JOURNAL OF VIROLOGY, Sept. 1991, p. 4741–4748
0022-538X/91/094741-08$02.00/0

Vol. 65, No. 9

Host-Derived 5′ Ends and Overlapping Complementary 3′ Ends of the Two mRNAs Transcribed from the Ambisense S Segment of Uukuniemi Virus

JAN FREDRIK SIMONS* AND RALF F. PETTERSSON

Stockholm Branch, Ludwig Institute for Cancer Research, Box 60202, S-10401 Stockholm, Sweden

Received 26 December 1990/Accepted 30 May 1991

Two mRNAs, coding for the N and NS_S proteins, are transcribed from the small (S) Uukuniemi virus RNA segment by an ambisense strategy (J. F. Simons, U. Hellman, and R. F. Pettersson, J. Virol. 64:247–255, 1990). In this report, we describe the analysis of the 5′ and 3′ ends of the two mRNAs. Primer extension as well as cloning and sequencing of individual mRNAs showed that the 5′ ends of both mRNAs contained nonviral sequences ranging from 7 to 25 residues in length (mean, 12 residues), indicating a cap-snatching mechanism similar to the one originally described for priming of influenza virus mRNA synthesis. In 35% of the cases, the first virion-specified nucleotide (an A residue) was substituted with a G residue. Between the translation termination codons of N and NS_S, there is a 74-residue-long noncoding intergenic region (Simons et al., J. Virol. 64:247–255, 1990). Nuclease protection assays using both RNA and DNA hybridization probes showed that the 3′ ends of the N and NS_S mRNAs overlap each other by about 100 nucleotides. The 3′ end of the NS_S mRNA extends into the coding sequence of the N mRNA, whereas the N mRNA is terminated just prior to the stop codon of NS_S. To our knowledge, this is the first example of overlapping complementary mRNAs in viruses with an ambisense coding strategy. No obvious transcription termination sequence was identified. However, because of a short palindromic sequence in the intergenic region, the 3′ ends of both mRNAs (and consequently also the template RNAs) can be folded into an A/U-rich hairpin structure. It remains to be determined whether this structure plays any role in transcription termination.

Although all members of the *Bunyaviridae* family have similar structures and modes of morphogenesis in the Golgi complex, significant differences exist in the organization and strategy of expression of the viral genes (3, 8, 14, 28). The genome of bunyaviruses, i.e., members of the *Bunyavirus*, *Hantavirus*, *Nairovirus*, *Phlebovirus*, and *Uukuvirus* genera, is divided into three RNA segments (L, M, and S) of negative polarity. The L segment encodes the RNA-dependent RNA polymerase, and the M segment encodes the two membrane glycoproteins, G1 and G2, which are cotranslationally cleaved from a precursor (8, 25, 30). In the case of members of the *Bunyavirus* and *Phlebovirus* genera, this precursor also contains a nonstructural polypeptide, NS_M (10, 17, 28), located either between G1 and G2 (*Bunyavirus*) or at the N terminus (*Phlebovirus*). The function of this protein, which is not present in the M RNA product of members of the other genera (28), is unknown.

The S segment shows the greatest variation in coding strategy. In members of the *Bunyavirus* genus, two proteins, the nucleocapsid protein N and a nonstructural protein NS_S, are translated from the same mRNA by using different and overlapping reading frames (11). Surprisingly, nairoviruses (31) and hantaviruses (27) do not encode any NS_S protein. The S RNA segment of phleboviruses (16) and uukuviruses (29), on the other hand, encodes both N and NS_S proteins translated from two subsegmental mRNAs transcribed from the S RNA by using an ambisense coding strategy. Thus, the N mRNA is transcribed from the 3′ half of the virion RNA (vRNA), whereas the NS_S mRNA is derived from the 3′ half of the full-length viral-complementary RNA (vcRNA). In the case of Uukuniemi (UUK) virus (*Uukuvirus*), both the vRNA and vcRNA strands (in a ratio of about 10:1) were found to be packaged into virions (29). The size and nucleotide sequence of the intergenic region (IR), located between the two coding regions, vary considerably between members of the phleboviruses and uukuviruses. Punta Toro (PT) virus (a phlebovirus) vRNA contains the longest IR so far found (359 residues) (9, 16). It is very A/U rich and can be folded into a large (about 200 residues) hairpin structure. Transcription of the two mRNAs terminates on both sides of the tip of this hairpin. The IRs of the vRNAs of the phleboviruses sandfly fever Sicilian virus (141 residues) (20), Rift Valley fever virus (82 residues), and Toscana virus (62 residues) (13) are relatively C rich and do not exhibit any clear secondary structure. In the case of UUK virus, the IR is also short (74 residues), contains 62% A+U, with short runs of A's and U's (29), and contains a short, almost perfect palindromic sequence that could form a hairpin structure. No sequences conserved between different bunyaviruses have been found within the intergenic or surrounding region. Thus, it is unclear what structural features determine transcription termination. The intergenic region of the S RNA segment of arenaviruses, which also exhibit an ambisense strategy, is rather short and can be formed into a fully base paired hairpin structure (1, 2).

The mRNAs of snowshoe hare (4), La Crosse (LAC) (22), and Germiston (5) viruses, all members of the *Bunyavirus* genus, have been found to contain short (10 to 18 nucleotides) nontemplated 5′-end sequences. These are assumed to be derived from the 5′ ends of host mRNAs by a mechanism similar to the cap-snatching originally described for influenza virus (18, 19). Bishop et al. (4) were able to clone and sequence the 5′ ends of five individual S RNA-sized molecules from snowshoe hare virus-infected cells. Three of these contained 13 to 14 nonviral nucleotides and had

* Corresponding author.

different sequences. Recently, Bouloy et al. (5) determined the sequences of the 5′ ends of several individual M and S mRNAs of Germiston virus and found 7 to 16 nonviral nucleotides. By primer extension, Patterson and Kolakofsky (22) also found 10 to 14 extra nucleotides at the 5′ end of LAC virus S mRNAs. Sequence determinations of a mixture of mRNAs indicated that these extra nucleotides had a heterogeneous sequence. Only one report on the analysis of the 5′ ends of a mRNA from a phlebovirus has been reported. Collett (7) found by primer extension analyses that the M mRNA of Rift Valley fever virus contained 12 to 14 extra nucleotides. Recently, nontemplated bases were also found at the 5′ ends of Tacaribe virus (an arenavirus) mRNAs. In this case, the nucleotides extended only 0 to 4 residues beyond the 3′ ends of the template RNA (12, 24).

To characterize the transcripts derived from the S RNA of UUK virus, which is distantly related to the phleboviruses (25, 29), we have analyzed the 5′ and 3′ ends of both mRNAs. Nonviral, probably host-specified, sequences were found at the 5′ ends of the two S RNA-derived mRNAs, indicating that these mRNAs also steal short sequences from host mRNAs to be used as primers for transcription initiation. We also found that the 3′ ends of the mRNAs overlap by about 100 residues, suggesting that they are able to self-anneal.

MATERIALS AND METHODS

Virus and cells. The origin and preparation of stock virus from the prototype strain S23 of UUK virus have been described previously (23). The titer of the virus stock was 3 $\times$ 10^8 to 4 $\times$ 10^8 PFU/ml. Cells were infected with a multiplicity of about 10 PFU per cell. The virus was grown in BHK-21 cells, clone 13, in plastic flasks or dishes in Eagle's minimal essential medium supplemented with 0.2% bovine serum albumin (BSA), 20 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES; pH 7.4), and antibiotics (23).

Synthetic oligonucleotide probes and primers. The position numbers of the nucleotides listed below refer to the virion complementary-sense RNA in reference 29. For the synthesis of single-stranded DNA probes by polymerase chain reaction (PCR), two primers were used: RP-58 (complementary to positions 993 to 1010) and RP-81 (555 to 572). For primer extension reactions, RP-125 (complementary to positions 26 to 43) was used for the N mRNA, and RP-130 (1680 to 1697) was used for the NS_S mRNA or the virion S RNA. For cloning of the 5′ ends of the N and NS_S mRNAs, primers RP-78 (complementary to positions 449 to 467) and RP-50 (1180 to 1198), respectively, were used for the cDNA synthesis step. For PCR amplification, primers RP-79 (complementary to positions 225 to 242) and RP-48 (1444 to 1461) or RP-130 (see above), respectively, were used in combination with RP-119 (5′-AGCGGAATTCTTTTTTTTTTTT-3′), which includes an *Eco*RI site.

Northern (RNA) blot analysis was done with four primers: RP-45 (complementary to positions 879 to 896), RP-104 (complementary to positions 838 to 855), RP-142 (727 to 766), and RP-160 (786 to 830).

RP-125 and RP-130 (see above) were used as primers for sequencing the 5′ ends of the cloned N and NS_S mRNAs, respectively.

Isolation of virion RNA and mRNAs. The procedures for isolating virion RNA from purified virus particles and mRNAs from virus-infected cells have been described in detail previously (29, 30).

Generation of riboprobes. Relevant cDNA fragments were inserted into the polylinker region of plasmid pGEM-3 (Promega Corp., Madison, Wis.), cleaved with appropriate restriction enzymes within the insert, and transcribed in the presence of ^{32}P-labeled CTP (Amersham Corp.), using either the SP6 or T7 promoter. The DNA template was subsequently removed by DNase treatment; following phenol-chloroform extraction, the riboprobe was precipitated with ethanol. In all of these steps, standard protocols either recommended by the manufacturer or described by Sambrook et al. (26) were used.

Generation of single-stranded DNA probes by PCR. Double-stranded DNA was first amplified under standard conditions by PCR, using appropriate pairs of primers. The amplified fragments were then separated on an 1% low-gelling-temperature agarose gel. The desired band was cut out, and the DNA was recovered by phenol-chloroform extraction followed by precipitation with ethanol. To generate radioactive single-stranded probes, 100 ng of gel-purified DNA fragments were then used as the template in a reaction mixture containing 10 pmol of only one of the primers, 50 mM KCl, 20 mM Tris-HCl (pH 8.4), 2.5 mM $MgCl_2$, 100 μg of BSA per ml, 4 μM each deoxynucleoside triphosphate, 25 μCi of [α-^{32}P]dCTP, and 4 U of *Taq* polymerase (Perkin-Elmer AB, Sundbyberg, Sweden) in a total volume of 100 μl. The reactions were run for 30 cycles, and the unincorporated nucleotides were removed by chromatography on a Sephadex G-50 column (Pharmacia-LKB, Uppsala, Sweden). The labeled probes thus obtained were precipitated with ethanol and used directly in the hybridization assays.

Mapping of N and NS_S mRNA 3′ ends with mung bean nuclease. Five micrograms of sucrose gradient-purified mRNA (29, 30) was precipitated with ethanol together with 150,000 cpm of single-stranded ^{32}P-labeled DNA probe, prepared as outlined above, and resuspended in 20 μl of hybridization solution [80% formamide, 40 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES; pH 6.4), 400 mM NaCl, 1 mM EDTA (pH 8.0)]. The samples were heated to 65°C for 10 min and hybridized overnight at 30°C. The volume was then increased to 320 μl by addition of nuclease buffer (0.28 M NaCl, 50 mM sodium acetate [pH 4.5], 4.5 mM $ZnSO_4$), and the single-stranded nucleic acids were digested with 300 U of mung bean nuclease (Pharmacia) at 30°C for 60 min. The reaction was terminated by addition of 80 μl of stop buffer (4 M ammonium acetate, 20 mM EDTA [pH 8.0], 40 μg of tRNA per ml). The material was precipitated with ethanol, resuspended in loading buffer containing formamide, and analyzed on an 8% sequencing gel.

Mapping of N and NS_S mRNA 3′ ends with RNases A and T_1. Five micrograms of mRNA was precipitated with ethanol and resuspended in 30 μl of hybridization buffer (see above). The riboprobe was resuspended in 50 μl of the same solution, and 1/50 was added to the test RNA, heated to 85°C for 5 min, and then incubated for 16 h at 45°C. Three hundred microliters of RNase digestion buffer (10 mM Tris-HCl [pH 7.5], 5 mM EDTA, 300 mM NaCl) containing RNase A (40 μg/ml) and RNase T_1 (2 μg/ml) (both enzymes supplied by Boehringer Mannheim) was added, and the mixture was incubated at 30°C for 60 min. The digestion was terminated by addition of 20 μl of 10% sodium dodecyl sulfate and 50 μg of proteinase K (Boehringer Mannheim) and incubated for 15 min at 37°C. The sample was then extracted with phenol and chloroform, ethanol precipitated in the presence of 20 μg of calf liver tRNA (Boehringer Mannheim), resuspended in loading buffer, and analyzed on an 8% sequencing gel.

Northern blot analysis. The RNA samples were denatured

in 50% formamide and 6.5% formaldehyde for 3 min at 70°C, run on a 1% agarose gel containing 6.6% formaldehyde, and subsequently transferred to Hybond-N filters (Amersham) (29). Hybridization was performed in the absence of formamide according to standard protocols.

Primer extensions. Template RNA was precipitated together with labeled primer (7.5×10^5 cpm/pmol; 150,000 cpm per reaction), resuspended in 30 μl of hybridization buffer (80% formamide, 40 mM PIPES [pH 6.4], 400 mM NaCl, 1 mM EDTA), and incubated for 40 h at 30°C. After precipitation, the annealed primer and template were dissolved in reverse transcriptase buffer (50 mM Tris-HCl [pH 8], 5 mM $MgCl_2$, 5 mM dithiothreitol, 50 mM KCl, 50 mg of BSA per ml, 550 μM each deoxynucleoside triphosphate), and the primer was extended with avian myeloblastosis virus reverse transcriptase (20 U; Promega Corp.) for 90 min at 42°C. The products were reprecipitated with ethanol and run on an 8% sequencing gel.

Cloning and sequencing of N and NS_S mRNA 5′ ends. Cytoplasmic RNA was extracted from UUK virus-infected cells and fractionated on a 15 to 30% sucrose gradient (29, 30). Synthesis of cDNA was performed under standard conditions (26, 29) following denaturation of the S mRNA template by 300 μM CH_3HgOH for 3 min at room temperature. The template RNA was subsequently denatured in 300 μM NaOH at 50°C for 30 min and incubated for 30 min at the same temperature after addition of acetic acid to a final concentration of 2%. The cDNA was then separated from the primers and other low-molecular-weight solutes by four washes in a Centricon 30 microconcentrator (Amicon), precipitated with ethanol, and poly(A) tailed by using a terminal deoxynucleotidyltransferase-tailing kit (Boehringer Mannheim). After phenol extraction and ethanol precipitation, the cDNA was PCR amplified by using an oligo(dT) primer containing a 5′-terminal *Eco*RI site and a primer complementary to a downstream viral sequence. The PCR-amplified material was made blunt end by DNA polymerase I (Klenow fragment) and cloned either into the *Hin*dII site of a pGEM-3 (Promega) vector (NS_S mRNA clones) or into the *Ava*I-*Eco*RI sites of a pGEM-3 Zf(−) vector (Promega) (N mRNA clones).

Sequence determination was performed by a modification of the dideoxy-chain termination method, using double-stranded cDNA inserts as the template and Sequenase (United States Biochemical Corp.). The sequencing primers used are listed above.

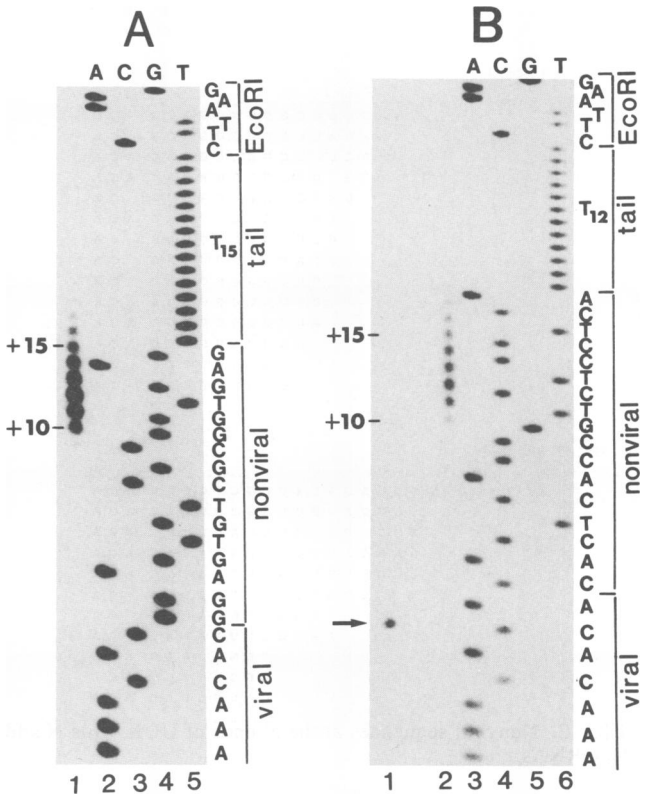

FIG. 1. Autoradiograms of primer extensions and sequence analyses of the 5′ ends of the N and NS_S mRNAs. Primer extensions were carried out by using N (panel A, lane 1) and NS_S (panel B, lane 2) mRNAs as templates and oligonucleotides complementary to the 5′ ends of the respective mRNAs as primers (see Materials and Methods). As a control, primer extension was also done by using the vRNA as the template (panel B, lane 1). The arrow indicates the position of this product and shows that it migrated to a position close to that corresponding to the 5′ end of the template vRNA. +10 and +15 indicate the number of the extra nucleotides found at the 5′ ends of the mRNAs. Also shown are two examples of sequences obtained for the cloned 5′ ends of an N (panel A, lanes 2 to 5) and an NS_S (panel B, lanes 3 to 6) mRNA species. The 5′ ends were cloned and sequenced as described in Materials and Methods. viral, virus-specified sequences; nonviral, probably host-specified sequences; tail, sequence derived from the poly(dA)-tailing procedure; *Eco*RI, linker sequence introduced to facilitate cloning.

RESULTS

Cloning and sequencing of the 5′ ends of N and NS_S mRNAs. To investigate the structures of the 5′ ends, a representative number of 5′ ends of the N and NS_S mRNAs were cloned and sequenced. Complementary DNA was synthesized by using as templates N and NS_s mRNAs, separated from M and L mRNAs by sucrose gradient centrifugation, and primer RP-78 or RP-50, 449 to 467 (N mRNA) or 523 to 541 (NS_S mRNA) nucleotides downstream from the 5′ end of the respective mRNAs. The cDNA was tailed with dA residues and PCR amplified by using an oligo(dT) primer containing an *Eco*RI site (RP-119) and virus-specific primers located approximately 250 nucleotides from the 5′ ends of the two mRNAs (RP-79 for N mRNA and RP-48 for NS_S mRNA). To increase the specificity of the PCR, these primers were different from those used for first-strand cDNA synthesis. The amplified DNA was cloned into pGEM-3 (NS_S mRNA clones) or pGEM-3 Zf(−) (N mRNA clones) vectors, and 16 N and 9 NS_S mRNA-specific clones were isolated for sequence analysis. Sequencing of the inserts (lanes 2 to 5 and 3 to 6 in Fig. 1A and B, respectively) showed that all contained extra 5′ nonviral nucleotides ranging from 7 to 25 residues, with 11 to 14 nucleotides being the most common lengths (Fig. 2). In nine cases (36%), the first virus-encoded nucleotide (an A residue), complementary to the 3′-terminal U residue of the template vRNA, was missing. In eight of these mRNAs, this nucleotide was replaced by a G residue. Comparison of the sequences with each other did not reveal any apparent similarities, although some nucleotides seemed to be favored in certain positions. In cases in which the first virus-encoded A was present, the last host-derived nucleotide was a C in 13 of 15 mRNAs. In addition, a C at position −3 was frequently observed in these cases (10 of 15). The base composition of the pooled sequenced 5′ ends was 16.9% A, 31.7% C, 27.3% G, and 24.1% U.

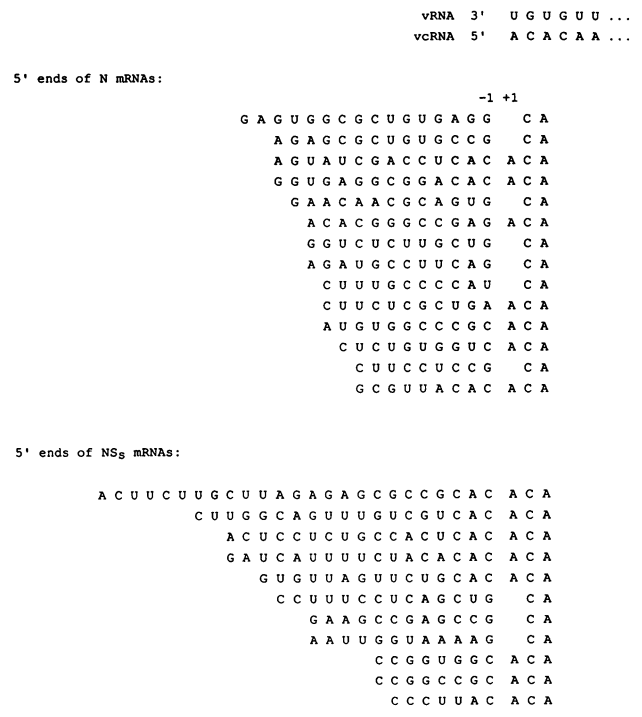

FIG. 2. Nonviral sequences at the 5′ ends of UUK virus N and NS$_S$ mRNAs.

Sequencing of the N mRNAs downstream from the virus-encoded 5′ end (data not shown) revealed an error in the previously published sequence (29). The error was due to the fact that a wrong nucleotide was included in the primer used for cDNA synthesis. The correct sequence of the extreme 3′ end of the S vRNA should read 5′-TTGGAGGTCTTTGT GT-3′, where the underlined G is a C in the original publication. This change results in a further increase of the stability of the base-paired 3′ and 5′ stem structure of the S vRNA segment.

Determination of the average length of the nonviral sequences. To determine the actual length distribution of the nonviral 5′ ends more accurately, total N and NS$_S$ mRNAs were analyzed by primer extension, using the same primers as for sequencing of the 5′ ends of the respective mRNAs. Lane 1 in Fig. 1A and lane 2 in Fig. 1B represent the primer extensions of the N and NS$_S$ mRNAs, respectively. As determined by densitometric scanning, about 85% of the mRNAs contained host-derived sequences of 10 to 16 nucleotides, the mean being 12 residues. The rest were either shorter or longer than this. A longer exposure of the gel revealed sequences as short as 8 and as long as 20 residues. Using vRNA as the template and RP-130 (see Materials and Methods) as the primer, a cDNA product (extended toward the 5′ end) with the expected length was obtained (Fig. 1B, lane 1). For an unknown reason, this product migrated between positions +1 and +2 compared with the marker (lanes 3 and 4).

Mapping of the 3′ ends of the N and NS$_S$ mRNAs by nuclease protection assays. The length of the N (about 800 residues) and NS$_S$ (about 850 residues) mRNAs (29) suggest that they terminate somewhere around the 74-nucleotide-long noncoding intergenic region. As a first step to investigate the mechanism of transcription termination, we mapped the locations of the 3′ ends of these mRNAs. To this end, nuclease protection assays were performed. Since the mobilities between single-stranded RNA (ss-RNA) and ssDNA differs somewhat and we used ssDNAs as markers, the mapping was carried out by using both ssRNA and ssDNA probes. The riboprobes were synthesized in vitro from pGEM-3 vectors containing the appropriate cDNA inserts by using either the SP6 or the T7 promoter. The ssDNA probes were generated by PCR. In the latter case, double-stranded DNA was first amplified and purified on an agarose gel and subsequently used as a template in a reaction performed in the presence of only one primer and ^{32}P-labeled dCTP.

By using ssDNA probes complementary to the 3′ ends of both mRNAs (Fig. 3D), mung bean nuclease-resistant, protected fragments of 236 (NS$_S$ mRNA; Fig. 3A, lane 3) and 313 to 315 (N mRNA; Fig. 3A, lane 4) nucleotides, respectively, were obtained. By using two different riboprobes complementary to the N mRNA (Fig. 3E and F), RNase A+T$_1$-resistant fragments of 70 to 73 and 150 to 153 nucleotides were obtained (Fig. 3B, lane 3; Fig. 3C, lane 3). A riboprobe complementary to the NS$_S$ mRNA (Fig. 3E) also gave a protected fragment of 70 to 73 nucleotides (Fig. 3B, lane 4). As is evident from Fig. 3A to C, the bands corresponding to the protected fragments are somewhat heterogeneous. Whether this is due to heterogeneous 3′ ends or to nibbling of the ends of the protected fragments is unclear. The results obtained with the two different kinds of probes were in good agreement with each other. In summary, the 3′ end of the N mRNA could be mapped to a eight-residue-long region, whereas that of the NS$_S$ mRNA was mapped to a five-residue-long region (Fig. 4). These results mean that the two mRNAs overlap each other by about 100 nucleotides.

Confirmation of the mRNA 3′ ends by Northern blot analyses. To confirm the results presented above, the N and NS$_S$ mRNAs were subjected to Northern blot analyses using as probes pairs of synthetic oligonucleotides corresponding to sequences just upstream (RP-104 and RP-160) and downstream (RP-45 and RP-142) of the predicted 3′ ends (Fig. 5B). The former probes hybridized to the N (Fig. 5A, lane 2) and NS$_S$ (lane 7) mRNAs, respectively, whereas the latter ones failed to do so (lanes 4 and 9, respectively), in full concordance with the results obtained from the nuclease protection assays. None of the four probes hybridized to RNA from uninfected cells (lanes 1, 3, 6, and 8). Lanes 5 and 10 show that probes RP-45 and RP-142, which failed to hybridize to the mRNAs, were able to detect the full-length S vcRNA and vRNA, respectively.

DISCUSSION

In this study, we have analyzed the structures of the 5′ ends and mapped the positions of the 3′ ends of the two subsegmental mRNAs transcribed from the S RNA segment of UUK virus. Our refined view on the strategy of expression of the S segment is summarized in Fig. 6. Two main conclusions can be drawn from the results. First, the 5′ ends of the N and NS$_S$ mRNAs contain nonviral sequences, about 10 to 15 nucleotides in length. Second, the 3′ ends of the two S segment-derived mRNAs extend across the 74-residue-long noncoding intergenic region, resulting in an about 100-residue overlap.

For both the N and NS$_S$ mRNAs, the 5′ nonviral sequences were heterogeneous, with a somewhat higher G+C (59%) than A+U (41%) content. Since this base composition may reflect that of the 5′ ends of host mRNAs in general, this difference may be without significance. Primer extension of

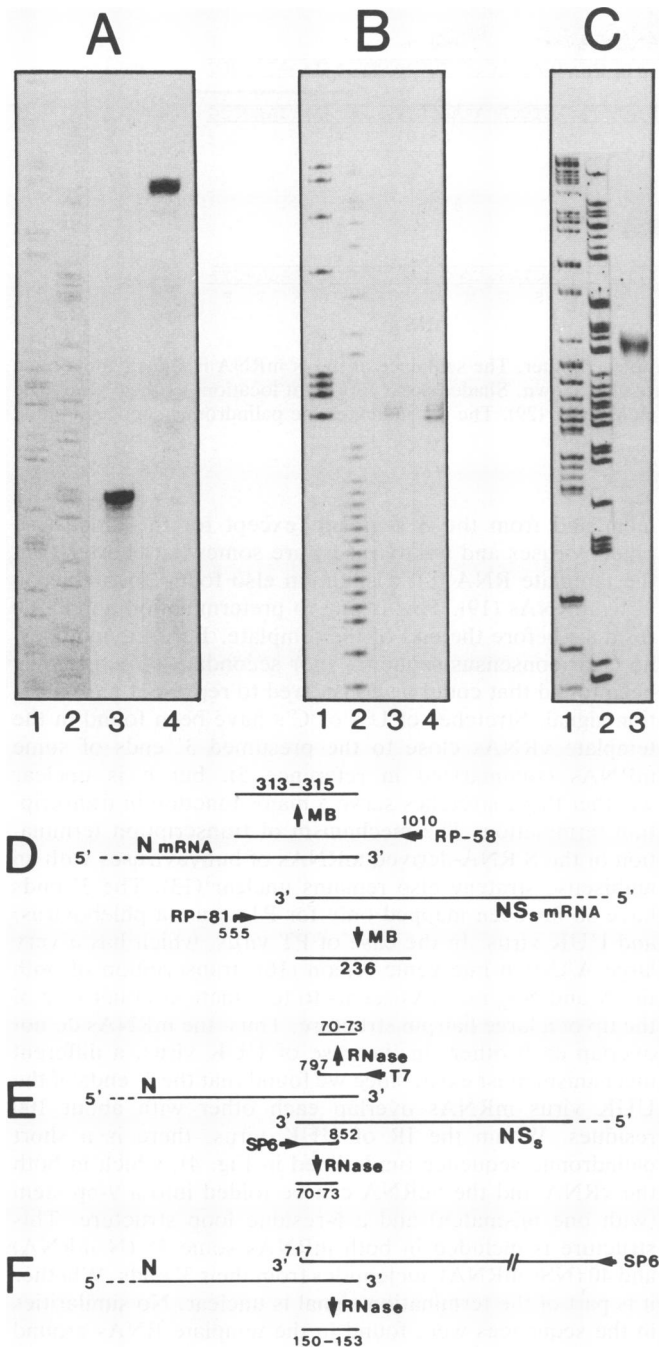

FIG. 3. Mapping of the 3′ ends of the N and NS$_S$ mRNAs by nuclease protection assays. (A to C) Autoradiograms of the protected fragments after analyses on an 8% sequencing gel together with different sequencing ladders used as size markers (lanes 1 and 2). (D to F) Schematic representations of the probes used and the protected fragments obtained after hybridization and nuclease treatment. (A and D) Mapping of the 3′ ends using ssDNA probes generated from primers RP-81 and RP-58 by PCR. The mung bean nuclease-resistant fragments (236 and 313 to 315 residues) obtained are shown in lanes 3 and 4, respectively. (B and E) Mapping of the 3′ ends by using riboprobes. Radioactive probes were transcribed in vitro from viral cDNA inserts, using either the SP6 or T7 promoter. The cDNA insert (in pGEM-3) was first linearized with restriction endonucleases at position 797 (N mRNA) or 852 (NS$_S$ mRNA) to generate specific 3′ termini. Following hybridization to the mRNAs, RNase A+T$_1$-resistant fragments of 70 to 73 nucleotides in length were obtained in both cases (lanes 3 and 4). To confirm these results for the N mRNA, a different riboprobe was transcribed from the SP6 promoter after cleavage of the insert at position 717 (C and F). A protected fragment of 150 to 153 nucleotides in length was obtained (lane 3). Numbering of the cleavage sites of the cDNA inserts corresponds to that of the full-length S vcRNA segment (29).

the 5′ ends indicated that the lengths of the nonviral sequences varied from 8 to 20 residues, with a majority being 10 to 15 residues (mean, 12). Determination of the sequences of individual 5′ ends showed an even wider length distribution (7 to 25 residues). Although the cloning strategy should have minimized the risk of losing nucleotides from the extreme 5′ ends, it is conceivable that some of the shorter clones lack one or several nucleotides. It is possible that reverse transcriptase in some cases is unable to copy the nucleotides close to the cap structure of the 5′ ends of some mRNAs. In analogy to influenza viruses (18, 19), we assume that the 5′ ends of UUK virus, and bunyaviruses in general, are derived from the 5′ ends of host mRNAs. De novo initiation of eukaryotic mRNA synthesis by polymerase II occurs in 50% of the cases with an A, in 25% with a G, and in the remaining 25% with a pyrimidine (C or U) residue (6). We observed, however, an equal representation of A (32%), G (32%), and C (33%) at the extreme 5′ ends of the N and NS$_S$ mRNAs. The cDNAs were tailed with A residues, and therefore it was not possible to estimate the frequency of U residues at the 5′ ends. Since we do not know whether these bases in all cases represent the penultimate nucleotide in the cap structure, it is not possible to tell whether this base distribution reflects that of the general population of mRNA 5′ ends in BHK-21 cells. The finding of a clone with as many as 25 nonviral nucleotides was surprising. Such a long sequence has not been described for any other viral mRNAs. Whether this 5′ end indeed corresponds to a cellular mRNA or is a cloning artifact is unclear. Using the corresponding synthetic oligonucleotide, we have been unable to detect any host BHK-21 cell mRNA species by Northern blot analysis (unpublished data).

As has been noticed for influenza (18), snowshoe hare (4) and Germiston (5) virus mRNAs, the first virus-encoded base (an A residue at position +1) was missing from some of the mRNAs in UUK virus. With the exception of one case, this nucleotide was replaced by a G residue, as found for the viruses mentioned above. The details of the biochemistry involved in cap-snatching and priming of transcription have been elegantly worked out for influenza virus (18, 19). The structural analyses done on the 5′ ends of bunyavirus mRNAs suggest that the mechanisms may be very similar. Thus, an endonucleolytic cleavage, preferentially after a purine, by a viral enzyme would first generate the capped oligonucleotide primer, followed by elongation with the 3′ end of the vRNA or vcRNA as the template. An endonucleolytic activity has indeed been found in LAC virus. Furthermore, capped and methylated (but not uncapped or unmethylated) RNAs were found to stimulate in vitro transcription by the LAC virion transcriptase (22). That the in vivo-synthesized LAC virus mRNAs are indeed capped was recently shown by Hacker et al. (15), using affinity selection with anticap antibodies. Whereas the role of the viral proteins responsible for the cap-snatching, priming, and chain elongation has to a large extent been defined for influenza virus (19), nothing is known about the bunyavirus proteins involved in these events. Likely candidates are the L and the

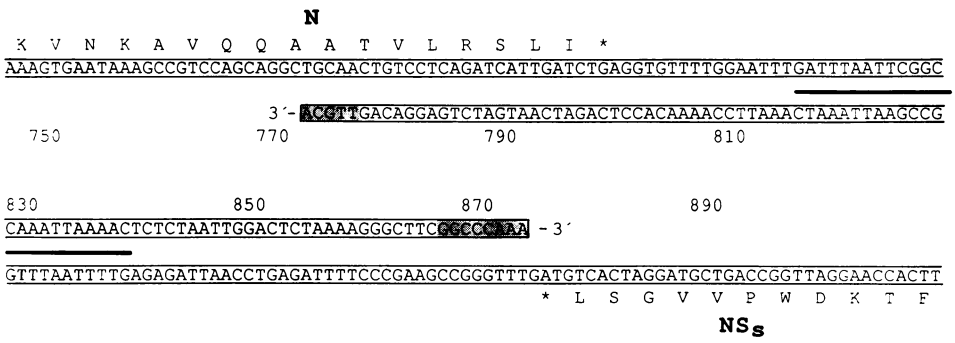

FIG. 4. Positions of the 3′ termini of the N and NS_S mRNAs relative to each other. The sequence of the N mRNA is shown above, and that of the NS_S is shown below. The C-terminal amino acid sequences are also shown. Shaded areas represent locations of the 3′ ends and are based on the data presented in Fig. 3. Numbering is according to Simons et al. (29). The bar indicates the palindromic sequence found in the intergenic region.

NS_S proteins, although the latter is missing from hantavirus- and the nairovirus-infected cells (3, 8, 28). It should also be pointed out that whereas influenza virus cap-snatching and transcription take place in the nucleus, transcription in bunyavirus-infected cells appears to be confined to the cytoplasm (8, 28). Since nontemplated bases have recently also been found at the 5′ ends of the mRNAs of an arenavirus (Tacaribe) (12, 24), it appears that transcription of all known negative-strand viruses with a segmented genome (orthomyxoviruses, arenaviruses, and bunyaviruses) is primed by short oligonucleotides derived from host mRNAs.

Where studied, the mRNAs transcribed from the M segment and from the S segment (except for the ambisense phleboviruses and uukuviruses) are somewhat shorter than the template RNA (8), a situation also found for influenza virus mRNAs (19). This is due to pretermination about 100 residues before the end of the template. For bunyaviruses, no clear consensus sequence (nor secondary structure) has been found that could be considered to represent a termination signal. Stretches of U's or C's have been found in the template vRNAs close to the presumed 3′ ends of some mRNAs (summarized in reference 5), but it is unclear whether these stretches serve a major function in transcription termination. The mechanism of transcription termination of the S RNA-derived mRNAs of bunyaviruses with an ambisense strategy also remains unclear (13). The 3′ ends have so far been mapped only for PT virus (a phlebovirus) and UUK virus. In the case of PT virus, which has a very large A/U-rich intergenic region (16), transcription of both the N and NS_S mRNAs seems to terminate at either side of the tip of a large hairpin structure. Thus, the mRNAs do not overlap each other. In the case of UUK virus, a different mechanism must exist, since we found that the 3′ ends of the UUK virus mRNAs overlap each other with about 100 residues. Within the IR of UUK virus, there is a short palindromic sequence (underlined in Fig. 4), which in both the vRNA and the vcRNA can be folded into a 9-bp stem (with one mismatch) and a 6-residue loop structure. This structure is included in both mRNAs some 35 (N mRNA) and 40 (NSs mRNA) nucleotides from their 3′ ends. Whether it is part of the termination signal is unclear. No similarities in the sequences were found in the template RNAs around the 3′ ends of the two mRNAs. The lack of possibilities to manipulate the sequences of bunyaviruses makes it very difficult to pinpoint regulatory sequences of this kind. The glycoproteins G1 and G2 and the N proteins of phleboviruses and uukuviruses show a low degree of homology, suggesting that they are distantly related (25, 29). Mapping the location of the 3′ ends of more phleboviruses and uukuvirus mRNAs, and perhaps finding conserved motifs may therefore shed some light on the mechanism of transcription termination. As noted in the introduction, the sequences of the IRs of sandfly fever Sicilian, Rift Valley fever, and Toscana viruses (for which the 3′ ends of the mRNAs have not yet been mapped) are, however, very different from each other and from those of PT and UUK viruses and do not display any clear secondary structures (13).

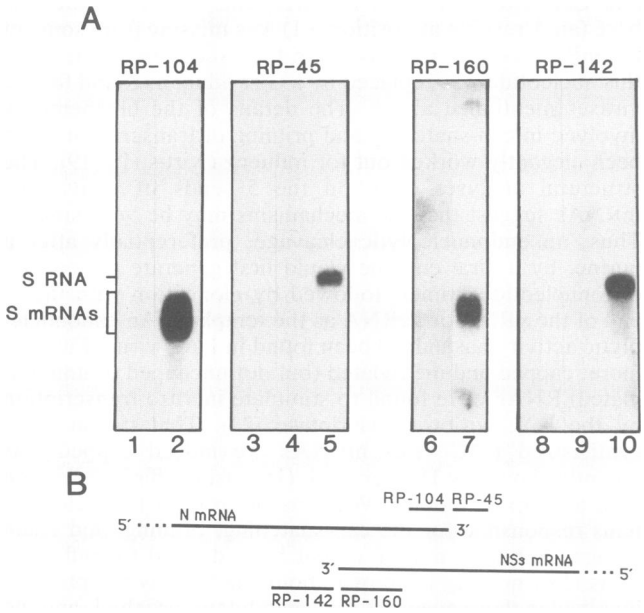

FIG. 5. Mapping of the 3′ ends by Northern blot analysis using strand-specific oligonucleotides. mRNA from infected or uninfected cells was isolated from CsCl pellets and analyzed by Northern blotting, using as probes the synthetic oligonucleotides RP-104 (panel A, lanes 1 and 2) and RP-45 (lanes 3 to 5) (complementary to the N mRNA-sense strand) and RP-142 (lanes 6 and 7) and RP-160 (lane 8 to 10) (complementary to the NS_S mRNA-sense strand) (see Materials and Methods). RP-45 and RP-142 were also hybridized to S vcRNA (lane 5) and vRNA (lane 10). (B) Positions of the probes relative to the 3′ ends of the two mRNAs.

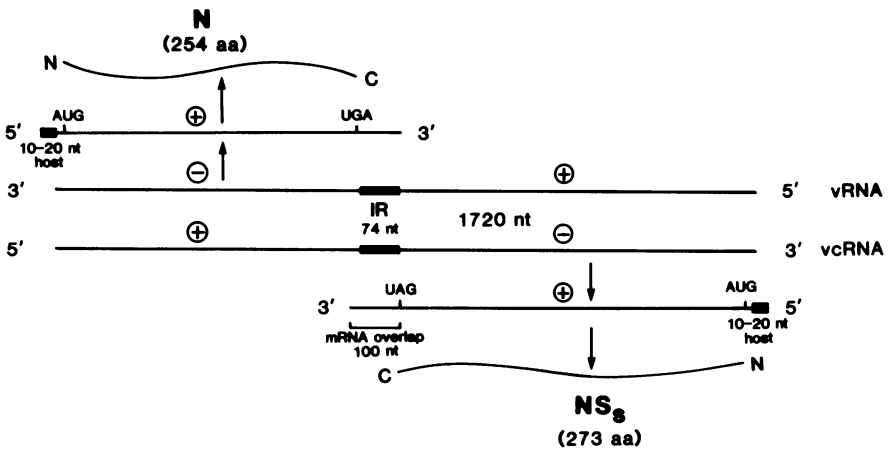

FIG. 6. Schematic representation of the coding strategy of the UUK virus S RNA segment.

ACKNOWLEDGMENTS

We thank Robert Persson, Yihai Cao, and Ulf Eriksson for stimulating discussions, as well as Maj-Greth Nilsson and Anita Bergström for excellent technical assistance.

REFERENCES

1. **Auperin, D. D., V. Romanowski, M. Galinski, and D. H. L. Bishop.** 1984. Sequencing studies of Pichinde arenavirus S RNA indicate a novel coding strategy, an ambisense viral S RNA. J. Virol. **52**:897–904.
2. **Auperin, D. D., D. R. Sasso, and J. B. McCormick.** 1986. Nucleotide sequence of the glycoprotein gene and the intergenic region of the Lassa virus S genome RNA. Virology **154**:155–167.
3. **Bishop, D. H. L.** 1990. Bunyaviridae and their replication, p. 1155–1173. *In* B. N. Fields et al. (ed.), Virology, 2nd ed. Raven Press, New York.
4. **Bishop, D. H. L., M. E. Gay, and Y. Matsuoko.** 1983. Nonviral hetergeneous sequences are present at the 5′ ends of one species of snowshoe hare bunyavirus S complementary RNA. Nucleic Acids Res. **11**:6409–6418.
5. **Bouloy, M., N. Pardigon, P. Vialat, S. Gerbaud, and M. Girard.** 1990. Characterization of the 5′ and 3′ ends of viral messenger RNAs isolated from BHK21 cells infected with Germiston virus (Bunyavirus). Virology **175**:50–58.
6. **Bucher, P., and E. N. Trifonov.** 1986. Compilation and analysis of eukaryotic POL II promoter sequences. Nucleic Acids Res. **14**:10009–10026.
7. **Collett, M. S.** 1986. Messenger RNA of the M segment RNA of Rift Valley fever virus. Virology **151**:151–156.
8. **Elliott, R. M.** 1990. Molecular biology of the Bunyaviridae. J. Gen. Virol. **71**:501–522.
9. **Emery, V. C., and D. H. L. Bishop.** 1987. Characterization of Punta Toro S mRNA species and identification of an inverted complementary sequence in the intergenic region of Punta Toro phlebovirus ambisense S RNA that is involved in mRNA transcription termination. Virology **156**:1–11.
10. **Fazakerly, J. F., F. Gonzalez-Scarano, J. Strickler, B. Dietzschold, F. Karush, and N. Nathanson.** 1988. Organization of the middle RNA segment of snowshoe hare bunyavirus. Virology **167**:422–432.
11. **Fuller, F., A. S. Brown, and D. H. L. Bishop.** 1983. Bunyavirus nucleoprotein, N, and a non-structural protein, NSs, are coded by overlapping reading frames in the S RNA. J. Gen. Virol. **64**:1705–1714.
12. **Garcin, D., and D. Kolakofsky.** 1990. A novel mechanism for the initiation of Tacaribe arenavirus genome replication. J. Virol. **64**:6196–6203.
13. **Giorgi, C., L. Accardi, L. Nicoletti, M. C. Gro, K. Takehara, C. Hilditch, and D. H. L. Bishop.** 1991. Sequences and coding strategies of the S RNAs of Toscana and Rift Valley fever viruses compared to those of Punta Toro, Sicilian sandfly fever and Uukuniemi viruses. Virology **180**:738–753.
14. **Gonzalez-Scarano, F., and N. Nathanson.** 1990. Bunyaviruses, p. 1195–1228. *In* B. N. Fields et al. (ed.), Virology, 2nd ed. Raven Press, New York.
15. **Hacker, D., S. Rochat, and D. Kolakofsky.** 1990. Anti-mRNAs in La Crosse bunyavirus-infected cells. J. Virol. **64**:5051–5057.
16. **Ihara, T., H. Akashi, and D. H. L. Bishop.** 1984. Novel coding strategy (ambisense genomic RNA) revealed by sequence analysis of Punta Toro phlebovirus S RNA. Virology **136**:293–306.
17. **Ihara, T., J. Smith, J. M. Dalrymple, and D. H. L. Bishop.** 1985. Complete sequences of the glycoproteins and M RNA of Punta Toro phlebovirus compared to those of Rift Valley fever virus. Virology **144**:246–259.
18. **Krug, R. M.** 1981. Priming of influenza viral RNA transcription by capped heterologous RNAs. Curr. Top. Microbiol. Immunol. **93**:125–149.
19. **Krug, R. M., F. V. Alonso-Caplen, I. Julkunen, and M. G. Katze.** 1989. Expression and replication of the influenza virus genome, p. 89–152. *In* R. M. Krug (ed.), The influenza viruses. Plenum Press, New York.
20. **Marriott, A. C., K. W. Vernon, and P. A. Nuttall.** 1989. The S RNA segment of Sandfly fever Sicilian virus: evidence for an ambisense genome. Virology **169**:341–345.
21. **Patterson, J. L., B. Holloway, and D. Kolakofsky.** 1984. La Crosse virions contain a primer-stimulated RNA polymerase and a methylated cap-dependent endonuclease. J. Virol. **52**:215–222.
22. **Patterson, J. L., and D. Kolakofsky.** 1984. Characterization of La Crosse virus small-genome transcripts. J. Virol. **49**:680–685.
23. **Pettersson, R. F., and L. Kääriäinen.** 1973. The ribonucleic acids of Uukuniemi virus, a non-cubical tick-borne arbovirus. Virology **56**:608–619.
24. **Raju, R., L. Raju, D. Hacker, D. Garcin, R. Compans, and D. Kolakofsky.** 1990. Nontemplated bases at the 5′ ends of Tacaribe virus mRNAs. Virology **174**:53–59.
25. **Rönnholm, R., and R. F. Pettersson.** 1987. Complete nucleotide sequence of the M RNA segment of Uukuniemi virus encoding the membrane glycoproteins G1 and G2. Virology **160**:191–202.
26. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **Schmaljohn, C. S., G. B. Jennings, J. Hay, and J. M. Dalrymple.** 1986. Coding strategy of the S genome segment of Hantaan virus. Virology **155**:633–643.
28. **Schmaljohn, C. S., and J. L. Patterson.** 1990. Bunyaviridae and their replication, p. 1175–1194. *In* B. N. Fields et al. (ed.),

Virology, 2nd ed. Raven Press, New York.
29. **Simons, J. F., U. Hellman, and R. F. Pettersson.** 1990. Uukuniemi virus S RNA segment: ambisense coding strategy, packaging of complementary strands into virions, and homology to members of the genus phlebovirus. J. Virol. **64**:247–255.
30. **Ulmanen, I., P. Seppälä, and R. F. Pettersson.** 1981. In vitro translation of Uukuniemi virus-specific RNAs: identification of a nonstructural protein and a precursor to the membrane glycoproteins. J. Virol. **37**:72–79.
31. **Ward, V. K., A. C. Marriott, A. A. El-Ghorr, and P. A. Nuttall.** 1990. Coding strategy of the S RNA segment of Dugbe virus (Nairovirus; Bunyaviridae). Virology **175**:518–524.