Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

VIROLOGY 245, 303–312 (1998)
ARTICLE NO. VY989164

Characterization of the Two Overlapping Papain-like Proteinase Domains Encoded in Gene 1 of the Coronavirus Infectious Bronchitis Virus and Determination of the C-Terminal Cleavage Site of an 87-kDa Protein

K. P. Lim and D. X. Liu[1]

Institute of Molecular Agrobiology, The National University of Singapore, 1 Research Link, Singapore 117604

Received December 10, 1997; returned to author for revision January 16, 1998; accepted March 26, 1998

In a previous report, we showed that proteolytic processing of an 87-kDa mature viral protein from the coronavirus infectious bronchitis virus (IBV) 1a and 1a/1b polyproteins was mediated by two putative overlapping papain-like proteinase domains (PLPDs) encoded within the region from nucleotides 4243 to 5553 of ORF 1a (Liu *et al.,* 1995). In this study, we demonstrate that only the first domain, PLPD-1, is responsible for this cleavage, as deletion of the second domain did not affect the formation of the 87-kDa protein. Site-directed mutagenesis studies further showed that a previously predicted nucleophilic cysteine residue (Cys^{1274}) and a histidine residue (His^{1437}) were essential for the proteinase activity, indicating that they may be important components of the catalytic center of the proteinase. Meanwhile, expression of a series of deletion mutants revealed that the 87-kDa protein was encoded by the 5′-most 2.6 kb of ORF 1a. Deletion and amino acid substitution mutation studies demonstrated that the Gly^{673}–Gly^{674} dipeptide bond was most likely the cleavage site responsible for releasing the C-terminus of the 87-kDa protein from the 1a and 1a/1b polyproteins.

INTRODUCTION

Avian infectious bronchitis virus (IBV) is the prototype of Coronaviridae, a family of enveloped viruses which possess a large continuous positive-stranded RNA genome. The genomic RNA is 27.6 kb in length and contains at least 10 distinct open reading frames (ORFs) (Boursnell *et al.,* 1987). Most of these ORFs do not appear to be translated from the genomic RNA but are expressed from a set of subgenomic mRNAs. Six mRNA species, including a genome-length mRNA (mRNA 1) and five subgenomic mRNA species (mRNA 2-6), are produced in virus-infected cells. These mRNAs range in length from about 2 to 27.6 kb and have been shown to form a 3′-coterminal "nested" structure (Stern and Sefton, 1984).

Approximately 70% of the coding potential of the IBV genome is contained within the 5′-terminal unique region of mRNA 1. Nucleotide sequencing has shown that mRNA 1 contains two large overlapping ORFs (1a and 1b), with ORF 1a having the potential to encode a polyprotein of 441 kDa and 1b a polyprotein of 300 kDa (Boursnell *et al.,* 1987). The downstream ORF 1b is produced as a fusion protein of 741 kDa with 1a by a ribosomal frameshifting mechanism (Brierley *et al.,* 1987, 1989). These polyproteins are expected to be cleaved by viral or cellular proteinases to produce functional products associated with viral replication.

Recently, products encoded by these two ORFs have been identified. These include 87-, 10-, and 100-kDa polypeptides (Liu *et al.,* 1994, 1995, 1997; Liu and Brown, 1995). Three ORF 1a-encoding proteinase domains, i.e., two overlapping papain-like proteinase domains (PLPDs) encoded by the IBV sequence between nucleotides 4243 and 5553 and a picornavirus 3C-like proteinase domain encoded between nucleotides 8937 and 9357, were demonstrated to be involved in the production of these mature viral proteins (Liu *et al.,* 1994, 1995, 1997; Liu and Brown, 1995). Proteolytic processing of the 1a and 1a/1b polyproteins by the equivalent proteinase activities of human coronavirus and mouse hepatitis virus has also been documented. So far, four mature products have been identified in cells infected by human coronavirus strain 229E. These include a 34-kDa protein representing the 3C-like proteinase, a 105-kDa protein corresponding to the 100-kDa protein of IBV, a 71-kDa protein with an ATPase activity, and a 41-kDa protein processed from the 1b polyprotein (Ziebuhr *et al.,* 1995; Grotzinger *et al.,* 1996; Heusipp *et al.,* 1997a, b). Similarly, three mature products, including two N-terminal cleavage products of 28 and 65 kDa and a 27-kDa protein identified as the 3C-like proteinase, have been reported for mouse hepatitis virus (Denison and Perlman, 1987; Denison *et al.,* 1995; Lu *et al.,* 1995).

In this communication, we report that of the two predicted PLPDs of IBV, only the first one, PLPD-1, encoded

[1] To whom reprint requests should be addressed. E-mail: liudx@ima.org.sg.

0042-6822/98 $25.00

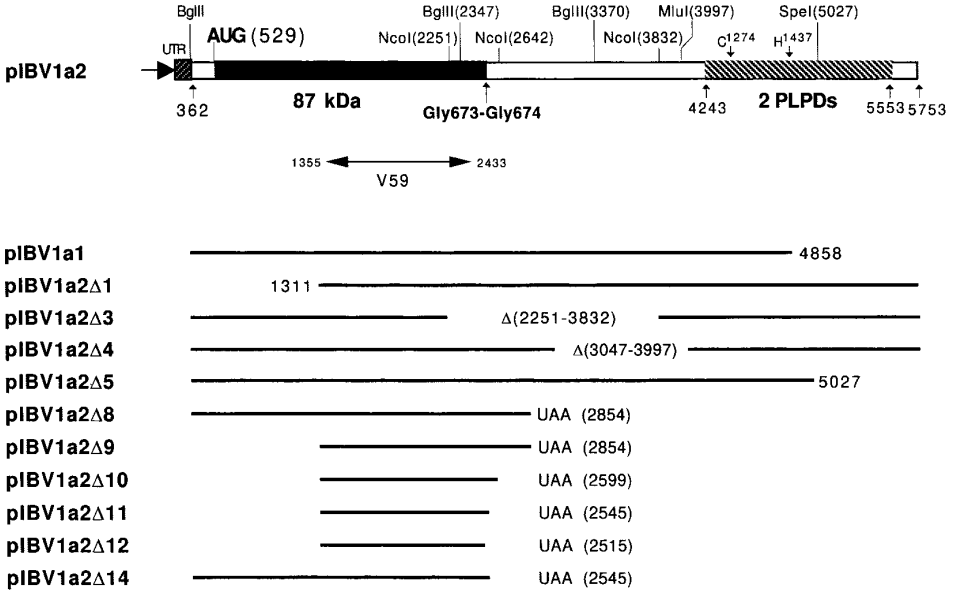

FIG. 1. Diagram of the structure of plasmid pIBV1a2, showing the locations of the 87-kDa protein, the two putative overlapping papain-like proteinase domains (PLPDs), and the restriction sites used in this study. The C-terminal cleavage site ($Gly^{673} \downarrow Gly^{674}$ dipeptide) of the 87-kDa protein, the catalytic residues of the PLPD-1 (Cys^{1274} and His^{1437}), and the IBV sequence recognized by antiserum V59 are indicated. Also shown are the IBV sequences present in plasmids pIBV1a1, pIBV1a2Δ1, pIBV1a2Δ3, pIBV1a2Δ4, pIBV1a2Δ5, pIBV1a2Δ8, pIBV1a2Δ9, pIBV1a2Δ10, pIBV1a2Δ11, pIBV1a2Δ12, and pIBV1a2Δ14.

between nucleotides 4243 and 5019, is required for cleavage of the 87-kDa protein at its C-terminus. Site-directed mutagenesis studies confirmed that the previously predicted Cys^{1274} and His^{1437} residues of PLPD-1 are essential for the proteinase activity, suggesting that they may be components of the catalytic center of the proteinase. Meanwhile, expression of a series of C-terminal deletion mutants indicated that the Gly^{673}–Gly^{674} dipeptide bond encoded by nucleotides 2545 to 2550 may be responsible for releasing the C-terminus of the 87-kDa protein. Site-directed mutagenesis and internal deletion studies further demonstrated that this dipeptide bond was indeed the C-terminal cleavage site of the 87-kDa protein. As no further N-terminal cleavage of the 1a polyprotein was observed, it is therefore established that the 87-kDa protein is encoded by the ORF 1a sequence between nucleotides 529 and 2547.

RESULTS

Further analysis of processing of the 1a and 1a/1b polyproteins to an 87-kDa protein by the two overlapping papain-like proteinase domains

We have previously demonstrated that expression of the IBV sequence from nucleotides 362 to 5753 (pIBV1a2) led to the detection of three protein species: a 250-kDa product representing the full-length product of the construct and two cleavage products of 160 and 87 kDa, respectively (Liu *et al.,* 1995). However, when the IBV sequence from nucleotides 362 to 4858 (pIBV1a1) was expressed, no processing of the full-length 220-kDa protein to smaller products was observed (Liu *et al.,* 1995). These results suggest strongly that the two putative overlapping papain-like proteinase domains are involved in processing of the 1a and 1a/1b polyproteins to the 87-kDa protein species. To further characterize the proteinase activity and to determine the cleavage sites of the 87-kDa protein, a series of N-terminal, internal, and C-terminal deletion mutants were constructed based on pIBV1a2. Plasmid pIBV1a2Δ1 contains an N-terminal deletion of the IBV sequence from nucleotides 362 to 1310; pIBV1a2Δ3 and pIBV1a2Δ4 contain internal deletions of the IBV sequences from nucleotides 2251 to 3832 and from 3047 to 3997, respectively; pIBV1a2Δ5 contains a C-terminal deletion of the IBV sequence from nucleotides 5027 to 5753 (see Figs. 1 and 2a). These constructs were then expressed in Vero cells using the vaccinia virus-T7 expression system (Fuerst *et al.,* 1986).

The N-terminal deletion construct pIBV1a2Δ1 was firstly expressed in Vero cells. As can be seen, expression of pIBV1a2 resulted in the detection, once again, of the 250-, 160-, and 87-kDa protein species (Fig. 2b). Similarly, three protein species of approximately 220, 160, and 58 kDa were detected from lysates prepared from pIBV1a2Δ1-transfected cells by immunoprecipitation with region-specific antiserum V59 (Fig. 2b). The 220-kDa species is probably the full-length product, while the 160- and 58-kDa protein species may represent products derived from a cleavage event. As the 160-kDa protein comigrated with the 160-kDa protein

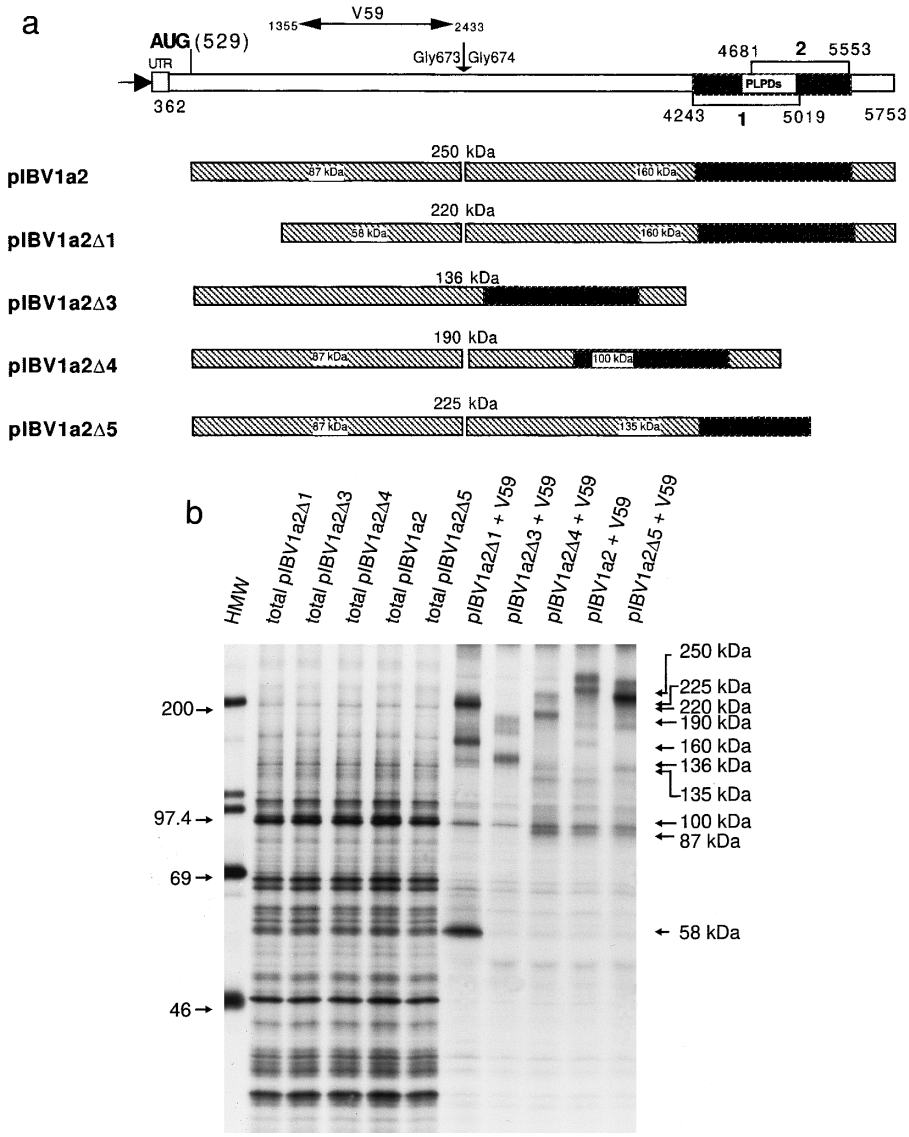

FIG. 2. (a) Diagram showing the full-length and cleavage products synthesized from plasmids pIBV1a2, pIBV1a2Δ1, pIBV1a2Δ3, pIBV1a2Δ4, and pIBV1a2Δ5. (b) Analysis of transiently expressed ORF 1a products from cells transfected with pIBV1a2, pIBV1a2Δ1, pIBV1a2Δ3, pIBV1a2Δ4, and pIBV1a2Δ5. Plasmid DNA was transiently expressed in Vero cells using the vaccinia virus/T7 expression system. Cells were labeled with [^{35}S]methionine, lysates were prepared, and polypeptides were either analyzed directly or immunoprecipitated with antiserum V59. Gel electrophoresis was performed on an SDS–10% polyacrylamide gel and polypeptides were detected by fluorography. HMW, high-molecular-mass markers (numbers indicate kilodaltons).

detected from pIBV1a2-transfected cells (Fig. 2b), this result suggests that it may represent the C-terminal cleavage product of the 220-kDa protein, and the 58-kDa protein may be an N-terminally truncated version of the 87-kDa protein.

In a previous study, we speculated that the 160-kDa species may be an intermediate cleavage product containing the 87-kDa protein, as both the 87- and 160-kDa proteins can be immunoprecipitated by the two N-terminal-specific antisera (V52 and V59) used (Liu *et al.,* 1995). It was therefore likely that a second cleavage site may be located between the regions coding for the 87-kDa protein and the two PLPDs. To investigate this possibility, plasmid pIBV1a2Δ3 was constructed and expressed. As can be seen, only the full-length protein of approximately 136 kDa was immunoprecipitated (Fig. 2b). No cleavage products were detected from the expression of this construct. This result hence rules out the possibility that the 160-kDa is an intermediate cleavage product covering the 87-kDa protein and a protein of 70 kDa (Fig. 2b). The reason for the coimmunoprecipitation of the 160-kDa by antiserum V59 is considered under Discussion.

Plasmid pIBV1a2Δ4 was next expressed to test the effect of internal deletion of the region between the sequences coding for the 87-kDa protein and PLPDs on processing to the 87-kDa protein. As shown in Fig. 2b,

expression of pIBV1a2Δ4 led to the detection of three products of approximately 190, 100, and 87 kDa. Both the 190- and 87-kDa protein species were efficiently immunoprecipitated by antiserum V59; only weak immunoprecipitation of the 100-kDa protein, however, was observed (Fig. 2b). This result demonstrates that the IBV sequence between nucleotides 3047 and 3997 is not essential for processing to the 87-kDa protein species and suggests that the C-terminal cleavage site of the 87-kDa protein is encoded between nucleotides 2251 and 3047.

Finally, the C-terminal cleavage construct pIBV1a2Δ5 was expressed to investigate if both of the two putative PLPDs were required for processing to the 87-kDa protein. As can be seen, expression of this construct led to the detection of three protein species of approximately 225, 135, and 87 kDa, respectively (Fig. 2b). Both the 225- and 87-kDa proteins were efficiently immunoprecipitated by antiserum V59 (Fig. 2b). This result indicates that PLPD-1 is responsible for releasing the 87-kDa protein.

In addition to the protein species described above, expression of pIBV1a2, pIBV1a2Δ3, pIBV1a2Δ4, and pIBV1a2Δ5 also resulted in the detection of protein species which migrated more slowly on SDS–PAGE than their respective full-length products (Fig. 2b). Furthermore, protein species that migrated between the cleavage products were also observed (Fig. 2b). The identities of these products are currently uncertain, but they may represent modified forms of the cleavage and full-length products.

The polypeptides detected from the expression of pIBV1a2, pIBV1a2Δ1, pIBV1a2Δ3, pIBV1a2Δ4, and pIBV1a2Δ5 are illustrated in Fig. 2a.

Mutagenesis studies of the putative catalytic center of PLPD-1

Data presented above showed that only the first of the two overlapping proteinase domains was required for processing to the 87-kDa product. Computer analysis of the proteinase domain suggested that it may belong to the papain-like class of thiol proteinase with a catalytic dyad of cysteine and histidine residues (Gorbalenya *et al.,* 1989, 1991). In fact, the Cys^{1274} and His^{1437} residues were predicted to be the catalytic dyad of the proteinase (Lee *et al.,* 1991) (see Fig. 1). To test this prediction, site-directed mutagenesis was carried out using pIBV1a2Δ4 as the template. By monitoring the production of the 87-kDa protein, the effect of the mutations on the proteinase activity could be tested.

The putative nucleophilic cysteine (Cys^{1274}) residue, which was encoded by nucleotides 4348 to 4350, was first mutated to a Ser residue, giving rise to the mutant construct pIBV1a2Δ4C^{1274}-S. As shown in Fig. 3, expression of pIBV1a2Δ4, once again, yielded the 190-kDa full-length product and the 87-kDa protein species. Expression of pIBV1a2Δ4C^{1274}-S, however, resulted in the

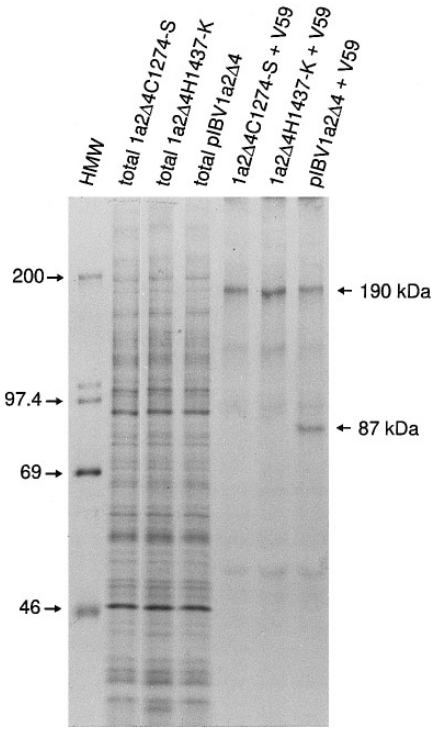

FIG. 3. Mutational analysis of the two putative catalytic residues (Cys^{1274} and His^{1437}) of PLPD-1. Vero cells were transfected with plasmids pIBV1a2Δ4, pIBV1a2Δ4C^{1274}-S, and pIBV1a2Δ4H^{1437}-K and labeled with [^{35}S]methionine and lysates were prepared. Polypeptides were either analyzed directly or immunoprecipitated with antiserum V59. Gel electrophoresis was performed on an SDS–10% polyacrylamide gel and polypeptides were detected by fluorography. HMW, high-molecular-mass markers (numbers indicate kilodaltons).

detection of only the 190-kDa full-length product. No 87-kDa protein was detected, indicating that this substitution disrupts the proteinase activity required for the release of the 87-kDa protein. Similarly, mutation of the predicted His^{1437} residue (encoded by nucleotides 4837 to 4839) to a Lys was introduced, giving rise to pIBV1a2Δ4H^{1437}-K. The mutant construct was then expressed in Vero cells. As can be seen, only the 190-kDa full-length product was immunoprecipitated by antiserum V59; no 87-kDa protein was detected (Fig. 3). These results support the computer prediction that both the Cys^{1274} and His^{1437} are essential for the catalytic activity of the proteinase.

Deletion analysis of the C-terminal cleavage site of the 87-kDa protein

Data obtained from the expression of pIBV1a2Δ3 and pIBV1a2Δ4 indicate that the C-terminal cleavage site of the 87-kDa protein is encoded between nucleotides 2251 and 3047. To define more precisely the cleavage site, a series of C-terminal deletion mutants were constructed by introducing a termination codon (UAA) into different positions of the IBV sequence (see Fig. 1). By comparing the migration of products expressed from the mutant

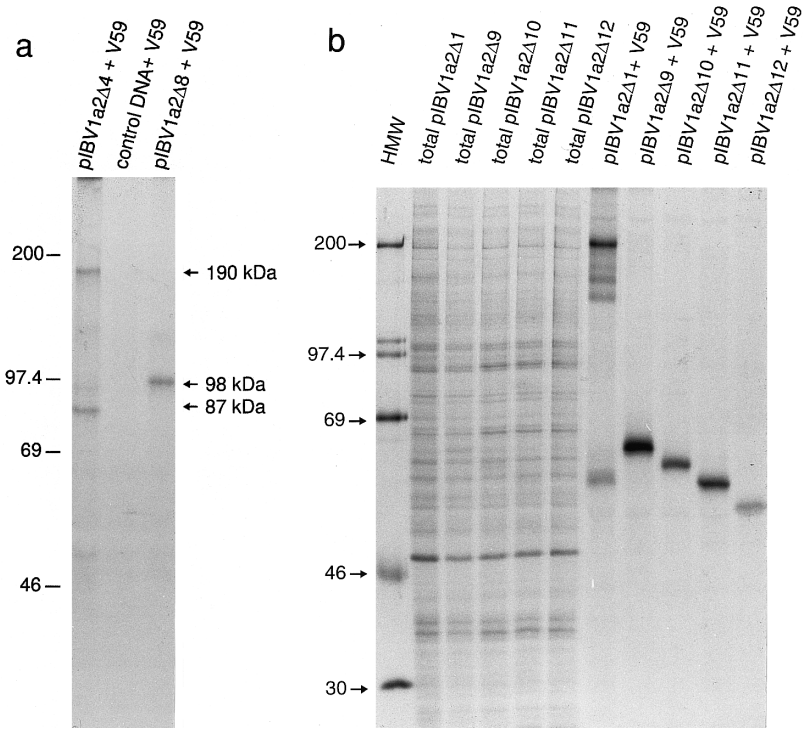

FIG. 4. Deletion analysis of the C-terminal cleavage site of the 87-kDa protein. Plasmids pIBV1a2Δ4 and pIBV1a2Δ8 (a) and plasmids pIBV1a2Δ1, pIBV1a2Δ9, pIBV1a2Δ10, pIBV1a2Δ11, and pIBV1a2Δ12 (b) were transiently expressed in Vero cells using the vaccinia virus/T7 expression system. Cells were labeled with [^{35}S]methionine and lysates were prepared. Polypeptides were either analyzed directly or immunoprecipitated with antiserum V59. Gel electrophoresis was performed on an SDS–10% polyacrylamide gel and polypeptides were detected by fluorography. Control DNA indicates that a plasmid not containing the IBV sequence was used. HMW, high-molecular-mass markers (numbers indicate kilodaltons).

constructs with the 87-kDa protein processed from the wild-type plasmids, it would allow us to locate more precisely the C-terminal cleavage site of the 87-kDa protein.

The first of such mutants, pIBV1a2Δ8, was made by introducing a UAA codon at the nucleotide position 2854. Expression of this plasmid led to the synthesis of a product migrating on SDS–PAGE more slowly than the 87-kDa protein expressed from pIBV1a2Δ4 (Fig. 4a). The apparent molecular mass of this protein is approximately 98 kDa, indicating that the actual cleavage site is located at least 250 bp upstream of this position. However, this interpretation was based on the assumption that the 87-kDa protein was encoded by the 5′-most part of ORF 1a and no N-terminal cleavage of the 1a polyprotein occurred before the 87-kDa protein.

To avoid this ambiguity, four deletion mutants, pIBV1a2Δ9, pIBV1a2Δ10, pIBV1a2Δ11, and pIBV1a2Δ12, were constructed based on pIBV1a2Δ1 by inserting the UAA triplet at nucleotide positions 2854, 2599, 2545, and 2515, respectively (see Fig. 1). Expression of pIBV1a2Δ9 and pIBV1a2Δ10 led to the detection of two protein species migrating more slowly on SDS–PAGE than the 58-kDa protein expressed from pIBV1a2Δ1. Their apparent molecular masses are approximately 65 and 62 kDa, respectively (Fig. 4b). Expression of pIBV1a2Δ11 led to the detection of a protein species which almost comigrates with the 58-kDa expressed from pIBV1a2Δ1 on SDS–PAGE (Fig. 4b). Expression of pIBV1a2Δ12, however, resulted in the detection of a product with an apparent molecular mass of approximately 56 kDa (Fig. 4b). These results indicate that the C-terminal cleavage site of the 87-kDa protein is in the vicinity of amino acid residue Gly^{673} encoded between nucleotides 2545 and 2547.

Further determination of the C-terminal cleavage site of the 87-kDa protein by site-directed mutagenesis and deletion analysis

From the observation that the C-terminus of the 87-kDa protein is probably located near nucleotide position 2550, it is tempting to speculate that this protein may be cleaved from the precursor at a potential Gly^{673}–Gly^{674} or Thr^{676}–Val^{677} dipeptide bond. By comparative analysis of the cleavage sites used by other viral papain-like proteinases, it seems likely that either of the two sites could be the scissile bond of PLPD-1 (Bonilla *et al.,* 1995, 1997; Dong and Baker, 1994; Hughes *et al.,* 1995). To explore these possibilities, substitution mutations of the Gly^{673} residue with an Ala and the Thr^{676} residue with a Ser were made, giving rise to the mutant constructs pIBV1a2Δ4G^{673}-A and pIBV1a2Δ4T^{676}-S. As can be seen, expression of pIBV1a2Δ4G^{673}-A led to the detection of the full-length 190-kDa protein; no processing to the

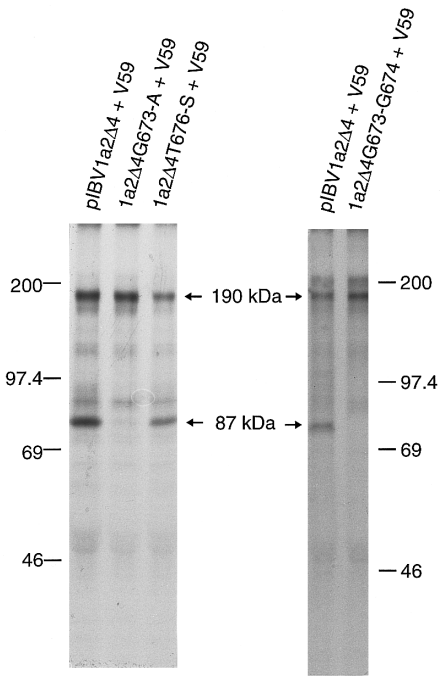

FIG. 5. Mutational and deletion analyses of the C-terminal cleavage site of the 87-kDa protein. Plasmids pIBV1a2Δ4, pIBV1a2Δ4G^{673}-A, pIBV1a2Δ4T^{676}-S, and pIBV1a2Δ4G^{673}-G^{674} were transiently expressed in Vero cells using the vaccinia virus/T7 expression system. Cells were labeled with [^{35}S]methionine and lysates were prepared. Polypeptides were immunoprecipitated with antiserum V59. Gel electrophoresis was performed on an SDS–10% polyacrylamide gel and polypeptides were detected by fluorography. Numbers indicate molecular masses in kilodaltons.

87-kDa protein was observed (Fig. 5). However, both the 190- and 87-kDa proteins were observed when pIBV1a2Δ4T^{676}-S was expressed (Fig. 5). Furthermore, the relative amount of the 87-kDa protein to the 190-kDa protein expressed from this mutant construct was very similar to that expressed from the wild-type construct pIBV1a2Δ4 (Fig. 5), indicating that the production of the 87-kDa protein was not affected by mutation of the Thr^{676} to a Ser residue. These results suggest strongly that the Gly^{673}–Gly^{674} dipeptide bond is the C-terminal cleavage site of the 87-kDa protein.

To support this conclusion further, the six nucleotides coding for the Gly^{673}–Gly^{674} scissile bond were deleted by PCR, giving rise to plasmid pIBV1a2Δ4ΔG^{673}–G^{674}. As can be seen, expression of pIBV1a2Δ4ΔG^{673}–G^{674} resulted in the detection of the 190-kDa full-length product only. No 87-kDa protein was detected (Fig. 5), confirming that deletion of the Gly^{673}–Gly^{674} dipeptide bond totally blocks the release of the 87-kDa protein from its precursor.

No N-terminal cleavage of the 1a polyprotein before the 87-kDa protein

To rule out the possibility that further N-terminal cleavage is required for releasing the 87-kDa protein from the 1a and 1a/1b polyprotein precursors, plasmid pIBV1a2Δ14 was made by extending the 5′ end of pIBV1a2Δ11 to cover the IBV sequence up to nucleotide 362 and was expressed in Vero cells. The mobility of the full-length product expressed from this plasmid was then compared with the 87-kDa protein processed from pIBV1a2Δ4. As expected, expression of pIBV1a2Δ14 resulted in the synthesis of a protein which comigrates on SDS–PAGE with the 87-kDa protein expressed from pIBV1a2Δ4 (data not shown). This result indicates that no N-terminal cleavage of the 1a polyprotein occurs before the 87-kDa protein.

Throughout this and our previous report (Liu *et al.,* 1995), this protein was referred to as 87 kDa based purely on its migration on linear SDS–PAGE. In an early communication, Brierley *et al.,* (1990) noted that Western blots of IBV-infected chicken kidney cells and Vero cells with antiserum V59 led to the detection of a virus-specific polypeptide of 75 kDa. This 75-kDa protein was subsequently shown to migrate on linear SDS–PAGE at the same position as the 87-kDa product (Liu *et al.,* 1995). It is likely that the gradient SDS–PAGE system used in the original report may reflect more accurately the real molecular mass of the protein. To test this possibility, lysates prepared from pIBV1a2Δ4- and pIBV1a2Δ14-transfected cells were immunoprecipitated with antiserum V59 and separated on an SDS–7.5 to 15% polyacrylamide gradient gel. Once again, comigration of the 87-kDa protein expressed from both plasmids was observed (Fig. 6). It was also noted that the 87-kDa protein and the full-length product migrated slightly more rapidly in this gel system; the estimated molecular masses are approximately 78 and 180 kDa, respectively. Obviously, the migration of the 87-kDa protein in this gel system is closer to its calculated molecular mass of 75 kDa.

DISCUSSION

Co- and posttranslational processing of polyproteins encoded by the 27.6-kb genome-length mRNA, mRNA 1, of the coronavirus IBV is an essential step in the viral replication cycle. Three viral proteinase domains, i.e., two overlapping PLPDs and a picornavirus 3C-like proteinase domain, have been predicted to be involved in these processing events (Gorbalenya *et al.,* 1989, 1991; Lee *et al.,* 1991). PLPD-1, the so-called "main" proteinase, is encoded by ORF 1a from nucleotides 4243 to 5019 and located downstream of the "X" domain (Lee *et al.,* 1991; Gorbalenya *et al.,* 1991). The second PLPD, which shares similarity with the streptococcus proteinase, is encoded by nucleotides 4681 to 5553 (Gorbalenya *et al.,* 1989, 1991). In a previous communication, we reported that the region coding for the two putative PLPDs was required for processing of the 1a and 1a/1b polyproteins to an 87-kDa mature viral product (Liu *et al.,* 1995). Evidence presented in this report demonstrates that only PLPD-1

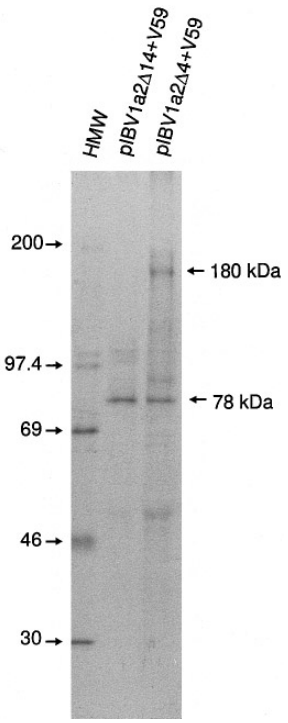

FIG. 6. Analysis of the migration in gradient SDS–PAGE of the 87-kDa protein transiently expressed from pIBV1a2Δ4 and pIBV1a2Δ14. Vero cells were transfected with plasmid DNAs and labeled with [^{35}S]methionine and lysates were prepared. Polypeptides were immunoprecipitated with antiserum V59. Gel electrophoresis was performed on an SDS–7.5–15% polyacrylamide gel and polypeptides were detected by fluorography. HMW, high-molecular-mass markers (numbers indicate kilodaltons).

is responsible for cleavage of the 87-kDa protein from its polyprotein precursors. This conclusion is supported by two observations. Firstly, production of similar amount of the 87-kDa protein was observed from expression of constructs with or without the second proteinase domain. Second, mutation of the two predicted catalytic residues of PLPD-1, the Cys^{1274} and His^{1437}, totally abolished the proteinase activity required for processing to the 87-kDa protein. As the second PLPD is still present in these mutants, this result reinforces the conclusion that only PLPD-1 is required for cleavage of the 1a and 1a/1b polyproteins to the 87-kDa protein species.

The observation that the PLPD-1 activity is inactivated by substitution mutations of the Cys^{1274} and His^{1437} confirms that they are most likely the catalytic residues of the proteinase. However, when the IBV sequence up to nucleotide 4858 (and therefore includes these two core residues) was expressed, no proteinase activity was detected (Liu *et al.,* 1995), indicating that besides the catalytic center of the enzyme, the region from nucleotides 4859 to 5027 is also indispensable for the activity of the proteinase. The significance of this region in the formation of the catalytic center, substrate binding, and overall folding of the proteinase remains to be understood. As for the second PLPD, no functional role has been reported so far. This lack of activity hampers its further characterization by mutagenesis study. In fact, no proteinase activity has been documented for the second PLPD of human coronavirus and mouse hepatitis virus strains JHM and A59 (Baker *et al.,* 1993; Bonilla *et al.,* 1995, 1997; Dong and Baker, 1994; Hughes *et al.,* 1995; Herold *et al.,* 1998). It is likely that the second PLPD of coronavirus may be functionally inactive.

Deletion and mutagenesis studies demonstrated convincingly that cleavage of the C-terminus of the 87-kDa protein may occur between the Gly^{673}–Gly^{674} dipeptide bond. Examination of cleavage sites recognized by other viral papain-like proteinases suggested that they may consist of two small amino acids with short, uncharged side chains; in most cases, the P1 site is a Gly residue and the P1′ a Gly, an Ala, or a Val (Bonilla *et al.,* 1997; Dong and Baker, 1994). According to these criteria, the Thr^{676}–Val^{677} dipeptide bond encoded by nucleotides 2554 to 2559 may also be a cleavage site of PLPD-1. However, as shown in Fig. 5, mutation of the Thr^{676} residue to a Ser did not affect processing to the 87-kDa protein. Another potential cleavage position is the Gly^{680}–Glu^{681} dipeptide bond encoded by nucleotides 2566 to 2571. This possibility is ruled out by deletion of the Gly^{673}–Gly^{674} dipeptide bond, which totally abolished the release of the 87-kDa protein (Fig. 5). As these two residues are located six and seven amino acids, respectively, upstream of the Gly^{680}, this deletion should render very little, if any, effect on the cleavability of the Gly^{680}–Glu^{681} dipeptide bond if it were a genuine cleavage site. These results virtually exclude the possibilities that either the Thr^{676}–Val^{677} or Gly^{680}–Glu^{681} dipeptide bond is the C-terminal cleavage site of the 87-kDa protein. Meanwhile, deletion studies also suggested that no N-terminal cleavage was required for releasing the 87-kDa protein from the 1a and 1a/1b polyproteins, indicating that it is encoded by ORF 1a from the authentic AUG initiation codon at nucleotide positions 529 to 531. Therefore, the 87-kDa protein is released from its polyprotein precursors via a single cleavage mediated by PLPD-1 at the Gly^{673}–Gly^{674} scissile bond; it is composed of 673 amino acid residues with a calculated molecular mass of 75 kDa. The reason for the increased apparent molecular masses of this protein (as well as the full-length product encoded by the first 5753 bp of ORF 1a) on SDS–PAGE is uncertain, but it may reflect the amino acid composition of the products encoded by this region. Examination of the deduced amino acid sequence shows that the region between nucleotides 2300 and 3800 encodes a product with unusually high content of negatively charged amino acid residues (21.8%). This may influence the mobility of the products on SDS–PAGE. To support this view, it was observed that deletion of this region resulted in synthesis of the full-length protein with an electrophoretic mobility on SDS–PAGE consistent with its calculated molecular mass (Fig. 2, lane pIBV1a2Δ3). As no functional

domain has been allocated to this region of the genome yet, the significance of such a high proportion of negatively charged amino acid composition remains unclear.

In summary, data presented here and in our previous communication demonstrate that IBV PLPD-1 resembles the first PLPD of human coronavirus 229E and mouse hepatitis virus strains JHM and A59 in terms of the composition of the catalytic residues and the cleavage specificity (Baker *et al.,* 1993; Bonilla *et al.,* 1995, 1997; Dong and Baker, 1994; Hughes *et al.,* 1995; Herold *et al.,* 1998). However, three major differences were also observed. First, IBV PLPD-1 exhibits no proteinase activity in cell-free expression systems. Extensive investigation showed that expression of ORF 1a up to nucleotide 5753 in rabbit reticulocyte lysate (Liu *et al.,* 1995) led to the detection of only the full-length product and several products derived from premature termination of translation. No cleavage of the full-length product to the 87-kDa protein was observed. Interestingly, no proteinase activity could be detected when the IBV 3C-like proteinase was expressed in cell-free expression systems (Liu *et al.,* 1994). Second, although the general recognition pattern of the three coronavirus PLPD described so far is similar, IBV PLPD-1 shows unique features. The most notable difference of the IBV PLPD-1 recognition site is in the P5 position. Instead of a basic amino acid, a Val residue was located at this position. This indicates that IBV PLPD-1 may be more distantly related to other coronavirus PLPD. In fact, the nucleotide sequences around the IBV PLPD-1 recognition site are not compatible with either alignment model proposed to explain the evolutionary relationship among proteins of human coronavirus 229E, mouse hepatitis virus, and transmissible gastroenteritis coronavirus (Herold *et al.,* 1998). Finally, no *trans*-cleavage activity has been detected so far for IBV PLPD-1 either in cell-free translation systems or in intact cells (unpublished observations). Once again, this is markedly different from PLPD of both human coronavirus and mouse hepatitis virus. It is therefore likely that IBV PLPD-1 may be a unique member among the coronavirus papain-like proteinase domains.

In a previous communication, we reported that a cellular proteinase activity presented in Vero cell lysates could mediate cleavage of *in vitro* synthesized ORF 1a products at a position closed to the C-terminal cleavage site of the 87-kDa protein (Liu *et al.,* 1995). Subsequently, we observed that the same cell lysate was able to mediate cleavage of the 1a and 1a/1b polyproteins at other positions (unpublished observations). However, none of the cleavage products released from these positions corresponds to genuine mature cleavage products, if compared with products generated from virus-infected or transfected cells. Cleavage at these positions does not therefore reflect real cleavage events that occurred in virus life cycles.

Deletion studies presented in this report clearly showed that cleavage at the Gly^{673}–Gly^{674} dipeptide bond may be the only proteolytic event occurring in the polyprotein encoded by the first 5753 bp of ORF 1a. This cleavage led to the production of the N-terminal 87-kDa and the C-terminal 160-kDa cleavage products. It was intriguing to note that the 160-kDa protein can also be efficiently immunoprecipitated by the N-terminal specific antiserum V59 (see Fig. 2). We have also observed that the 87-kDa protein can be coimmunoprecipitated by a region-specific antiserum recognizing the IBV sequence encoded between nucleotides 4868 and 5576 (unpublished observations). The coimmunoprecipitation of the 160-kDa protein by antiserum V59 is likely due to the interaction between the N- and C-terminal cleavage products. Further investigation is underway to characterize this interaction.

MATERIALS AND METHODS

Transient expression of IBV sequences in Vero cells using a vaccinia virus-T7 expression system

IBV sequences placed under the control of a T7 promoter were transiently expressed in eukaryotic cells using the system described by Fuerst *et al.,* (1986). Briefly, 60–80% confluent monolayers of Vero cells grown on 35-mm dishes were infected with 10 PFU/cell of a recombinant vaccinia virus (vTF7-3) which expresses bacteriophage T7 RNA polymerase. The cells were then transfected with 2.5 μg plasmid DNA (prepared with Qiagen plasmid Midi kits) by using DOSPER (or DOTAP) liposomal transfection reagent according to the instructions of the manufacturer (Boehringer Mannheim). As determined by monitoring the expression of a reporter plasmid coding for the green fluorescent protein from the jellyfish *Aequorea victoria,* 25–35% of the transfection efficiency was usually achieved by this procedure.

After incubation at 37°C for 5 h, the cells were washed twice with methionine-free medium and were labeled with 25 μCi/ml [^{35}S]methionine. The radiolabeled cells were harvested at 18 h postinfection.

Polymerase chain reaction (PCR)

Amplification reactions of the respective template DNAs with appropriate primers were performed with Pfu DNA polymerase (Stratagene) under standard buffer conditions with 2 mM $MgCl_2$. PCR reaction conditions were 30 cycles of 95°C for 45 s, *X*°C for 45 s, and 72°C for *X* min. The annealing temperature (*X*°C) and the extension time (*X* min) were subjected to adjustments according to the melting temperature of the primers employed and the length of the PCR fragments synthesized.

Site-directed mutagenesis

Site-directed mutagenesis was carried out, as previously described, by two rounds of PCR with two pairs of primers (Liu *et al.,* 1997).

Radioimmunoprecipitation

Plasmid DNA-transfected Vero cells were lysed with radioimmunoprecipitation assay buffer [50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 1% sodium deoxycholate, 1% NP-40, 0.1% sodium dodecyl sulfate (SDS)] and precleared by centrifugation at 12,000 rpm for 5 min at 4°C in a microfuge. Immunoprecipitation was performed as described previously (Liu *et al.,* 1991).

SDS–polyacrylamide gel electrophoresis

SDS–polyacrylamide gel electrophoresis (SDS–PAGE) of viral polypeptides was performed with 10% polyacrylamide gels (Laemmli, 1970). The ^{35}S-labeled polypeptides were detected by autoradiography of the dried gels.

TABLE 1

Oligonucleotides Used for PCR Amplification and Site-Directed Mutagenesis

Name	Nucleotide sequence	Position
LDX-3	5′-ACGCGGATCCACCATGGGTTCTAAG-3′	1304–1323
LDX-4	5′-TCTGTTTGCAAGTTACATCG-3′	4060–4041
LKP-1	5′-TCTCAGGGATCCCCCCACATACC-3′	370–383
LKP-2	5′-ACGCTAACGCGTTCATCAAGAGGCAG-3′	3047–3022
LKP-3	5′-CTCTCATAGACGCGTTAGATCAAATGGC-3′	2870–2843
LKP-4	5′-ATAGGCCCGGGTTAAGGTGGTGGTATCT-3′	2612–2584
LKP-5	5′-CAGTCCCGGGTTATGCTTTGCAAACCAC-3′	2557–2530
LKP-6	5′-CATTAATCCCGGGTTATTGAGACATTGGTG-3′	2530–2501
LKP-7	5′-CGATGGAAATAGCTGGATTAGTTCAGC-3′	4338–4364
LKP-8	5′-GAACTAATCCAGCTATTTCCATCGCG-3′	4361–4336
LKP-9	5′-CTAATAGTGGCAAGTGTTATACACAAGC-3′	4826–4853
LKP-10	5′-GTATAACACTTGCCACTATTAGTAGAAC-3′	4847–4820
LKP-11	5′-GTTTGCAAAGCAGCTGGTAAGACTGT-3′	2533–2558
LKP-12	5′-GTGACAGTCTTACCAGCTGCTTTGCA-3′	2561–2537
LKP-13	5′-GGTAAGAGCGTCACCTTTGGAGAAA-3′	2548–2572
LKP-14	5′-GGTGACGCTCTTACCGCCTGCTTTG-3′	2562–2538
LKP-15	5′-TTGCAAAGCAAAGACTGTCACCTTTG-3′	2535–2560
LKP-16	5′-AGGTGACAGTCTTTGCTTTGCAAACC-3′	2532–2557

Construction of plasmids

Plasmids pIBV1a1 and pIBV1a2 (formerly called pKT1a1 and pKT1a2), which cover the IBV sequences from nucleotides 362 to 4858 and from 362 to 5753, respectively (Liu *et al.,* 1995), were used to construct pIBV1a2Δ1, pIBV1a2Δ3, pIBV1a2Δ4, pIBV1a2Δ5, pIBV1a2Δ8, and pIBV1a2Δ9 (see Fig. 1).

Plasmid pIBV1a2Δ1, which covers the IBV sequence from nucleotides 1311 to 5753 and contains an artificial AUG initiation codon in an optimal context (ACCAUGG) located immediately upstream of the viral sequence, was constructed by cloning a *Bam*HI/*Mlu*I-digested PCR fragment into *Bgl*II/*Mlu*I-digested pIBV1a2 (Fig. 2a). *Bgl*II digests pIBV1a2 at a position immediately upstream of the viral sequence and *Mlu*I digests the IBV sequence at nucleotide position 3997. The sequences of the upstream primer (LDX-3) and downstream primer (LDX-4) used were indicated in Table 1.

Plasmid pIBV1a2Δ3 was made by deletion of the IBV sequence between nucleotides 2251 and 3832 from pIBV1a2. This was achieved by digestion of pIBV1a2 with *Nco*I, which cuts the IBV sequences at nucleotide positions 2251, 2642, and 3832, respectively, and religation with T4 DNA ligase (Fig. 1).

Plasmid pIBV1a2Δ4 was constructed by cloning a 2685-bp *Bam*HI/*Mlu*I-digested PCR fragment, containing the IBV sequence from nucleotides 362 to 3047, into *Bgl*II/*Mlu*I-digested pIBV1a2, resulting in the deletion of the region between nucleotides 3047 and 3997. The upstream and downstream primers used were LKP-1 and LKP-2, respectively (Table 1).

Plasmid pIBV1a2Δ5, which covers the IBV sequence from nucleotides 362 to 5027, was made as follows. A 1030-bp restriction fragment covering the IBV sequence between nucleotides 3997 and 5027 was obtained by digestion of pIBV1a2 sequentially with *Mlu*I and *Spe*I and was then inserted into *Mlu*I/*Xba*I-digested pIBV1a1, giving rise to pIBV1a2Δ5. *Spe*I cuts the IBV sequence at nucleotide position 5027, and *Xba*I digests pIBV1a1 at a position 120 nucleotides downstream of the viral sequence presented in this plasmid.

Plasmid pIBV1a2Δ8, which contains a UAA stop codon at nucleotide position 2854, was constructed by inserting a *Bam*HI/*Mlu*I-digested PCR fragment into *Bgl*II/*Mlu*I-digested pIBV1a2. The sequences of the upstream primer LKP-1 and downstream primer LKP-3 are shown in Table 1. As can be seen, the UAA codon was introduced by the downstream primer LKP-3. Plasmid pIBV1a2Δ9, which covers the IBV sequence from nucleotides 1311 to 5753 and contains a UAA termination codon at nucleotide position 2854, was constructed as follows. A 1571-bp PCR fragment was generated with pIBV1a2Δ8 as the template and LDX-3 and LKP-3 as primers (Table 1). This PCR fragment was then digested with *Bam*HI and *Mlu*I and ligated into *Bgl*II- and *Mlu*I-digested pIBV1a2, giving rise to pIBV1a2Δ9.

Plasmids pIBV1a2Δ10, pIBV1a2Δ11, and pIBV1a2Δ12, which contain the IBV sequences starting from nucleotide 1311 and ending at nucleotides 2598, 2544, and 2514, respectively, were constructed by cloning *Bam*HI- and *Sma*I-digested PCR fragments into *Bgl*II- and *Sma*I-digested pKTO (Liu *et al.*, 1994). The PCR fragments used to construct these plasmids were generated with pIBV1a2Δ1 as template and LDX-3 as the upstream primer. The downstream primers used were LKP-4, LKP-5, and LKP-6, respectively (Table 1). Again, a UAA termination codon inserted immediately downstream of

the viral sequences in these constructs was introduced by the downstream primers (see Fig. 1).

Two mutants, pIBV1a2Δ4C^{1274}-S and pIBV1a2Δ4H^{1437}-K, with alterations at the putative catalytic residues Cys^{1274} and His^{1437}, respectively, were created by PCR with pIBV1a2Δ4 as templates. Plasmid pIBV1a2Δ4C^{1274}-S contains a substitution mutation of the Cys^{1274} residue to a Ser and pIBV1a2Δ4H^{1437}-K contains a mutation of the His^{1437} residue to a Lys. The pairs of primers used to introduce these mutations were LKP-7/LKP-8 and LKP-9/LKP-10, respectively (Table 1).

Similarly, plasmids pIBV1a2Δ4G^{673}-A, pIBV1a2Δ4T^{676}-S, and pIBV1a2Δ4ΔG^{673}-G^{674} were constructed by two rounds of PCR based on plasmid pIBV1a2Δ4. Plasmid pIBV1a2Δ4G^{673}-A contains a substitution mutation of the Gly^{673} residue to an Ala, and pIBV1a2Δ4T^{676}-S contains a mutation of the Thr^{676} residue to a Ser. The sequences of the mutation primers, LKP-11/LKP-12 and LKP-13/LKP-14, are shown in Table 1 Plasmid pIBV1a2Δ4ΔG^{673}-G^{674} contains a deletion of six nucleotides coding for Gly^{673} and Gly^{674} residues. The pair of primers used was LKP–15/LKP–16 (Table 1).

All the mutant constructs were selected and confirmed by automated sequencing.

REFERENCES

Baker, S. C., Yokomori, K., Dong, S. H., Carlisle, R., Gorbalenya, A. E., Koonin, E. V., and Lai, M. C. (1993). Identification of the catalytic sites of a papain-like cysteine proteinase of murine coronavirus. *J. Virol.* 67, 6056–6063.

Bonilla, P. J., Hughes, S. A., Pinon, J. D., and Weiss, S. R. (1995). Characterization of the leader papain-like proteinase of MHV-A59: Identification of a new *in vitro* cleavage site. *Virology* 209, 489–497.

Bonilla, P. J., Hughes, S. A., and Weiss, S. R. (1997). Characterization of a second cleavage site and demonstration of activity in *trans* by the papain-like proteinase of the murine coronavirus mouse hepatitis virus Strain A59. *J. Virol.* 71, 900–909.

Boursnell, M. E. G., Brown, T. D. K., Foulds, I. J., Green, P. F., Tomley, F. M., and Binns, M. M. (1987). Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *J. Gen. Virol.* 68, 57–77.

Brierley, I., Boursnell, M. E. G., Binns, M. M., Bilimoria, B., Blok, V. C., Brown, T. D. K., and Inglis, S. C. (1987). An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J.* 6, 3779–3785.

Brierley, I., Digard, P., and Inglis, S. C. (1989). Characterization of an efficient coronavirus ribosomal frameshifting signal: Requirement for an RNA pseudoknot. *Cell* 57, 537–547.

Brierley, I., Boursnell, M. E. G., Binns, M. M., Bilimoria, B., Rolley, N. J., Brown, T. D. K., and Inglis, S. C. (1990). Products of the polymerase-encoding region of the coronavirus IBV. *Adv. Exp. Med. Biol.* 276, 275–281.

Denison, M., and Perlman, S. (1987). Identification of putative polymerase gene product in cells infected with murine coronavirus A59. *Virology* 157, 565–568.

Denison, M. R., Hughes, S. A., and Weiss, S. R. (1995). Identification and characterization of a 65-kDa protein processed from the gene 1 polyprotien of the murine coronavirus MHV-A59. *Virology* 207, 316–320.

Dong, S. H., and Baker, S. C. (1994). Determination of the p28 cleavage site recognized by the first papain-like cysteine proteinase of murine coronavirus. *Virology* 204, 541–549.

Fuerst, T. R., Niles, E. G., Studier, F. W., and Moss, B. (1986). Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc. Natl. Acad. Sci. USA* 83, 8122–8127.

Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1989). Coronavirus genome: Prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Res.* 17, 4847–4861.

Gorbalenya, A. E., Koonin, E. V., and Lai, M. M. C. (1991). Putative papain-related thiol proteases of positive-strand RNA viruses. *FEBS Lett.* 288, 201–205.

Grotzinger, C., Heusipp, G., Ziebuhr, J., Harms, U., Suss, J., and Siddell, S. G. (1996). Characterization of a 105-kDa polypeptide encoded in gene 1 of the human coronavirus HCV 229E. *Virology* 222, 227–235.

Herold, J., Gorbalenya, A. E., Thiel, V., Schelle, B., and Siddell, S. G. (1998). Proteolytic processing at the amino terminus of human coronavirus 229E gene 1-encoded polyproteins: Identification of a papain-like proteinase and its substrate. *J. Virol.* 72, 910–918.

Heusipp, G., Harms, U., Siddell, S. G., and Ziebuhr, J. (1997a). Identification of an ATPase activity associated with a 71-kilodalton polypeptide encoded in gene 1 of the human coronavirus 229E. *J. Virol.* 71, 5631–5634.

Heusipp, G., Grotzinger, C., Herold, J., Siddell, S. G., and Ziebuhr, J. (1997b). Identification and subcellular localization of a 41 kDa, polyprotein lab processing product in human coronavirus 229E-infected cells. *J. Gen. Virol.* 78, 2789–2794.

Hughes, S. A., Bonilla, P. J., and Weiss, S. R. (1995). Identification of the murine coronavirus p28 cleavage site. *J. Virol.* 69, 809–813.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.

Lee, H-J., Shieh, C-K., Gorbalenya, A. E., Koonin, E. V., Monica, N. L., Tuler, J., Bagdzhadzhyan, A., and Lai, M. M. C. (1991). The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* 190, 567–582.

Liu, D. X., Cavanagh, D., Green, P., and Inglis, S. C. (1991). A polycistronic mRNA specified by the coronavirus infectious bronchitis virus. *Virology* 184, 531–544.

Liu, D. X., Brierley, I., Tibbles, K. W., and Brown, T. D. K. (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.

Liu, D. X., and Brown, T. D. K. (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.

Liu, D. X., Tibbles, K. W., Cavanagh, D., Brown, T. D. K., and Brierley, I. (1995). Identification, expression and processing of an 87 kDa polypeptide encoded by ORF 1a of the coronavirus infectious bronchitis virus. *Virology* 208, 48–57.

Liu, D. X., Xu, H. Y., and Brown, T. D. K. (1997). Proteolytic processing of the coronavirus infectious bronchitis virus 1a polyprotein: Identification of a 10-kilodalton polypeptide and determination of its cleavage sites. *J. Virol.* 71, 1814–1820.

Lu, Y., Lu, X., and Denison, M. R. (1995). Identification and characterization of a serine-like proteinase of the murine coronavirus MHV-A59. *J. Virol.* 69, 3554–3559.

Stern, D. F., and Sefton, B. M. (1984). Coronavirus multiplication: Location of genes for virion proteins on the avian infectious bronchitis virus genome. *J. Virol.* 50, 22–29.

Ziebuhr, J., Herold, J., and Siddell, S. G. (1995). Characterization of a human coronavirus (strain 229E) 3C-like proteinase activity. *J. Virol.* 69, 4331–4338.