JOURNAL OF VIROLOGY, Sept. 1992, p. 5277–5283
0022-538X/92/095277-07$02.00/0

Vol. 66, No. 9

The Nucleocapsid Protein Gene of Bovine Coronavirus Is Bicistronic

SAVITHRA D. SENANAYAKE, MARTIN A. HOFMANN,† JOANNE L. MAKI,‡ AND DAVID A. BRIAN*

Department of Microbiology, The University of Tennessee, Knoxville, Tennessee 37996-0845

Received 27 April 1992/Accepted 28 May 1992

For animal RNA viruses that replicate through an RNA intermediate, reported examples of bicistronic mRNAs with overlapping open reading frames in which one cistron is contained entirely within another have been made only for those with negative-strand or double-stranded genomes. In this report, we demonstrate for the positive-strand bovine coronavirus that an overlapping open reading frame potentially encoding a 23-kDa protein (names the I [for internal open reading frame] protein) and lying entirely within the gene for the 49-kDa nucleocapsid phosphoprotein is expressed during virus replication from a single species of unedited mRNA. The I protein was specifically immunoprecipitated from virus-infected cells with an I-specific antipeptide serum and was shown to be membrane associated. Many features of I protein synthesis conform to the leaky ribosomal scanning model for regulation of translation. This, to our knowledge, is the first example of a bicistronic mRNA for a cytoplasmically replicating, positive-strand animal RNA virus in which one cistron entirely overlaps another.

Sequence analysis of the gene for the 448-amino-acid nucleocapsid (N) protein of bovine coronavirus (BCV) revealed a second open reading frame (ORF) situated entirely within the N gene (in the +1 reading frame) and potentially encoding a moderately hydrophobic protein of 207 amino acids (23 kDa; named the I [for internal ORF] protein; Fig. 1) (32). An internal ORF of identical size and nearly identical position was found within the N gene of the antigenically related mouse hepatitis virus strain A59 (5). Except for one interrupting point mutation in the N gene of mouse hepatitis virus strain JHM (AAA→UAA at amino acid 17 in comparison with BCV) (46) and two in the N gene of human coronavirus strain OC43 (AUG→AUC and CAG→UAG at amino acids 1 and 61, respectively, in comparison with BCV) (24), internal ORFs of identical size were found in these antigenically related viruses as well. No such internal ORFs were found, however, in the N genes of the more distantly related avian infectious bronchitis virus (7), porcine transmissible gastroenteritis virus (25), feline infectious peritonitis virus (12), and human coronavirus 229E (43).

Because the internal ORF is conserved in size between BCV and mouse hepatitis virus A59, and a significant (45%) amino acid sequence identity is also found between their putative I proteins, we initiated experiments to determine whether the I protein is produced during virus replication and, if so, to determine how it is generated and what its intracellular location might be. The existence of this protein, or its homolog, from any coronavirus has not yet been demonstrated to our knowledge. In this report, we describe experiments that demonstrate expression of the I protein in virus-infected cells and document that it is probably made by downstream initiation of translation on the N mRNA rather than from an edited mRNA molecule. The I protein is membrane associated, but the mechanism of this association is not yet completely established. To our knowledge, this is the first example of a bicistronic mRNA in a cytoplasmically replicating, positive-strand animal RNA virus in which one cistron entirely overlaps another.

MATERIALS AND METHODS

Construction of plasmids and in vitro expression analysis of the N and potential I genes. For in vitro expression analysis of the N and potential I genes from the same synthetic mRNA molecule, capped transcripts having the leader and body sequence of the N mRNA were generated in vitro and translated in wheat germ lysate. Plasmid pSP6LN was used in all experiments in which the N gene was analyzed (Fig. 2). Transcripts of plasmid did not totally mimic intracellular N mRNA structure, however, since the 5′ terminus of the leader was incomplete and the poly(A) tail along with the first 82 nucleotides (nt) of the 3′ noncoding region were missing. We have recently determined the 5′ leader on the BCV N mRNA to be 5′-GAUUGUGAGCGAUUUGCGUG CGUGCAUCCCGCUUCACUGAUCUCUUGUUAGAUC UUUUUAUAAUCUAAACUUUAAGGAUG (GenBank accession number M62375 [21a]), in which nt 63 is the point of divergence between the genome and the leader sequence and the underlined AUG is the initiator codon for N. In the pSP6LN construct, the 5′ seven nucleotides of the leader (GAUUGUG) were replaced by the sequence 5′ cggaauucg-guaccuUUGCG, in which the first 15 nt (lowercase) provided *Eco*RI and *Kpn*I restriction endonuclease sites and the UUGCG (uppercase) was our perception of the leader 5′ end at the time the construct was made.

pSP6LN was constructed in three steps. (i) The BCV leader sequence was cloned by amplifying an extended primer from the 5′ end of the N mRNA by the polymerase chain reaction (PCR) and cloning the amplified fragment into pGEM3Z (Promega). Primers used for PCR were 5′-CCAG AACGATTTCCAAAGGACGCTCT-3′ (primer 1), which anneals to nt 31 through 56 downstream from the start codon

* Corresponding author.
† Present address: Institut fuer Virusforschung und Immunoprophylaxe, CH-4000 Basel, Switzerland.
‡ Present address: Rhone Merieux, Athens, GA 30601.

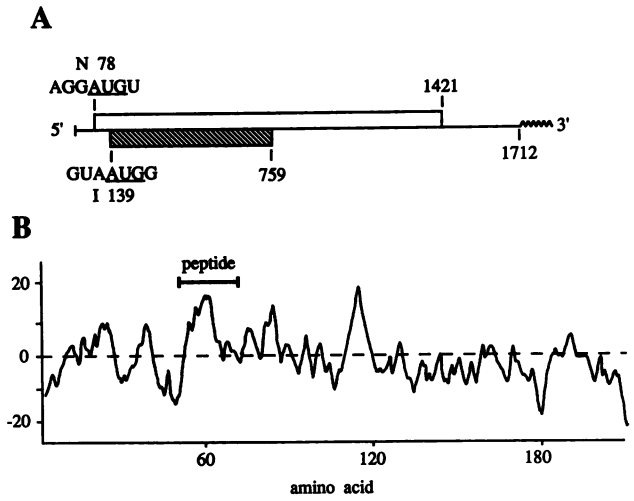

FIG. 1. (A) Schematic diagram of the N and I protein ORFs on the N mRNA of BCV. Symbols: solid line, the mRNA molecule, including the 5′ leader and the 3′ noncoding region; vertical bar, the 5′ cap; wavy line, the 3′ poly(A) tail; open rectangle, the N ORF; hatched rectangle, the I ORF. The sequences surrounding the initiator codons are shown, and the base positions are indicated. (B) Hydrophobicity profile of the I protein as determined by the algorithm of Kyte and Doolittle (30), using a seven-amino-acid window. Hydrophobic regions are below the hatched line. The region used for deriving a synthetic peptide is shown.

on N mRNA and contains a convenient *Xmn*I site, and 5′-CGGAATTCGGTACCTTTGCGAGCGATTTGCGTG CG-3′ (primer 2), which anneals near the 3′ end of the leader minus strand and contains the additional 15 nt with the *Eco*RI and *Kpn*I restriction endonuclease sites. The amplified cDNA fragment was cut with *Eco*RI and *Xmn*I, and the *Eco*RI site was blunt ended by filling in with Klenow enzyme. The blunt-ended fragment was cloned into the *Eco*RI site of the pGEM3Z vector that had been filled in with

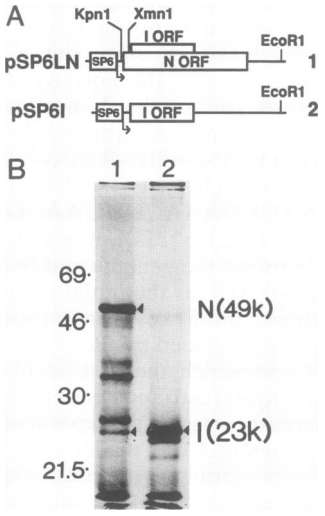

FIG. 2. (A) Plasmid constructs used for making synthetic transcripts containing both the N and I ORFs (pSP6LN) or the I ORF alone (pSP6I). (B) Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of translation products of the N gene transcript (lane 1) and the I gene transcript (lane 2).

Klenow enzyme, recreating *Kpn*I sites at the points of ligation. The resulting plasmid was pBL7. (ii) The *Hpa*I (which cuts 20 nt upstream of the matrix protein start codon)-*Sty*I [which cuts 82 nt upstream from the base of the poly(A) tail and within the 3′ noncoding region] fragment of clone MA5 (32) was blunt ended with mung bean nuclease and subcloned into the *Sma*I site of the pGEM4Z (Promega) vector. The resulting plasmid was pBMN. (iii) The *Kpn*I-*Xmn*I leader-containing fragment from pBL7, the *Xmn*I-*Eco*RI N gene-containing fragment from pBMN, and the *Kpn*I- and *Eco*RI-linearized, dephosphorylated pGEM3Z vector were directionally joined in a three-way ligation reaction and used to transform *Escherichia coli* JM109. The resulting clone (pSP6LN) was confirmed by DNA sequencing at the joined regions.

For in vitro expression of only the internal ORF, plasmid pSP6I was prepared as described in step (iii) above except that only the *Xmn*I-*Eco*RI fragment of clone pBMN (Fig. 2A) and the *Sma*I-*Eco*RI-linearized, dephosphorylated pGEM3Z vector were used in a two-way ligation reaction.

For preparation of transcripts for in vitro translation, plasmids pSP6LN and pSP6I were linearized with *Eco*RI and transcribed with SP6 RNA polymerase.

Sequencing analyses used to seek evidence of edited mRNA molecules. For sequencing of extended primers prepared from bulk N mRNA, cytoplasmic RNA was extracted from infected cells at 6 h postinfection as previously described (22), and 5′-end-labeled primer 1 was annealed and extended with reverse transcriptase, purified, and sequenced by the method of Maxam and Gilbert (38) as previously described (44).

For dideoxynucleotide termination sequencing of asymmetrically amplified cDNA of the N mRNA 5′ end, cytoplasmic RNA from infected cells was first used to prepare single-stranded cDNA by priming the reaction with primer 3 [5′-GGGGGGGGGGACGCGTTTTTTTTTT-3′; a primer which contains an *Mlu*I site and primes within the poly(A) tail] and extending it with reverse transcriptase, using conditions that we have previously described (21). The N mRNA cDNA was amplified symmetrically with 30 cycles of PCR using primers 2 and 3, and the N mRNA-length product was purified by agarose gel electrophoresis and electroeluted by standard procedures (21, 42). From this, a portion was used to inoculate a 60-cycle asymmetric PCR in which primer 2 was the limiting primer and either primer 1 or primer 4 (5′-GTCCCGATCGAGAATGTCAGCCGGGG-3′; a primer which binds to nt 409 through 434 downstream from the start of the I gene) was the nonlimiting primer. Asymmetric PCR and dideoxynucleotide termination sequencing with 5′-end-labeled primer 2 were performed as previously described (20).

Preparation of rabbit antipeptide serum. A synthetic peptide of 21 amino acids (NH_2-KKERSLNLQRDKVCLL HQESQ-COOH) representing amino acid positions 53 through 73, a hydrophilic region of the I protein (Fig. 1B), was synthesized. One milligram was suspended in 1 ml of water, emulsified with an equal volume of Freund's complete adjuvant and 0.2 mg of *N*-acetylmuramyl-L-alanyl-D-isoglutamine (Calbiochem-Behring), and injected subcutaneously at eight sites over the back and sides of a 2-month-old rabbit. The rabbit was boosted three times at 2-week intervals by similarly injecting 1 mg of peptide emulsified in Freund's incomplete adjuvant, and the rabbit was bled 2 weeks after the last boost.

Analysis of the N and I proteins synthesized in vitro. For analysis of radiolabeled proteins synthesized in vitro, ap-

proximately 1 μg of transcript was translated in 50 μl of wheat germ lysate translation mix (Promega) containing 5 mCi of [^{35}S]methionine (>800 Ci/mmol; ICN) per ml. For immunoprecipitation, 10 μl of rabbit serum (preimmune or antipeptide) or 10 μl of bovine serum (preimmune or hyperimmune; a kind gift from L. J. Saif, Ohio State University) was used per 50-μl translation reaction volume as described by Anderson and Blobel (4). When rabbit serum was used, Sephadex-conjugated staphylococcal protein A was used to concentrate the antigen-antibody complex; when bovine serum was used, Sephadex-conjugated staphylococcal protein G was used. Radiolabeled proteins were analyzed by electrophoresis in gels of 10% polyacrylamide by the method of Laemmli (31) and were analyzed as previously described (32). For quantitation, dried gels were scanned with the Ambis Radioanalytic Imaging System (Ambis, San Diego, Calif.). ^{14}C-labeled molecular weight marker proteins were myosin (200 kDa), phosphorylase *b* (92.5 kDa), bovine serum albumin (69 kDa), soybean trypsin inhibitor (21.5 kDa), and lysozyme (14.3 kDa) (Amersham) and were included in all gels for protein analysis.

For analyzing membrane association and translocation, proteins were synthesized in the presence of canine pancreatic microsomes (Promega) by the method of Walter and Blobel (50). Removal by washing in the presence of EDTA and measurement of resistance to trypsin digestion were done by the method of Morimoto et al. (39), and membrane association by extraction at pH 11 was done by the method of Fujiki et al. (14).

Analysis of the I protein synthesized in vivo. For immunoprecipitation analysis of the I protein made in infected cells, confluent monolayers of human rectal tumor cells in 35-mm-diameter dishes were infected with a multiplicity of 10 PFU per cell and incubated in the presence of 10 mCi of [^{35}S]methionine (Translabel; ICN) per ml in methionine-free Dulbecco's minimum essential medium to radiolabel the viral proteins. At 48 h postinfection, cells were lysed in phosphate-buffered saline (PBS; 250 μl per dish) containing 1% Nonidet P-40, 100 μg of phenymethylsulfonyl fluoride (Sigma) per ml, and 1 μg of aprotinin (Sigma) per ml, the lysate was clarified, and proteins were immunoprecipitated with 20 μl of rabbit antipolypeptide serum per 200 μl of lysate, using the method of Anderson and Blobel (4).

For immunofluorescence analysis of the I protein, human rectal tumor cells at 20% confluency on Labtek chamber slides (Nunc) were infected with a multiplicity of approximately 5 PFU per cell and at 24 h postinfection were rinsed and fixed either with absolute ethanol to examine intracytoplasmic fluorescence or with 4% paraformaldehyde in PBS to examine surface fluorescence. For staining, fixed cells were incubated with a 1:20 dilution of preimmune or peptide-specific antiserum and then with a 1:50 dilution of fluorescence-conjugated goat anti-rabbit immunoglobulin G.

RESULTS

The I protein is made during virus replication and becomes membrane associated. In vivo labeling and RNA hybridization experiments have identified the N mRNA as the smallest of the 3′ coterminal mRNAs for BCV (22), suggesting that if the I protein is made, it is synthesized from the N mRNA species. To test whether the N mRNA can serve as a template for synthesis of I, SP6 RNA polymerase-generated transcripts of plasmid pSP6LN were translated in vitro. In addition to N, which represented 43% of the product synthesized, a protein the predicted size of I (23 kDa) which represented 12% of the total product was also made (Fig. 2B, lane 1). The 23-kDa protein species synthesized from the pSP6LN transcript migrated with the same mobility as did the major product from the pSP6I transcript on which I synthesis begins at the first available methionine codon (Fig. 2B, lane 2). Transcripts from a modified pSP6LN construct in which the predicted I initiator codon had been destroyed by mutagenesis failed to make the 23-kDa protein, thus confirming this codon as the start of I synthesis (data not shown). In addition to N and the 23-kDa protein, three other proteins intermediate in size between N and the 23-kDa species were produced from the pSP6LN transcript, and together these represented approximately 45% of the total product (Fig. 2B, lane 1). These proteins apparently are products of premature termination of N translation or translation products of prematurely terminated transcripts, since all three of the proteins intermediate in size between N and the 23-kDa species immunoprecipitated with polyclonal N-specific antiserum (data not shown). They are unlikely to be products of internal initiation on the N ORF, since the next possible start codon within the N ORF is at amino acid position 232 (32), and this would yield a predicted protein smaller than 22 kDa.

To test whether transcripts with the precise structure of N mRNA would also synthesize the I protein, N mRNA was separately cloned from infected cells, and a clone, pLN, confirmed by sequencing to have the proper termini [i.e., a 5′-terminal GAUUGUG. . . , the N ORF, the complete 3′ noncoding sequence, and a 3′ poly(A) tail of 21 nt] was used to prepare transcripts for translation. The results of this translation were similar to those depicted in Fig. 2B, lane 1, except that the species between N and I were diminished in amount (representing less than 20% of the total product), and N and I represented 50 and 20%, respectively, of the product.

Three approaches were used to determine whether I is made during virus replication. In the first, the I-specific antipeptide serum was used to test for immunoprecipitation of the 23-kDa protein from infected cells. Antipeptide serum prepared in a rabbit, but not preimmune serum, immunoprecipitated I protein made in vitro (Fig. 3A, lanes 2 and 3), thus establishing the I specificity of the antipeptide serum. The antipeptide serum but not the preimmune serum immunoprecipitated a product of the same size from infected cells (Fig. 3A, lanes 5 and 6), indicating that I is made during virus replication and probably undergoes no posttranslational modification that would alter its electrophoretic mobility, such as N-linked glycosylation at amino acid 24 or signal peptide cleavage (32). In the second approach, hyperimmune calf serum known to immunoprecipitate N protein made in vitro (Fig. 3B, lane 2) was used to test for immunoprecipitation of I made in vitro. Hyperimmune but not preimmune calf serum immunoprecipitated I (Fig. 3B, lanes 5 and 6), indicating that the I protein not only is made during virus infection but also induces an antibody response. In the third approach, the antipeptide serum was used to test for immunofluorescence of BCV-infected cells. Figure 4 illustrates an internal diffuse cytoplasmic fluorescence in cells fixed with ethanol (Fig. 4B) and a surface fluorescence on cells fixed with paraformaldehyde only (Fig. 4D). No immunofluorescence was observed when preimmune serum was used (Fig. 4A and C) or when uninfected cells were used with the immune serum (data not shown). Taken together, these results indicate that the I protein is made during virus replication, probably does not undergo posttranslational modification such as N-linked glycosylation or signal peptide

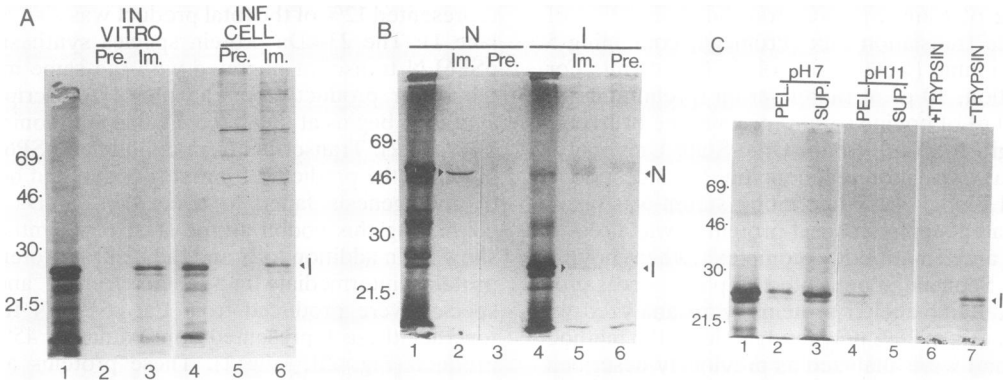

FIG. 3. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of N and I proteins. (A) I protein made in vitro from transcripts of pSP6I (lanes 1 and 4) was immunoprecipitated with antipeptide serum (lane 3) or preimmune serum (lane 2). Radiolabeled proteins from infected cell lysates were immunoprecipitated with antipeptide serum (lane 6) or preimmune serum (lane 5). (B) N protein made in vitro from transcripts of pSP6LN (lane 1) was immunoprecipitated with hyperimmune bovine serum (lane 2) or preimmune serum (lane 3). I protein made in vitro from transcripts of pSP6I (lane 4) was immunoprecipitated with hyperimmune bovine serum (lane 5) or preimmune serum (lane 6). (C) I protein made in vitro in the presence of microsomes was subjected to the different treatments indicated before analysis (lanes 2 through 7). I protein made in vitro in the absence of microsomes (lane 1) served as a marker.

cleavage, and stimulates an antibody response in the infected animal.

The I protein is moderately hydrophobic overall, since 79 (38%) of its amino acids are hydrophobic (32). With some hydrophobic amino acids concentrated on the termini of the protein (Fig. 1B), it is possible that one of the termini or both become transmembrane domains and serve to anchor I to the membrane. To test whether I associates with membranes during synthesis in vitro, transcripts of pSP6I were translated in the presence of canine pancreatic microsomes, and the association of I with pelleted microsomal membranes was examined. Figure 3C illustrates that I remains associated with pelleted membranes both before (lane 2) and after (lane 4) microsomal lysis, indicating either that there is a strong nonanchored attraction between I and the membrane or that there is anchorage as the result of partial translocation or possibly from addition of a fatty acid side chain. As with synthesis of the I protein in cells, there was no change in the electrophoretic mobility of I after synthesis in the presence of microsomes, suggesting that there is no glycosylation or signal peptide cleavage that would indicate membrane translocation.

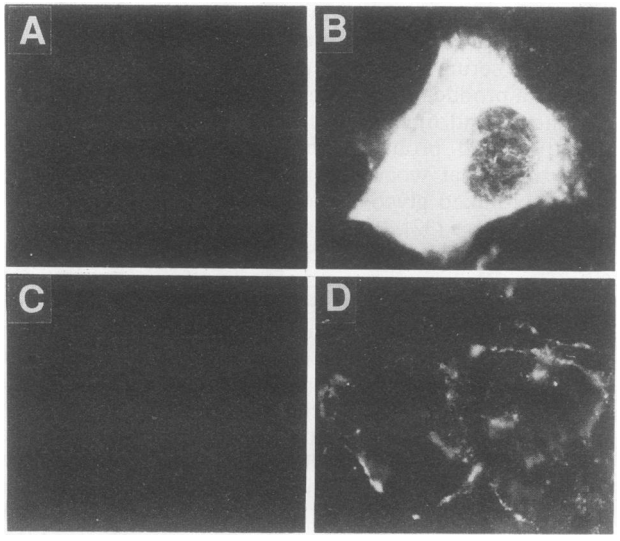

FIG. 4. Immunofluorescence of I protein in cell culture by indirect staining with I-specific peptide antiserum. Cells fixed with ethanol for internal fluorescence were incubated with preimmune (A) or immune (B) serum. Cells fixed with paraformaldehyde for surface fluorescence were incubated with preimmune (C) or immune (D) serum.

Of the possible mechanisms for the association of the I protein with microsomes, we can rule out divalent cation-mediated attraction, since washing of microsomes at pH 7 in the presence of a chelating agent (25 mM EDTA) removed only a small fraction of I from the membranes (data not shown). We cannot rule out simple hydrophobic attraction, however. Regarding the possibility of translocation, we can rule out the possibility that a majority of I had become translocated into the lumen of the microsome to become anchored as a type I membrane protein (i.e., anchored by its carboxy terminus) or as a type II membrane protein (i.e., anchored by its amino terminus), since no portion of I larger than a 12-kDa fragment could be identified by gel electrophoresis following trypsin treatment in microsomes (Fig. 3C, lane 6). Our experiments do not, however, rule out the possibility that only a small portion of I had become translocated, leaving a majority of the protein on the cytoplasmic surface of the microsome. Further study is required to determine the mechanism of membrane association.

An unedited mRNA is used for the expression of I. Despite the fact that I can be synthesized in vitro from the N mRNA, from precedent in other animal RNA virus systems (27, 40, 47–49), it is possible that the N mRNA species undergoes editing in vivo to yield a subpopulation of mRNA molecules from which I is put in frame with N to yield a fusion protein. In this molecule, the initiator codon for N would be used to synthesize I. Alternatively, the N initiator codon could be ablated by editing, leaving the I initiator codon as the first methionine codon on the transcript. The first possibility is unlikely, since I was not immunologically detected in studies using N-specific polyclonal antiserum (data not shown). Nevertheless, two experimental approaches were taken to seek evidence for either kind of mRNA editing. In the first, an oligonucleotide primer binding just downstream of the N initiator codon was 5′ end labeled, extended with reverse transcriptase, and sequenced by the method of Maxam and

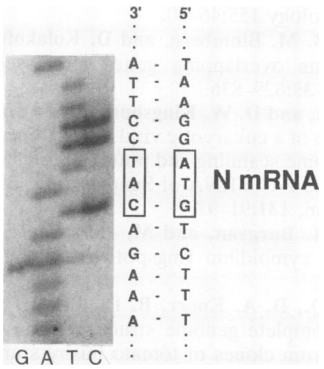

FIG. 5. Maxam and Gilbert sequencing of radiolabeled extended primer through the region of the initiator codon of N mRNA.

Gilbert (38). No ambiguity in the sequence was observed, suggesting that all or at least the vast majority of N-specific mRNA molecules are unedited (data not shown). Figure 5 shows the results of this experiment in the region of the N initiator codon and illustrates the lack of ambiguity and thus the intactness of the N initiator codon. In a second approach, single-stranded DNA molecules prepared by asymmetric PCR amplification of N mRNA cDNA were sequenced by the dideoxynucleotide chain termination method. No ambiguity in this sequence was observed regardless of whether the reaction was primed from just downstream of the N start codon (with primer 1) or from just downstream of the I start codon (with primer 4), indicating the absence of any editing between the 5′ end of the mRNA and the I start codon (data not shown). In addition, no modified N initiator codons were found in any of 138 separate cDNA clones of N mRNA prepared as part of another study (21a).

DISCUSSION

We have demonstrated that a large overlapping ORF located within the N gene of BCV is expressed during virus replication, resulting in the synthesis of a 23-kDa protein that we name the I (for internal ORF) protein. Our data show that I is expressed from the unedited N mRNA molecule by downstream initiation of translation. Although bicistronic mRNAs for animal RNA minus-strand viruses (myxoviruses [45], paramyxoviruses [3, 6, 11, 15–17, 35, 37, 45], and bunyaviruses [2]) and double-stranded viruses (reoviruses [13]) in which one ORF completely overlaps another have been described, to our knowledge, this is the first example of such a bicistronic mRNA for a cytoplasmically replicating, positive-strand animal RNA virus. A functional bicistronic mRNA of similar structure has recently been reported for members of the positive-strand tombusviruses of plants, however (18, 19, 41).

The function of the I protein is not apparent, nor can it be deduced from sequence similarities to proteins of known function identified by a computer search of GenBank. Significant similarities were found with only the I homolog in the N genes of mouse hepatitis virus (45% amino acid sequence identity) and human coronavirus OC43 (91% identity). The absence of an I homolog in other coronaviruses would indicate that I does not perform a universally required function for coronavirus survival. It may, however, be critical for replication in the animal for those coronaviruses closely related to BCV, which, interestingly, happen to be those also possessing a hemagglutinin-esterase protein or a nonfunctional hemagglutinin-esterase gene (23, 26, 36, 51). I protein function may not be required for replication in cell culture, and thus disruptions in the I genes of mouse hepatitis virus strain JHM and human coronavirus strain OC43 may have occurred during cell culture passage prior to the time of molecular cloning (24, 46).

The mechanism by which the I protein associates with membranes is unresolved. Although we have shown that a majority of the I protein molecule probably does not translocate to become a type I or a type II membrane protein, we have not ruled out the possibility that a small portion containing the hydrophilic peptide region translocated to the luminal surface and remain anchored. Such an anchored protein would most readily explain the occurrence of some of the I protein on the cell surface. The fact that there is a diffuse distribution of the I protein throughout the cytoplasm (Fig. 4B), however, would suggest that I, like N, is synthesized on free ribosomes in the cytoplasm. Although this pattern is inconsistent with membrane proteins that are undergoing anchorage or translocation, it does not rule out the possibility that a subpopulation of the I protein is being synthesized at the endoplasmic reticulum. Further work is needed to determine the exact nature of the membrane association of the I protein and whether it becomes part of the virion.

The question of how ribosomes access the downstream start codon of the I ORF have not been rigorously answered, but the observed pattern of I synthesis from an unedited N mRNA and the nucleotide sequence of the mRNA conform to the leaky scanning model for ribosomal accession of downstream ORFs (28, 29). In the N mRNA molecule, the start codon for N (AGGAUGU, the first available methionine codon) is in a suboptimal context for initiation of translation, since the nucleotide at the +4 position is a pyrimidine rather than a purine and the start codon for I (GUAAUGG, the second available methionine codon), with a purine at both the −3 and +4 positions, is in an optimal context. This pattern, according to the leaky scanning model, would allow passage of some ribosomes past the first codon to initiate translation at the second.

In light of recent evidence showing internal ribosomal entry to be important for synthesis of downstream overlapping ORFs on the polycistronic paramyxovirus P mRNA (3, 9, 10, 11, 16), it will be important to characterize the mechanisms of I synthesis further. Theoretically, I protein synthesis by a leaky scanning mechanism would be under greater influence of upstream ORF changes than would synthesis by an internal entry mechanism, and such changes may be an important means of regulating the amount of I synthesized. We have noted, for example, that leader sequence variations exist in BCV and that the predominant leader variant found during persistent infection contains an intraleader ORF of 33 nt (21a). We are currently investigating whether such an intraleader ORF causes variation in the relative rates of I synthesis.

Overlapping ORFs within the 5′ unique region of polycistronic coronavirus mRNAs have been identified (8, 33–34a), and these too appear to be translated from a single species of mRNA. In these cases, however, the overlaps are relatively short, ranging from 17 nt between the 3b and 3c protein genes in mRNA 3 of avian infectious bronchitis virus (34) to 5 nt between the 13- and 9.5-kDa protein genes in mRNA 5 of mouse hepatitis virus (8). Interestingly, the homologs for the 13- and 9.5-kDa proteins in BCV are apparently synthesized from separate mRNA species (1), suggesting that short

overlaps in mouse hepatitis virus may have arisen from a recent mutational event that combined two formerly separated ORFs. Because of its paucity of sequences from coronaviruses or other viruses showing homology to the I protein, it is difficult at this time to see a potential lineage for the origin of the I ORF.

ACKNOWLEDGMENTS

We thank Phiroze B. Sethna for many helpful discussions.

This work was supported by Public Health Service grant AI14367 from the National Institutes of Health. S.D.S. and J.L.M. were supported in part by postdoctoral fellowships from the Tennessee Center of Excellence Program in Livestock Diseases and Human Health, University of Tennessee College of Veterinary Medicine, and M.A.H. was supported by a postdoctoral fellowship from the Swiss National Science Foundation.

REFERENCES

1. **Abraham, S., T. E. Kienzle, W. E. Lapps, and D. A. Brian.** 1990. Sequence and expression analysis of potential nonstructural proteins of 4.9, 4.8, 12.7, and 9.5 kDa encoded between the spike and membrane protein genes of the bovine coronavirus. Virology **177**:488–495.
2. **Akashi, H., and D. H. L. Bishop.** 1983. Comparison of the sequences and coding of LaCrosse and snowshoe hare bunyavirus S RNA species. J. Virol. **45**:1155–1158.
3. **Alkhatib, G., B. Massie, and D. J. Briedis.** 1988. Expression of bicistronic measles virus P/C mRNA by using hybrid adenoviruses: levels of C protein synthesized in vivo are unaffected by the presence or absence of the upstream initiator codon. J. Virol. **62**:4059–5009.
4. **Anderson, D. J., and G. Blobel.** 1983. Immunoprecipitation of protein from cell-free translations. Methods Enzymol. **96**:111–120.
5. **Armstrong, J., S. Smeekens, and P. Rottier.** 1983. Sequence analysis of the nucleocapsid gene from murine coronavirus MHV-A59. Nucleic Acids Res. **11**:883–891.
6. **Bellini, W. J., G. Englund, S. Rozenblatt, H. Arnheiter, and C. D. Richardson.** 1985. Measles virus P gene codes for the two proteins. J. Virol. **53**:908–919.
7. **Boursnell, M. E. G., M. M. Binns, I. J. Foulds, and T. D. K. Brown.** 1985. Sequences of the nucleocapsid genes from two strains of avian infectious bronchitis virus. J. Gen. Virol. **66**:573–580.
8. **Budzilowicz, C. J., and S. Weiss.** 1987. In vitro synthesis of two polypeptides from a nonstructural gene of coronavirus mouse hepatitis virus strain A59. Virology **157**:509–515.
9. **Curran, J., and D. Kolakofsky.** 1989. Scanning independent ribosomal initiation of the Sendai virus Y proteins in vitro and in vivo. EMBO J. **8**:521–526.
10. **Curran, J. A., and D. Kolakofsky.** 1988. Scanning independent ribosomal initiation of the Sendai virus X protein. EMBO J. **7**:2869–2874.
11. **Curran, J. A., C. Richardson, and D. Kolakofsky.** 1986. Ribosomal initiation at alternate AUGs on Sendai virus P/C mRNA. J. Virol. **57**:684–687.
12. **DeGroot, R. J., A. C. Andeweg, M. C. Horzinek, and W. J. M. Spaan.** 1988. Sequence analysis of the 3′ end of the feline coronavirus FIPV 79-1146 genome: comparison with the genome of porcine coronavirus TGEV reveals large insertions. Virology **167**:370–376.
13. **Ernst, H., and A. J. Shatkin.** 1985. Reovirus hemagglutinin mRNA codes for two polypeptides in overlapping reading frames. Proc. Natl. Acad. Sci. USA **82**:48–52.
14. **Fujiki, Y., A. L. Hubbard, S. Fowler, and P. B. Lazarow.** 1982. Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. J. Cell Biol. **93**:97–102.
15. **Galinski, M. S., M. A. Mink, D. M. Lambert, S. L. Wechsler, and M. W. Pons.** 1986. Molecular cloning and sequence analysis of the human parainfluenza 3 virus mRNA encoding the P and C proteins. Virology **155**:46–60.
16. **Giorgi, C., B. M. Blumberg, and D. Kolakofsky.** 1983. Sendai virus contains overlapping genes expressed from a single mRNA. Cell **35**:829–836.
17. **Gupta, K. C., and D. W. Kingsbury.** 1985. Translational modulation in vitro of a eukaryotic viral mRNA encoding overlapping genes: ribosome scanning and potential roles of conformational changes in the P/C mRNA of Sendai virus. Biochem. Biophys. Res. Commun. **131**:91–97.
18. **Grieco, F., J. Burgyan, and M. Russo.** 1989. The nucleotide sequence of cymbidium ringspot virus RNA. Nucleic Acids Res. **17**:6383.
19. **Hearne, P. Q., D. A. Knorr, B. I. Hillman, and T. J. Morris.** 1990. The complete genome structure and synthesis of infectious RNA from clones of tomato bushy stunt virus. Virology **177**:141–151.
20. **Hofmann, M. A., and D. A. Brian.** 1991. Sequencing PCR DNA amplified directly from a bacterial colony. BioTechniques **11**: 30–31.
21. **Hofmann, M. A., and D. A. Brian.** 1991. A PCR-enhanced method for sequencing the 5′ end of messenger RNAs. PCR Methods Appl. **1**:43–45.

21a. **Hofmann, M. A., and D. A. Brian.** Unpublished data.

22. **Hofmann, M. A., P. B. Sethna, and D. A. Brian.** 1990. Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture. J. Virol. **64**:4108–4114.
23. **Hogue, B. G., B. King, and D. A. Brian.** 1984. Antigenic relationship between proteins of the bovine coronavirus, human coronavirus OC43, and mouse hepatitis virus A59. J. Virol. **51**:384–388.
24. **Kamahora, T., L. H. Soe, and M. M. C. Lai.** 1989. Sequence analysis of nucleocapsid gene and leader RNA of human coronavirus OC43. Virus Res. **12**:1–9.
25. **Kapke, P. A., and D. A. Brian.** 1986. Sequence analysis of the porcine transmissible gastroenteritis coronavirus nucleocapsid protein gene. Virology **151**:41–49.
26. **Kienzle, T. E., S. Abraham, B. G. Hogue, and D. A. Brian.** 1990. Structure and orientation of expressed bovine coronavirus hemagglutinin-esterase protein. J. Virol. **64**:1834–1838.
27. **Kondo, K., H. Bando, M. Tsurudome, M. Kawano, M. Nishio, and Y. Ito.** 1990. Sequence analysis of the phosphoprotein (P) genes of human parainfluenza type 4A and 4B viruses and RNA editing at transcript of the P genes: the number of G residues added in imprecise. Virology **178**:321–326.
28. **Kozak, M.** 1986. Bifunctional messenger RNAs in eukaryotes. Cell **47**:481–483.
29. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell Biol. **108**:229–241.
30. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
31. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
32. **Lapps, W., B. G. Hogue, and D. A. Brian.** 1987. Sequence analysis is of the bovine coronavirus nucleocapsid and matrix protein genes. Virology **157**:47–57.
33. **Leibowitz, J. L., S. Perlman, G. Weinstock, J. R. DeVries, C. Budzilowicz, J. M. Weissemann, and S. R. Weiss.** 1988. Detection of a coronavirus nonstructural protein encoded in a downstream open reading frame. Virology **164**:156–164.
34. **Liu, D. X., D. Cavanagh, P. Green, and S. C. Inglis.** 1991. A polycistronic mRNA specific by the coronavirus infectious bronchitis virus. Virology **184**:531–544.

34a. **Liu, D. X., and S. C. Inglis.** 1992. Identification of two new polypeptides encoded by mRNAs of the coronavirus infectious bronchitis virus. Virology **186**:342–347.

35. **Luk, D., A. Sanchez, and A. K. Banerjee.** 1986. Messenger RNA encoding the phosphoprotein (P) gene of human parainflenze virus 3 is bicistronic. Virology **153**:318–325.
36. **Luytjes, W., P. J. Bredenbeek, A. F. S. Noten, M. C. Horzinek, and W. J. M. Spaan.** 1988. Sequence of mouse hepatitis virus A59 mRNA2: indications for RNA recombination between

coronaviruses and influenza C virus. Virology **166:**415–422.
37. **Matsuoka, Y., J. Curran, T. Pelet, D. Kolakofsky, R. Ray, and R. W. Compans.** 1991. The P gene of human parainfluenza virus type 1 encodes P and C proteins, but not a cysteine-rich V protein. J. Virol. **65:**3406–3410.
38. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65:**499–560.
39. **Morimoto, T., M. Arpin, and S. Gaetani.** 1983. Use of proteases for the study of membrane insertion. Methods Enzymol. **96:** 121–150.
40. **Paterson, R. G., and R. A. Lamb.** 1990. RNA editing by G-nucleotide insertion in mumps virus P-gene mRNA transcripts. J. Virol. **64:**4137–4145.
41. **Rochon, D. M., and J. C. Johnston.** Infectious transcripts from cloned cucumber necrosis virus cDNA: evidence for a bifunctional subgenomic mRNA. 1991. Virology **181:**656–665.
42. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
43. **Schreiber, S. S., T. Kamahora, and M. M. C. Lai.** 1989. Sequence analysis of the nucleocapsid protein gene of human coronavirus 229E. Virology **169:**142–151.
44. **Sethna, P. B., M. A. Hofmann, and D. A. Brian.** 1991. Minus-strand copies of replicating coronavirus mRNAs contain antileaders. J. Virol. **65:**320–325.
45. **Shaw, M. W., P. W. Choppin, and R. A. Lamb.** 1983. A previously unrecognized influenza B virus glycoprotein from a bicistronic mRNA that also encodes the viral neuraminidase. Proc. Natl. Acad. Sci. USA **80:**4879–4883.
46. **Skinner, M. A., and S. G. Siddell.** 1983. Coronavirus JHM: nucleotide sequence of the mRNA that encodes nucleocapsid protein. Nucleic Acids Res. **11:**5045–5054.
47. **Thierry, P., J. Curran, and D. Kolakofsky.** 1991. The P gene of bovine parainfluenza virus 3 expresses all three reading frames from a single mRNA editing site. EMBO J. **10:**443–448.
48. **Thomas, S. M., R. A. Lamb, and R. G. Paterson.** 1988. Two mRNAs that differ by two nontemplated nucleotides encode the amino coterminal proteins P and V of the paramyxovirus SV5. Cell **54:**891–902.
49. **Vidal, S., J. Curran, and D. Kolakofsky.** 1990. Editing of the Sendai virus P/C mRNA by G insertion occurs during mRNA synthesis via a virus-encoded activity. J. Virol. **64:**239–246.
50. **Walter, P., and G. Blobel.** 1983. Preparation of microsomal membranes for cotranslational protein translocation. Methods Enzymol. **96:**84–93.
51. **Yokomori, K., N. La Monica, S. Makino, C.-K. Shieh, and M. M. C. Lai.** 1989. Biosynthesis, structure, and biological activities of envelope protein gp65 of murine coronavirus. J. Virol. **63:**5310–5318.