proteins from extracts of RS virus-infected and uninfected cells.

Translation of total infected-cell mRNA in reticulocyte lysates and wheat germ extracts yielded similar patterns of [^{35}S]methionine-labeled polypeptides when analyzed by SDS-PAGE (Fig. 3). Eight readily detectable in vitro polypeptide products, the 9.5K, 11K, 14K, 24K, M (26K), P

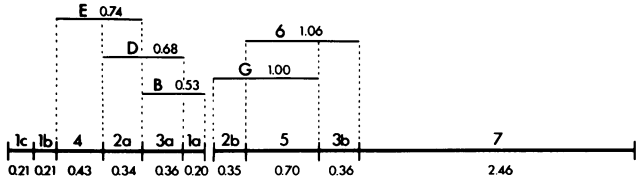

FIG. 2. RS virus transcriptional map, including the deduced position for the gene encoding mRNA 2b. Horizontal placements represent the sequence homologies among the 10 unique mRNAs and five of the polycistronic RNAs. Other polycistronic RNAs analyzed for construction of the map are shown elsewhere (8). The length of the line (in megadaltons) representing each gene or transcript is proportional to the MW of the RNA minus a contribution of 125 adenylate residues (the average size of polyadenylate estimated for Newcastle disease virus [36]).

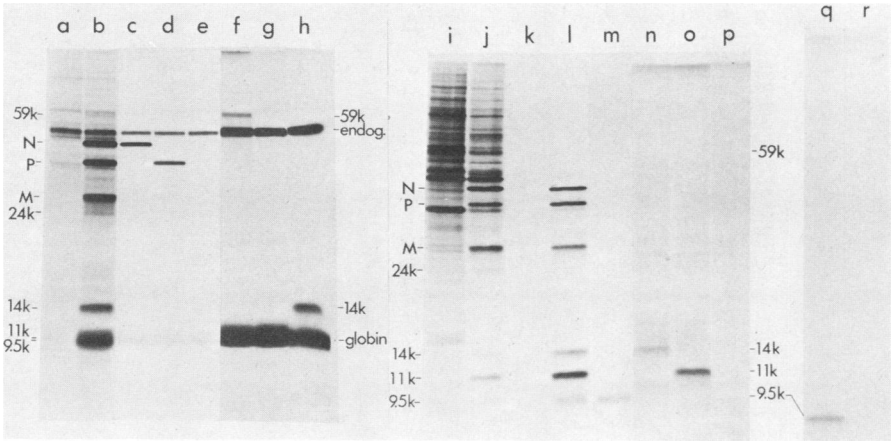

FIG. 3. SDS-PAGE of [^{35}S]methionine-labeled polypeptides synthesized in reticulocyte lysates (lanes a through h) and wheat germ extracts (lanes k through r) in response to mRNAs purified by hybridization to cDNA clones of RNA 4 (c), 2a (d), 5 (f), 1c (h and n), 1a (m and q), and 1b (o). Control lanes show the products synthesized in the absence of added mRNA (e, g, and p), in response to total mRNA from uninfected (a and k) and RS virus-infected (b and l) HEp-2 cells and in response to mRNA selected from uninfected-cell mRNA by hybridization to an RNA 1a cDNA clone (r). Also shown are labeled polypeptides extracted from uninfected (i) and RS virus-infected (j) cells. The polypeptides were separated on 13% (a through h) or 15% (i through r) polyacrylamide gels and were analyzed by fluorography. Film exposures: 4 to 8 h, lanes a through e and i through m; ca. 30 h, lanes f through h and n through r. The positions of viral and nonviral (globin and endogenous [endog.]) in vitro products are marked. The 9.5K and 11K proteins were separated in lanes i through r but not in lanes a through h; this was due to differences in the acrylamide-bisacrylamide ratio.

(34K), N (42K), and 59K proteins, were specific to mRNA extracted from RS virus-infected cells (Fig. 3, cf. lane l and lane k, which contains polypeptides synthesized in response to mRNA from species comigrated with polypeptides specific to RS virus-infected cells [cf. lanes l and j]). Three additional viral polypeptides found previously in extracts of infected cells, the F and G glycoproteins and the 200K L protein, did not comigrate with in vitro products. The L protein presumably was not detected among the in vitro products due to both the limitation of cell-free systems for translating large mRNAs and the apparently low level of its mRNA. The 59K polypeptide, detected in vitro but not in vivo, is a candidate for the unprocessed polypeptide chain of one of the two glycoproteins (21); an in vitro product that could account for the second glycoprotein will be described below.

Translation of mRNAs purified by hybrid selection with cDNA clones of RNAs 2a, 4, and 5 yielded the P, N, and 59K polypeptides, respectively (Fig. 3, lanes d, c, and f, respectively). Cloned cDNAs of RNAs 1a, 1b, and 1c provided purification of mRNAs encoding the 9.5K (lanes m and q), 11K (lane o), and 14K (lanes h and n) polypeptides, respectively. Because the 9.5K polypeptide has been reported previously only by our laboratory (21) and migrated near the position of endogenous cellular polypeptides (cf. lanes i and j, Fig. 1), a cDNA clone of RNA 1a was tested for possible selection of mRNA from a preparation of uninfected-cell mRNA. However, the only detectable translation products were the endogenous products of the wheat germ lysate (Fig. 3, lane r). This result confirmed that the 9.5K protein was encoded by mRNA specific to RS virus-infected cells, consistent with our previous demonstration (8) that RNA 1a is virus specific.

Translation in vitro of mRNAs purified by hybridization to cloned cDNAs of RNAs 3a and 3b encoded the M and 24K polypeptides, respectively (Fig. 4, lanes e and k).

Finally, cloned cDNAs of RNA 2b provided purification of mRNA encoding two polypeptides of approximate MW 36,000 to 38,000 and 32,000 to 34,000 (designated 36K and 32K, respectively) (Fig. 5, lane e). These species have not been reported previously. cDNA clones of mRNA 2b did not select detectable mRNA activity from uninfected-cell mRNA (not shown). This result, together with the demonstration that mRNA 2b is virus specific (Fig. 1), confirmed that the 32K and 36K proteins were virus specific. The 36K protein, which was more abundant than the 32K protein, could be detected as a minor band among the translation products of total infected-cell mRNA (Fig. 5, lane d). Peptide mapping by limited proteolysis confirmed this identification (not shown). A polypeptide band with an electrophoretic mobility similar to that of the 36K protein was also present in the SDS-PAGE pattern of polypeptides extracted from RS virus-infected cells (lane b). However, peptide mapping (not shown) showed that this infected-cell polypeptide was not related to the 36K protein, but rather to the cellular protein marked with an arrow in lane a. Thus, the 36K and 32K proteins were virus specific but did not have detectable counterparts in vivo.

The relative methionine, cysteine, and leucine contents of the 59K, 36K, 32K, N, P, 24K, and M proteins were compared qualitatively by in vitro translation of hybrid-selected and total infected-cell mRNA with [^{35}S]methionine, [^{35}S]cysteine, and [^{3}H]leucine (not shown). The P protein was deficient in cysteine, and the N protein had a relatively low cysteine content. The 36K and 32K proteins were both markedly low in methionine relative to their leucine and cysteine contents; a low methionine content has been reported previously for the RS virus G glycoprotein (12, 29).

Peptide mapping. The 9.5K, 11K, 14K, 24K, M, P, and N proteins synthesized in vitro with hybrid-selected mRNAs comigrated with authentic viral proteins extracted from infected cells (Fig. 3, 4, and 5). However, comigration of a hybrid-selected translation product with an intracellular viral protein was not proof of relatedness. To confirm relatedness and thereby confirm these identifications, the 9.5K, 11K, 14K, 24K, M, P, and N proteins synthesized with hybrid-selected mRNAs were compared with the comigrating intracellular proteins by limited digest peptide mapping (Fig. 6A).

The 9.5K protein was not digested detectably by papain (10 μg/ml), chymotrypsin (200 μg/ml), or *Staphylococcus*

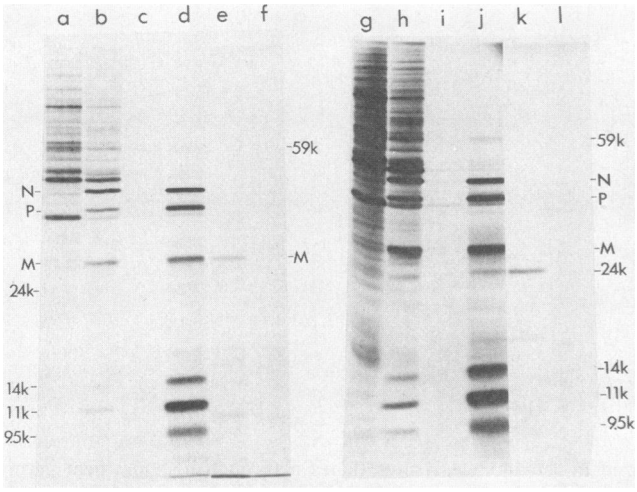

FIG. 4. SDS-PAGE of in vitro translation products corresponding to RNAs 3a (lane e) and 3b (lane k). Wheat germ extracts were programmed with no added mRNA (f and l), with total poly(A)$^+$ RNA from uninfected (c and i) and RS virus-infected (d and j) HEp-2 cells, and with mRNAs selected by hybridization to cDNA clones of RNAs 3a (e) and 3b (k). The [^{35}S]methionine-labeled products were analyzed in parallel with [^{35}S]methionine-labeled polypeptides extracted from uninfected (a and g) and RS virus-infected (b and h) cells by SDS-PAGE on 15% polyacrylamide gels and by fluorography.

aureus protease (150 μg/ml) (not shown). It is not known whether this is a characteristic of the 9.5K protein or simply a function of its small size. In the other comparisons, the pattern of fragments for each of the in vitro polypeptides matched closely that of its counterpart extracted from infected cells. This confirmed the identities of the polypeptides encoded by the hybrid-selected mRNAs.

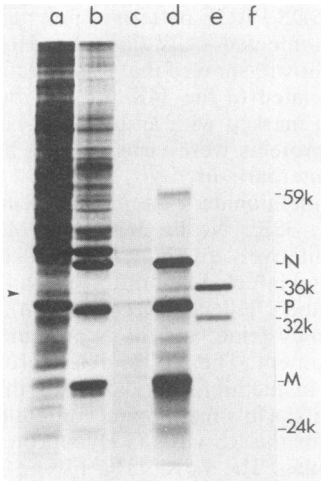

FIG. 5. SDS-PAGE of [^{35}S]methionine-labeled polypeptides synthesized in wheat germ extracts in response to mRNA purified by hybridization with cDNA clones of RNA 2b (lane e). Control lanes contain the products synthesized in vitro in the absence of added mRNA (f) and in response to total mRNA from uninfected (c) and RS virus-infected (d) cells. Other lanes contain labeled polypeptides extracted from uninfected (a) and RS virus-infected (b) cells. The polypeptides were separated on an 11.5% polyacrylamide gel, fixed, and analyzed by fluorography. Arrow, Position of polypeptide bands in lanes a, b, and d that appeared to comigrate with the 36K protein (see the text).

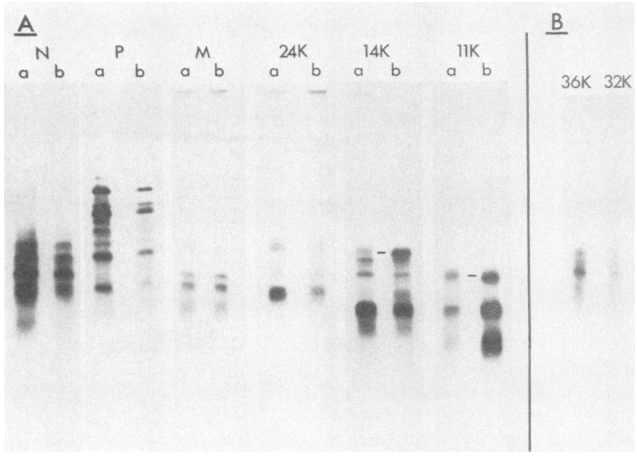

FIG. 6. Analysis of RS virus proteins by limited proteolysis and SDS-PAGE. (A) The N, P, M, 24K, 14K, and 11K proteins synthesized in vitro with hybrid-selected mRNAs (lanes a) were compared with their authentic counterparts extracted from RS virus-infected cells (lanes b). The M protein was digested with 50 μg of *S. aureus* protease per ml; the N, P, 24K, 14K, and 11K proteins were treated with 5 μg of papain per ml. The [^{35}S]methionine-labeled fragments were separated on a 17% polyacrylamide gel, fixed, and analyzed by fluorography. The bars in the 11K and 14K protein lanes indicate the positions of the indigested proteins. (B) [^{35}S]cysteine-labeled 36K and 32K proteins were compared by digestion with 5 μg of papain per ml, SDS-PAGE (17% polyacrylamide), and fluorography.

Peptide mapping was also used to compare the 32K and 36K proteins, which were both encoded by mRNA selected with clones of mRNA 2b as described above. Comparison of the [^{35}S]cysteine-labeled 36K and 32K proteins by limited digest peptide mapping with papain indicated that the two proteins were related (Fig. 6B). However, digestion of the [^{35}S]methionine-labeled or [^{35}S]cysteine-labeled 36K and 32K proteins with *S. aureus* protease yielded single protease-resistant fragments (not shown). Because single fragments are not sufficient for comparison, it will be important to confirm the relatedness of the 36K and 32K proteins by tryptic peptide mapping, immunological cross-reactivity, or sequencing of cDNA clones of mRNA 2b.

DISCUSSION

A 10th RS virus mRNA, mRNA 2b, was detected by cDNA cloning and RNA blot hybridization, and its transcriptional map position was determined. Hybrid-selected translation experiments identified the polypeptides encoded by the newly identified mRNA and by eight of the nine previously reported mRNAs. These results identified the viral mRNAs, were further proof that these viral transcripts are unique mRNAs, and provided the basis for identifying 10 unique RS virus proteins.

Translation of hybrid-selected mRNAs determined directly the polypeptide coding assignments for 9 of the 10 RS virus poly(A)$^+$ RNAs. The coding assignments were as follows: RNA 1a (MW ca. 0.24×10^6), 9.5K protein; RNA 1b (MW ca. 0.26×10^6), 11K protein; RNA 1c (MW ca. 0.26×10^6), 14K protein; RNA 2a (MW ca. 0.38×10^6), the 34K P protein; RNA 3a (MW ca. 0.40×10^6), the 26K M protein; RNA 3b (MW ca. 0.40×10^6), 24 K protein; RNA 4 (MW ca. 0.47×10^6), the 42K N protein; and RNA 5 (MW ca. 0.74×10^6), 59K protein. RNA 2b (MW ca. 0.39×10^6) encoded two

methionine-poor polypeptides, the 36K and 32K proteins, which appeared to be related.

These results confirmed and extended our previous identifications based on translation of mRNAs separated by gel electrophoresis (21). The coding assignment of the RNA 7 (MW 2.50×10^6) was not investigated, but by size considerations and by analogy to other paramyxoviruses and VSV, RNA 7 is the only appropriate candidate for encoding the 200K L protein. Assuming that RNA 7 encodes the L protein, each of the 10 unique poly(A)$^+$ RNAs specified by RS virus is an mRNA, and each encodes a single unique polypeptide chain. However, the formal possibility exists that one or more of these mRNAs may contain additional translational reading frames encoding additional proteins that were not detected in these experiments. This possibility is being examined in detail by sequencing cDNA clones.

The cDNA clones used for the hybrid selections previously have been shown to be copies of virus-specific RNAs (Fig. 1) (8) and did not hybridize detectably with uninfected-cell mRNAs, whether tested by RNA blot hybridization (Fig. 1) (8; unpublished data) or by control hybrid selection experiments with uninfected-cell mRNA (Fig. 3 and unpublished data). Therefore, the polypeptides synthesized in response to the hybrid-selected mRNAs are virus specific. Because the 9.5K, 11K, 14K, 24K, M, P, 36K, N, and 59K proteins were shown to be encoded by different unique mRNAs, we conclude that these nine proteins are all unique.

The N, P, M, and L proteins are structural components of the virion and have counterparts among the other paramyxoviruses and VSV (29, 37; unpublished data). The 9.5K, 11K, and 14K proteins are not abundant in preparations of purified virions (4, 11, 28–30, 33, 34, 37; unpublished data) and therefore are candidates for nonstructural proteins. A 24K to 25K protein has been reported as a component of purified virions (4, 11, 28–30, 33, 37), and peptide mapping showed that this structural protein is the same as the 24K protein described here (unpublished data). Thus, the 24K protein is a candidate for a novel structural protein that lacks known counterparts among other paramyxoviruses and VSV. Finally, the 36K and 59K proteins were detected by translation of viral mRNAs in vitro but were not detected in vivo (21; unpublished data). These proteins might be undetectable in infected cells due to inefficient synthesis or rapid degradation. However, because the cell-free translation systems that we used lacked glycosylation activities, we suggest that the 36K and 59K proteins are the unprocessed polypeptide chains of the two viral glycoproteins; in vivo these would have higher MWs due to carbohydrate additions. The 36K protein, but not the 59K protein, was markedly deficient in methionine (data not shown), a characteristic of the G glycoprotein (11, 30; unpublished data). Although the 84K G glycoprotein is much larger in apparent MW than the 36K protein, carbohydrate contributions of such a magnitude have been reported (e.g., reference 22), and the G glycoprotein does appear to be heavily glycosylated (2, 11, 29, 30; unpublished data). We are currently investigating possible relationships between the 59K and 36K proteins and the G and F glycoproteins by peptide mapping and immunological cross-reactivity.

Based on the mRNA coding assignments identified here, the viral polypeptides were assigned to positions in the transcriptional map (Fig. 7). By comparison with the paramyxoviruses Sendai virus (14) and Newcastle disease virus (7) and with VSV (1), the transcriptional map for RS virus has three novel features. (i) The gene encoding the N protein, rather than being first in the gene order, is preceded

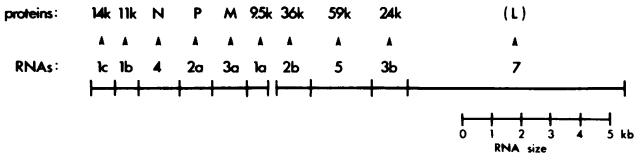

FIG. 7. Assignment of polypeptides to the mRNAs and genome transcriptional map of RS virus. The RS virus transcriptional map was deduced from analysis of polycistronic RNAs as outlined in the text and described in previous work (8). The genes are identified by the mRNAs they encode and are drawn to scale by the estimated mRNA molecular sizes (shown on the scale in kilobases [kb]). The in vitro polypeptide encoded by each mRNA is indicated; the L protein is in parentheses to emphasize that the identification of the RNA 7 coding assignment was indirect and is tentative (see the text).

instead by two genes encoding proteins thought to be nonstructural (the 14K and 11K proteins). (ii) The M protein gene and the genes thought to encode the glycoproteins (represented by RNAs 2b and 5), rather than being contiguous, are separated by a gene encoding a protein thought to be a third nonstructural protein (the 9.5K protein). (iii) The genes thought to encode the glycoproteins and the L protein, rather than being contiguous, are separated by a gene encoding a novel structural protein (the 24K protein). At present, the deduced transcriptional map is incomplete, because none of the known polycistronic RNAs links the two groups of genes (Fig. 2). Preliminary results from UV mapping studies provided evidence for a single RS virus promoter (Dickens et al., Abstr. Annu. Meet. Am. Soc. Microbiol. 1983), suggesting that this apparent gap was not due to an additional promoter site. There is presently no evidence to suggest that the gap represents an additional, unidentified gene(s). We are currently investigating both the nature of the gap and the relationship between the transcriptional map and the 3′ to 5′ physical map of the RS virus genome.

The work presented here showed that the genome organization and number of gene products of RS virus are more complex than those of the prototypical paramyxovirus, as exemplified by Newcastle disease virus and Sendai virus. In addition to the differences in gene order described above, RS virus encodes at least 10 unique proteins, compared with the 6 reported previously for Newcastle disease virus and mumps virus (9, 18) and the 7 reported previously for Sendai virus, simian virus 5, and canine distemper virus (12, 15, 26). The 10 RS virus proteins include 4 novel small proteins (the 24K, 14K, 11K, and 9.5K proteins). It is possible that additional analysis of other paramyxoviruses may identify previously unknown counterparts to some of the novel RS virus proteins.

We assume that these additional RS virus proteins have functional or structural significance. One speculative interpretation is that the functions of these additional proteins might be unique to RS virus. Alternatively, RS virus and other paramyxoviruses might have generally similar enzymatic activities and polypeptide functional domains (apart from the known differences in glycoprotein biological activities) that are distributed among different numbers of polypeptide chains.

ACKNOWLEDGMENTS

We thank Steve Loechel for technical assistance, Joyce Bradshaw and Kathleen Snipes for typing, Dennis Hruby, Andrew Ball, and John Patton for advice on hybrid selection, and the members of our lab for their interest and encouragement.

This work was supported by Public Health Service grants AI-12464 and AI-15135 from the National Institute of Allergy and Infectious Disease and by grants CA-09156 and CA-19014 from the National Cancer Institute. P.L.C. received support from Public Health Service grant AI-06956 from the National Institute of Allergy and Infectious Disease.

LITERATURE CITED

1. **Ball, L. A.** 1980. Gene order, p. 61–73. *In* D. H. L. Bishop (ed.), Rhabdoviruses. CRC Press, Inc., Cleveland, Ohio.
2. **Bernstein, J. M., and J. F. Hruska.** 1981. Respiratory syncytial virus proteins: identification by immunoprecipitation. J. Virol. **38**:278–285.
3. **Cash, P., C. R. Pringle, and C. M. Preston.** 1979. The polypeptides of human respiratory syncytial virus: products of cell-free protein synthesis and post-translational modifications. Virology **92**:375–384.
4. **Cash, P., W. H. Wunner, and C. R. Pringle.** 1977. A comparison of the polypeptides of human and bovine respiratory syncytial viruses and murine pneumonia virus. Virology **82**:369–379.
5. **Choppin, P. W., and R. W. Compans.** 1975. Reproduction of paramyxoviruses, p. 98–178. *In* H. Fraenkel-Conrat and R. R. Wagner (ed.), Comprehensive virology. Plenum Publishing Corp., New York.
6. **Cleveland, D. W., S. G. Fischer, W. M. Kirschner, and U. K. Laemmli.** 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. **252**:1102–1106.
7. **Collins, P. L., L. E. Hightower, and L. A. Ball.** 1980. Transcriptional map for Newcastle disease virus. J. Virol. **35**:682–693.
8. **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.
9. **Collins, P. L., G. W. Wertz, L. A. Ball, and L. E. Hightower.** 1982. Coding assignments of the five smaller mRNAs of Newcastle disease virus. J. Virol. **43**:1024–1031.
10. **Davis, N. L., and G. W. Wertz.** 1982. Synthesis of vesicular stomatitis virus negative-strand RNA in vitro: dependence on viral protein synthesis. J. Virol. **41**:821–832.
11. **Dubovi, E. J.** 1982. Analysis of proteins synthesized in respiratory syncytial virus-infected cell. J. Virol. **42**:372–378.
12. **Etkind, P. R., R. K. Cross, R. A. Lamb, D. C. Merz, and P. W. Choppin.** 1980. In vitro synthesis of structural and non-structural proteins in Sendai and SV5 viruses. Virology **100**:22–33.
13. **Fernie, B. F., and J. L. Gerin.** 1982. Immunochemical identification of viral and nonviral proteins of the respiratory syncytial virus virion. Infect. Immun. **37**:243–249.
14. **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.
15. **Hall, W. W., R. A. Lamb, and P. W. Choppin.** 1980. Polypeptides of canine distemper virus: synthesis in infected cells and relatedness to the polypeptides of other Morbilliviruses. Virology **100**:433–449.
16. **Herman, R. C., S. Adler, R. J. Colonno, A. K. Banerjee, R. A. Lazzarini, and H. Westphal.** 1978. Intervening polyadenylate sequences in RNA transcripts of vesicular stomatitis virus. Cell **15**:587–596.
17. **Herman, R. C., M. Schubert, J. D. Keene, and R. A. Lazzarini.** 1980. Polycistronic vesicular stomatitis virus RNA transcripts. Proc. Natl. Acad. Sci. U.S.A. **77**:4662–4665.
18. **Herrler, G., and R. W. Compans.** 1982. Synthesis of mumps virus polypeptides in infected Vero cells. Virology **119**:430–438.
19. **Holmes, D. S., and M. Quigley.** 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. **114**:193–197.
20. **Huang, Y. T., and G. W. Wertz.** 1982. The genome of respiratory syncytial virus is a negative-stranded RNA that codes for at least seven mRNA species. J. Virol. **43**:150–157.
21. **Huang, Y. T., and G. W. Wertz.** 1983. Respiratory syncytial virus mRNA coding assignments. J. Virol. **46**:667–672.
22. **Hunter, E., E. Hill, M. Hardwick, A. Bhown, D. E. Schwartz, and R. Tizard.** 1983. Complete sequence of the Rous sarcoma virus *env* gene: identification of structural and functional regions of its product. J. Virol. **46**:920–936.
23. **Kingsbury, D. W., M. A. Bratt, P. W. Choppin, R. P. Hansen, Y. Hosaka, V. ter Meulen, E. Norrby, W. Plowright, R. Rott, and W. H. Wunner.** 1978. Paramyxovirology. Intervirology **10**:137–152.
24. **Kofatos, F. C., C. W. Jones, and A. Efstratiadis.** 1979. Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. Nucleic Acids Res. **7**:1541–1552.
25. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
26. **Lamb, R. A., and P. W. Choppin.** 1978. Determination by peptide mapping of the unique polypeptides in Sendai virions and infected cells. Virology **84**:469–478.
27. **Lambert, D. M., M. W. Pons, G. Mbuy, and K. Dorsch-Hasler.** 1980. Nucleic acids of respiratory syncytial virus. J. Virol. **36**:837–846.
28. **Levine, S.** 1977. Polypeptides of respiratory syncytial virus. J. Virol. **21**:427–431.
29. **Peeples, M., and S. Levine.** 1979. Respiratory syncytial virus polypeptides: their location in the virion. Virology **95**:137–145.
30. **Pringle, C. R., P. V. Shirodaria, H. B. Gimenez, and S. Levine.** 1981. Antigen and polypeptide synthesis by temperature-sensitive mutants of respiratory syncytial virus. J. Gen. Virol. **54**:173–183.
31. **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.
32. **Rigby, P. W. J., M. Dieckmann, 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.
33. **Ueba, O.** 1980. Purification and polypeptides of respiratory syncytial virus. Microbiol. Immunol. **24**:361–364.
34. **Venkatesan, S., N. Elango, and R. M. Chanock.** 1983. Construction and characterization of cDNA clones for four respiratory syncytial viral genes. Proc. Natl. Acad. Sci. U.S.A. **80**:1280–1284.
35. **Walsh, E. E., and J. Hruska.** 1983. Monoclonal antibodies to respiratory syncytial virus proteins: identification of the fusion protein. J. Virol. **47**:171–177.
36. **Weiss, S. R., and M. A. Bratt.** 1974. Polyadenylate sequences on Newcastle disease virus mRNA synthesized in vivo and in vitro. J. Virol. **13**:1220–1230.
37. **Wunner, W. H., and C. R. Pringle.** 1976. Respiratory syncytial virus proteins. Virology **73**:228–243.