JOURNAL OF VIROLOGY, Mar. 1997, p. 1814–1820 Vol. 71, No. 3
0022-538X/97/$04.00+0

Proteolytic Processing of the Coronavirus Infectious Bronchitis Virus 1a Polyprotein: Identification of a 10-Kilodalton Polypeptide and Determination of Its Cleavage Sites

D. X. LIU,[1,2]* H. Y. XU,[1] AND T. D. K. BROWN[2]

Institute of Molecular Agrobiology, National University of Singapore, Singapore 118240,[1] and Division of Virology, Department of Pathology, University of Cambridge, Cambridge CB2 1QP, United Kingdom[2]

Received 30 July 1996/Accepted 12 November 1996

Proteolytic processing of the polyprotein encoded by mRNA 1 is an essential step in coronavirus RNA replication and gene expression. We have previously reported that an open reading frame (ORF) 1a-specific proteinase of the picornavirus 3C proteinase group is involved in processing of the coronavirus infectious bronchitis virus (IBV) 1a/1b polyprotein, leading to the formation of a mature viral protein of 100 kDa. We report here the identification of a novel 10-kDa polypeptide and the involvement of the 3C-like proteinase in processing of the ORF 1a polyprotein to produce the 10-kDa protein species. By using a region-specific antiserum, V47, raised against a bacterial-viral fusion protein containing IBV sequence encoded between nucleotides 11488 and 12600, the 10-kDa polypeptide was detected in lysates from both IBV-infected and plasmid DNA-transfected Vero cells. Coexpression, deletion, and mutagenesis studies showed that this novel polypeptide was encoded by ORF 1a from nucleotide 11545 to 11878 and was cleaved from the 1a polyprotein by the 3C-like proteinase domain. Evidence presented suggested that a previously predicted Q-S ($Q^{3783}S^{3784}$) dipeptide bond encoded by ORF 1a between nucleotides 11875 and 11880 was responsible for the release of the C terminus of the 10-kDa polypeptide and that a novel Q-N ($Q^{3672}N^{3673}$) dipeptide bond encoded between nucleotides 11542 and 11547 was responsible for the release of the N terminus of the 10-kDa polypeptide.

Six mRNA species are produced in cells infected with the prototype virus of the *Coronaviridae*, avian infectious bronchitis virus (IBV). These include the genome-length mRNA (mRNA 1), of 27.6 kb, and five subgenomic mRNA species (mRNAs 2 to 6), with sizes ranging from 2 to 7 kb. The evidence available suggests that the four virion structural proteins, i.e., the spike (S), membrane (M), nucleocapsid (N), and small envelope (E) proteins, are encoded by subgenomic mRNA 2, mRNA 4, mRNA 6, and the third open reading frame (ORF) of mRNA 3, respectively (11, 21). Nucleotide sequencing of the genomic RNA of IBV has shown that the 5′-terminal unique region of mRNA 1 contains two large ORFs (ORFs 1a and 1b), with ORF 1a having the potential to encode a polyprotein of 441 kDa and ORF 1b having the potential to encode a polyprotein of 300 kDa (1) (Fig. 1). The downstream ORF 1b protein is produced as a fusion protein of 741 kDa with 1a by a ribosomal frameshift (2, 3, 10). The 1a/1b fusion polyprotein is expected to be cleaved by viral or cellular proteinases to produce functional products associated with viral RNA replication (Fig. 1).

Recently, we have reported the identification of products encoded by mRNA 1 in both virus-infected and plasmid DNA-transfected cells by using region-specific antisera raised in rabbits against bacterial-viral fusion proteins. This includes the identification of an 87-kDa protein encoded by the 5′-most part of ORF 1a and a 100-kDa protein encoded by ORF 1b (13, 14). Both viral papain-like proteinase domains encoded by the ORF 1a sequence from nucleotide 4242 to 5553 (Fig. 1) and cellular proteinases have been demonstrated to be involved in the proteolytic processing of the 1a polyprotein to the 87-kDa protein, and a picornavirus 3C-like proteinase (3C-like proteinase) encoded by the ORF 1a sequence between nucleotides 8937 and 9357 (Fig. 1) has been shown to be responsible for the proteolytic processing of the 1a/1b fusion polyprotein to the 100-kDa protein species (13, 14). Meanwhile, characterization of the 3C-like proteinase domain by use of internal deletion and substitution mutations of its predicted catalytic center and the cleavage sites has been carried out. These studies confirmed that the predicted nucleophilic cysteine residue (Cys^{2922}) and the histidine 2820 (His^{2820}) residue are essential for the proteinase activity and that two putative Q-S dipeptide bonds flanking the 100-kDa protein-encoding region are the target cleavage sites of the 3C-like proteinase.

We report here the identification of a novel 10-kDa polypeptide in IBV-infected Vero cells by using region-specific antiserum V47, raised against a bacterial-viral fusion protein containing the IBV sequence encoded between nucleotides 11488 and 12600. Coexpression, deletion, and mutagenesis studies showed that this novel polypeptide was encoded by ORF 1a between nucleotides 11545 and 11877 and was cleaved from the 1a polyprotein by the 3C-like proteinase domain. A previously predicted Q-S ($Q^{3783}S^{3784}$) dipeptide bond encoded by ORF 1a between nucleotides 11875 and 11880 was demonstrated to be responsible for releasing the C terminus of the 10-kDa protein, and a novel Q-N ($Q^{3672}N^{3673}$) dipeptide bond specified by the IBV sequence between nucleotides 11542 and 11547 was responsible for the release of the N terminus of the 10-kDa protein. The first identification of a Q-N dipeptide bond as an authentic cleavage site of the IBV 3C-like proteinase presented here illustrates the diversity of the cleavage specificities of the proteinase and suggests that proteolytic cleavage of the 1a/1b polyprotein is a more complicated process than the previously predicted.

* Corresponding author. Mailing address: Institute of Molecular Agrobiology, National University of Singapore, 59A the Fleming, 1 Science Park Dr., Singapore 118240. Phone: 65 7719823. Fax: 65 7742857. E-mail: imaliudx@leonis.nus.sg.

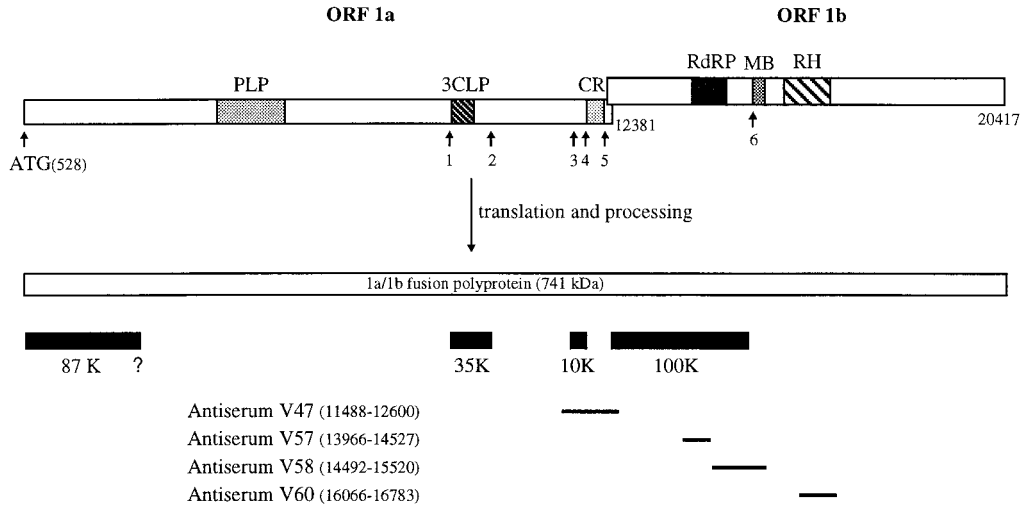

FIG. 1. Diagram showing the structures of ORFs 1a and 1b and the locations of four cleavage products, the 87-kDa (87K), 35-kDa (35K), 10-kDa (10K), and 100-kDa (100K) proteins, of the 1a/1b polyprotein. The locations of several putative functional domains, including the papain-like proteinase (PLP), the 3C-like proteinase (3CLP), the cysteine-rich region (CR), the RNA-dependent RNA polymerase (RdRP), the metal binding region (MB) and the RNA helicase (RH), are shown. Also shown are the nucleotide positions of the IBV sequences used to raise antisera V47, V57, V58, and V60 and the Q-S(N) cleavage sites responsible for the release of the 35-, 10-, and 100-kDa proteins. The positions of the last nucleotides of the triplets encoding the Q residues of the Q-S(N) dipeptide bonds are as follows: 1, 8865; 2, 9786; 3, 11544; 4, 11877; 5, 12312; and 6, 15131. The two Q-S cleavage sites flanking the 35-kDa protein containing the 3C-like proteinase domain were determined by in vitro expression and mutagenesis studies (23).

MATERIALS AND METHODS

Virus and cells. The Beaudette strain of IBV was used in this analysis. Virus stocks were prepared by infecting Vero cells at a multiplicity of approximately 0.1 PFU/cell and incubating for 48 h. The virus was assayed by plaquing on Vero cells.

Vero cells were grown at 37°C in 5% CO_2 and maintained in Glasgow's modified minimal essential medium supplemented with 10% fetal calf serum.

Labelling of IBV-infected cells with [^{35}S]methionine. Confluent monolayers of Vero cells were infected with IBV at a multiplicity of approximately 2 PFU/cell. Prior to being labelled, the cells were incubated in methionine-free medium for 30 min. After 120 min of labelling with [^{35}S]methionine (60 μCi/ml), the cells were scraped off the dishes in phosphate-buffered saline, recovered by centrifugation, and stored at −70°C.

Radioimmunoprecipitation. IBV-infected and plasmid DNA-transfected Vero cells were lysed with radioimmunoprecipitation assay buffer (50 mM Tris HCl [pH 7.5], 150 mM NaCl, 1% sodium deoxycholate, 0.1% sodium dodecyl sulfate [SDS]) and precleared by centrifugation at 15,000 × *g* for 5 min at 4°C. Radioimmunoprecipitation with polyclonal rabbit antisera was carried out as described previously (12).

SDS-PAGE. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of virus polypeptides was carried out with a range of polyacrylamide concentrations (12.5 to 17.5%) (7). Labelled polypeptides were detected by autoradiography or fluorography of dried gels.

Cell-free transcription and translation. Plasmid DNA was transcribed and translated in vitro by using the TnT T7-coupled reticulocyte lysate system (TnT system) according to the instructions of the manufacturer (Promega). Reaction products were separated by SDS-PAGE and detected by autoradiography.

PCR. Appropriate primers and template DNAs were used in amplification reactions with *Pfu* DNA polymerase (Stratagene) under standard buffer conditions with 2 mM $MgCl_2$. The PCR conditions were 30 cycles of 92°C for 30 s, 56°C for 30 s, and 72°C for 6 min.

Site-directed mutagenesis. Site-directed mutagenesis was carried out by two rounds of PCR with two pairs of primers. The first pair was referred to as the cloning primers (primers 1 and 2). Primer 1 is the upstream primer, which represents the 5′ end sequence of the PCR fragment, and primer 2 is the downstream primer, representing the complementary sequence of the 3′ end of the PCR fragment. The second pair was referred to as the mutation primers (primers 3 and 4). Primer 3 is the upstream primer containing the desired mutation, and primer 4 is the downstream primer with sequence complementary to primer 3.

The first round of PCR was carried out with appropriate template and primer combinations of primer 1-primer 4 and primer 3-primer 2. The two PCR fragments were then gel purified. Equal amounts of DNAs were mixed and used for the second round of PCR together with primer 1 and primer 2. The PCR fragment was then cloned into an appropriate vector, and the mutant was selected by nucleotide sequencing.

Construction of plasmids. Plasmid pIBV3C, which covers the IBV sequence from nucleotide 8871 to 9786 with an AUG initiation codon in an optimal context (ACCAUGG) located immediately upstream of the viral sequence and a UAG termination codon located immediately downstream of the viral sequence, was constructed by cloning an *Nco*I/*Bam*HI-digested PCR fragment into *Nco*I/*Bam*HI-digested pKT0 (14). The sequence of the upstream PCR primer is 5′-AGATTACCATGGGTTTTAAGAAAC-3′, and that of the downstream primer is 5′-TTAACAGGATCCTATTGTAATCTAAC-3′.

Plasmid pBP5, which contains the IBV sequence from nucleotide 10752 to 12600 and was originally called pKTBP5 (13), was used to construct three deletion plasmids. An *Hin*dIII/*Sna*BI restriction fragment covering the IBV sequence from nucleotide 11306 to 12227, was cloned into *Hin*dIII/*Sma*I-digested pKT0, giving plasmid pBP5Δ2.

Plasmid pBP5Δ4 was made by cloning an *Nco*I/*Bam*HI-digested PCR fragment covering the IBV sequence from nucleotide 11568 to 12312 into *Nco*I/*Bam*HI-digested pKT0. The sequence of the upstream PCR primer is 5′-GCC ACCCATGGTTAAAACAAAGGC-3′, and that of the downstream primer is 5′-GCTCCAGGATCCTATTGAACAGAAGA-3′.

Plasmid pBP5Δ5 was made by cloning an *Nco*I/*Bam*HI-digested PCR fragment covering the IBV sequence from nucleotide 11306 to 11877 into *Nco*I/*Bam*HI-digested pKT0. The upstream primer is a 20-mer oligonucleotide corresponding to the T7 promoter sequence (5′-TAATACGACTCACTATAGGG-3′), and the sequence of the downstream primer is 5′-GCTCCAGGATCCTACTGTAAGACAAC-3′.

Plasmid pBP5Δ4Q^{3783}-E was created by PCR with pBP5Δ4 as the template. The sequence of the mutation primer is 5′-GTTGTTGTCTTAGAGTCTAAAGG-3′, and the sequence of the complementary primer is 5′-CCTTTAGACTCTAAGACAACAAC-3′.

Plasmid pBP5Δ5Q^{3672}-E was constructed by PCR with pBP5Δ5 as the template. The sequence of the mutation primer is 5′-GATGTTGTTTTGGAGAATAATGAGCTT-3′, and the sequence of the complementary primer is 5′-CTCATTATTCTCCAAAACAACATCAAC-3′.

RESULTS

Further identification of gene products encoded by mRNA 1 in IBV-infected Vero cells. We have previously reported the application of four region-specific antisera (V47, V57, V58, and V60) for detection of viral products encoded by the 3′ region of ORF 1a, the junction of 1a/1b, and the 5′ region of ORF 1b (13). Immunoprecipitation studies of IBV-infected and plasmid-transfected Vero cells with antisera V57 and V58 led to the identification of a 100-kDa polypeptide, which was subsequently identified to be encoded by the IBV sequence from nucleotide 12313 to 15131 and cleaved from the 1a/1b polyprotein by the 3C-like proteinase (10, 13). However, no viral products could be detected by immunoprecipitation with

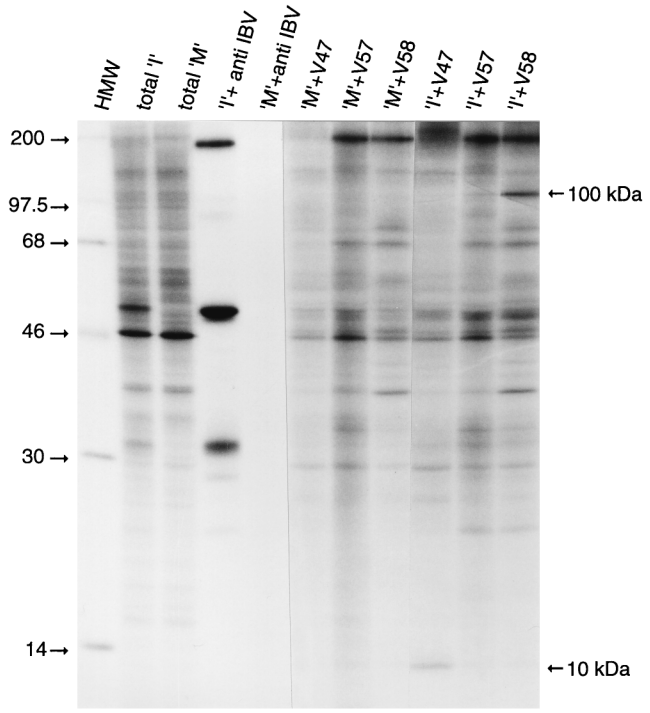

FIG. 2. Detection of polypeptides encoded by mRNA 1 in IBV-infected ('I') and mock-infected ('M') Vero cells by immunoprecipitation with region-specific antisera (see text). Cells were labelled with [^{35}S]methionine, lysates were prepared, and polypeptides were either analyzed directly or immunoprecipitated with the antiserum indicated above each lane. Polypeptides were separated by SDS–17.5% PAGE and detected by fluorography. HMW, high-molecular-mass markers (numbers indicate kilodaltons).

the other two antisera (V47 and V60), although immunoprecipitation of in vitro-synthesized 1a/1b proteins showed that antiserum V47, raised in rabbits against a bacterial-viral fusion protein containing the IBV sequence encoded between nucleotides 11488 and 12600, could efficiently precipitate its target products (13). The failure to detect any specific viral products with this antiserum may either reflect the low abundance of the corresponding cleavage products in virus-infected cells or simply indicate that unexpected cleavages may occur in this region, leading to the formation of products too small to be detected by the SDS-PAGE system used. To investigate these possibilities, confluent monolayers of Vero cells were infected with IBV at a higher multiplicity, and the precipitated products were separated by SDS–17.5% PAGE. Figure 2 shows the results of immunoprecipitation of cell lysates prepared from Vero cells infected with IBV at a multiplicity of approximately 2 PFU per cell. As can be seen, antiserum V58 efficiently precipitated the 100-kDa protein from IBV-infected-cell lysates (Fig. 2). The same protein species was also detected by immunoprecipitation with antiserum V57 (Fig. 2). These results are consistent with the data previously reported (13), although more efficient detection of the 100-kDa protein was observed with both antisera. Immunoprecipitation of the same virus-infected-cell lysates with antiserum V47, however, led to specific detection of a polypeptide with an apparent molecular mass of approximately 10 kDa (Fig. 2), which was not observed in our previous study (13).

Reactivity of antiserum V47 to the in vitro-synthesized target products. The detection of a 10-kDa polypeptide from IBV-infected cells with antiserum V47 was unexpected, as no cleavage sites in the V47-recognizing region were predicted to be used by the 3C-like proteinase to release a 10-kDa product (Fig. 1). In order to locate precisely the region encoding the 10-kDa protein, the reactivity of antiserum V47 to its target products was determined by immunoprecipitation of in vitro-synthesized proteins. For this purpose, plasmid pBP5, which contains the IBV sequence from nucleotide 10752 to 12600 (13), was linearized at different positions with restriction enzymes and transcribed and translated in vitro with the TnT transcription-coupled translation system (Promega). As shown in Fig. 3a, three restriction enzymes were used to linearize pBP5: *Acc*I digests the IBV sequence at nucleotide 11739, *Sna*BI digests the IBV sequence at position 12227, and *Bam*HI incises the vector sequence immediately downstream of the viral sequence. Transcription and translation of *Acc*I-digested and *Sna*BI-digested pBP5 led to the synthesis of two protein species with molecular masses of approximately 38 and 56 kDa, respectively, representing the full-length products encoded by these constructs (Fig. 3b). Transcription and translation of *Bam*HI-digested pBP5 resulted in the synthesis of two protein species of approximately 60 and 72 kDa (Fig. 3b), representing the 1a termination and the 1a/1b frameshifting products, respectively (13). Immunoprecipitation showed that all four protein species can be efficiently precipitated by antiserum V47 (Fig. 3b). As antiserum V47 was raised against a bacterial-viral fusion protein containing the product of the IBV sequence from nucleotide 11488 to 12600, these results indicate that an antigenic determinant recognized by this serum is located in the 1a polyprotein encoded between nucleotides 11488 and 11739.

Coexpression of pBP5 with pIBV3C in a eukaryotic system. Plasmid pBP5 was then expressed either on its own or together with pIBV3C in Vero cells by using the vaccinia virus-T7 expression system (5). Plasmid pIBV3C covers the whole 3C-like proteinase-encoding region from nucleotide 8871 to 9786 with an artificial AUG codon in an optimal context (ACCAUGG) located immediately upstream of the viral sequence, and this proteinase has been shown to be able to cleave the 1a/1b polyprotein in *trans* (data not shown). As shown in Fig. 4, expression of pBP5 in Vero cells led to the synthesis of the 60- and 72-kDa protein species; no processing of the two proteins to smaller polypeptides was observed. Cotransfection of pBP5 with pIBV3C, however, resulted in complete processing of the 60- and 72-kDa proteins (Fig. 4). Instead of the 60- and 72-kDa protein species, a polypeptide with an apparent molecular mass of 10 kDa was detected (Fig. 4). This 10-kDa protein species was shown to comigrate in SDS-PAGE with the 10-kDa polypeptide detected in IBV-infected Vero cells (data not shown). No other processing products were detected (Fig. 4).

Determination of the region encoding the 10-kDa polypeptide by deletion analysis. The results obtained from the coexpression of pBP5 and pIBV3C (Fig. 4) confirm that the 10-kDa polypeptide is encoded within the IBV sequence from nucleotide 10752 to 12600. To define further the sequence encoding the 10-kDa polypeptide, constructs covering different parts of the V47-recognizing region were made by ligation of different restriction fragments obtained from pBP5 or PCR fragments into a T7 promoter-based expression vector (Fig. 5a) and expressed either alone or together with pIBV3C in Vero cells by using the vaccinia virus-T7 system.

Expression of pBP5Δ2, a plasmid which covers the IBV sequence from nucleotide 11306 to 12227 and therefore contains the putative $Q^{3783}S^{3784}$ cleavage site encoded by the IBV sequence from nucleotide 11875 to 11880 (Fig. 5a), resulted in the synthesis of a protein of approximately 35 kDa, representing the full-length product encoded by this plasmid (Fig. 5b).

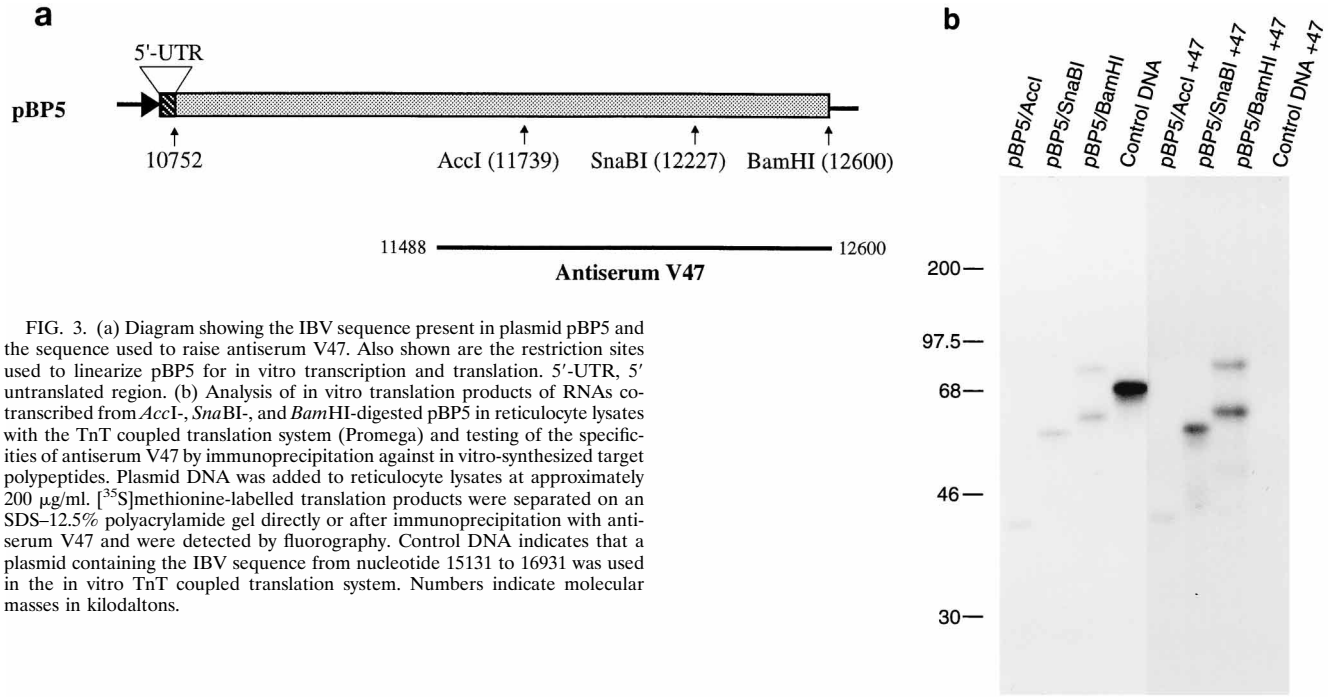

FIG. 3. (a) Diagram showing the IBV sequence present in plasmid pBP5 and the sequence used to raise antiserum V47. Also shown are the restriction sites used to linearize pBP5 for in vitro transcription and translation. 5′-UTR, 5′ untranslated region. (b) Analysis of in vitro translation products of RNAs co-transcribed from *Acc*I-, *Sna*BI-, and *Bam*HI-digested pBP5 in reticulocyte lysates with the TnT coupled translation system (Promega) and testing of the specificities of antiserum V47 by immunoprecipitation against in vitro-synthesized target polypeptides. Plasmid DNA was added to reticulocyte lysates at approximately 200 μg/ml. [^{35}S]methionine-labelled translation products were separated on an SDS–12.5% polyacrylamide gel directly or after immunoprecipitation with antiserum V47 and were detected by fluorography. Control DNA indicates that a plasmid containing the IBV sequence from nucleotide 15131 to 16931 was used in the in vitro TnT coupled translation system. Numbers indicate molecular masses in kilodaltons.

Coexpression of pBP5Δ2 with pIBV3C led to the detection of the 10-kDa protein (Fig. 5b), indicating that the 10-kDa protein is encoded within this region of the IBV sequence and that the putative $Q^{3783}S^{3784}$ dipeptide bond might be recognized and cleaved by the 3C-like proteinase to release the 10-kDa protein.

Expression of pBP5Δ4, which covers the IBV sequence from nucleotide 11568 to 12312 and contains a termination codon (UAG) immediately downstream of the viral sequence (Fig. 5a), resulted in the synthesis of the full-length protein of approximately 30 kDa (Fig. 5c). Coexpression of pBP5Δ4 with pIBV3C led to the detection of a novel product migrating in SDS-PAGE slightly faster than the 10-kDa protein (Fig. 5c). The predicted size of this product is approximately 9 kDa. It is therefore possible that the 9-kDa polypeptide may represent an N-terminally truncated version of the 10-kDa protein and that its C terminus may be released by cleavage at the predicted $Q^{3783}S^{3784}$ dipeptide bond.

Determination of the C-terminal cleavage site of the 10-kDa protein by mutational analysis. Our previous report showed that substitution mutation of the Q residue of the predicted Q-S(G) cleavage site with an E residue totally abolished cleavage at this dipeptide bond (10). Mutation of the S residue to several different residues, however, had no obvious effect (10). As the data presented above strongly suggest that the predicted $Q^{3783}S^{3784}$ cleavage site may be responsible for release of the C terminus of the 10-kDa protein, substitution mutation of the Q residue of the $Q^{3783}S^{3784}$ dipeptide bond with an E residue was used to define the cleavage site. Site-directed mutagenesis was therefore carried out to introduce this Q-to-E mutation, giving the construct pBP5Δ4Q^{3783}-E. Expression of pBP5Δ4Q^{3783}-E in Vero cells resulted in the detection of the 30-kDa full-length product (Fig. 6). Coexpression of pBP5Δ4Q^{3783}-E with pIBV3C showed that processing of the full-length 30-kDa protein to the 9-kDa polypeptide was significantly inhibited; only a trace amount of the 9-kDa species was detected by immunoprecipitation with antiserum V47 (Fig. 6).

Determination of the N-terminal cleavage site of the 10-kDa protein. The mutagenesis studies presented above demon-

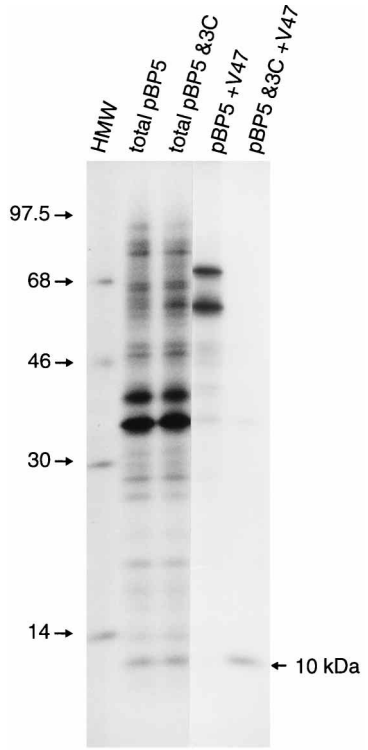

FIG. 4. Analysis of transiently expressed mRNA 1 products from cotransfection of plasmids pBP5 and pIBV3C, using the system described by Fuerst et al. (5), in which Vero cells were infected with a recombinant vaccinia virus expressing the T7 RNA polymerase and subsequently transfected with plasmid DNA. Cells were labelled with [^{35}S]methionine, lysates were prepared, and polypeptides were either analyzed directly or immunoprecipitated with antiserum V47. Polypeptides were separated by SDS–17.5% PAGE and detected by fluorography. HMW, high-molecular-mass markers (numbers indicate kilodaltons).

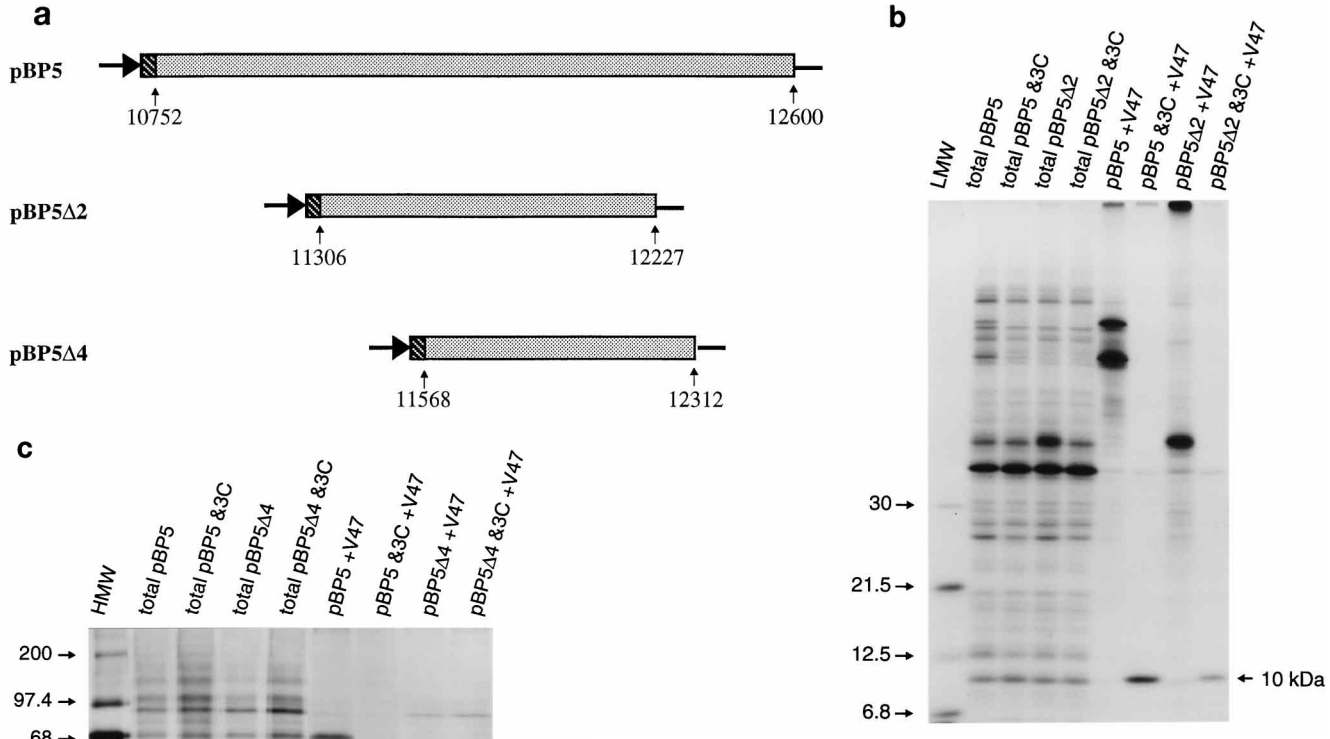

FIG. 5. (a) Diagram showing IBV sequences present in plasmids pBP5, pBP5Δ2, and pBP5Δ4. (b and c). Determination of the region encoding the 10-kDa polypeptide by deletion analysis. Plasmids pBP5 and either pBP5Δ2 (b) or pBP5Δ4 (c) were transiently expressed in Vero cells by transfection either alone or together with pIBV3C, using the recombinant vaccinia virus-T7 expression system as described in the legend for Fig. 4. The transfected cells were labelled with [^{35}S]methionine, lysates were prepared, and polypeptides were analyzed directly or immunoprecipitated with antiserum V47. Polypeptides were separated by SDS–17.5% PAGE and detected by fluorography. LMW and HMW, low- and high-molecular-mass markers, respectively (numbers indicate kilodaltons).

strated that the $Q^{3783}S^{3784}$ dipeptide bond was the cleavage site responsible for the release of the C terminus of the 10-kDa protein. Furthermore, coexpression of the IBV sequence from nucleotide 11568 to 12312 (pBP5Δ4) with the 3C-like proteinase domain led to detection of an N-terminally truncated version of the 10-kDa protein. These results suggest that the N-terminal cleavage site of the 10-kDa polypeptide is encoded between nucleotides 11306 and 11568. To facilitate mutagenesis study of the N-terminal cleavage site of the 10-kDa protein, plasmid pBP5Δ5, which covers the IBV sequence from nucleotide 11306 to 11877 and contains a termination codon (UAG) immediately downstream of the viral sequence, was constructed. Expression of pBP5Δ5 resulted in the synthesis of a polypeptide of approximately 25 kDa, representing the full-length product encoded by this construct (Fig. 7). Coexpression of pBP5Δ5 with pIBV3C led to the detection of the 10-kDa protein species, which comigrates with the 10-kDa protein detected from coexpression of pBP5 and pIBV3C (Fig. 7). As the C terminus of the polyprotein encoded by pBP5Δ5 is Q^{3783}, these results reinforce the conclusion that the $Q^{3783}S^{3784}$ dipeptide bond is the C-terminal cleavage site of the 10-kDa protein and demonstrate that the 10-kDa protein is encoded within the IBV sequence from nucleotide 11306 to 11877.

Site-directed mutagenesis was then carried out to identify the N-terminal cleavage site. Based on the expression and processing patterns of pBP5Δ5, it was assumed that one of the Q-X dipeptide bonds located within the 1a polyprotein encoded from nucleotide 11306 to 11877 was the candidate cleavage site. Examination of the amino acid sequence of this region indicated that two Q-X dipeptide bonds might be recognized and cleaved by the 3C-like proteinase to release the N terminus of the 10-kDa protein. The first potential site is the $Q^{3692}A^{3693}$ dipeptide encoded by the IBV sequence from nucleotide 11602 to 11607. Cleavage at this position would result in the synthesis of a polypeptide with a calculated molecular mass of 11.4 kDa, which is consistent with the predicted molecular mass of the 10-kDa protein. However, data presented in Fig. 5c showed that coexpression of pBP5Δ4 with pIBV3C led to the synthesis of a polypeptide migrating in SDS-PAGE slightly faster than the 10-kDa protein, ruling out the possibility that this site is responsible for the release of the N terminus of the 10-kDa protein. The second potential cleavage site is the $Q^{3672}N^{3673}$ dipeptide bond encoded between nucleotides 11542 and 11547. Cleavage at this position would result in the synthesis of a polypeptide with a calculated molecular mass of 13.9 kDa. To investigate whether this Q-N dipeptide is responsible for release of the N terminus of the 10-kDa protein,

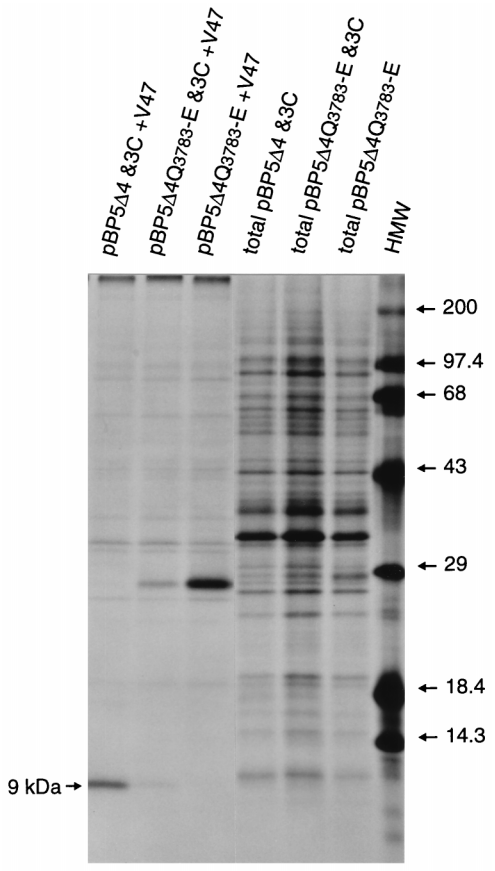

FIG. 6. Mutational analysis of the predicted $Q^{3783}S^{3784}$ cleavage site responsible for release of the C terminus of the 10-kDa polypeptide. The mutants were transiently expressed in Vero cells, using the recombinant vaccinia virus-T7 expression system. The transfected cells were labelled with [^{35}S]methionine, lysates were prepared, and polypeptides were either analyzed directly or immunoprecipitated with antiserum V47. Polypeptides were separated by SDS–17.5% PAGE and detected by fluorography. HMW, high-molecular-mass markers (numbers indicate kilodaltons).

site-directed mutagenesis was carried out to mutate the Q^{3672} residue to E, giving the construct pBP5Δ5Q^{3672}-E. As shown in Fig. 7, expression of plasmid pBP5Δ5Q^{3672}-E led to the synthesis of the full-length 25-kDa protein. Coexpression of plasmid pBP5Δ5Q^{3672}-E with pIBV3C resulted in the synthesis of only the full-length 25-kDa protein; no formation of the 10-kDa product was observed (Fig. 7). These results demonstrate that the $Q^{3672}N^{3673}$ dipeptide bond is the cleavage site responsible for the release of the N terminus of the 10-kDa protein.

DISCUSSION

Coronavirus gene expression involves proteolytic processing of the mRNA 1-encoded polyproteins by viral and cellular proteinases. Recently, we have demonstrated that an ORF 1a-encoded proteinase domain of the picornavirus 3C proteinase group is responsible for proteolytic processing of the IBV 1a/1b fusion polyprotein to a 100-kDa protein species (10, 13). Subsequently, the 100-kDa protein was identified to be encoded by the IBV sequence from nucleotide 12313 to 15131, and the two putative Q-S dipeptide bonds flanking the 100-kDa protein were confirmed to be recognized and cleaved by the 3C-like proteinase to release the 100-kDa protein (10). In this communication, we report the identification of a novel ORF 1a-specific polypeptide of 10 kDa and the involvement of the 3C-like proteinase in processing of the 1a polyprotein. Evidence presented above showed that this novel polypeptide was encoded by ORF 1a between nucleotides 11545 and 11877 and was cleaved from the 1a polyprotein by the 3C-like proteinase domain. Deletion and mutagenesis studies revealed that a previously predicted Q-S ($Q^{3783}S^{3784}$) dipeptide bond encoded by ORF 1a from nucleotide 11875 to 11880 was responsible for the release of the C terminus of the 10-kDa protein and that a novel Q-N ($Q^{3672}N^{3673}$) dipeptide bond encoded between nucleotides 11542 and 11547 was responsible for the release of the N terminus of the 10-kDa protein.

The demonstration here that a Q-N dipeptide bond is one of the authentic cleavage sites for the IBV 3C-like proteinase suggests a great diversity of the cleavage specificities of the proteinases in this group. Our previous mutagenesis studies showed that replacement of the $Gln^{891(1b)}$ residue with a Glu completely blocked the processing of the 1a/1b polyprotein to the 100-kDa protein (10). However, alterations of the $Ser^{892(1b)}$ residue to either Ala or Gly had no effect on this processing, suggesting that both Gln-Ala and Gln-Gly dipeptide bonds can be recognized and cleaved efficiently by the IBV 3C-like proteinase (10). For picornaviruses, the 3C proteinases also show diversity in recognizing their target cleavage sites. It was reported that most cleavages mediated by picoronavirus 3C proteinases occurred at the Gln-Gly peptide bond; less commonly,

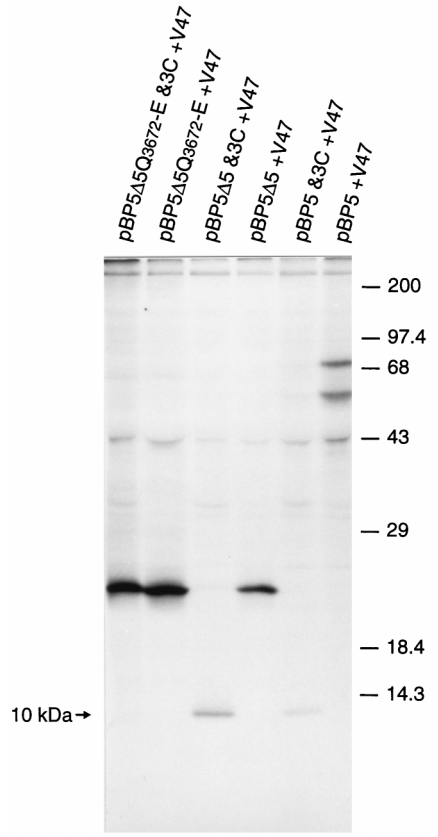

FIG. 7. Determination of the N-terminal cleavage site of the 10-kDa polypeptide by deletion and mutational analyses. Plasmids pBP5, pBP5Δ5, and pBP5Δ5Q^{3672}-E were transiently expressed in Vero cells, using the recombinant vaccinia virus-T7 expression system. The transfected cells were labelled with [^{35}S]methionine, lysates were prepared, and polypeptides were immunoprecipitated with antiserum V47, separated by SDS–17.5% PAGE, and detected by fluorography. Numbers indicate molecular masses in kilodaltons.

cleavages between Gln-Ser, Gln-Ala, Glu-Ser, or Glu-Gly pairs were also observed (16). Some Q-X dipeptide bonds are authentic cleavage sites in a certain virus but are inactive in others. For example, Q-T and Q-I are cleavage sites in foot-and-mouth disease virus but cannot be recognized by the 3C proteinase of encephalomyocarditis virus (4, 17). Systematic substitution mutations are required to provide a complete picture of the cleavage specificity of the IBV 3C-like proteinase.

It is intriguing that some internal Q-S (Q-G or Q-A) dipeptide bonds could not be recognized and cleaved by the 3C-like proteinase domain. A typical example here is the $Q^{3692}A^{3693}$ dipeptide bond encoded by the IBV sequence from nucleotide 11602 to 11607. The mechanism involved in this discrimination is currently unknown. Since no obvious sequence homology was found around the authentic Q-X cleavage sites, the primary amino acid sequence context is unlikely to be a predominant determinant. Interestingly, examination of the amino acid sequence of the 1a/1b polyprotein showed that if a Q-X dipeptide bond was located in a functional domain, this site was usually inactive. Examples include two Q-G dipeptide bonds [$Q^{566(1b)}G^{567(1b)}$ and $Q^{732(1b)}G^{733(1b)}$], encoded by the IBV sequence from nucleotide 14154 to 14159 and from nucleotide 14652 to 14657, respectively; a Q-A dipeptide bond [$Q^{657(1b)}A^{658(1b)}$], encoded between nucleotides 14427 and 14432 within the polymerase domain; and two Q-G dipeptide bonds [$Q^{1172(1b)}G^{1173(1b)}$ and $Q^{1430(1b)}G^{1431(1b)}$], encoded from nucleotide 15972 to 15977 and from nucleotide 16746 to 16751, respectively, within the helicase domain (10, 14a). This suggests that local folding of a certain functional domain might mask a suitable Q-X dipeptide bond located in the region and block the cleavage occurring at that position. Further investigations are required to address this issue.

During the manipulation of plasmid pBP5 and its derivatives, we noted that a segment of viral cDNA in this region is either unstable in or highly toxic to all of the *Escherichia coli* host strains used. The yields of these plasmids were very low, and no bacterial colonies could be recovered by transformation of plasmid DNA into bacteria (data not shown). We also observed a spontaneous deletion of 240 bp of IBV sequence from nucleotide 11385 to 11625 when site-directed mutagenesis was carried out by using uracil-containing single-stranded DNA templates prepared from pBP5Δ2 (data not shown). Bacteria harboring this deletion plasmid exhibited normal growth properties, and normal transformation efficiencies was observed with the deletion plasmid. Unstable and/or toxic viral cDNA fragments were also reported for other plant and animal viruses. For example, multiplication of full-length cDNA clones of yellow fever virus (19) and Japanese encephalitis virus (22) in bacteria always led to introduction of mutations in the viral sequences. Similar problems were also observed during construction of cDNA clones from dengue virus type 4 (8), RNA 2 of beet necrotic yellow vein virus (18), pea early browning virus (15), and eggplant mosaic virus (20).

Throughout this report, the 10-kDa polypeptide has been referred to as such based on its migration in SDS-PAGE. However, the amino acid sequence of this protein suggests that it has a calculated molecular mass of 13.9 kDa. The decreased apparent molecular mass of the protein in SDS-PAGE might be a reflection of its amino acid composition. Examination of the deduced amino acid sequence of this protein showed that it contained a very high proportion of valine residues (16.22%). As no functional domains have so far been identified in this region of the viral genome by computer-aided methods (6, 9), it is not clear if this unusual amino acid composition might reflect some functional significance of this protein in the virus infection cycle.

ACKNOWLEDGMENT

This work was partially supported by the Agriculture and Food Research Council, United Kingdom.

REFERENCES

1. **Boursnell, M. E. G., T. D. K. Brown, I. J. Foulds, P. F. Green, F. M. Tomley, and M. M. Binns.** 1987. Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J. Gen. Virol. **68:**57–77.
2. **Brierley, I., M. E. G. Boursnell, M. M. Binns, B. Bilimoria, V. C. Blok, T. D. K. Brown, and S. C. Inglis.** 1987. An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. EMBO. J. **6:**3779–3785.
3. **Brierley, I., P. Digard, and S. C. Inglis.** 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell **57:**537–547.
4. **Carroll, A. R., D. J. Rowlands, and B. E. Clarke.** 1984. The complete nucleotide sequence of the RNA coding for the primary translation product of foot-and-mouth disease virus. Nucleic Acids Res. **12:**2461–2472.
5. **Fuerst, T. R., E. G. Niles, F. W. Studier, and B. Moss.** 1986. Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. Proc. Natl. Acad. Sci. USA **83:**8122–8127.
6. **Gorbalenya, A. E., E. Y. Koonin, A. P. Donchenko, and V. M. Blinov.** 1989. Coronavirus genome: prediction of putative functional domains in the nonstructural polyprotein by comparative amino acid sequence analysis. Nucleic Acids Res. **17:**4847–4860.
7. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
8. **Lai, C. J., B. Zhao, H. Hori, and M. Bray.** 1991. Infectious RNA transcribed from stably cloned full-length cDNA of dengue type-4 virus. Proc. Natl. Acad. Sci. USA **88:**5139–5143.
9. **Lee, H.-J., C.-K. Shieh, A. E. Gorbalenya, E. V. Koonin, N. L. Monica, J. Tuler, A. Bagdzhadzhyan, and M. M. C. Lai.** 1991. The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. Virology **180:**567–582.
10. **Liu, D. X., and T. D. K. Brown.** 1995. Characterization and mutational analysis of an ORF-1a-encoding proteinase domain responsible for proteolytic processing of the infectious bronchitis virus 1a/1b polyprotein. Virology **209:**420–427.
11. **Liu, D. X., and S. C. Inglis.** 1991. Association of the infectious bronchitis virus 3C protein with the virion envelope. Virology **185:**911–917.
12. **Liu, D. X., D. Cavanagh, P. Green, and S. C. Inglis.** 1991. A polycistronic mRNA specified by the coronavirus infectious bronchitis virus. Virology **184:**531–544.
13. **Liu, D. X., I. Brierley, K. W. Tibbles, and T. D. K. Brown.** 1994. A 100-kilodalton polypeptide encoded by open reading frame (ORF) 1b of the coronavirus infectious bronchitis virus is processed by ORF 1a products. J. Virol. **68:**5772–5780.
14. **Liu, D. X., K. W. Tibbles, D. Cavanagh, T. D. K. Brown, and I. Brierley.** 1995. Identification, expression and processing of an 87 kDa polypeptide encoded by ORF1a of the coronavirus infectious bronchitis virus. Virology **208:**48–57.

14a. **Liu, D. X., et al.** Unpublished observations.

15. **MacFarlane, S. A., C. V. Wallis, S. C. Taylor, M. G. Goulden, K. R. Wood, and J. W. Davies.** 1991. Construction and analysis of infectious transcripts synthesized from full-length clones of both genomic RNAs of pea early browning virus. Virology **182:**124–129.
16. **Palmenberg, A. C.** 1990. Proteolytic processing of picornaviral polyprotein. Annu. Rev. Microbiol. **44:**603–623.
17. **Parks, G. D., J. C. Baker, and A. C. Palmenberg.** 1989. Proteolytic cleavage of encephalomyocarditis viral capsid region substrates by precursors to the 3C enzyme. J. Virol. **63:**1054–1058.
18. **Quillet, L., H. Guilley, G. Jonard, and K. Richards.** 1989. In vitro synthesis of biologically active beet necrotic yellow vein virus RNA. Virology **172:**293–301.
19. **Rice, C. M., A. Grakoui, R. Galler, and T. J. Chambers.** 1989. Transcription of infectious yellow fever RNA from full-length cDNA templates produced by in vitro ligation. New Biol. **1:**285–296.
20. **Skotnicki, M. L., S. W. Ding, A. M. MacKenzie, and A. J. Gibbs.** 1993. Infectious eggplant mosaic tymovirus and ononis yellow mosaic tymovirus from cloned cDNA. Arch. Virol. **131:**47–60.
21. **Stern, D. F., and B. M. Sefton.** 1984. Coronavirus multiplication: location of genes for virion proteins on the avian infectious bronchitis virus genome. J. Virol. **50:**22–29.
22. **Sumiyoshi, H., C. H. Hoke, and D. W. Trent.** 1992. Infectious Japanese encephalitis virus RNA can be synthesized from in vitro-ligated cDNA templates. J. Virol. **66:**5425–5431.
23. **Tibbles, K. W., I. Brierley, D. Cavanagh, and T. D. K. Brown.** 1996. Characterization in vitro of an autocatalytic processing activity associated with the predicted 3C-like proteinase domain of the coronavirus avian infectious bronchitis virus. J. Virol. **70:**1923–1930.