Virology **263,** 471–484 (1999)
Article ID viro.1999.9954, available online at http://www.idealibrary.com on IDEAL®

Further Requirements for Cleavage by the Murine Coronavirus 3C-like Proteinase: Identification of a Cleavage Site within ORF1b

Josefina D. Piñón, Henry Teng, and Susan R. Weiss[1]

Department of Microbiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6076

Received May 25, 1999; returned to author for revision July 1, 1999; accepted August 4, 1999

The coronavirus mouse hepatitis virus strain A59 (MHV-A59) encodes a 3C-like proteinase (3CLpro) that is proposed to be responsible for the majority of the processing events that take place within the replicase polyproteins pp1a and pp1ab. In this study we demonstrate that the Q939↓S940 peptide bond, located between the polymerase and Zn-finger regions of pp1ab (the POL↓Zn site), is processed by the 3CLpro, albeit inefficiently. Mutagenesis of the POL↓Zn site, as well as the previously identified HD1↓3C site in the 1a region of pp1a and pp1ab, demonstrated that the amino acid residues at the P2 and P1 positions of the cleavage site, occupied by L and Q, respectively, were important determinants of 3CLpro substrate specificity. Finally, a direct comparison of the 3CLpro-mediated cleavages at the HD1↓3C and POL↓Zn sites was made by determining the rate constants using synthetic peptides. The results show that while a larger polypeptide substrate carrying the HD1↓3C site was processed more efficiently than a polypeptide substrate carrying the POL↓Zn site, cleavage of the synthetic peptide substrates containing these two cleavage sites occurred at similar efficiencies. This indicates that the overall conformation of a large polyprotein substrate is important in the accessibility of the cleavage site to the proteinase.

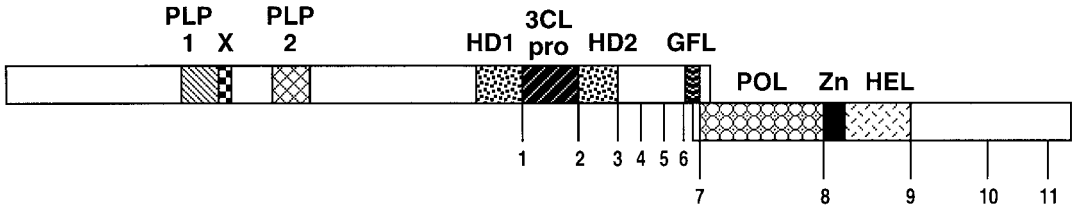
INTRODUCTION

The family Coronaviridae is composed of a group of viruses that cause a variety of diseases in different animal hosts. The murine coronavirus, mouse hepatitis virus (MHV), causes a range of diseases in mouse, including enteritis, hepatitis, encephalitis, and a demyelinating disease (Holmes and Lai, 1996; Houtman and Fleming, 1996; Lai, 1990). Coronaviruses, along with the arteriviruses, are classified under the newly established order Nidovirales based on the similarities in their genome organization and replication strategy (Cavanagh, 1997; de Vries *et al.,* 1997). The name Nidovirales originates from the Latin word *nidus,* meaning "nest," and refers to the 3′ nested set of subgenomic mRNAs that is produced during viral infection (de Vries *et al.,* 1997). As with all positive-strand RNA viruses, entry of the viral genome into the cytoplasm is followed by the translation of the positive-strand RNA genome, resulting in the expression of viral proteins. The coronavirus genome is organized into seven genes that are separated by stop codons and intergenic sequences (Holmes and Lai, 1996; Lai, 1990). Thus, translation of the viral genome results only in the expression of gene 1 proteins. The rest of the viral genome is expressed through subgenomic mRNAs that are transcribed by the viral RNA-dependent RNA polymerase encoded in gene 1. Replication of the viral genome also requires the replicase proteins encoded in gene 1. Thus, for both viral replication and subgenomic mRNA transcription to take place, the expression of gene 1 products is essential.

The replicase gene (gene 1) (Fig. 1) of coronaviruses, spanning 20–22 kb, is organized into two overlapping open reading frames, ORF1a and ORF1b (Bonilla *et al.,* 1994; Lee *et al.,* 1991). The expression of the downstream ORF1b is mediated by a ribosomal frameshift event that is aided by the formation of a pseudoknot structure within the overlapping region (Bredenbeek *et al.,* 1990; Brierley *et al.,* 1987; Herold and Siddell, 1993). Thus, two polypeptides, pp1a and pp1ab, are expressed from gene 1, with the translation of pp1ab being only 25–40% as efficient as that of pp1a in *in vitro* studies (Bredenbeek *et al.,* 1990; Brierley *et al.,* 1987). Through a series of intricate cotranslational and posttranslational processing events, these polyproteins are converted into a functional complex that in turn is responsible for both genomic RNA replication and subgenomic mRNA transcription (de Vries *et al.,* 1997). Responsible for these processing events are at least two or three viral proteinases encoded within the ORF1a region of gene 1 (Fig. 1). Two of these proteinase domains, by sequence analysis, share similarities with the cellular proteinase papain. A third proteinase, resembling the poliovirus 3C proteinase, has also been identified (Gorbalenya *et al.,* 1989; Lee *et al.,* 1991). The coronavirus 3C-like proteinase

[1] To whom correspondence and reprint requests should be addressed at 203A Johnson Pavilion, 3610 Hamilton Walk, Philadelphia, PA 19104-6076. Fax: (215) 573 4858. E-mail: weisssr@mail.med.upenn.edu.

0042-6822/99 $30.00



cleavage site	putative protein next to C-terminus of cleavage site	cleavage site sequence										
		P5	P4	P3	P2	P1	↓	P1'	P2'	P3'	P4'	P5'
1	3C	T	S	F	L	Q	↓	S	G	I	V	K
2	HD2	G	V	K	L	Q	↓	S	K	R	T	R
3	?	V	S	Q	I	Q	↓	S	R	L	T	D
4	p22	L	Q	A	L	Q	↓	S	E	F	V	N
5	?	T	V	V	L	Q	↓	N	N	E	L	M
6	GFL	T	V	R	L	Q	↓	A	G	T	A	T
7	POL	G	S	Q	F	Q	↓	S	K	D	T	N
8	HEL	S	A	V	L	Q	↓	S	V	G	A	C
9	?	N	P	R	L	Q	↓	C	T	T	N	L
10	?	F	T	R	L	Q	↓	S	L	E	N	V
11	?	Y	P	R	L	Q	↓	A	A	A	D	W

FIG. 1. Map of ORF1a and 1b showing the locations of the predicted functional domains. The following domains are shown: papain-like proteinase (PLP-1 and PLP-2), X domain (X), poliovirus 3C-like proteinase (3CLpro), hydrophobic domains (HD1 and HD2), growth factor-like domain (GFL), RNA-dependent RNA polymerase (POL), zinc-finger domain (Zn), and helicase (HEL). The predicted 3CLpro cleavage sites are indicated by the numbers and the cleavage site sequences from the P5 to the P5′ position are listed in table form (arrow indicates site of cleavage) (Bonilla *et al.,* 1994; Bredenbeek *et al.,* 1990; Gorbalenya *et al.,* 1989; Lee *et al.,* 1991).

(3CLpro), flanked on either side by hydrophobic, possibly membrane-spanning regions (HD1 and HD2), is believed to be the prinicipal viral proteinase responsible for the processing events leading to the formation of the viral replicase complex, with as many as 11 potential cleavage sites identified throughout pp1ab (Gorbalenya *et al.,* 1989; Lee *et al.,* 1991) (see Fig. 1). The presence of the 3CLpro is conserved in all coronavirus genomes studied to date (Bonilla *et al.,* 1994; Boursnell *et al.,* 1987; Eleouet *et al.,* 1995; Herold *et al.,* 1993; Lee *et al.,* 1991). The 3CLpro of MHV-A59 has been identified as a 29-kDa protein (p29) both in *in vitro* study and in MHV-A59-infected cells (Piñón *et al.,* 1997). (Lu *et al.* (1995) reported a molecular weight of 27 kDa for the same polypeptide.) The catalytic residues of the MHV-A59 3CLpro, His3374, and Cys3478 have also been identified (Lu *et al.,* 1995). Treatment of infected cells with E-64d, a known inhibitor of the 3CLpro, results in the inhibition of viral RNA replication in these cells (Kim *et al.,* 1995), demonstrating the importance of the action of the 3CLpro in the events leading to viral replication. Van Dinten *et al.* (1999) demonstrated the importance of 3CLpro cleavages using an infectious clone of the related arterivirus EAV; introduction of mutations into the candidate ORF 1b 3CLpro cleavage sites had drastic effects on RNA synthesis and virus replication. These fndings indicate that this proteinase is a good potential target for antiviral therapy.

The cleavage sites of the coronavirus 3CLpro (Fig. 1) conform to the consensus XQ↓Z (arrow indicates site of cleavage), with X being a hydrophobic residue, usually L, although the amino acids I, M, V, and F are also found in this position (de Vries *et al.,* 1997). At the P1′ position, Z is usually a small uncharged residue such as S, A, G, or C (de Vries *et al.,* 1997), with S being the most common residue at this position. Data recently obtained for the avian infectious bronchitis virus (IBV), the human coronavirus (HCV-229E), and MHV-A59 demonstrated that many of these predicted cleavage sites are functional. For all three strains, many of the cleavage sites in pp1a have been identified, including a noncanonical LQ↓N site that had not previously been predicted (Lu *et al.,* 1995; 1998; Liu and Brown, 1995; Liu *et al.,* 1997; Ng and Liu, 1998; Piñón *et al.,* 1997; Tibbles *et al.,* 1996; Ziebuhr *et al.,* 1995; Ziebuhr and Siddell, 1999). In addition, several of the processing sites in ORF1b have also been identified in both IBV and HCV-229E (Grötzinger *et al.,* 1996; Heusipp *et al.,* 1997a,b; Liu *et al.,* 1994, 1998). However, for MHV, there has been greater difficulty in demonstrating processing by the 3CLpro at any of the predicted ORF1b cleavage sites. In this study we demonstrate processing by the 3CLpro at the putative ORF1b

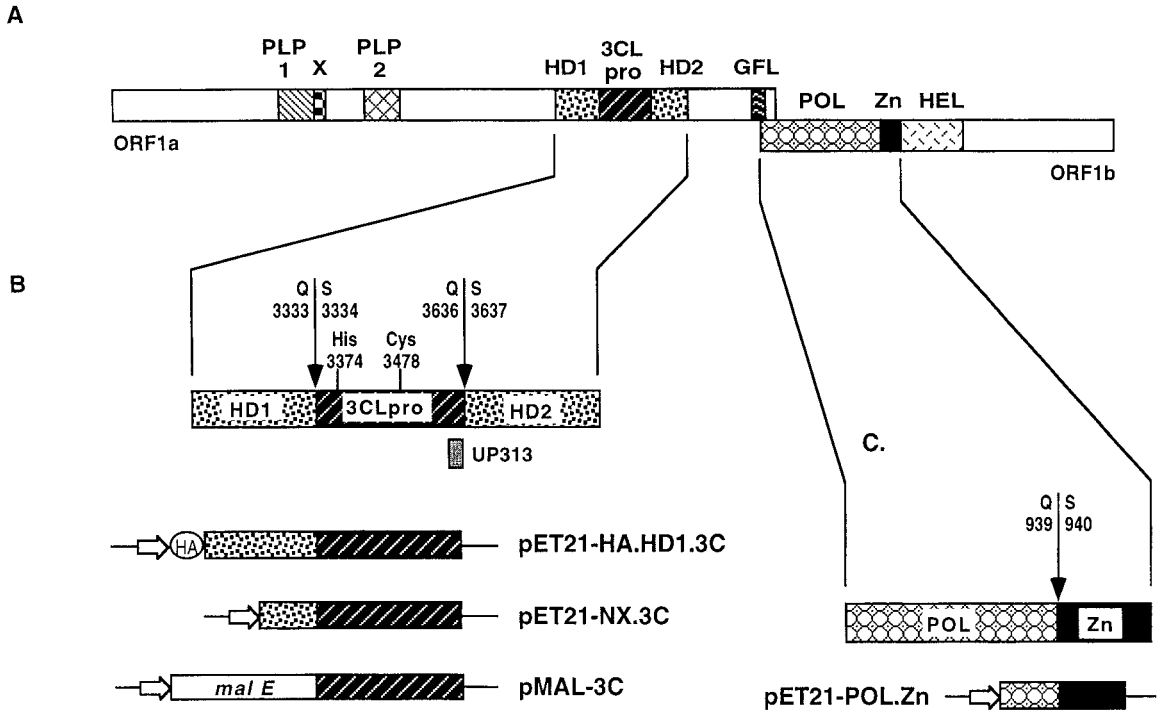


FIG. 2. Functional domains of MHV-A59 gene 1 and diagram of plasmid constructs. (A) Location of the predicted functional domains in ORF1a and 1b. (B) Enlarged map of the HD1.3C.HD2 region and schematic representation of plasmids derived from this region. The 3CLpro lies within the region from S3334 and Q3636 (Bonilla *et al.,* 1994; Bredenbeek *et al.,* 1990; Lee *et al.,* 1991). The positions of the catalytic residues, His3374 and Cys3478, are shown. pET 21-HA-HD1.3C encodes MHV-A59 amino acids from S3149 to G3636 directly downstream of the influenza hemagglutinin (HA) tag under the control of the T7 promoter (✣). pET21-NX.3C encodes MHV-A59 amino acids from K3258 to Q3636. pMal-3C.wt encodes MHV-A59 amino acid sequences from S3334 to G3636 fused to the *mal E* gene encoding the maltose-binding protein (MBP). The construct pMAL-3C is used for the overexpression of the 3CLpro in *Escherichia coli.* (C) Enlarged map of POL and Zn region of ORF1b and schematic representation of plasmids derived from this region. The position of the Q939↓S940 cleavage site at the junction between POL and Zn is shown. pET21-POL.Zn encodes MHV-A59 ORF1b amino acids from L714 to Q1201 under the control of the T7 promoter.

cleavage site situated between the putative polymerase (POL) and zinc finger (Zn) domains (the POL↓Zn site) (Fig. 2). Furthermore, identification of this site allowed us to make comparisons between the efficiencies of processing at this ORF1b site and a previously identified site in ORF1a located at the junction between hydrophobic domain 1 (HD1) and the 3CLpro (the HD1↓3C site).

RESULTS

Identification of a 3CLpro cleavage site in ORF1b

We have previously demonstrated processing at an ORF1a site in MHV-A59, located at the junction between HD1 and the 3CLpro (HD1↓3C), by a recombinant MHV-A59 3CLpro expressed as a fusion protein with the maltose binding protein (MBP) (Piñón *et al.,* 1997). The plasmid pET21-NX.3C C3478A, encoding the carboxy-terminal 98 amino acids of HD1 and the inactivated 3CLpro, is *in vitro* transcribed and translated to yield a 40-kDa substrate NX.3C (Fig. 3A, lane 1). As previously demonstrated (Piñón *et al.,* 1997), upon addition of the recombinant MBP-3CLpro, this substrate is efficiently processed into the 29-kDa proteinase (3C) and the 11-kDa HD1-derived product (NX) (lane 2). Other studies using similar methods have led to the identification of ORF1a cleavage sites downstream of the 3CLpro, including a previously unidentified, noncanonical LQ↓N site (Lu *et al.,* 1998). However, the demonstration of processing at ORF1b sites has proven to be difficult. Our work with papain-like proteinase 1 (PLP-1) showed that substrate length, and possibly substrate conformation, played an important role in the ability of a substrate to be cleaved efficiently by the proteinase (Teng *et al.,* 1999). We therefore created several substrates of various lengths, encoding different putative cleavage sites in ORF1b, in order to investigate processing by the recombinant MBP-3CLpro. Of these, only the substrate expressed from pET21-POL.Zn, encoding MHV-A59 ORF1b amino acids L714–Q1201, proved to be useful in our investigations. *In vitro* transcription–translation of the construct pET21-POL.Zn resulted in the expression of a 55-kDa full-length substrate, POL.Zn (predicted molecular weight 56 kDa) (Fig. 3B, lane 1). Upon incubation of this substrate with the recombinant MBP-3CLpro enzyme, cleavage between Q939 and S940, would give rise to two products, an N-terminal product with a predicted molecular weight

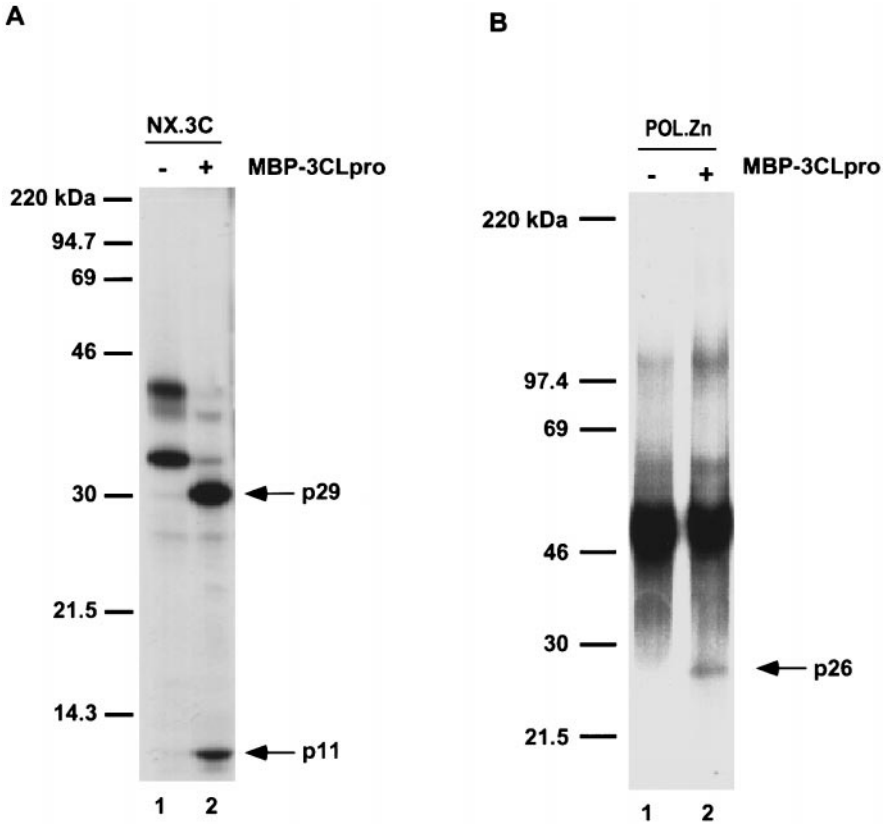


FIG. 3. Demonstration of *trans* processing by the recombinant MBP-3CLpro. (A) *trans* processing at the HD1↓3C site. Radiolabeled, *in vitro* transcribed, and translated substrate from pET21-NX.3C was incubated with MBP-3CLpro (lane 2) or an equal volume of column buffer/20% glycerol (lane 1) and the processed products were separated on a 15% SDS–PAGE gel. The arrows on the right of the panel indicate the electrophoretic migration of the 29-kDa 3CLpro- and the 11-kDa HD1-derived cleavage products (NX). (B) *trans* processing at the POL↓Zn site. The plasmid pET21-POL.Zn was *in vitro* transcribed and translated. Radiolabeled substrate was incubated either with MBP-3CLpro or with an equal volume of column buffer/20% glycerol (denoted by plus or minus signs above the lanes, respectively). Processed products were analyzed on a 12% SDS gel. The electrophoretic migration of p26 is indicated by an arrow on the right of the panel. The molecular weight in kilodaltons of prestained protein markers is indicated on the left of each panel.

of 26 kDa (POL) and a C-terminal product with a predicted molecular weight of 30 kDa (Zn). In Fig. 3B, lane 2, the addition of the recombinant MBP-3CLpro to the POL.Zn substrate resulted in the production of a 26-kDa protein which presumably corresponds to the N-terminal processing product. Mutagenesis of Q939 to K (lane 4) or R (lane 6) abolished this processing, suggesting that the cleavage is occurring at the predicted site and that, in view of the substrate specificity demonstrated, the observed processing event is in fact due to the action of the recombinant MBP-3CLpro.

We were unable to detect the C-terminal 30-kDa product by SDS–PAGE analysis. One explanation could be that p30 cannot be resolved from p26 in our gel system. (There have been previous reports of viral proteins migrating with electrophoretic mobilities different from that expected. The MHV-A59 3CLpro (p29) itself migrates with an electrophoretic mobility faster than its predicted molecular weight of 33 kDa (Piñón *et al.,* 1997; Lu *et al.,* 1995)). In addition, the predicted cleavage product p30 has approximately half the methionine content of p26, which may contribute to the difficulty in its detection.

Characterization of 3CLpro cleavage sites

The identification of these two cleavage sites, one in ORF1a (HD1↓3C) and one in ORF1b (POL↓Zn), allowed us to further define the amino acids required for efficient processing by the 3CLpro. For these experiments, we chose to use the pET21-HA.HD1.3C construct, rather than other plasmids encoding 3CLpro, because the HA.HD1.3C polypeptide can be efficiently cleaved both *in cis* and *in trans* and does not require membranes for its cleavage (Piñón *et al.,* 1997).

Several sets of mutations, from the P3 to the P3′ position, were introduced into the construct pET21-HA.HD1.3C by PCR mutagenesis using the mutagenesis primers outlined in Table 1. The effect of these cleavage site mutations on the autocatalytic *cis* release of the 29-kDa 3CLpro was assayed by the expression of the

TABLE 1

Primers for PCR Amplification and Mutagenesis

Primer name	Nucleotide sequence (5′–3′)[a]	MHV-A59 gene 1[b]
ORF1a		
FMP F3331A	CTCTGTTACTACATCA**GCT**TTACAGTCTGGTATA	10184–10217
RMP F3331A	TATACCAGACTGTAA**AGC**TGATGTAGTAACAGAG	10217–10184
FMP F3331H	CTCTGTTACTACATCA**CAT**TTACAGTCTGGTATA	10184–10217
RMP F3331H	TATACCAGACTGTAA**ATG**TGATGTAGTAACAGAG	10217–10184
FMP F3331W	TCTGTTACTACATCA**TGG**TTACAGTCTGGTATAG	10185–10218
RMP F3331W	CTATACCAGACTGTAA**CCA**TGATGTAGTTACAGA	10218–10185
FMP L3332I	GTTACTACATCATTT**ATA**CAGTCTGGTATAGTG	10188–10220
RMP L3332I	CACTATACCAGACTG**TAT**AAATGATGTAGTAAC	10220–10188
FMP L3332S	GTTACTACATCATTT**TCA**CAGTCTGGTATAGTG	10188–10220
RMP L3332S	CACTATACCAGACTG**TGA**AAATGATGTAGTAAC	10220–10188
FMP Q3333A	TACTACATCATTTTTA**GCG**TCTGGTATAGTGAAG	10190–10223
RMP Q3333A	CTTCACTATACCAGA**CGC**TAAAAATGATGTAGTA	10223–10190
FMP Q3333K	ACTACATCATTTTTA**AAA**TCTGGTATAGTGAAG	10191–10223
RMP Q3333K	CTTCACTATACCAGA**TTT**TAAAAATGATGTAGT	10223–10191
FMP Q3333R	ACTACATCATTTTTA**CGG**TCTGGTATAGTGAAG	10191–10223
RMP Q3333R	CTTCACTATACCAGA**CCG**TAAAAATGATGTAGT	10223–10191
FMP S3334A	ACATCATTTTTACAG**GCA**GGTATAGTGAAGATG	10194–10226
RMP S3334A	CATCTTCACTATACC**TGC**CTGTAAAAATGATGT	10226–10194
FMP S3334C	ACATCATTTTTACAG**TGT**GGTATAGTGAAGATG	10194–10226
RMP S3334C	CATCTTCACTATACC**ACA**CTGTAAAAATGATGT	10226–10194
FMP G3335A	TCATTTTTACAGTCT**GCT**ATAGTGAAGATGGTG	10197–10229
RMP G3335A	CACCATCTTCACTAT**AGC**AGACTGTAAAAATGA	10229–10197
FMP G3335P	ATCATTTTTACAGTCT**CCT**ATAGTGAAGATGGTG	10196–10229
RMP G3335P	CACCATCTTCACTAT**AGG**AGACTGTAAAAATGAT	10229–10196
FMP I3336L	ATTTTTACAGTCTGGT**CTA**GTGAAGATGGTGTCG	10199–10232
RMP I3336L	CGACACCATCTTCAC**TAG**ACCAGACTGTAAAAAT	10232–10199
ORF1b		
FMP V937S	TATTTAAGAAGTGCA**TCG**CTGCAAAGCGTTGG	16331–16362
RMP V937S	CCAACGCTTTGCAG**CGA**TGCACTTCTTAAATA	16362–16331
FMP L938I	GAAGTGCAGTG**ATC**CAAAGCGTTGG	16338–16362
RMP L938I	CCAACGCTTTG**GAT**CACTGCACTTC	16362–16338
FMP L938M	GAAGTGCAGTG**ATG**CAAAGCGTTG	16338–16361
RMP L938M	CAACGCTTTG**CAT**CACTGCACTTC	16361–16338
FMP Q939K	GAAGTGCAGTGCTG**AAA**AGCGTTGGTGCC	16338–16366
RMP Q939K	GGCACCAACGCT**TTT**CAGCACTGCACTTC	16366–16338
FMP Q939R	GAAGTGCAGTGCTG**CGA**AGCGTTGGTGCCTG	16338–16368
RMP Q939R	CAGGCACCAACGCT**TCG**CAGCACTGCACTTC	16368–16338
FMP S940A	GTGCAGTGCTGCAA**GCC**GTTGGTGCCTGCG	16341–16370
RMP S940A	CGCAGGCACCAAC**GGC**TTGCAGCACTGCAC	16370–16341
FMP S940N	GCAGTGCTGCAA**AAC**GTTGGTGCCTGC	16343–16369
RMP S940N	GCAGGCACCAAC**GTT**TTGCAGCACTGC	16369–16343

[a] Mutated codons are in boldface.
[b] Reverse position of MHV-A59 gene 1 indicates negative-strand primer.

mutated substrates using *in vitro* transcription–translation, followed by SDS–PAGE analyses of the protein products (Fig. 4). We observed that the L3332 and Q3333 residues, at the P2 and P1 positions, respectively, were most sensitive to mutations. Any mutation at either one of these positions inhibited the autocatalytic *cis* processing by the 3CLpro (Fig. 4, lanes 6–10). In contrast, mutations at any of the other positions studied were tolerated and the expression of precursor proteins harboring mutations at these sites still resulted in the autocatalytic release of the 29-kDa 3CLpro. Exceptions are the S3334C substitution at position P1′ (lane 12) and the G3335P substitution at position P2′ (lane 14). These mutations also abolish the *cis* processing by the 3CLpro. The effect of the G3335P mutation, however, is expected since the introduction of a P at this site could result in a drastic change in the conformation at the cleavage site.

To investigate whether the cleavage sequence requirements for *trans* cleavage at the HD1↓3C site by the recombinant MBP-3CLpro parallel that observed for *cis* cleavage, the same set of cleavage site mutations were introduced into the construct pET21-HA.HD1.3C C3478A,

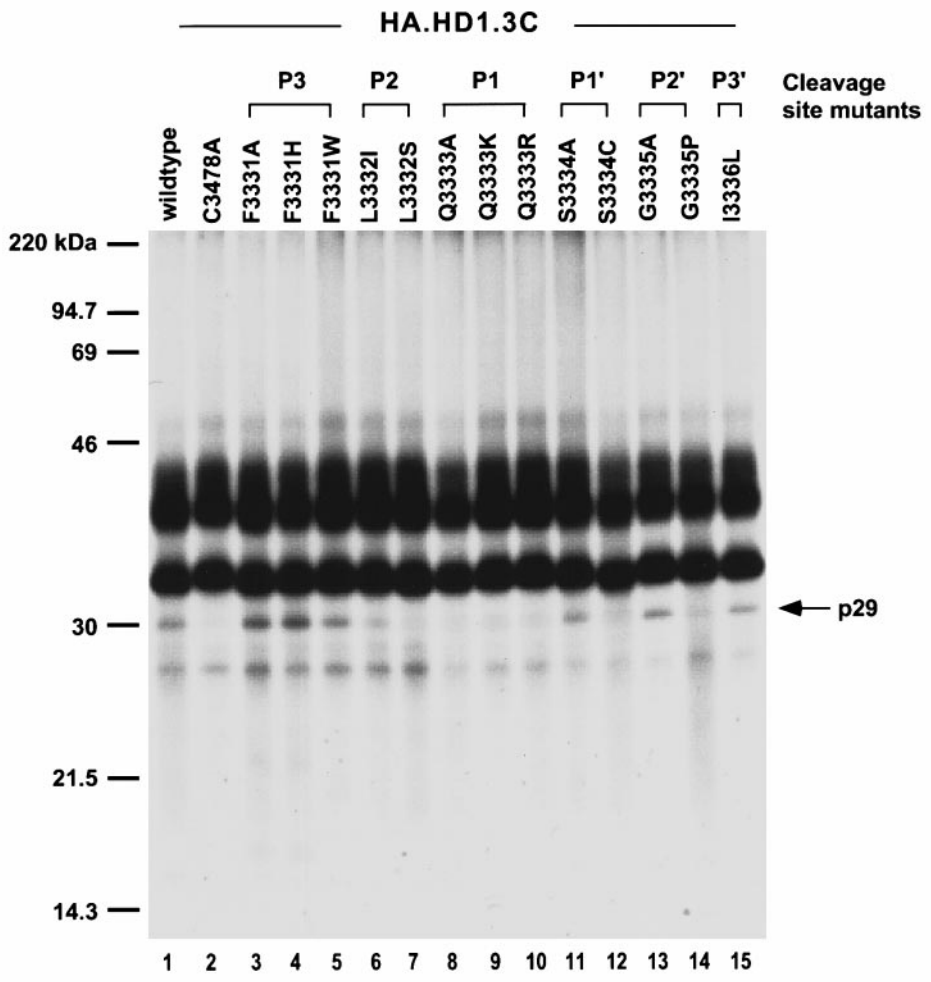


FIG. 4. HD1↓3C cleavage site mutagenesis: *cis* processing at the HD1↓3C site. The autocatalytic release of the 29-kDa 3CLpro from precursors harboring mutations at the HD1↓3C cleavage site in the plasmid pET21-HA.HD1.3C was assayed by *in vitro* transcription and translation of the wildtype and mutant plasmids, followed by separation of precursor and products on a 12% SDS gel. The electrophoretic migration of p29 is indicated by an arrow to the right of the panel. Molecular weight markers are indicated on the left.

which also carries an inactivating mutation in the catalytic cysteine residue of the proteinase. The release of the 29-kDa 3CLpro from this precursor can only be accomplished by incubation with the recombinant MBP-3CLpro. Figure 5 showed that the effect of these mutations on *trans* cleavage paralleled the effects on *cis* cleavage. Those mutations centering around the P2 and P1 positions of the cleavage site (L3332 and Q3333, respectively) affected *trans* processing the most.

For the POL↓Zn cleavage site in ORF1b, similar mutagenesis studies were conducted in order to determine the cleavage specificity requirements at this site (Fig. 6). Mutations from the P3 to the P1′ position were introduced into the plasmid pET21-POL.Zn. The mutant substrates were expressed using an *in vitro* transcription–translation system and then incubated with the recombinant MBP-3CLpro. The effect of the mutations on the production of p26 was assayed by SDS–PAGE and compared to the processing of the wildtype POL.Zn substrate (Fig. 6, lanes 1 and 2). The results were similar to that observed with the ORF1a HD1↓3C site in that the L938 and Q939 residues, at the P2 and P1 positions respectively, were most sensitive to mutations. The L938I mutation (Fig. 6, lanes 5 and 6) resulted in the inhibition of p26 processing, indicating the sensitivity of this site to a conservative change. However, the L938M substitution (Fig. 6, lanes 7 and 8) did not abolish processing of p26. Mutation of Q939 to either K (lanes 9 and 10) or R (lanes 11 and 12) abolished processing at this site. Substitution of S940 with an A (lanes 13 and 14) was tolerated, whereas mutating S940 to the bulkier N (lanes 15 and 16) resulted in the inhibition of p26 processing.

The results of mutagenesis studies on both the HD1↓3C and POL↓Zn cleavage sites demonstrate that the P1 and P2 positions at the cleavage site are the primary determinants of cleavage specificity by the 3CLpro in both *cis* and *trans* processing.

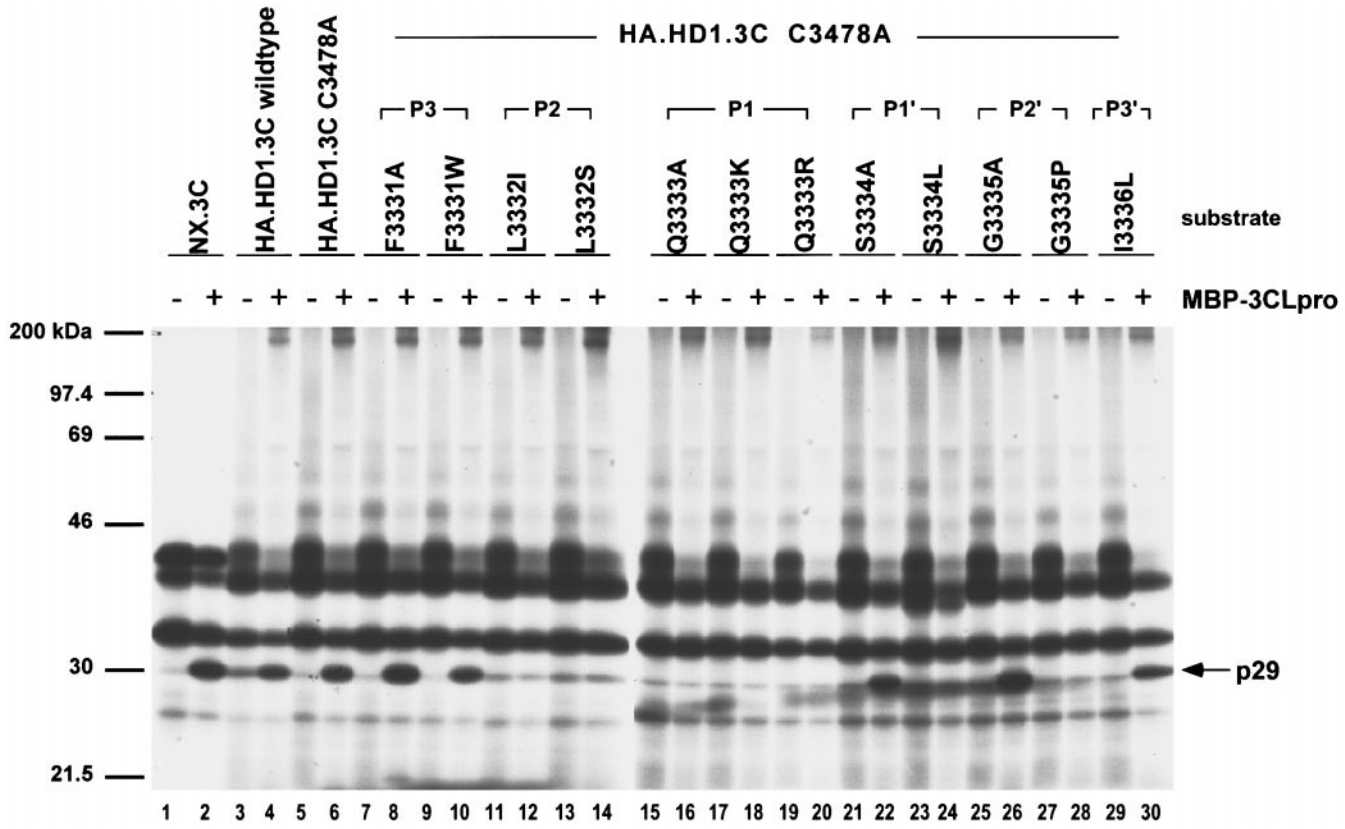


FIG. 5. HD1↓3C cleavage site mutagenesis: *trans* processing at the HD1↓3C site. The same cleavage site mutants used in the study of *cis* processing were introduced into pET21-HA.HD1.3C C3478A, which also harbors a mutation at the catalytic cysteine residue of the proteinase. Equivalent counts per minute of radiolabeled substrates expressed from these mutant plasmids by TnT were incubated with MBP-3CLpro or an equivalent volume of column buffer/20% glycerol (+/− MBP-3CLpro). Cleavage products were then separated on a 10% SDS–gel. The electrophoretic migration of p29 is indicated by an arrow on the right. Molecular weight markers are indicated on the left.

Comparison of ORF1a and ORF1b sites

We investigated the efficiency of processing at the ORF1a site compared to the ORF1b site. The substrates, NX.3C and POL.Zn, were incubated with a fixed amount of recombinant proteinase for increasing periods of time. We observed that the ORF1a substrate, NX.3C, is readily processed within 1 h, with the substrate completely converted into product by 20 h (Fig. 7A). Processing of the ORF1b substrate, however, is not observed until after 4 h of incubation with the recombinant proteinase. Furthermore, cleavage of POL.Zn occurs at a much lower level and does not reach completion even after 30 h (Fig. 7B). This inefficiency of processing of the POL↓Zn site compared to the HD1↓3C site may be explained in two ways. First, although the primary sequences of the two sites reveal no obvious reason why one is processed more efficiently than the other, it is possible that the subtle differences in the sequences of the two sites are enough to make the HD1↓3C site a more efficient substrate than the POL↓Zn site. Alternatively, the difference may not be inherent to the primary sequence of the cleavage sites, but rather to the conformation of the entire substrate as a whole. It is possible that the substrate conformation of NX.3C allows the recombinant proteinase better access to the cleavage site. In the POL.Zn substrate, the cleavage site may be more obscured, and in this manner the virus regulates both when and how much of its encoded proteins are produced.

To determine whether the difference in processing efficiencies at the HD1↓3C and POL↓Zn sites can be attributed to the primary sequences of the cleavage sites, we synthesized a 14-mer peptide, representing the P6–P8′ residues of the HD1↓3C cleavage site, and a 15-mer peptide, representing the P6–P9′ amino acids of the POL↓Zn cleavage site. Each peptide was then incubated with the recombinant MBP-3CLpro and the resulting cleavage products were separated from each other and from the substrate by reverse-phase chromatography. For the ORF1a peptide, reverse-phase chromatography of the reaction at zero time resulted in a single peak, representing the peptide substrate, in the elution profile (Fig. 8A). Incubation with MBP-3CLpro followed by chromatography resulted in two additional peaks, representing the cleavage products, and a reduction in the

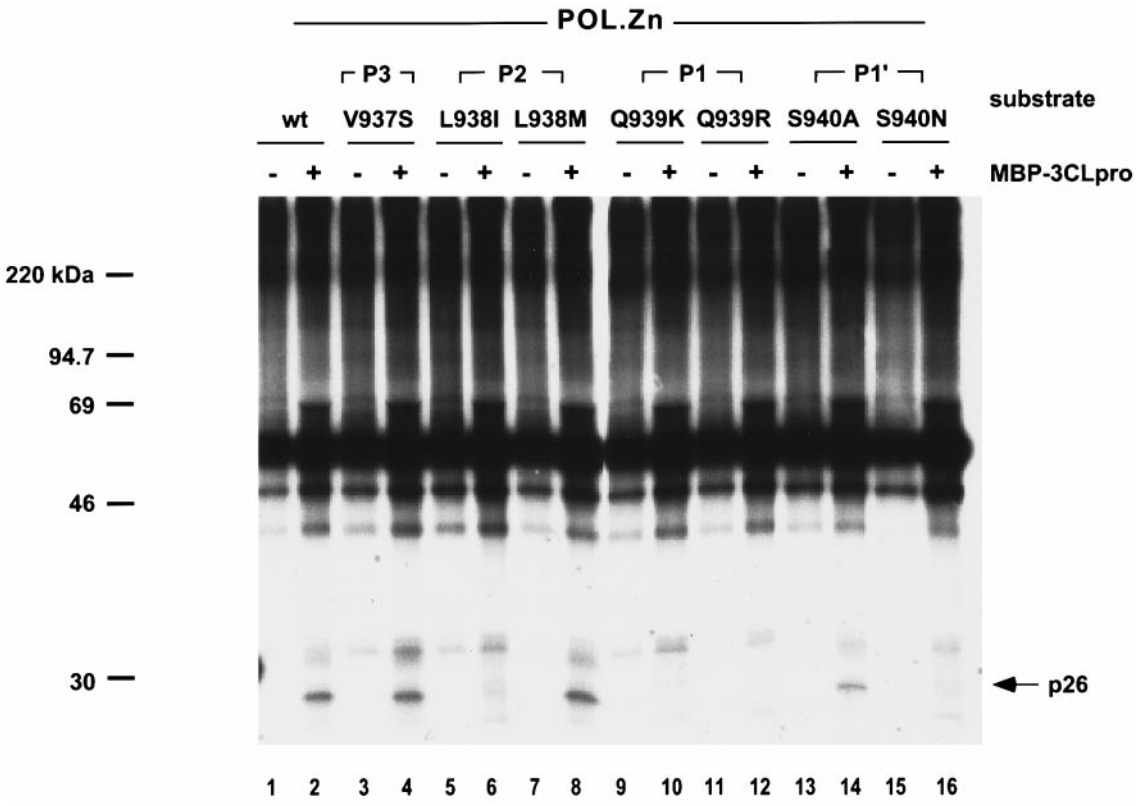


FIG. 6. POL↓Zn cleavage site mutagenesis. Recombinant MBP-3CLpro was used in posttranslational *trans* cleavage assays with radiolabeled substrates generated from pET21-POL.Zn that encoded a wildtype cleavage sequence or harbored mutations around the POL↓Zn site. Substrate volumes containing equivalent counts per minute were incubated with MBP-3CLpro (denoted by plus sign above the lanes) or an equivalent volume of column buffer/20% glycerol (denoted by a minus sign above the lanes). The electrophoretic migration of p26 is indicated by the arrow to the right of the panel. Molecular weight markers are indicated on the left.

substrate peak. Microsequencing of the cleavage products confirmed that cleavage had occurred between Q3333 and S3334.

For the ORF1b peptide similar results were observed in that the peptide substrate eluted as a single peak at zero time (Fig. 8B). Incubation with MBP-3CLpro prior to separation resulted in two new peaks, representing the two cleavage products. A reduction in the substrate peak was also observed. Again, the authenticity of cleavage of the synthetic peptide was confirmed by microsequencing of the cleavage products, which showed that cleavage had occurred between Q939 and S940.

The level of cleavage of the ORF1a peptide was not any different from, and in some cases was less than, that of the ORF1b peptide. We observed that under identical reaction conditions no more than 30% of the ORF1a peptide was cleaved by the MBP-3CLpro, whereas with the ORF1b peptide the level of cleavage was between 30 and 50%.

In order to allow direct comparison between the cleavage efficiencies of the NX↓3C and POL↓Zn sites, we determined the k_{cat} and K_m values for the reactions with the synthetic peptides described above. Results show that cleavage of the ORF1a peptide by the recombinant MBP-3CLpro yielded a k_{cat} of 0.015 s^{-1}, and a K_m of 2.32 $\pm$ 0.50 mM. With the ORF1b peptide the k_{cat} was slightly slower, with a value of 0.006 s^{-1}, and the K_m was calculated at 0.27 $\pm$ 0.07 mM.

Our results therefore show that cleavage (k_{cat}) of the ORF1a peptide occurred at only a slightly higher rate (2.5-fold) than that of the ORF1b peptide. The lack of substantial difference between the k_{cat} values is not surprising given that the sequences of the peptides are very similar. Interestingly, when the catalytic efficiency (k_{cat}/K_m) was taken into consideration, the ORF1b peptide ($k_{cat}/K_m = 2.1 \times 10^{-5}$ $M^{-1}s^{-1}$) was a slightly better substrate than the ORF1a peptide ($k_{cat}/K_m = 6.4 \times 10^{-6}$ $M^{-1}s^{-1}$). Taken together, the peptide cleavage results presented here suggest that the cleavage efficiencies of the two peptides are similar.

DISCUSSION

The action of viral-encoded proteinases is essential to viral replication (Dougherty and Semler, 1993). This makes viral-encoded proteinases potentially good targets for antiviral drugs. In the murine coronavirus, two such proteinases are under continued investigation in order to better understand the manner in which these proteinases function. Papain-like proteinase 1 (PLP-1)

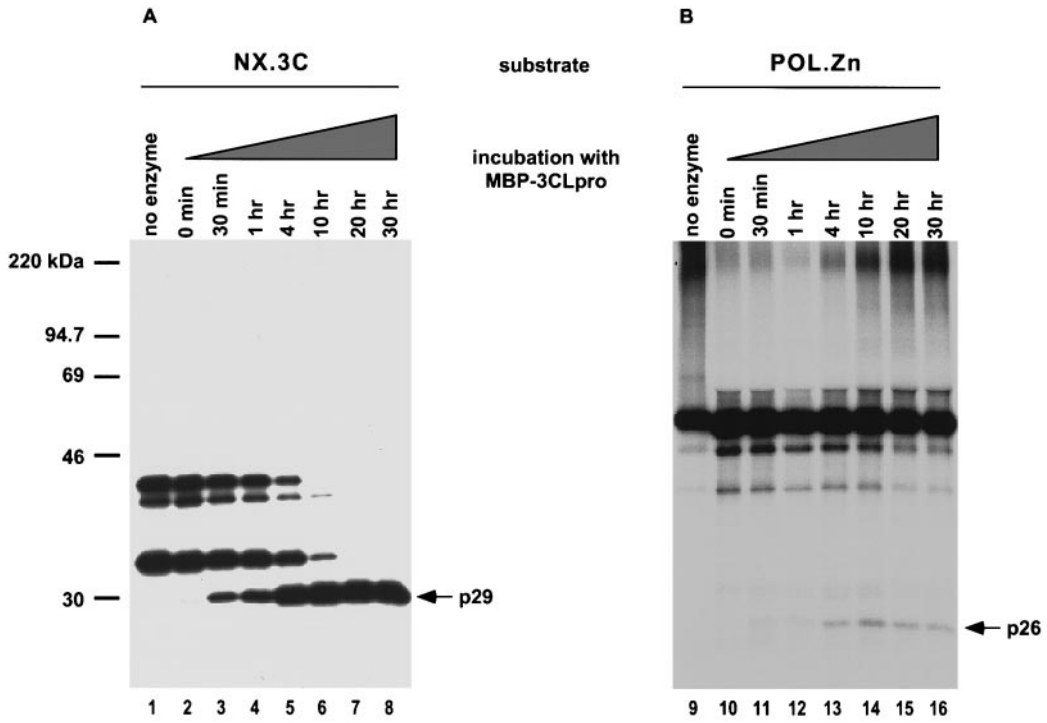


FIG. 7. Cleavage efficiencies at the HD1↓3C and POL↓Zn sites. Substrate volumes containing equivalent counts per minute of either substrate NX.3C (A) or POL.Zn (B) were incubated with MBP-3CLpro for increasing periods of time. The *trans* processing reaction was terminated at the time point specified above each lane by the addition of 2× SDS–sample buffer. The electrophoretic migration of p29 and p26 are indicated to the right of A and B, respectively. Molecular weight protein markers are shown on the left of A.

has been linked to the processing of several nonstructural proteins encoded in the 5′ end of the viral genome (Baker *et al.,* 1989; Bonilla *et al.,* 1995; Denison *et al.,* 1992; 1995; Hughes *et al.,* 1995). These cleavage products, p28 and p65, are not yet linked with any known viral function. However, those viral proteins with presumed functions in viral replication and viral RNA transcription, such as the RNA-dependent RNA polymerase (POL) and the zinc-finger/helicase (Zn-HEL) proteins, are believed to be processed by the 3C-like proteinase of the virus. The 3CLpro is predicted to cleave at, at least, 11 sites in pp1ab. Many of the cleavage sites located in the 1a region of pp1ab have been shown to be functional cleavage sites and processing by the 3CLpro at these sites has been demonstrated in the coronaviruses IBV (Liu *et al.,* 1997; Liu and Brown, 1995; Ng and Liu, 1998; Tibbles *et al.,* 1996), HCV-229E (Ziebuhr *et al.,* 1995; Ziebuhr and Siddell, 1999), and MHV-A59 (Lu *et al.,* 1995, 1998; Piñón *et al.,* 1997). Some of the mature viral products resulting from these cleavages have also been identified in infected cells (Liu *et al.,* 1997; Lu *et al.,* 1998; Ng and Liu, 1998; Piñón *et al.,* 1997; Ziebuhr and Siddell, 1999; Denison *et al.,* 1999). According to computer predictions, further processing at the Q939↓S940, Q1539↓C1540, Q2060↓S2061, and Q2434↓A2435 sites located in the 1b region of MHV-A59 pp1ab would result in mature viral products of 106, 67, 59, 42, and 33 kDa, respectively, corresponding to POL, Zn-HEL, and the three C-terminal-most cleavage products. In HCV-229E, viral products of 105, 71, and 41 kDa, corresponding to the POL, Zn-HEL, and the second C-terminal-most proteins, have been identified in infected cells and the role of the 3CLpro in the processing of these products has been authenticated *in vitro* (Grötzinger *et al.,* 1996; Heusipp *et al.,* 1997a,b). Similarly, in the case of IBV, viral proteins of 100, 39, and 35 kDa have been identified in infected cells (corresponding to POL and the two C-terminal-most proteins) and cotransfection experiments have implicated the 3CLpro in the processing of these viral products (Liu *et al.,* 1994, 1998). For MHV-A59, however, demonstration of processing at any MHV cleavage site in the 1b region of pp1ab has lagged behind that of HCV and IBV. Here we report the first demonstration of processing at the site between POL and Zn in pp1ab by the MHV-A59 3CLpro.

The cleavage at the POL↓Zn site is highly inefficient compared to the processing observed at the HD1↓3C site, as evidenced by the time course assays illustrated in Fig. 7. Phosphorimager analyses indicate that the POL.Zn substrate is cleaved fivefold less efficiently than the NX.3C substrate (data not shown). The differences in these efficiencies, however, could not be explained by the subtle differences in the primary sequences of the cleavage sites alone. In fact, when presented to the enzyme in the context of a 14- or 15-mer peptide substrate, there was no substantial difference in cleavage efficiency between these two peptides, an observation

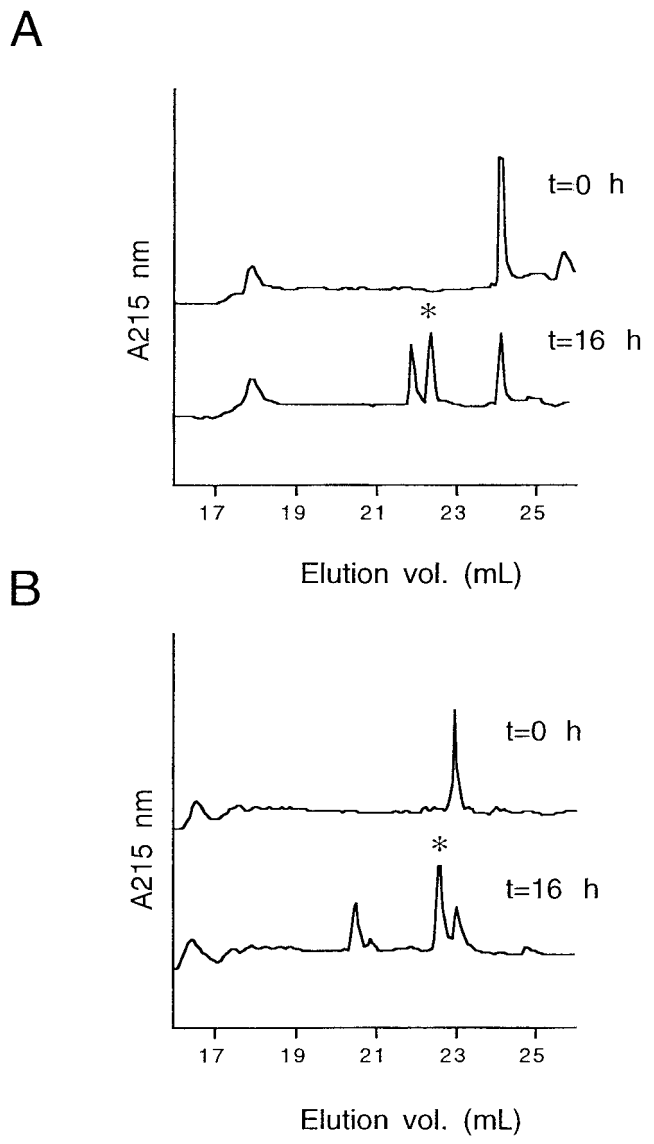


FIG. 8. Chromatograms of cleavage products of the synthetic peptides corresponding to ORF1a amino acids and ORF1b amino acids. A and B show chromatograms of ORF1a peptide (1 mM) and ORF1b peptide (1 mM), respectively, at 0 h and after 16 h of incubation with MBP-3CLpro (17.5 μM enzyme) at 30°C. In both panels the chromatograms for $t = 0$ h are offset to allow comparison between the two time points. The asterisks (*) indicate the cleavage products that were used in peptide sequencing in order to confirm the sites of cleavage.

that would seem to contradict the results obtained with the larger polyprotein substrates. Recently, Ziebuhr and Siddell (1999) investigated the efficiency of processing at several HCV-229E 3CLpro sites located at the C-terminus of pp1a or the central region of pp1ab. They observed that several viral products were produced less efficiently than others, with reduced cleavage activity at two ORF 1a sites, between V-Q3546↓S3547 and L-Q3824↓N3825, compared with cleavage at the sites flanking the 3Clpro domain. The corresponding LQ↓N site in MHV-A59 has also been shown to be less efficiently cleaved than the LQ↓S sites flanking the 3CLpro domain (Lu *et al.,* 1998). In the case of L-Q3824↓N3825 in HCV-229E, additional peptide cleavage data demonstrated that the properties of the cleavage sequence itself, rather than the overall conformation of the polypeptide and the accessibility of the cleavage site, contribute to the observed inefficiency of processing at this site (Ziebuhr and Siddell, 1999). While our results do not contradict theirs, they do demonstrate that, *in vivo,* the conformation of a larger polypeptide substrate is likely to be as important a determinant of cleavage as is the primary structure and sequence of the cleavage site. Our results clearly show that, at least with the case of the POL↓Zn site of MHV-A59, the observed inefficiency of cleavage was likely due to the overall conformation of the polypeptide, which may directly translate into the accessibility of the cleavage site, rather than the primary sequence of the site. Taken together, both sets of results highlight important regulatory mechanisms employed by the virus to coordinate the temporal production and the accumulation of the various replicase proteins.

Thus, the very slow *in vitro* processing at the ORF 1b cleavage site, compared with that at the ORF 1a site, generally correlates with the levels of ORF 1a and ORF 1b polypeptides found in infected cells. However, there are inherent differences between *in vitro* cleavage reactions with either recombinant proteins or peptide substrates and *in vivo* processing; these include the lengths of the substrates, the concentration of the enzyme and substrates, as well as the subcellular localization of replication complexes *in vivo* (Denison *et al.,* 1999; Ziebuhr and Siddell, 1999). These differences may all contribute to the extended lengths of time necessary for *in vitro* cleavages.

The substrate specificity of the coronavirus 3CLpro has been determined mainly through the identification of functional cleavage sites and a visual inspection of these cleavage site sequences. Mutagenesis has been done primarily to verify the authenticity of the cleavage site and most mutagenesis studies have not extended beyond the Q residue that is absolutely conserved at the P1 position in all 3CLpro cleavage sites identified to date. We have extended our mutagenesis studies to cover the P3 to P3′ positions of the cleavage site sequence. Our results demonstrated that the substrate sequence specificity of the 3CLpro is primarily influenced by the amino acid residues present at the P2 and P1 positions of the cleavage site. Substitution of the Q residue at the P1 position with any other amino acid has resulted in substrates that could not be cleaved by the 3CLpro, demonstrating the importance of this residue in the substrate. We could not detect any processing in substrates containing mutations at this position (Figs. 4–6). The P2 position of the cleavage site is most often occupied by an L; however, in some cleavage sites identified in HCV-229E, this position is occupied by a V or an I (Grötzinger

et al., 1996; Ziebuhr and Siddell, 1999), suggesting that this position is not as strictly conserved as the P1 position and may thus be able to tolerate some mutations. However, our results demonstrated otherwise, in that even a conservative change to I resulted in a reduction of cleavage (Figs. 4–6), to below 20%, as measured by phosphorimager analysis (data not shown). This suggests that perhaps an I at the P2 position of the cleavage site can be functional only when compensated for elsewhere in the substrate sequence. The only change tolerated at this position was a change to M in the P2 position at the POL↓Zn site (Fig. 6). Interestingly, the JHM strain of MHV encodes an M instead of an L at this position (Lee *et al.,* 1991). Thus such a mutation resulted in a wildtype JHM POL↓Zn site and a cleavage efficiency equivalent to wildtype levels. This provides indirect evidence that the POL↓Zn site in JHM is a functional 3CLpro site. Although the S found at the P1′ position of both the HD1↓3C and POL↓Zn sites is not as sensitive to mutations, we observed that some mutations, such as S3334C (in the HD1↓3C site) (Fig. 4) and S940N (in the POL↓Zn site) (Fig. 6), are not tolerated. Additionally, both S3334A and S940A substitutions, though tolerated, resulted in a reduction, but not complete inhibition of cleavage (Figs. 4–6), suggesting that the residue at the P1′ position also plays a role in substrate recognition, although to a lesser extent than those at the P1 and P2 positions.

Aside from the amino acid residues at the P2, P1, and P1′ positions, the amino acid sequences surrounding the scissile Q↓S(A,G) peptide bonds that are recognized by the 3CLpro do not share any other significant primary structure similarity. Furthermore, the presence of an LQA tripeptide not cleaved by the proteinase (Lu *et al.,* 1998) would again suggest the existence of a common conformational determinant shared by all 3CLpro susbstrates that is necessary for 3CLpro-mediated processing.

The POL↓Zn site is the first functional cleavage site identified in the ORF1b region of the MHV-A59 pp1ab. Further work remains to be done in order to identify other functional cleavage sites in ORF1b. Additionally, the mature viral products resulting from these processing events remain to be identified in infected cells. A direct comparison of the cleavage efficiencies of each site will help to elucidate the complex posttranslational processing pattern of the MHV-A59 gene 1 polyprotein, as well as provide insight into the regulatory mechanisms employed by the virus to maintain the production of its proteins under control.

MATERIALS AND METHODS

Plasmids

The parental plasmids used in this study are illustrated in Fig. 2. The plasmids pET21-NX.3C, pET21-HA.HD1.3C, and pMAL-3C.wt have all been described elsewhere (Piñón *et al.,* 1997). pET21-NX.3C encompasses MHV-A59 nucleotides 9912–10661, encoding the last 98 amino acids of HD1 and the entire 303-amino-acid region encompassing the 3CLpro (from K3235 to Q3636). pET21-NX.3C C3478A is the same as pET21-NX.3C but carries the inactivating C3478A mutation in the catalytic cysteine residues of the proteinase. The plasmid pET21-HA.HD1.3C encodes MHV-A59 nucleotides 9654–10661, encoding HD1 and the 3CLpro from S3149 to Q3636, directly behind the influenza hemagglutinin (HA) epitope under the control of the T7 promoter in pET21a. The plasmid pET21-HA.HD1.3C C3478A contains the 3CLpro inactivating mutation C3478A in the background of the parental pET21-HA.HD1.3C plasmid. The plasmid pMAL-3C.wt encodes MHV-A59 nucleotides 10209–10661, corresponding to the 3CLpro region from S3334 to Q3636, behind the *mal E* gene in the pMALc2 vector (New England Biolabs). This plasmid encodes the 3CLpro domain fused to the coding sequence of the maltose binding protein and is used for overexpression of the recombinant MBP-3CLpro enzyme.

A region of MHV-A59 gene 1, from nucleotides 15677 to 17140, corresponding to ORF1b amino acids L714 to Q1201, was PCR amplified from a plasmid encoding the entire ORF1b sequence using the primers F1bP 2383–2394 (5′-TTCGAATTCCCCGGGGGATCCCTTATGGCATGCAATGGACAC-3′) and R1bP 3846–3835 (5′-CGAATTCCTCTAGAAAGCTTGCTGAAACGTCTCAGGCACACT-3′). The resulting PCR fragment was digested with *Bam*HI and *Hin*dIII (denoted by the underlined sequences in the primers) and cloned into the corresponding sites of pET21a, resulting in the plasmid pET21-POL.Zn.

Site-directed mutagenesis

The pET21-HA.HD1.3C cleavage site mutants L3332I, L3332S, Q3333K, Q3333R, and S3334A were created by two rounds of PCR as described previously (Bonilla *et al.,* 1995; Hughes *et al.,* 1995) using the FMP and RMP primers listed in Table 1 and the primers FIJ31 (5′-TGGCTTGTCATGTATGGTGC-3′) and RSP 10661–10644 (5′-AACATATCCTACAGAACC-3′). The resulting mutant fragments were digested with *Kpn*I and *Bam*HI and cloned into the same sites in pET21-HA.HD1.3C. All other cleavage site mutants in pET21-HA.HD1.3C were created using the QuikChange mutagenesis kit (Stratagene) following the manufacturer's protocols. The mutagenic primers used are listed in Table 1. Following PCR amplification, the amplified plasmids were digested with the restriction enzyme *Dpn*I, which digests methylated and hemimethylated DNA, thus destroying the parental plasmid and any hybrids containing one parental strand and one mutated strand. *Escherichia coli* XL1-Blue supercompetent cells (Stratagene) were then transformed with the mutated plasmids. The presence of the desired mutation was verified by sequencing. The fragments containing

the HD1.3C cleavage site mutations were also subcloned into pET21-HA.HD1.3C C3478A using the *Nde*I and *Bam*HI sites in order to create plasmids carrying both the inactivating C3478A mutation and mutations at the cleavage site. These plasmids were used to express substrates used in *trans* cleavage assays.

All pET21-POL.Zn cleavage site mutants V937S, L938I, L938M, Q939K, Q939R, S940A, and S940N were created using the QuikChange mutagenesis kit as described above. The primers used in creating these mutants are also listed in Table 1.

In vitro transcription and translation

Cell-free expression of plasmid DNAs was carried out using the TnT rabbit reticulocyte lysate-coupled transcription–translation system (Promega) at 30°C for 2 h, as previously described (Piñón *et al.,* 1997). The incorporation of [^{35}S]methionine into acid precipitable counts was used as an indicator of protein synthesis. Equivalent amounts of acid precipitable counts were directly analyzed by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) or used in posttranslational proteolytic assays as indicated. Radioimmunoprecipitations were carried out as described previously (Bonilla *et al.,* 1995; Denison *et al.,* 1991).

Overexpression and partial purification of the recombinant 3CLpro

Expression of the recombinant 3CLpro from pMAL-3C.wt (NEB), which expresses 3CLpro as an MBP-3CL fusion protein, and the purification of the fusion protein was carried out according to the manufacturer's protocol and as described by Herold *et al.* (1996). Briefly, *E. coli* TB1 cells transformed with pMAL-3C.wt were grown at 37°C in the presence of ampicillin (100 μg/ml) until the A_{600nm} reached 0.8, at which point the cells were induced with isopropylthio-β-D-galactoside at a final concentration of 0.5 mM for 4 h at 25°. Cells were harvested and then resuspended in 10 ml of column buffer [20 mM Tris–Cl (pH 7.5)], 200 mM NaCl, 1 mM EDTA, 1 mM DTT] per gram of cells. The cell suspension was then lysed by sonication. Cell debris was pelleted by centrifugation at 9000*g* for 30 min. The crude lysates were diluted 1:5 in column buffer and then loaded onto an amylose column (bed volume 6 ml), preequilibrated with column buffer, at a flow rate of 1 ml/min. The column was then washed with 12 column volumes of column buffer and the MBP-3CLpro fusion protein was eluted with column buffer containing 10 mM maltose. Fractions of 1 ml were collected and those containing the 72-kDa recombinant MBP-3CLpro were identified by analyzing 10-μl aliquots by SDS–PAGE. Fractions containing the recombinant proteinase were pooled and the concentration of the fusion proteinase was determined using the Bradford assay against known concentrations of bovine serum albumin. The recombinant proteinase was stored at −80°C in column buffer supplemented with 20% glycerol.

Posttranslational proteolytic assays

Radiolabeled substrates containing the cleavage sequences of HD1.3C or POL.Zn (Fig. 2) were generated using the TnT rabbit reticulocyte lysate system. Lysate volumes containing equivalent counts per minute were incubated with approximately 1–10 μg of recombinant proteinase or an equivalent volume of column buffer/20% glycerol at 30°C for 12–16 h or, where applicable, the specified lengths of time. The processed products were analyzed by SDS–PAGE followed by autoradiography. Phosphorimager analysis was carried out as previously described (Hughes *et al.,* 1995; Teng *et al.,* 1999).

Peptide assays

Recombinant MBP-3CLpro enzyme (1.27 mg/ml in elution buffer supplemented with 20% glycerol and 2 mM DTT, 17.5 μM enzyme) was incubated with synthetic peptide ORF1a (0.02 to 0.6 mM in 10% DMSO), with the sequence H_2N-Thr–Thr–Ser–Phe–Leu-Gln↓Ser–Gly–Ile–Val–Lys–Met–Val–Ser-COOH, corresponding to ORF1a amino acids 3328 to 3341 (arrow indicates cleavage site) or ORF1b peptide (0.02 to 0.25 mM in 10% DMSO), with the sequence H_2N-Arg-Ser-Ala-Val-Leu-Gln↓Ser-Val-Gly-Ala-Cys-Val-Val-Cys-Ser-COOH, corresponding to ORF1b amino acids 934 to 948 (arrow indicates cleavage site) in a final reaction volume of 100 μl. The reactions were allowed to proceed at 30°C for 60–120 min, at which time the reactions were quenched by addition of trichloroacetic acid (TCA) to 1% final concentration. The samples were then chilled on ice, and the denatured protein was precipitated by centrifugation. For zero time point samples, the enzyme was mixed with TCA prior to the addition of peptide substrates and the reactions were carried out as described above.

Separation of cleavage product from substrate was carried out with the ÄKTApurifier System (Amersham Pharmacia Biotech) equipped with a Sephasil Peptide C18 5-μm ST 4.6/100 reverse-phase column (Amersham Pharmacia Biotech). Elution was performed with a linear gradient of 92% eluent A (0.1% aqueous TFA)/8% eluent B (0.1% TFA in 90% acetonitrile/10% water) up to 100% eluent B (Seybert *et al.,* 1997) over 26 min (flow rate 0.8 ml/min, detection at 215 nm). With the ORF1a peptide the level of cleavage was no more than 30%, whereas with the ORF1b peptide the level of cleavage was between 30 and 50%.

The data obtained were fitted to the Michaelis–Menten equation and the k_{cat} and K_m values were obtained using Kaleidagraph 3.08 (Synergy Software).

ACKNOWLEDGMENTS

The authors thank Ravi Mayreddy for the construction of the ORF1b-encoding plasmids. Peptide sequencing was provided by the Protein Chemistry Laboratory of the School of Medicine (University of Pennsylvania), supported by core grants of the Diabetes and Cancer Centers (DK-19525 and CA-16520). This work was supported by NIH Grant AI-17418.

REFERENCES

Baker, S. C., Shieh, C. K., Chang, M. F., Vannier, D. M., and Lai, M. M. C. (1989). Identification of a domain required for autoproteolytic cleavage of murine coronavirus gene A polyprotein. *J. Virol.* **63,** 3693–3699.

Bonilla, P. J., Gorbalenya, A. E., and Weiss, S. R. (1994). Mouse hepatitis virus strain A59 RNA polymerase gene ORF 1a: heterogeneity among MHV strains. *Virology* **198,** 736–740.

Bonilla, P. J., Hughes, S. A., Piñon, 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.

Boursnell, M. E., Brown, T. D., 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.

Bredenbeek, P. J., Pachuk, C. J., Noten, A. F., Charite, J., Luytjes, W., Weiss, S. R., and Spaan, W. J. (1990). The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59; a highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. *Nucleic Acids Res.* **18,** 1825–1832.

Brierley, I., Boursnell, M. E. G., Binns, M. M., Billimoria, 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.

Cavanagh, D. (1997). Nidovirales: A new order comprising *Coronaviridae* and *Arteriviridae. Arch. Virol.* **142,** 629–633.

Denison, M. R., Hughes, S. A., and Weiss, S. R. (1995). Identification and characterization of a 65-kDa protein processed from the gene 1 polyprotein of the murine coronavirus MHV-A59. *Virology* **20,** 316–320.

Denison, M. R., Spaan, W. J. M., van der Meer, Y., Gibson, C. A., Sims, A. C., Prentice, E., and Lu, X. T. (1999). The putative helicase of the coronavirus mouse hepatitis virus is processed from the replicase gene polyprotein and localizes in complexes that are active in viral RNA synthesis. *J. Virol.* **73,** 6862–6871.

Denison, M. R., Zoltick, P. W., Hughes, S. A., Giangreco, B., Olson, A. L., Perlman, S., Leibowitz, J. L., and Weiss, S. R. (1992). Intracellular processing of the N-terminal ORF 1a proteins of the coronavirus MHV-A59 requires multiple proteolytic events. *Virology* **189,** 274–284.

Denison, M. R., Zoltick, P. W., Leibowitz, J. L., Pachuk, C. J., and Weiss, S. R. (1991). Identification of polypeptides encoded in open reading frame 1b of the putative polymerase gene of the murine coronavirus mouse hepatitis virus A59. *J. Virol.* **65,** 3076–3082.

de Vries, A. A. F., Horzinek, F. M. C., Rottier, P. J. M., and de Groot, R. J. (1997). The genome organization of the Nidovirales: Similarities and differences between arteri-, toro-, and coronaviruses. *Semin. Virol.* **8,** 33–47.

Dougherty, W. G., and Semler, B. L. (1993). Expression of virus-encoded proteinases: Functional and structural similarities with cellular enzymes. *Microbiol. Rev.* **57,** 781–822. [review]

Eleouet, J. F., Rasschaert, D., Lambert, P., Levy, L., Vende, P., and Laude, H. (1995). Complete sequence (20 kb) of the polyprotein-encoding gene 1 of transmissible gastroenteritis virus. *Virology* **206,** 817–822.

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.

Grötzinger, C., Heusipp, G., Ziebhur, 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., Rabbe, T., Schelle-Prinz, B., and Siddell, S. G. (1993). Nucleotide sequence of the human coronavirus 229E RNA polymerase locus. *Virology* **195,** 680–691.

Herold, J., and Siddell, S. G. (1993). An "elaborated" pseudoknot is required for high frequency frameshifting during translation of HCV 229E polymerase mRNA. *Nucleic Acids Res.* **21,** 5838–5842.

Herold, J., Siddell, S. G., and Ziebuhr, J. (1996). Characterization of coronavirus RNA polymerase gene products. *Methods Enzymol.* **275,** 68–89.

Heusipp, G., Grötzinger, C., Herold, J., Siddell, S. G., and Ziebuhr, J. (1997a). Identification and subcellular localization of a 41 kDa polyprotein 1ab processing product in human coronavirus 229E-infected cells. *J. Gen. Virol.* **78,** 2789–2794.

Heusipp, G., Harms, U., Siddell, S. G., and Ziebuhr, J. (1997b). 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.

Holmes, K. V., and Lai, M. M. C. (1996). Coronaviridae: The viruses and their replication. *In* "Fields Virology", (B. N. Fields, D. M. Knipe, and P. M. Howley, Eds.), 3rd ed., Vol. 2, pp. 1075–1093. Lippincott–Raven, Philadelphia.

Houtman, J. J., and Fleming, J. O. (1996). Pathogenesis of mouse hepatitis virus-induced demyelination. *J. Neurovirol.* **6,** 361–376.

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

Kim, J. C., Spence, R. A., Currier, P. F., Lu, X., and Denison, M. R. (1995). Coronavirus protein processing and RNA synthesis is inhibited by the cysteine proteinase inhibitor E64d. *Virology* **208,** 1–8.

Lai, M. M. C. (1990). Coronavirus: Organization, replication and expression of genome. *Annu. Rev. Microbiol.* **44,** 303–333. [review]

Lee, H. J., Shieh, C. K., Gorbalenya, A. E., Koonin, E. V., LaMonica, N., 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* **180,** 567–582.

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 coronavirus infectious bronchitis virus is processed by ORF1a products. *J. Virol.* **68,** 5772–5780.

Liu, D. X., and Brown, T. D. (1995). Characterization and mutational analysis of an ORF1a-encoding proteinase domain responsible for proteolytic processing of the infectious bronchitis virus 1a/1b polyprotein. *Virology* **209,** 420–427.

Liu, D. X., Shen, S., Xu, H. Y., and Wang, S. F. (1998) Proteolytic mapping of the coronavirus infectious bronchitis virus 1b polyprotein: Evidence for the presence of four cleavage sites of the 3C-like proteinase and identification of two novel cleavage products. *Virology* **246,** 288–297.

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.

Lu, X. T., Sims, A. C., and Denison, M. R. (1998). Mouse hepatitis virus 3C-like protease cleaves a 22-kilodalton protein from the open reading frame 1a polyprotein in virus-infected cells and *in vitro. J. Virol.* **72,** 2265–2271.

Ng, L. F. P., and Liu, D. X. (1998). Identification of a 24-kDa polypeptide processed from the coronavirus infectious bronchitis virus 1a polyprotein by the 3C-like proteinase and determination of its cleavage sites. *Virology* **243,** 388–395.

Piñón, J. D., Mayreddy, R. R., Turner, J. D., Khan, F. S., Bonilla, P. J., and Weiss, S. R. (1997). Efficient autoproteolytic processing of the MHV-

A59 3C-like proteinase from the flanking hydrophobic domains requires membranes. *Virology* **230,** 309–322.

Seybert, A., Ziebuhr, J., and Siddell, S. G. (1997). Expression and characterization of a recombinant murine coronavirus 3C-like proteinase. *J. Gen. Virol.* **78,** 71–75.

Sims, A. C., Lu, X. T., and Denison, M. R. (1998). Expression, purification, and activity of recombinant MHV-A59 3CLpro. *Adv. Exp. Med. Biol.* **440,** 129–134.

Teng, H., Piñón, J. D., and Weiss, S. R. (1999). Expression of murine coronavirus recombinant papain-like proteinase: Efficient cleavage is dependent on the lengths of both the substrate and the proteinase polypeptides. *J. Virol.* **73,** 2658–2666.

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

van Dinten, L. C., Sietske, R., Gorbalenya, A. E., and Snijder, E. J. (1999). Proteolytic processing of the open reading frame 1b-encoded part of the arterivirus replicase is mediated by nsp4 serine protease and is essential for virus replication. *J. Virol.* **73,** 2027–2037.

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

Ziebuhr, J., and Siddell, S. G. (1999). Processing of the human coronavirus 229E replicase polyproteins by the virus-encoded 3C-like proteinase: Identification of proteolytic products and cleavage sites common to pp1a and pp1ab. *J. Virol.* **73,** 177–185.