Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 5134–5138, August 1987
Biochemistry

Gene overlap and site-specific attenuation of transcription of the viral polymerase L gene of human respiratory syncytial virus

(negative-strand RNA viruses/transcriptive signals/mode of transcription/alternative transcripts)

PETER L. COLLINS*, ROBERT A. OLMSTED*, MELANIE K. SPRIGGS*, PHILIP R. JOHNSON*†, AND ALICIA J. BUCKLER-WHITE*

*Laboratory of Infectious Diseases, Building 7, Room 100, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892; and †Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, TN 37232

Communicated by Robert M. Chanock, March 27, 1987

ABSTRACT **Sequence analysis of the gene encoding polymerase L protein of respiratory syncytial virus showed that L-gene transcription initiates within its upstream neighbor, the gene encoding the 22-kDa protein 22K. This is an exception to the canonical mode of sequential transcription of nonoverlapping genes described for other nonsegmented negative-strand RNA viruses. As a consequence of the gene overlap, the termination/polyadenylylation signal for the 22K gene is located at nucleotides 56–68 within the L gene. L-gene transcription yielded two RNAs: an abundant, truncated, polyadenylylated transcript resulting from termination at the internal signal and a markedly less abundant large polyadenylylated transcript representing the complete L gene. This result showed that the internal termination/polyadenylylation signal is an attenuator of L-gene transcription.**

Human respiratory syncytial (RS) virus, a paramyxovirus, is an important cause of pediatric respiratory tract disease (1). The genome (vRNA) of RS virus, a single negative-sense strand of RNA of ≈15,000 nucleotides (nt), encodes 10 mRNAs that each encode a single major polypeptide (2). RS virus proteins that have counterparts in other paramyxoviruses are the major nucleocapsid protein N, phosphoprotein P, and large protein L of the viral nucleocapsid and the matrix protein M, putative attachment glycoprotein G, and fusion glycoprotein F of the virion envelope. Other RS virus proteins appear to lack direct counterparts in most other paramyxoviruses: the 22-kDa envelope-associated protein 22K, the nonstructural proteins 1C and 1B, and the small hydrophobic protein 1A (2–4).

Complete cDNAs and nucleotide sequences have been obtained for 9 of the 10 RS virus mRNAs, excluding only the L mRNA (refs. 1, 2, and 5–10 and references cited therein). Synthetic oligonucleotides designed from these sequences were used to direct sequencing of intergenic regions in vRNA (2). This determined that the viral gene order is 3′–1C–1B–N–P–M–1A–G–F–22K–L–5′ and demonstrated that the intergenic regions between the first nine genes are nonconserved in sequence and length. The latter results were unexpected because the intergenic regions determined for the paramyxoviruses Sendai, human parainfluenza 3, and measles viruses are highly conserved within and among viruses (3, 11–13). In general, the gene products, genome organization, and mode of gene expression of these viruses appear to contain elements of unexpected diversity.

In the present work, the organization of RS virus vRNA was investigated further by sequence analysis of the L gene and L-gene transcripts. These studies showed that the L gene initiates within its upstream neighbor, the 22K gene. mRNA mapping studies demonstrated that this arrangement of overlapping genes results in site-specific attenuation of L-gene transcription. These studies provided additional insights into the structure and mode of transcription of nonsegmented negative-strand viral genomes.

MATERIALS AND METHODS

Fractionation of RS Virus RNA on Sucrose Gradients. RS virus mRNA [poly(A)$^+$ RNA, selected with oligo(dT)-cellulose, from forty 175-cm^2 flasks of HEp-2 cells] and total RS virus RNA (nonselected, from twenty 175-cm^2 flasks, to provide marker RNAs) were fractionated in parallel on linear gradients of 10–30% (wt/wt) sucrose in 10 mM Tris·HCl, pH 7.6/100 mM NaCl/2 mM EDTA/0.5% $NaDodSO_4$ by centrifugation in a Beckman SW41 rotor at 39,000 rpm for 2.5 hr at 20°C; 1-ml fractions were collected and RNA was precipitated with ethanol.

Nuclease Mapping. A 128-base-pair (bp) cDNA clone spanning the 22K–L gene junction (nt 47–174, Fig. 1) was constructed from two overlapping synthetic 80-nt oligonucleotides, representing nt 47–126 and the complement of nt 95–174 in Fig. 1. The oligonucleotides were hybridized, made blunt-ended with reverse transcriptase, inserted into the *Sma* I site of pGEM-4 (Promega Biotec, Madison, WI) and cloned in *Escherichia coli* strain RR1. The cDNA clone used here was in the counterclockwise orientation and its sequence was confirmed. vRNA-sense RNA was synthesized using *Xba* I-cut plasmid as template for bacteriophage SP6 polymerase, and mRNA-sense RNA was synthesized using *Eco*RI-cut plasmid as template for phage T7 polymerase, following the supplier's (Promega Biotec) protocols. Hybridization and subsequent RNase treatment were at room temperature with excess probe and excess (50 μg/ml) RNase T1 (Pharmacia P-L Biochemicals) according to Promega Biotec protocols.

RESULTS

Partial Nucleotide Sequence of the L Gene. Synthetic oligonucleotide primers were constructed from the sequence of the penultimate 22K gene and used to direct dideoxynucleotide sequencing of vRNA extracted from purified virions. Fig. 1 shows the mRNA-sense sequence of the last 153 nt of the 22K gene followed by an additional 183 nt determined here from vRNA.

A synthetic oligodeoxynucleotide, oligonucleotide L, was constructed to be complementary to nucleotides 206–217 in Fig. 1. Oligonucleotide L hybridized to the L mRNA on an RNA blot (Fig. 2*a*), showing that the L gene follows the 22K gene in the genetic map. However, the start of the L gene could not be predicted from the sequence alone because this

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: RS virus, respiratory syncytial virus; vRNA, viral genomic RNA; nt, nucleotide(s).

```
                                                                                                         84
TAT ACA ATA TAT ATA TTA GTG TCA TAA CAC TCA ATT CTA ACA CTC ACC ACA TCG TTA CAT TAT TAA TTC AAA CAA TTC AAG TTG
  L start                                           oligo 22K-L         22K end                         168
TGG GAC AAA ATG GAT CCC ATT ATT AAT GGA AAT TCT GCT AAT GTT TAT CTA ACC GAT AGT TAT TTA AAA GGT GTT ATC TCT TTC
             MET Asp Pro Ile Ile Asn Gly Asn Ser Ala Asn Val Tyr Leu Thr Asp Ser Tyr Leu Lys Gly Val Ile Ser Phe
                                                    oligo L                                              252
TCA GAG TGT AAT GCT TTA GGA AGT TAC ATA TTC AAT GGT CCT TAT CTC AAA AAT GAT TAT ACC AAC TTA ATT AGT AGA CAA AAT
Ser Glu Cys Asn Ala Leu Gly Ser Tyr Ile Phe Asn Gly Pro Tyr Leu Lys Asn Asp Tyr Thr Asn Leu Ile Ser Arg Gln Asn
                                                                                                         336
CCA TTA ATA GAA CAC ATG AAT CTA AAG AAA CTA AAT ATA ACA CAG TCC TTA ATA TCT AAG TAT CAT AAA GGT GAA ATA AAA TTA
Pro Leu Ile Glu His MET Asn Leu Lys Lys Leu Asn Ile Thr Gln Ser Leu Ile Ser Lys Tyr His Lys Gly Glu Ile Lys Leu
```

FIG. 1. Nucleotide sequence (mRNA-sense) of RS virus vRNA spanning the junction of the 22K and L genes. Sequence shown represents the last 153 nt of the 22K gene (positions 1–153) and the first 251 nt of the L gene (positions 86–336). Mapping of the 22K gene boundary was described previously (9, 10). L gene-start and 22K gene-end sequences are boxed. Predicted amino-terminal, partial amino acid sequence for the L protein is shown; the 22K gene sequences shown do not encode protein. Sequences complementary to oligonucleotides L and 22K/L are underlined. Additional sequence of 824 nt extending further into the L gene was determined by successive primer-extension reactions with additional oligonucleotides (unpublished data); the translational reading frame shown remained open for the length of the available sequence.

region of vRNA lacked the 9-nt conserved gene-start sequence 3′ CCCCGUUUA 5′ (vRNA-sense) at the start of each of the nine previously sequenced RS virus genes (2).

In order to map and sequence the 5′ end of the L mRNA and thereby map the start of the L gene, oligonucleotide L was used for primer extension on intracellular RS virus mRNA (Fig. 3*a*). The extension products contained a major strong-stop doublet (labeled 1) representing the 5′ end of the L mRNA and also included a second strong-stop doublet (labeled 2) whose significance is discussed below.

Comparison of the sequence obtained from L mRNA (Fig. 3*a*) with that obtained from vRNA (Fig. 1) mapped the start of the L gene to nt 86 in Fig. 1. Surprisingly, this showed that the first 68 nt of the L gene are the same as the last 68 nt of the 22K gene. Confirmatory sequencing of this region of vRNA using synthetic primers eliminated the possibility that this 68-nt sequence was a repeated sequence. Thus, the 22K and L genes overlap for 68 nt. This also showed that the L gene has a gene-start sequence, 3′ CCCUGUUUU 5′ (vRNA-sense), that contains two differences (underlined) from the sequence conserved among the other RS virus genes.

Oligonucleotide 22K/L (see Fig. 1 for map position) was constructed to be complementary to transcripts of the 68-nt overlap. In RNA blot hybridization experiments (Fig. 2*b*), oligonucleotide 22K/L hybridized to the relatively more abundant 22K mRNA, to the relatively less abundant L mRNA, and to the F-22K mRNA, a previously described readthrough transcript of the F and 22K genes (5, 8, 9). In addition, oligonucleotide 22K/L was used for primer extension on intracellular RS virus mRNA (Fig. 3*b*). Consistent with the RNA blot hybridization experiments, the greater fraction of primer hybridized to the 22K mRNA, as evidenced by the long, 22K-specific sequencing ladder. A smaller fraction of primer hybridized to L mRNA, as evidenced by the presence of the characteristic strong-stop doublets 1 and 2. Whereas oligonucleotide L hybridized only to L mRNA (Figs. 2*a* and 3*a*), oligonucleotide 22K/L hybridized to both the 22K and the L mRNAs (Figs. 2*b* and 3*b*). This confirmed the existence and the map position of the common sequence in the two mRNAs.

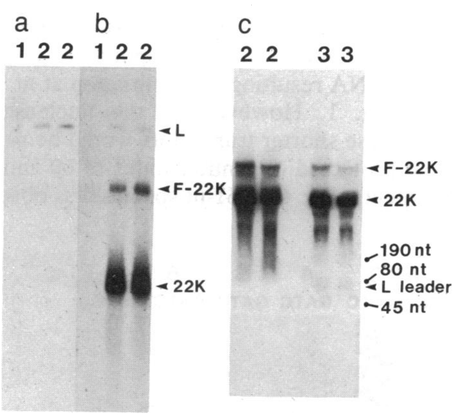

FIG. 2. Blot hybridization of RS virus mRNAs with oligonucleotides specific to the 22K and L genes. (*a* and *b*) mRNAs from uninfected (lanes 1) and RS virus-infected (lanes 2) cells were electrophoresed in a 0.8% agarose gel containing 2.2 M formaldehyde (14), blotted to nitrocellulose, and hybridized to 5′ end-labeled oligonucleotides L (*a*) and 22K/L (*b*). (*c*) Additionally, the following RS virus mRNAs were separated in a 1.5% agarose gel containing 2.2 M formaldehyde: RS virus mRNA (lanes 2) and RS virus mRNA that was pretreated with RNase H in the presence of oligo(dT) to remove poly(A) (lanes 3) as described by Wilde and Morrison (15). The RNAs were electroblotted onto Zeta-Probe paper (Bio-Rad), which efficiently binds small RNAs, and hybridized to 5′-end-labeled oligonucleotide 22K/L. The positions of oligodeoxynucleotides of 190, 80, and 45 nt electrophoresed in parallel are shown. Each pair of RS virus mRNA samples represents two independent preparations.

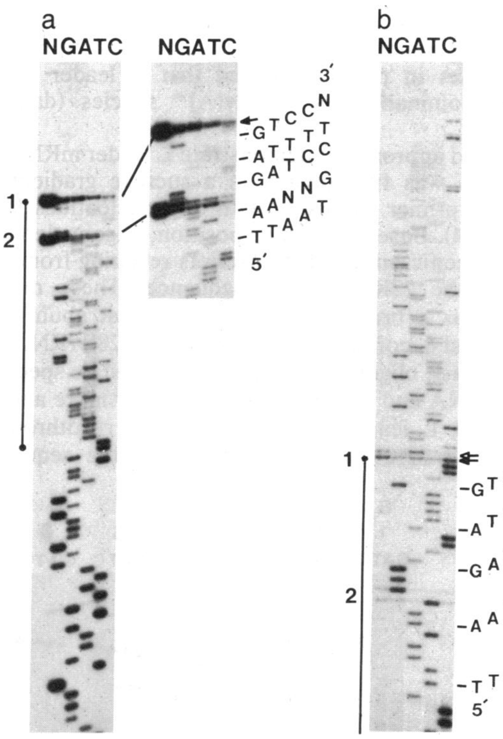

FIG. 3. Mapping and sequencing of mRNAs of the 22K and L genes. Oligonucleotide L (*a*) and oligonucleotide 22K/L (*b*) were 5′-end-labeled, hybridized to RS virus mRNA, and extended by reverse transcription in the presence of dideoxynucleotides (lanes G, A, T, and C) and without dideoxynucleotides (lanes N). Portions of sequencing gels of the reaction products are shown. Band 1 represents the 5′ end of the L mRNA, and band 2 represents a product of apparent premature transcriptional termination by reverse transcriptase (see the text). Deduced nucleotide sequences (vRNA-sense) are shown. Solid arrows indicate bands representing the exact 5′ end of the L mRNA; the open arrow indicates an artifactual end-addition product that is characteristic of primer extension on mRNA (8, 12).

Transcription of the L Gene. As is characteristic of nonsegmented negative-strand viruses (3, 11, 12, 16–18), the genes of RS virus end with semiconserved 12- to 13-nt gene-end sequences that are thought to direct polyadenylylation and transcriptional termination (2). Because transcription of the L gene initiates within the 22K gene, the 22K gene-end signal (3′ UCAAUAAAUUUU 5′, nt 142–153 in Fig. 1) is located within the L gene. Thus the synthesis of full-length L mRNA presumably can occur only if the polymerase transcribes across the 22K gene-end signal without terminating or synthesizing polyadenylate.

Some mechanism might exist that inhibits utilization of the 22K gene-end signal during L-gene transcription. Alternatively, the RS virus polymerase is known to occasionally fail to utilize gene-end sequences, resulting in the synthesis of polycistronic transcripts that are faithful copies of two or more adjacent genes and their intergenic sequences (2, 5, 7, 9, 10). This same mechanism could also generate full-length L mRNA. If so, the most abundant product of L-gene transcription would be predicted to be a 68-nt (exclusive of possible 3′ polyadenylate) RNA, designated L leader RNA, resulting from termination at the internal gene-end signal. Full-length L mRNA, resulting from occasional readthrough, would be much less abundant.

To test these possibilities, intracellular RS virus mRNA was analyzed by blot hybridization using oligonucleotide probes and a matrix that efficiently binds small RNAs (Fig. 2*c*). A candidate L leader mRNA was identified when the RS virus mRNA had been pretreated with RNase H in the presence of oligo(dT) to remove 3′ polyadenylate, which characteristically is heterogeneous in length and could obscure detection of a species containing a unique sequence of only 68 nt. Examination of poly(A)$^-$ and total [non-poly(A)-selected] RNA in parallel showed that L leader mRNA existed predominantly as a poly(A)$^+$ species (data not shown).

As a second approach to demonstrate L leader mRNA, RS virus mRNA was fractionated in a sucrose gradient and analyzed by primer extension with oligonucleotides L and 22K/L (Fig. 4). Based on its map position, oligonucleotide L would be specific only to L transcripts resulting from readthrough of the 22K gene-end sequence. These mRNAs sedimented as a broad peak and were most abundant in fractions 5 and 6, coincident with the peak of 28S rRNA (Fig. 4*a*). In contrast, oligonucleotide 22K/L would be specific to the 22K mRNA and to L mRNAs of 55 nt or more and thus would detect L leader mRNA as well as readthrough L mRNAs. The 22K mRNA (identified by the sequencing ladder extending above doublet 1) was distributed among fractions 2–4, with a peak in fraction 3 (Fig. 4*b*). L mRNAs (identified by doublets 1 and 2) were distributed throughout fractions 1–7, with one peak in fractions 2 and 3 and a second in fractions 5 and 6. Based on the results described above for oligonucleotide L, the L mRNA in fractions 4–6 was readthrough L mRNA. However, most of the L-specific transcripts detected by oligonucleotide 22K/L were contained in fractions 2 and 3, coincident with 4S RNA. Thus, this material was of the appropriate size and sequence to be L leader mRNA.

As a third approach to analyzing L-gene transcripts, a 128-nt synthetic cDNA representing the 22K–L gene junction was cloned in a plasmid containing SP6 and T7 promoters (Fig. 5) for the synthesis of vRNA-sense and mRNA-sense ^{32}P-labeled RNA probes. The length and sequence of the cloned cDNA were designed (Fig. 5) so that the vRNA-sense probe would, upon hybridization to mRNA and digestion with single-strand-specific RNase T1, give rise to protected oligonucleotides of different sizes for the 22K, L, and L leader mRNAs.

As shown in Fig. 6 (lanes e and f), the protected oligonucleotides included (*i*) a small amount of full-length probe that probably represents hybridization to template DNA; (*ii*) species of the predicted sizes, 115, 102, and 84 nt, to represent the 22K, L, and L leader mRNAs, respectively; and (*iii*) an additional abundant 116-nt species as well as several minor bands. To confirm the identities of the protected oligonucleotides, nuclease mapping was also performed with RS virus mRNAs that had been size-fractionated in sucrose gradients as described for Fig. 4. These results (Fig. 6, lanes 1–8) confirmed that the 84-, 115-, and 102-nt oligonucleotides were specific to mRNAs in the 4S (L leader mRNA), 18S (22K mRNA), and 28S (L mRNA) size classes, respectively. In addition, on the basis of these results the 116-nt species appeared to be specific to the 22K mRNA and probably resulted from use of an alternative cleavage site 1 nt from the predicted site (see Fig. 5 legend).

The additional minor protected species shown in Fig. 6 have not been analyzed further and raise the formal possibility of a low level of transcriptional initiation or termination at alternative sites. Additionally, strong-stop doublet 2 in Fig. 3 could have represented a shorter subpopulation of L mRNA and L leader mRNA resulting from initiation at nt 103 of the sequence in Fig. 1. However, in the nuclease-mapping experiments, these shorter transcripts would be represented by additional protected oligonucleotides of 60 and 78 nt. In some experiments (Fig. 6), a 61-nt species was observed, but

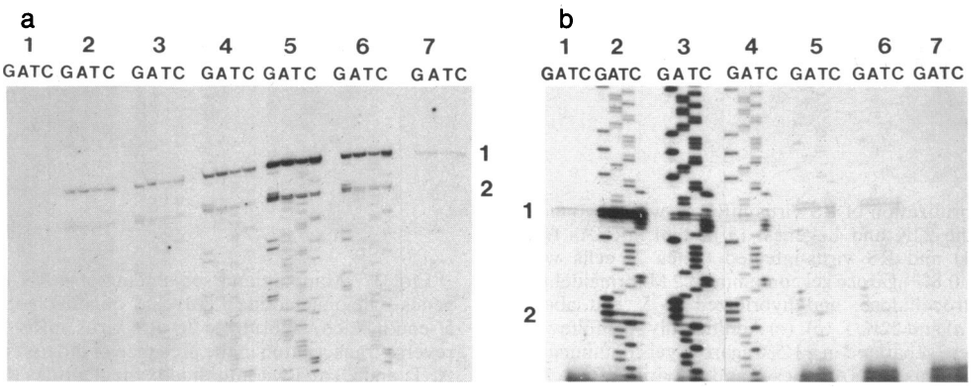

FIG. 4. Analysis of size-fractionated RS virus mRNA by primer extension. RS virus mRNAs were separated into fractions 1–7 by sedimentation in a sucrose gradient. Aliquots of mRNA from each fraction were hybridized with 5′-end-labeled oligonucleotide L (*a*) and oligonucleotide 22K/L (*b*) and extended with reverse transcriptase in the presence of dideoxynucleotides (lanes G, A, T, and C). The extension products were electrophoresed in sequencing gels, and autoradiograms are shown. Bands 1 and 2 are strong-stop doublets that are specific to L-gene transcripts. In a parallel gradient, RNA size markers of 4S, 18S, and 28S were contained, respectively, in fractions 2, 3–5, and 5–7 (data not shown).

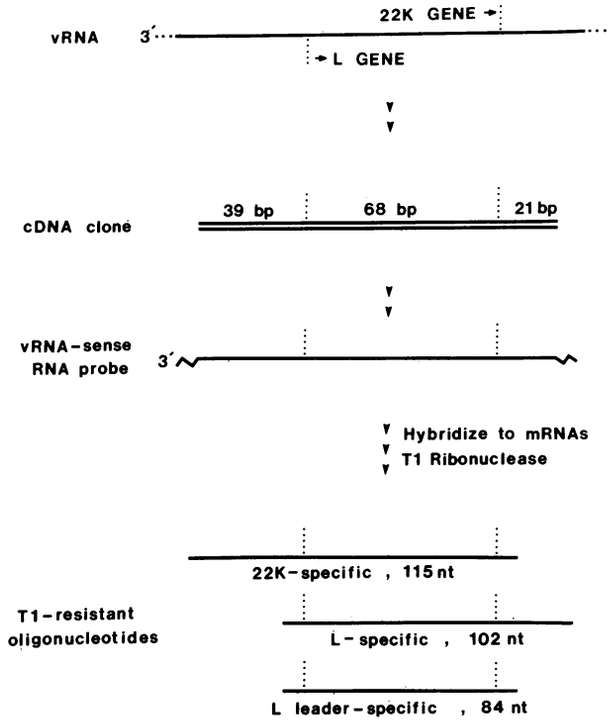

FIG. 5. Strategy for mapping transcripts of the L and 22K genes by hybridization and nuclease digestion. A 128-bp cDNA was constructed to contain nt 47–174 of the sequence shown in Fig. 1, spanning the junction of the 22K and L genes. The cDNA was designed so that the central 68 bp represented the overlap of the 22K and L genes, whereas the flanking 39-bp and 21-bp sequences were specific to the 22K gene alone and the L gene alone, respectively. The cDNA was placed under the control of flanking SP6 and T7 phage promoters in a plasmid vector (*Materials and Methods*). Transcription *in vitro* by SP6 polymerase yielded a radiolabeled 169-nt RNA probe that contained the 128-nt RS virus sequence (in vRNA sense) flanked by vector-specific sequences; transcription by the T7 polymerase yielded a 190-nt RNA (not shown above) that contained the 128-nt RS virus sequence (in mRNA sense) and represented the negative control. In the vRNA-sense probe, the vector-specific sequences immediately flanking the complement of position 47 were 3' CTCCTAGGGGG (the complement of position 47 is underlined); vector sequences immediately flanking the complement of position 174 were 3' CCCCATGGCTC (the complement of position 174 is underlined). The RNA probe was hybridized to RS virus mRNA and digested with an excess of RNase T1, which cleaves single-stranded RNA at the 3' side of guanosine residues. The predicted protected oligonucleotides are as follows: RNA probe that was hybridized to 22K mRNA would be cleaved at the 3' side of the complements of positions 47 and 162, resulting in a 115-nt protected oligonucleotide. As shown in Fig. 6 and described in the text, in practice an approximately equal amount of a 116-nt oligonucleotide was also observed. A probable explanation is that cleavage at the 3' site of the complement of position 47 was inefficient because that nucleotide was part of the hybridized duplex. The next available site for cleavage would be 1 nt away, in the vector-specific sequence; cleavage at this alternate site would generate a 116-nt oligonucleotide. RNA probe that was hybridized with L mRNA would be cleaved at the 3' side of the complement of position 78 and within the vector-specific flanking sequence 5 nt from the end of the cDNA at position 174. This would result in a 102-nt protected oligonucleotide. RNA probe that was hybridized to L leader mRNA would be cleaved at the 3' sides of the complements of positions 78 and 162, resulting in an 84-nt protected oligonucleotide.

the failure to reproducibly detect both predicted species suggested that nt 86 (Fig. 1) was the sole major transcriptional start site for the L gene. Strong-stop doublet 2 presumably was an artifact of the primer-extension reaction.

In summary, several independent methods (Figs. 2–6) identified two L-gene transcripts: the relatively nonabundant full-length L mRNA and the more abundant L leader mRNA that is generated by polyadenylylation and termination at the internal gene-end signal.

DISCUSSION

For all nonsegmented negative-strand RNA viruses examined to date, the viral polymerase initiates at the 3' vRNA end and transcribes the genes sequentially in their 3' to 5' order (17, 19). mRNAs are transcribed from nonoverlapping cistrons that contain transcriptional start and stop signals at each end and are separated by short intergenic regions (3, 11, 12, 18). Gene overlap does exist for the P and C genes of Sendai, measles, canine distemper, and parainfluenza type 3 viruses, but this is not an exception to the general transcriptional strategy, as both genes are expressed as a single mRNA containing two open reading frames (4, 13, 20). In contrast, here we showed that RS virus vRNA contains two overlapping genes, the 22K and L genes, that are expressed as separate mRNAs containing a common sequence.

In previous work, UV target-size analysis confirmed that the first nine genes in the RS virus genetic map are transcribed sequentially (21). The transcriptional map position for the tenth gene, the L gene, could not be determined unambiguously, but the low molar abundance of L mRNA and the pattern of readthrough mRNAs suggested that the L gene was last in the order of transcription (5). However, the finding here that the 22K and L genes overlap but are expressed as separate mRNAs is incompatible with the canonical model of linear, sequential transcription (17, 19).

One possibility was that both genes were transcribed into a single precursor that was cleaved by alternative pathways to generate the 22K and L mRNAs: cleavage at the L gene-start sequence would generate full-length L mRNA and a 3'-truncated 22K mRNA, and cleavage at the 22K gene-end sequence would generate full-length 22K mRNA and a 5'-truncated L mRNA. However, neither truncated mRNA was detected here. This provided evidence against a precursor-processing model for RS virus transcription. A second possibility was that transcription of the L gene was initiated from a second, internal polymerase-entry site (promoter), which would be unprecedented among negative-strand RNA viruses. However, if a separate promoter does exist for the L gene, it must be substantially less active than the promoter for the other nine genes because the transcripts of the L gene, including L leader mRNA, were less abundant on a molar basis than were transcripts of the 22K gene (Figs. 2, 3*b*, 4, and 6) or any other viral gene (3, 7).

The L gene-start and 22K gene-end signals, segregated in RS virus vRNA from adjoining intergenic, gene-start, and gene-end sequences, appeared to be sufficient alone for precise transcriptional initiation and termination/polyadenylylation, respectively. Thus, these sequences appear to function independently as signals for the polymerase.

Inefficient recognition of gene-end signals occurs at most RS virus genes, as evidenced by readthrough mRNAs representing various combinations of tandem genes (5, 7). Readthrough transcription is generally assumed to represent polymerase error, giving rise to nonabundant, nonessential RNAs. However, for RS virus, this now appears to be the only pathway for the synthesis of full-length L mRNA.

For nonsegmented negative-strand viruses in general, attenuation of sequential transcription appears to be the principal mechanism responsible for the nonequimolar accumulation of the various viral mRNAs and proteins (22). The presence of the 22K gene-end signal within the RS virus L gene appears to be a novel mechanism for obtaining additional attenuation of L-gene transcription. Also, the L gene-start signal contains two base differences from the otherwise

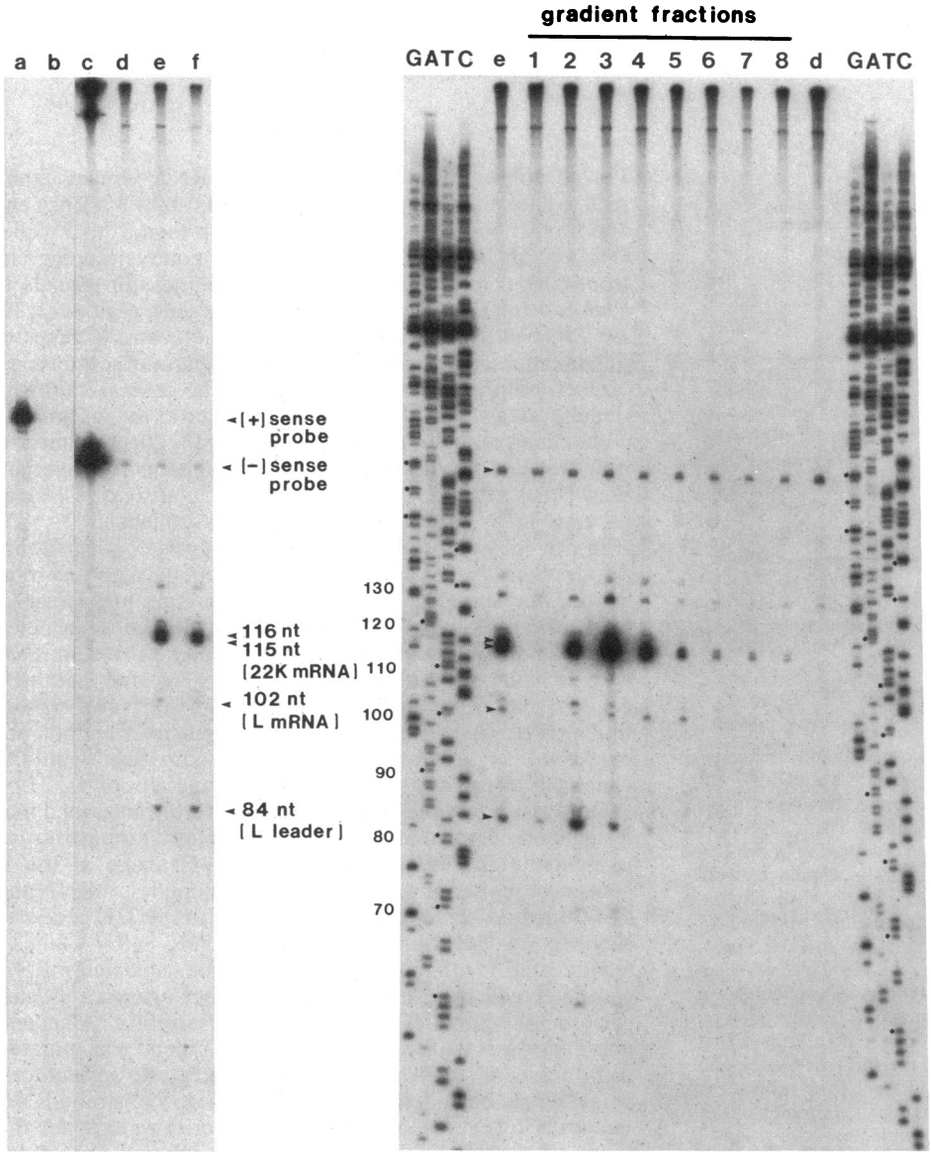

FIG. 6. Mapping of transcripts of the 22K and L genes by their ability to protect a ^{32}P-labeled RNA probe from digestion with RNase T1. Using the strategy outlined in Fig. 5, the 169-nt vRNA-sense RNA probe was hybridized with the following RNA samples, digested with RNase T1, and analyzed by electrophoresis in sequencing gels. Lanes: d, uninfected-cell mRNA; e, RS virus mRNA; f, total [non-poly(A)-fractionated] intracellular RS virus RNA; 1–8, size-fractionated RS virus mRNAs of increasing size representing gradient fractions 1–8 shown in Fig. 4. Control lanes: c, the undigested 169-nt vRNA-sense probe; a, the 190-nt ^{32}P-labeled RNA containing the 128-nt RS virus cDNA sequence (in mRNA sense) with flanking vector sequences; b, the mRNA-sense probe that had been hybridized with RS virus mRNA and digested with RNase T1. Lanes G, A, T, and C: molecular length markers generated by dideoxynucleotide sequencing of a heterologous cDNA clone. These were electrophoresed in parallel, and the calculated nucleotide lengths of the bands are marked. Because these markers were generated using a primer that lacks 5′ phosphate, the migration of the bands would be expected to be reduced by the equivalent of 0.5–1.0 nt, and the size estimates for the protected oligonucleotides were adjusted accordingly. Note that the gel deformed during transfer and drying so that the positions of bands were displaced downwards slightly across the gel from left to right.

conserved RS virus gene-start sequence and might also be a factor affecting L-gene expression.

The existence of an attenuator (the 22K gene-end signal) within the L gene and the possible existence of a separate promoter for the L gene raise the possibility that its transcription can be modulated independently of the other viral genes. For example, the frequency of readthrough of the attenuator in the L gene could be altered by viral or cellular factors or by changes in the efficiency of the viral polymerase due to the accumulation of mutations (23). Any increase or decrease in the availability of the polymerase L protein could have a dramatic effect on levels of transcription, replication, and virion production, altering the balance between acute and persistent infection. Although L leader mRNA might be a nonfunctional by-product of transcriptional attenuation, it does contain two open translational reading frames and therefore could be a functional mRNA.

1. McIntosh, K. & Chanock, R. M. (1985) *Virology*, ed. Fields, B. N. (Raven, New York), pp. 1285–1304.
2. Collins, P. L., Dickens, L. E., Buckler-White, A., Olmsted, R. A., Spriggs, M. K., Camargo, E. & Coelingh, K. V. W. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4594–4598.
3. Spriggs, M. K. & Collins, P. L. (1986) *J. Virol.* **59**, 646–654.
4. Spriggs, M. K. & Collins, P. L. (1986) *J. Gen. Virol.* **67**, 2705–2719.
5. Collins, P. L. & Wertz, G. W. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3208–3212.
6. Venkatesan, S., Elango, N. & Chanock, R. M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1280–1284.
7. Collins, P. L., Huang, Y. T. & Wertz, G. W. (1984) *J. Virol.* **49**, 572–578.
8. Collins, P. L., Huang, Y. & Wertz, G. W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7683–7687.
9. Collins, P. L. & Wertz, G. W. (1985) *J. Virol.* **54**, 65–71.
10. Elango, N., Satake, M. & Venkatesan, S. (1985) *J. Virol.* **55**, 101–110.
11. Rose, J. K. (1980) *Cell* **19**, 415–421.
12. Gupta, K. C. & Kingsbury, D. W. (1984) *Nucleic Acids Res.* **12**, 3829–3841.
13. Bellini, W. J., Englund, G., Rozenblatt, S., Arnheiter, H. & Richardson, C. D. (1985) *J. Virol.* **53**, 908–919.
14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
15. Wilde, A. & Morrison, T. (1984) *J. Virol.* **51**, 71–76.
16. Schubert, M. & Lazzarini, R. A. (1981) *Virology* **38**, 256–262.
17. Ball, L. A. & Wertz, G. W. (1981) *Cell* **26**, 143–144.
18. Shioda, T., Iwasaki, K. & Shibuta, H. (1986) *Nucleic Acids Res.* **14**, 1545–1563.
19. Banerjee, A. K., Abraham, G. & Colonno, R. J. (1977) *J. Gen. Virol.* **34**, 1–8.
20. Giorgi, C., Blumberg, B. M. & Kolakofsky, D. (1983) *Cell* **35**, 829–836.
21. Dickens, L. E., Collins, P. L. & Wertz, G. W. (1985) *J. Virol.* **52**, 364–369.
22. Iverson, L. E. & Rose, J. K. (1981) *Cell* **23**, 477–484.
23. Schubert, M., Harmison, G. G. & Meier, E. (1984) *J. Virol.* **51**, 505–514.