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 204, 541-549 (1994)

Determinants of the p28 Cleavage Site Recognized by the First Papain-like Cysteine Proteinase of Murine Coronavirus

SHANGHONG DONG* AND SUSAN C. BAKER*,†,[1]

*Molecular Biology Program and †Department of Microbiology and Immunology, Loyola University of Chicago, Stritch School of Medicine, Maywood, Illinois 60153

Received February 17, 1994; accepted July 14, 1994

The murine coronavirus polymerase gene is 22 kb in length with the potential to encode a polyprotein of approximately 750 kDa. The polyprotein has been proposed to encode three proteinase domains which are responsible for the processing of the polyprotein into mature proteins. The proteolytic activity of the first proteinase domain has been characterized and resembles the papain family of cysteine proteinases. This proteinase domain acts autoproteolytically to cleave the amino terminal portion of the polymerase polyprotein, releasing a 28-kDa protein designated p28. To identify the cleavage site of this papain-like cysteine proteinase, we isolated the peptide adjacent to p28 and determined the amino terminus sequence by Edman degradation reaction. We report that proteolysis occurs between the Gly-247 and Val-248 dipeptide bond. To determine the role of the amino acid residues surrounding the cleavage site, we introduced a total of 42 site-specific mutations at the residues spanning the P5 to P3′ positions and assessed the effects of the mutations on the processing of p28 in an *in vitro* transcription and translation system. The substitutions of Gly-247 at the P1 position or Arg-246 at the P2 position resulted in a dramatic decrease of proteolytic activity, and the mutations of Arg-243 at P5 position also led to considerable reduction in p28 cleavage. In contrast, the substitutions of amino acids Gly-244 (P4), Tyr-245 (P3), Val-248 (P1′), Lys-249 (P2′), and Pro-250 (P3′) had little or no effect on the amount of p28 that was released. This work has identified Gly-247–Val-248 as the cleavage site for the release of p28, the amino-terminal protein of the murine coronavirus polymerase polyprotein. Additionally, we conclude that the Gly-247 and Arg-246 are the major determinants for the cleavage site recognition by the first papain-like cysteine proteinase of murine coronavirus.

INTRODUCTION

The murine coronavirus mouse hepatitis virus (MHV) is an enveloped, plus-stranded RNA virus. The 31-kb RNA genome encodes eight genes that function in the replication and assembly of viral particles (Spaan *et al.*, 1988; Lai, 1990). Upon MHV infection, the viral RNA-dependent RNA polymerase is translated from the first gene which is located at the 5′ end of the RNA genome. This RNA polymerase mediates the unique discontinuous transcription of the coronavirus mRNAs (Lai *et al.*, 1983; Spaan *et al.*, 1983) and may play an important role in the high level of RNA recombination detected in MHV infection (Makino *et al.*, 1986). To better understand the RNA polymerase that functions in the complex MHV replication strategy, we must elucidate the mechanism of MHV RNA polymerase expression and proteolytic maturation.

The MHV RNA polymerase is encoded by the largest of the viral genes, gene 1. Gene 1 is 22 kb and has been completely cloned and sequenced (Pachuk *et al.*, 1989; Lee *et al.*, 1991; Bonilla *et al.*, 1994). Gene 1 contains two open reading frames designated ORF1a (13 kb) and ORF1b (9 kb). ORF1a and ORF1b overlap by 75 nt and can be translated into a single polyprotein of approximately 750 kDa via a ribosomal frameshifting mechanism (Brierly *et al.*, 1987, 1989; Bredenbeek *et al.*, 1990; Lee *et al.*, 1991). Computer analysis of the deduced amino acid sequences of the large polymerase polyprotein suggests that there are many functional domains, such as a polymerase motif, helicase domain, zinc-finger domains, transmembrane regions, cysteine-rich domains, and three proteinase domains (Gorbalenya *et al.*, 1989; Bredenbeek *et al.*, 1990; Lee *et al.*, 1991). The three proteinase domains, two papain-like cysteine proteinases (PCP-1 and PCP-2) and a polio virus 3C-like proteinase (3C-Pro), are proposed to process the polymerase polyprotein into mature protein products. Indeed, several small proteins ranging in size from 28 to 290 kDa have been detected from *in vitro* translation of MHV genomic RNA or in the infected cells using antiserum directed against viral peptides or fusion proteins (Denison and Perlman, 1986; Denison *et al.*, 1991, 1992). However, our knowledge of how these proteins are processed from the polymerase polyprotein is incomplete.

The first protein to be identified as a proteolytic processing product of the polymerase polyprotein was a 28-kDa protein, designated p28. Denison and Perlman demonstrated that *in vitro* translation of MHV genomic

[1] To whom correspondence and reprint requests should be addressed. Fax: (708) 216-9574.

0042-6822/94 $6.00

RNA produced two prominent proteins, p28 and p220 (Denison and Perlman, 1986). The appearance of p28 could be inhibited by the addition of the proteinase inhibitor $ZnCl_2$, suggesting that p28 was processed from a precursor polyprotein (Denison and Perlman, 1986). P28 is also detected in MHV-infected cells indicating that proteolytic processing is taking place *in vivo* (Denison and Perlman, 1987). Translation of the 5′-end sequences of MHV gene 1 demonstrated that p28 was indeed the amino-terminal portion of the polymerase polyprotein (Soe *et al.*, 1987). Furthermore, by translating RNAs which represented up to 5.3 kb of the 5′-end of the MHV genome, Baker *et al.* (1989) showed that the processing of p28 was an autoproteolytic event in which a downstream sequence was required for the cleavage of p28. This downstream domain was shown to be a papain-like cysteine proteinase (PCP-1) with Cys-1137 and His-1288 essential for the catalytic activity (Baker *et al.*, 1993). The function of PCP-1 appears to be to release p28 from the amino terminus of the polymerase polyprotein. Although PCP-1 activity has been demonstrated to occur in *cis* to cleave p28, it may also act in *trans* at additional sites in the polymerase polyprotein. MHV PCP-1 may be analogous to the papain-like cysteine proteinase of Sindbis virus, nsP2, which mediates both *cis* and *trans* cleavages of its nonstructural polyprotein (Hardy and Strauss, 1989; Shirako and Strauss, 1990; Strauss and Strauss, 1990). As a first step toward determining the cleavage specificity of MHV PCP-1, we have characterized the cleavage motif recognized by PCP-1 for the release of p28.

In this study, we have identified the cleavage site recognized by the MHV PCP-1 as Gly-247 (P1)–Val-248 (P1′) by amino-terminal sequencing of the protein adjacent to p28. We have performed site-directed mutagenesis analysis of the residues between the P5 and P3′ positions to determine the amino acid sequence requirement which defines the functional cleavage site. Our results showed that Gly-247 (P1) and Arg-246 (P2) are essential for p28 cleavage by PCP-1. This is the first report of the identification and characterization of a cleavage site in the MHV polymerase polyprotein.

MATERIALS AND METHODS

Plasmids and site-directed mutagenesis

Plasmid pT7-N27 contains MHV-JHM genomic cDNA sequence from nucleotide 187 to nucleotide 4434 with a 747-nt deletion of nucleotide 2064–2810 (see Fig. 1). We have demonstrated that the coding region of pS-N27 is translated to produce a 128-kDa protein which is autoproteolytically processed to two protein products p28 and a 100-kDa C-terminal protein (Baker *et al.*, 1993). For this study, the MHV-JHM coding region from pT7-N27 was removed by digestion with *Pvu*II and *Eco*RI and ligated into pSELECT-1 (Promega Biotech) that had been digested with *Sma*I and *Eco*RI. The resulting plasmid was designated pS-N27. Site-directed mutagenesis of pS-N27 was performed using the method of Hutchinson *et al.* (1978). Briefly, single-strand pS-N27 DNA was isolated from *Escherichia coli* strain JM109 after infection with helper phage R408 (Vieira and Messing, 1987). The single strand DNA served as the template for site-specific mutagenesis using synthetic oligonucleotides listed in Table 1 in conjunction with an ampicillin repair oligonucleotide (Lewis and Thompson, 1990). These oligonucleotides contained degenerate sequences at positions which would result in specific amino acid substitutions at the putative cleavage site and its flanking sequences (according to the nomenclature of Schechter and Berger, 1967) including residues: Arg-243 (P5), Gly-244 (P4), Tyr-245 (P3), Arg-246 (P2), Gly-247 (P1), Val-248 (P1′), Lys-249 (P2′), and Pro-250 (P3′). Additional mutagenesis was performed at the previously suggested cleavage sites Tyr-257 and Gly-258 (Soe *et al.*, 1987). The mutagenic oligonucleotide and ampicillin repair oligonucleotide were linked in the same plasmid during DNA synthesis, and these mutated plasmids were amplified in a repair deficient *E. coli* strain (BHM71-18mutS). Plasmid DNA isolated from BHM71-18mutS cells was transformed into *E. coli* strain JM109. Mutant plasmids were isolated from individual ampicillin resistant colonies and sequenced by double-stranded DNA sequencing (Sequenase, USB) using a forward primer B67 (5′-GGTAACAAAGG-GTCTGTG-3′, nucleotide 799 to 817 of MHV-JHM). Specific mutations found are listed in Table 1.

In vitro transcription–translation and immunoprecipitation

Plasmid DNA pS-N27 was linearized by *Eco*RI digestion and transcribed *in vitro* with T7 RNA polymerase and translated in rabbit reticulocyte lysates in a coupled transcription/translation system (TNT lysates, Promega Biotech) as previously described (Baker *et al.*, 1993). Standard reactions were performed in a total volume of 25 μl using 0.5 μg of linearized pS-N27 (or mutant pS-N27) in the presence of [^{35}S]methionine (20 μCi/translation reaction; Amersham) for 90 min at 30°. [^{35}S]Met incorporation in the translation products was quantitated by trichloroacetic acid (TCA) precipitation (Maniatis *et al.*, 1982). Equal TCA-precipitable counts of the primary protein products (typically 2×10^5 cpm) were immunoprecipitated using 3 μl of p28-specific antiserum as previously described (Baker *et al.*, 1993). The immunoprecipitated protein products or primary translation products were separated by electrophoresis on a 10% polyacrylamide gel containing 0.1% SDS (SDS–PAGE) (Laemmli, 1970). Following electrophoresis, the gel was fixed and enhanced with Entensify A and B (NEN Research Products) for 45 min, respectively, before being dried and exposed to Kodak X-ray film at −70°.

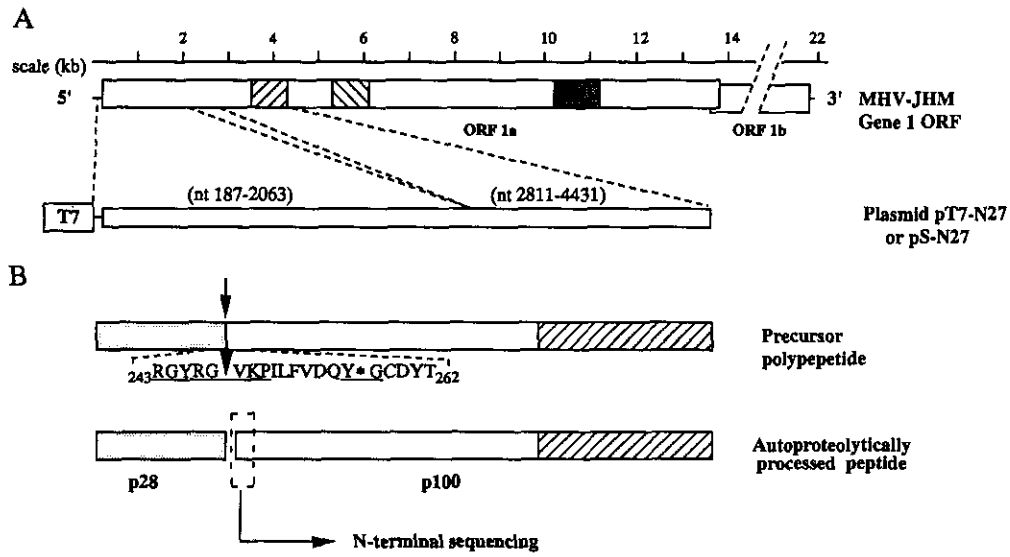

FIG. 1. (A) Schematic diagram of MHV-JHM gene 1 open reading frame and the structure of cDNA clone in the plasmid pT7-N27 and pS-N27. The two overlapping reading frames (ORF) are shown in the boxes with only ORF1a drawn to scale. The three proposed proteinase domains (Lee *et al.*, 1991) are indicated. The first papain-like cysteine proteinase domain (hatched box, between 3.5 and 4.3 kb) has been characterized (Baker *et al.*, 1993). Plasmid pT7-N27 (Baker *et al.*, 1993) contains MHV-JHM gene 1 sequence from nucleotide 187 (*Nar*I site) to 2063 (*Nsi*I site), and from nucleotide 2811 (*Kpn*I site) to 4431 (*Acc*I site). (B) Proteolytic processing of the polypeptide translated from RNA derived from plasmid pT7-N27. The precursor polypeptide (128 kDa) synthesized from linearized pT7-N27 in the *in vitro* T7 RNA polymerase transcription and rabbit reticulocyte lysate translation system is autoproteolytically processed to two products: N-terminal p28 and C-terminal peptide (100 kDa) (Baker *et al.*, 1993). The p28 region is indicated by the stippled box and the papain-like cysteine proteinase domain by the hatched box. The arrow represents the cleavage site of p28. The radiolabeled C-terminal 100-kDa peptide was isolated for N-terminal microsequencing to determine the cleavage site. Partial amino acid sequence of the potential cleavage site region is shown, in which the cleavage site predicted from microsequencing analysis is indicated by a black triangle and the previously proposed cleavage site (Soe *et al.*, 1987) is marked by an asterisk (*). The amino acids mutated in this study are underlined.

Protein microsequencing

To obtain significant amounts of translation products for protein sequencing, large scale *in vitro* translation reactions were performed. *In vitro* translation reactions were performed in a total volume of 100 μl using 4.0 μg of pT7-N27 DNA linearized with *Eco*RI, and in the presence of 80 μCi of [^{35}S]Met (Amersham), 40 μCi [^{3}H]Leu (IC Biomed.), or 40 μCi [^{3}H]Val (ICN Biomed.) for 90 min at 30°. The *in vitro* translation products were mixed with equal volume of 2× Laemmli sample buffer and separated by electrophoresis on a 7.5% SDS–PAGE gel for 24 hr at 50 V. Following electrophoresis, the translation products were electrotransferred to Immobilon polyvinylidene difluoride (PVDF) membrane (Millipore) at 40 V at 4° for at least 8 hr in buffer containing 25 m*M* Tris–base, 192 m*M* glycine, 20% (v/v) methanol, and 0.01% (w/v) SDS. The PVDF membrane was air-dried and exposed to Kodak X-ray film at −70°. The 100-kDa band of [^{35}S]Met-labeled proteins was identified by autoradiography and the corresponding proteins labeled with [^{35}S]Met, [^{3}H]Leu, or [^{3}H]Val were excised from the membrane. The excised protein samples were subjected to an Applied Biosystem gas-phase sequencer for partial N-terminal amino acid sequence analysis based on Edman degradation reaction (Matsudaira, 1987). Radioactivity released from each cycle was quantitated by scintillation counting for 10 min in a Beckman scintillation counter.

RESULTS

Determining the amino-terminal amino acid sequence of the protein adjacent to p28

To identify the cleavage site recognized by PCP-1 of MHV, we have isolated the protein adjacent to the cleavage site and microsequenced the amino-terminal portion of the protein. The protein used in these studies was generated by translation of the open reading frame of pT7-N27, which is the smallest in-frame deletion construct that maintains active proteolytic processing of p28 (Fig. 1). We have previously shown that *in vitro* translation of pT7-N27 generates a precursor polyprotein of 128 kDa which is autoproteolytically cleaved by the proteinase domain (PCP-1) to yield p28 and a 100-kDa protein (Baker *et al.*, 1993). We found that pT7-N27 served as an excellent template for efficient translation of large quantities of protein products. In addition, the 128-kDa precursor and 100-kDa processed products were easily resolved on a 7.5% SDS–polyacrylamide gel, facilitating the isolation of the 100-kDa protein. *In vitro* translation products of pT7-N27,

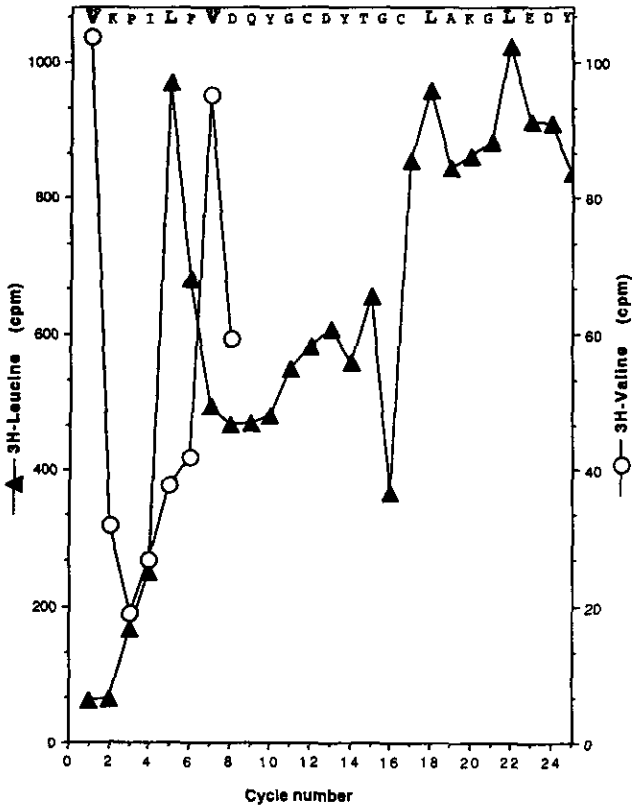

FIG. 2. N-terminal sequence analysis of C-terminal 100-kDa peptide derived from *in vitro* transcription/translation of plasmid pT7-N27. The protein products were labeled with [^{35}S]methionine, [^{3}H]Leu, or [^{3}H]Val, separated on the 7.5% SDS–PAGE gel, and subsequently electrotransferred to the Immobilon polyvinylidene difluoride (PVDF) membrane as described under Materials and Methods. The radiolabeled samples on the PVDF membrane were localized by autoradiography and the 100-kDa band was excised and subjected to microsequencing analysis based on Edman-degradation reaction. Radioactivity released at each cycle was quantitated using a Beckman scintillation counter and counts per minute (cpm) was plotted. The amino acid sequence in the single-letter code shown at the top of the graph corresponds to amino acid 248 to 272.

labeled with [^{3}H]Leu, [^{3}H]Val, or [^{35}S]methionine, were synthesized, separated by electrophoresis on a polyacrylamide gel, and transferred to PVDF membrane. The 100-kDa protein was identified and subjected to microsequencing as described under Materials and Methods. The sequencing profile generated from samples labeled with [^{3}H]Leu indicated that Leu occupied amino acid positions 5, 18, and 22 from the cleavage site (Fig. 2). Microsequencing of the 100-kDa protein labeled with [^{3}H]Val revealed Val in positions 1 and 7. Microsequencing of the 100-kDa protein labeled with [^{35}S]methionine indicated that there were no Met in the first 25 positions from the cleavage site (data not shown). By aligning this profile with the amino acid sequence deduced from the MHV-JHM nucleotide sequence (Soe *et al.*, 1987; Lee *et al.*, 1991), we identified the cleavage site for the generation of p28 as the Gly-247 (P1)–Val-248 (P1′) dipeptide bond.

Effect of site-specific mutations surrounding the cleavage site on the processing of p28

To determine the amino acid sequence requirements for PCP-1 cleavage of p28, we performed site-specific mutagenesis at positions P5 to P3′ and analyzed the effect of amino acid substitutions on proteolytic processing in an *in vitro* transcription/translation system. Site-specific mutagenesis was performed by the method of Hutchinson and co-workers using single-strand pS-N27 DNA and oligonucleotides containing degenerate sequences at the specific nucleotides selected for mutation (see Table 1 and Materials and Methods). Individual mutants at each site were identified by DNA sequencing. Plasmid DNAs encoding amino acid substitutions of each position were linearized by digestion with *Eco*RI and transcribed and translated *in vitro*, and the resulting protein products were analyzed by electrophoresis on a 10% polyacrylamide gel. As shown in Fig. 3, substitution of Gly-247 at the P1 position by Ala, Asp, Asn, or Val resulted in a dramatic decrease of the yield of p28 generated by autoproteolytic processing of the mutant polyprotein. However, the mutation at the P1 position did not completely abolish proteinase activity as is seen in a mutant of the catalytic site of PCP-1 (Fig. 3, lane C). When the products of the *in vitro* translation reaction were immunoprecipitated with anti-sera to p28, we detected two p28-specific bands (Fig. 4, P1 position), indicating that a low level of aberrant processing may be taking place in the P1 mutants. Such aberrant autoprocessing of a polypeptide encoding mutations in the cleavage site has also been reported for the poliovirus 2A proteinase (Hellen *et al.*, 1992). This aberrant processing may be due to an alteration in the presentation of the cleavage site to the proteinase domain.

The translation products generated from *in vitro* transcription/translation of plasmid DNAs encoding mutants at positions P5 to P3′ were analyzed directly (data not shown) and after immunoprecipitation with anti-p28 sera (Fig. 4). Autoradiography of the primary *in vitro*-translated products revealed that the total precursor and processed products bands were essentially equivalent for all the samples shown in Fig. 4. Mutation of the P2 Arg-246 position resulted in the most dramatic reduction in cleavage of p28. Clearly, both the P1 and P2 positions are critical for efficient cleavage of p28. In contrast, the surrounding residues at P3, P4, P1′, P2′, and P3′ positions had little or no effect on cleavage efficiency. Polyproteins encoding a mutation at these positions generally maintained equivalent cleavage of p28 as compared to the wild-type polyprotein. Mutation of the P5 position, Arg-243, did have some impact on the cleavage of p28. Mutation of Arg to Ala, Ile, or Thr resulted in a significant reduction in p28. Substitution of Arg with Lys, another positively charged amino acid, also reduced the cleavage of p28. This result is somewhat surprising because Lys

TABLE 1

OLIGONUCLEOTIDES USED FOR SITE-DIRECTED MUTAGENESIS OF MHV-JHM p28 CLEAVAGE SITE AND FLANKING AMINO ACID SEQUENCES

Amino acid (position)	Substitution (single-letter code)	Mutant codon	Mutagenic oligonucleotide (5′ to 3′)[a]
Gly-247 (P1)	Ala (A)	GCT	AGGGGCTATCGCATTGTTAAG
	Asp (D)	GAT	GA
	Val (V)	GTT	C
	Asn (N)	AAT	
Arg-246 (P2)	Pro (P)	CCC	CTTAGGGGCTATCCCGGTGTTAAG
	Leu (L)	CTC	TT
	His (H)	CAC	A
	Ser (S)	TCC	CTTAGGGGCTATTCCGGTGTTAAGC
	Tyr (Y)	TAC	T
			A
Tyr-245 (P3)	Arg (R)	CGT	CTTAGGGGCCCTCGCGGTGTTT
	Ser (S)	TCT	TT
	Phe (F)	TTT	G
	Cys (C)	TGT	
	Pro (P)	CCT	
Gly-244 (P4)	Ala (A)	GCC	GCTCTTCTTAGGACCTATCGCGGTGT
	Val (V)	GTC	GT
	Asp (D)	GAC	A
	Tyr (Y)	TAC	GCTCTTCTTAGGTCCTATCGCGGTGT
	Phe (F)	TTC	CT
	Ser (S)	TCC	A
	Leu (L)	CTC	
	His (H)	CAC	
Arg-243 (P5)	Lys (K)	AAA	GCTCTTCTTAAAGGCTATCGCGGA
	Ile (I)	ATA	GT
	Thr (T)	ACA	C
	Ala (A)	GCA	
Val-248 (P1′)	Ala (A)	GCT	GGCTATCGCGGTAATAAGCCC
	Asp (D)	GAT	GC
	Gly (G)	GGT	G
	Ser (S)	AGT	
	Thr (T)	ACT	
	His (H)	CAT	GGCTATCGCGGTCATAAGCCC
	Arg (R)	CGT	TG
	Leu (L)	CTT	T
	Phe (F)	TTT	
Lys-249 (P2′)	Ala (A)	GCG	TCGCGGTGTTGGGCCCATCCTG
	Arg (R)	AGG	AT
	Met (M)	ATG	C
	Thr (T)	ACG	
Pro-250 (P3′)	Arg (R)	CGC	TCGCGGTGTTAAGTGCATCCTGTT
	His (H)	CAC	CA
	Leu (L)	CTC	T
Tyr-257[b]	Leu (L)	CTT	GTGGACCAGCTTGGTTGCGAC
	Cys (C)	TGT	TG
	Arg (R)	CGT	G
	Phe (F)	TTT	
Gly-258[b]	Ala (A)	GCT	GGACCAGTATGATTGCGACTATAC
	Asp (D)	GAT	TC
			C

[a] The complimentary sequences of mutagenic oligonucleotides are shown with degenerate sequences in bold.
[b] Tyr-257 and Gly-258 dipeptide bond was previously predicted to be the p28 cleavage site (Soe *et al.*, 1987).

is the wild-type amino acid at this position in the A59 strain of MHV. The above results indicate that the P5 position may play some role in maintaining the appropriate conformation of the cleavage site.

Previous studies on the proteolytic processing of p28 suggested that the cleavage site for p28 generation may be between the Tyr-257 and Gly-258 dipeptide bond (Soe *et al.*, 1987). This suggestion was based primarily on a search of the N-terminal region of polymerase polyprotein for homology to a limited number of known viral

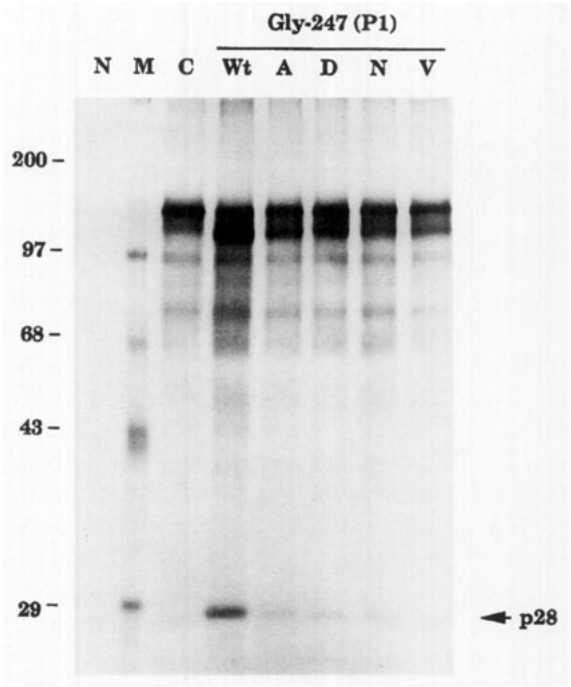

FIG. 3. *In vitro* translation protein products of wild type and mutants at P1 (Gly-247) position. Site-specific mutations are introduced by degenerate oligonucleotide mutagenesis as described under Materials and Methods. Linearized plasmids were translated in the T7 RNA polymerase coupled rabbit reticulocyte lysates system in the presence of [^{35}S]methionine and translation products were analyzed by 10% SDS–PAGE. Specific mutations at the position 247 are indicated at the top of the corresponding lanes, with additional lanes: M, molecular weight marker; N, no RNA; Wt, wild-type (Gly-247) pS-N27 translation products; C, control polypeptide with inactive PCP-1 in which the catalytic residue of the proteinase, cysteine-1137, was mutated to Ser (Baker *et al.*, 1993).

proteinase cleavage sites. Certain Tyr–Gly (or Y–G) sites (among others) are recognized by enterovirus 2A proteinases (Palmenberg, 1990). Substitution of Tyr-257 to either Ala or Asp and of Gly-258 to Ala, Asp, Phe, or Leu had no effect on p28 cleavage nor any effect on the mobility of the p28 protein (data not shown). These results indicated that Tyr-257 and Gly-258 are not essential for proteolytic processing of p28. In addition, the substitutions of residues at positions 257 and 258 do not appear to have any effect on the presentation of the nearby authentic cleavage site.

Our results demonstrate that p28 processing occurs between the Gly-247 and Val-248 dipeptide bond. Furthermore, Gly-247 at the P1 position and Arg-248 at the P2 position are the major determinants for efficient cleavage of p28 by MHV PCP-1.

DISCUSSION

Proteolytic processing is an essential event in the replication cycle of positive-stranded RNA viruses. Viral polyprotein precursors are processed into functional subunits by viral and/or host cell proteinases (Krausslich and Wimmer, 1988; Dougherty and Semler, 1993). During replication of MHV, three viral-encoded proteinase domains, PCP-1, PCP-2, and 3C-Pro, are proposed to be responsible for separating the 750-kDa polymerase polyprotein into functional protein products. To date, the activity of the first papain-like cysteine proteinase domain, PCP-1, has been demonstrated (Baker *et al.*, 1989, 1993). PCP-1 acts to cleave the amino-terminal portion of the MHV polymerase polyprotein, releasing p28 (Baker *et al.*, 1989, 1993). In this study, we have identified the cleavage site for p28 as the Gly-247 (P1) and Val-248 (P1′) dipeptide bond by N-terminal amino acid sequencing analysis of the protein product immediately downstream of p28. Site-specific mutagenesis of eight residues surrounding the cleavage site from the P5 to P3′ positions revealed that Gly-247 at the P1 position and Arg-246 at the P2 position are the major determinants of the p28 cleavage site recognized by MHV PCP-1. Additionally, surrounding residues such as Arg-243 (P5), may also play an important role in maintaining the correct conformation for the cleavage site interacting with the functional domain of PCP-1.

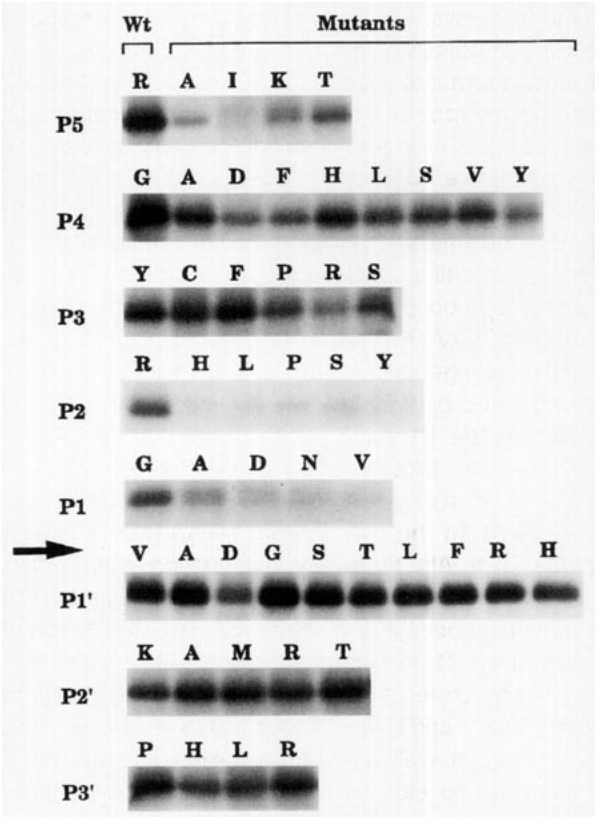

FIG. 4. Effect of mutation at residues between the P5 and P3′ positions on the cleavage of p28. Site-specific mutagenesis on the positions from P5 to P3′ was performed as described under Materials and Methods. Translation products were generated by *in vitro* transcription/translation of various pS-N27 plasmid DNAs. Equivalent cpm of each translation reaction were immunoprecipitated by p28 antiserum. Immunoprecipitation products were analyzed by 10% SDS–PAGE. The region corresponding to p28 is shown for each mutant. The amino acid at the site examined is shown above each lane.

Viral Proteinase		P5	P4	P3	P2	P1	P1'	P2'	P3'	P4'
MHV-JHM PCP1		R	G	Y	R	G	V	K	P	I
EAV nsP1		A	G	N	Y	G	G	Y	N	P
TEV HC-Pro		T	Y	N	V	G	G	M	N	R
HAV p29		L	A	R	I	G	G	R	L	N
p48		D	I	L	V	G	A	E	E	G
SV nsP2	(nsP1-nsP2)	A	D	I	G	A	A	L	V	E
	(nsP2-nsP3)	D	G	V	G	A	A	P	S	Y
	(nsP3-nsP4)	T	G	V	G	G	Y	I	F	S

FIG. 5. Comparison of viral papain-like cysteine proteinase cleavage sites. Viral proteinases are indicated, and protein junctions cleaved by Sindbis virus nsP2 proteinase are denoted. Amino acid sequences are shown in single letter codes and the arrow indicates the dipeptide bond being cleaved. MHV stands for mouse hepatitis virus, EAV for equine arteritis virus, TEV for tobacco etch virus, HAV for hypovirulence-associated virus, and SV for Sindbis virus.

The p28 cleavage site, the Gly–Val dipeptide bond, is similar to cleavage sites used by other viral papain-like cysteine proteinases (see Fig. 5). The cleavage site for the amino-terminal product of the equine arteritis virus (EAV) polymerase polyprotein has been defined (Snijder *et al.*, 1992). EAV is a member of the coronavirus superfamily and, like MHV, synthesizes leader-containing, 3′-coterminal, nested-set mRNAs (den Boon *et al.*, 1991). The EAV polymerase polyprotein is processed by a *cis*-acting papain-like cysteine proteinase (PCP) at a Gly–Gly dipeptide bond to release an amino-terminal protein of 30 kDa. EAV PCP and MHV PCP-1 appear to have similar cleavage requirements. Snijder and co-workers demonstrated by site-directed mutagenesis that the Gly residue at the P1 position was important for efficient autoproteolytic processing whereas the Gly residue at the P1′ position was more flexible to substitution (Snijder *et al.*, 1992). In the EAV system, when the P1 Gly was replaced with Ala, there was a reduction in the amount of p30 detected; when the P1 Gly was replaced with Val, little or no processing of p30 was detected. However, the Gly in the P1′ position was tolerant of mutation to Ala and Val and proteolytic processing of p30 was detected.

MHV PCP-1 and EAV PCP also share sequence homology with other well studied cysteine proteinases such as the tobacco etch virus helper component protease (TEV HC-Pro) and the proteinases of hypovirulence-associated virus (HAV) of the chestnut blight fungus. The cleavage site of TEV HC-Pro is also at a Gly–Gly dipeptide bond and both residues (P1 and P1′) are critical for efficient cleavage (Carrington *et al.*, 1989; Carrington and Herndon, 1992). There are two PCP proteinases (p29 and p48) reported in HAV. HAV p29 and p48 utilize Gly–Gly and Gly–Ala as their cleavage sites, respectively. The P1 Gly residue for both HAV cleavage sites is required for efficient cleavage (Choi *et al.*, 1991a,b; Shapira and Nuss, 1991). Overall, a Gly at the P1 position is required by MHV, EAV, TEV, and HAV cysteine proteinases for efficient cleavage activity; but the P1′ residue is somewhat flexible for MHV PCP-1, EAV PCP, and HAV p48 recognition and processing.

Viral cysteine proteinases such as EAV PCP, TEV HC-Pro, and HAV p29 described above have been termed "leader" viral papain-like proteinases (Gorbalenya *et al.*, 1991). These leader proteinases are located at the amino-termini of the polyprotein and mediate a single cleavage event in *cis* (or intramolecularly) at their own C-termini. In these polyproteins, the cleavage sites are located 50–100 amino acids downstream of the catalytic cysteine residues (Snijder *et al.*, 1992; Carrington *et al.*, 1989; Choi *et al.*, 1991a,b). MHV PCP-1 may be more similar to Sindbis virus nsP2 (classified as a "main" proteinase by Gorbalenya *et al.*, 1991) which cleaves the polyprotein precursor at multiple sites both in *cis* and in *trans* (Hardy and Strauss, 1989; de Groot *et al.*, 1990; Shirako and Strauss, 1990). MHV PCP-1 has been well characterized in an *in vitro* transcription/translation system (Baker *et al.*, 1993). The proteinase domain belongs to the cysteine family of proteinases with Cys-1137 and His-1288 essential for catalytic activity. The proteinase functions in *cis* and requires the p28 cleavage site (spaced almost 900 amino acid residues away from the proteinase domain) to fold into close proximity to the catalytic nucleophile Cys-1137 to allow efficient cleavage. We believe p28 cleavage may be a cotranslational event, in which the transient folding of the nascent polypeptide allows the cleavage site to interact with the functional domain of proteinase. This is supported by the evidence of the rapid and efficient cleavage of p28 in both cell-free system and infected cells. To date, the only known function of MHV PCP-1 is to release p28 from the N-terminus of the polymerase polyprotein. However, it is possible that mature polymerase polyprotein product may posses several sites which may be cleaved in *trans* by MHV PCP-1.

Viral proteinases have been shown to be very specific in their recognition of cleavage sites (Dougherty and Semler, 1993). The characterization of the *cis*-cleavage site of p28 for MHV PCP-1 should help us predict additional cleavage sites in the polyprotein. Indeed, there are 16 other Arg–Gly (or RG) sites in MHV-JHM polymerase polyprotein which may serve as potential targets for *trans*-cleavage by the mature PCP-1 product. However, the ability of the proteinase to recognize the cleavage site relies on the overall presentation of the cleavage site to the proteinase domain. Indeed, the RG site located at position P5 and P4 of the cleavage site (Fig. 4) does not seem to be efficiently recognized by PCP-1. The folding of the polyprotein, which we are still unable to predict without the aid of X-ray crystallography,

determines the accessibility of a proteinase domain to a putative cleavage site. Further studies involving the identification of proteolytic products of the MHV polyprotein and their cleavage sites will help us define the complete proteolytic processing pathway of the MHV polymerase polyprotein.

ACKNOWLEDGMENTS

We thank Dr. Basam Wakim for the microsequencing and Juliet Glomb and Dolph Ellefson for technical assistance. This work was supported by Public Health Service Research Grant AI32065 from the National Institutes of Health and a Junior Faculty Research Award from the American Cancer Society.

REFERENCES

BAKER, S. C., YOKOMORI, K., DONG, S., CARLISLE, R., GORBALENYA, A. E., KOONIN, E. V., and LAI, M. M. C. (1993). Identification of the catalytic sites of a papain-like cysteine proteinase of murine coronavirus. *J. Virol.* **67,** 6056–6063.

BAKER, S. C., SHIEH, C. K., SOE, L. H., 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.

BREDENBEEK, P. J., PACHUK, C. J., NOTEN, A. F. H., CHARITE, J., LUYTJES, W., WEISS, S. R., and SPAAN, W. J. M. (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,** 1825–1932.

BRIERLY, 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 frameshifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J.* **6,** 3779–3785.

BRIERLY, I., DIGARD, P., and INGLIS, S. C. (1989). Characterization of an efficient coronavirus ribosomal frameshift signal: Requirement for an RNA pseudoknot. *Cell* **57,** 537–547.

CARRINGTON, J. C., and HERNDON, K. L. (1992). Characterization of the potyviral HC-Pro autoproteolytic cleavage site. *Virology* **187,** 308–315.

CARRINGTON, J. C., CARY, S. M., PARKS, T. D., and DOUGHERTY, W. G. (1989). A second proteinase encoded by a plant potyvirus genome. *EMBO J.* **8,** 365–370.

CHOI, G. H., SHAPIRA, R., and NUSS, D. L. (1991a). Cotranslational autoproteolysis involved in gene expression from a double-stranded RNA genetic element associated with hypovirulence of the chestnut blight fungus. *Proc. Natl. Acad. Sci. USA* **88,** 1167–1171.

CHOI, G. H., PAWLYK, D. M., and NUSS, D. L. (1991b). The autocatalytic proteinase p29 encoded by a hypovirulence-associated virus of the chestnut blight fungus resembles the potyvirus-encoded proteinase HC-Pro. *Virology* **183,** 747–752.

DE GROOT, R., HARDY, W. R., SHIRAKO, Y., and STRAUSS, J. H. (1990). Cleavage-site preferences of Sindbis virus polyproteins containing the non-structural proteinase. Evidence for temporal regulation of polyprotein processing in vivo. *EMBO J.* **9,** 2631–2638.

DEN BOON, J. A., SNIJDER, E. J., CHIRNSIDE, E. D., DE VRIES, A. A. F., HORZINEK, M. C., and SPAAN, W. J. M. (1991). Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J. Virol.* **65,** 2910–2920.

DENISON, M. R., and PERLMAN, S. (1986). Translation and processing of mouse hepatitis virus virion RNA in a cell-free system. *J. Virol.* **60,** 12–18.

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

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.

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.

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.

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.

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.

HARDY, W. R., and STRAUSS, J. H. (1989). Processing the nonstructural polyproteins of Sindbis virus: nonstructural proteinase is in the C-terminal half of nsP2 and functions both in cis and in trans. *J. Virol.* **63,** 4653–4664.

HELLEN, C. U. T., LEE, C.-K., and WIMMER, E. (1992). Determinants of substrate recognition by poliovirus 2A proteinase. *J. Virol.* **66,** 3330–3338.

HUTCHISON, C. A., PHILLIPS, S., EDGELL, M. H., GILLIAM, S., JAHNKE, P., and SMITH, M. (1978). Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* **253,** 6551–6560.

KRAUSSLICH, H.-G., and WIMMER, E. (1988). Viral proteinases. *Annu. Rev. Biochem.* **57,** 701–754.

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

LAI, M. M. C. (1990). Coronaviruses: Organization, replication and expression of genome. *Annu. Rev. Microbiol.* **44,** 303–333.

LAI, M. M. C., PATTON, C. D., BARIC, R. S., and STOHLMAN, S. A. (1983). Presence of leader sequences in the mRNA of mouse hepatitis virus. *J. Virol.* **46,** 1027–1033.

LEE, H.-J., SHIEH, C.-K., GORBALENYA, A. E., KOONIN, E. V., LA MONICA, 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 polymerases. *Virology* **180,** 567–582.

LEWIS, M. K., and THOMPSON, D. V. (1990). Efficient site directed in vitro mutagenesis using ampicillin selection. *Nucleic Acids Res.* **18,** 3439–3443.

MAKINO, S., KECK, J. G., STOHLMAN, S. A., and LAI, M. M. C. (1986). High frequency RNA recombination of murine coronaviruses. *J. Virol.* **57,** 729–737.

MANIATIS, T. E., FRITSCH, E. F., and SAMBROOK, J. (1982). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

MATSUDAIRA, P. (1987). Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* **262,** 10,035–10,038.

PACHUK, C. J., BREDENBEEK, P. J., ZOLTICK, P. W., SPAAN, W. J. M., and WEISS, S. R. (1989). Molecular cloning of the gene encoding the putative polymerase of mouse hepatitis coronavirus, strain A59. *Virology* **171,** 141–148.

PALMENBERG, A. C. (1990). Proteolytic processing of picornaviral polyprotein. *Annu. Rev. Microbiol.* **44,** 603–623.

SCHECHTER, I., and BERGER, A. (1967). On the active site of proteinases. III. Papain. *Biochem. Biophys. Res. Commun.* **32,** 898–902.

SHAPIRA, R., and NUSS, D. L. (1991). Gene expression by a hypovirulence-

associated virus of the chestnut blight fungus involves two papain-like proteinase activities. *J. Biol. Chem.* **266,** 19419–19425.

SHIRAKO, Y., and STRAUSS, J. H. (1990). Cleavage between nsP1 and nsP2 initiates the processing pathway of Sindbis virus nonstructural polyprotein P123. *Virology* **177,** 54–64.

SNIJDER, E. J., WASSENAAR, A. L. M., and SPAAN, W. J. M. (1992). The 5′ end of the equine arteritis virus replicase gene encodes a papainlike cysteine protease. *J. Virol.* **66,** 7040–7048.

SOE, L. H., SHIEH, C.-K., BAKER, S. C., CHANG, M.-F., and LAI, M. M. C. (1987). Sequence and translation of the murine coronavirus 5′-end genomic RNA reveals the N-terminal structure of the putative RNA polymerase. *J. Virol.* **61,** 3968–3976.

SPAAN, W. J. M., CAVANAGH, D., and HORZINEK, M. C. (1988). Coronaviruses: Structure and genome expression. *J. Gen. Virol.* **69,** 2939–2952.

SPAAN, W., DELIUS, H., SKINNER, M., ARMSTRONG, J., ROTTIER, P., SMEEKENS, S., VAN DER ZEIJST, B. A., and SIDDELL, S. G. (1983). Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO J.* **2,** 1839–1944.

STRAUSS, E. G., and STRAUSS, J. H. (1990). Alphavirus proteinases. *Sem. Virol.* **1,** 347–356.

VIEIRA, J., and MESSING, J. (1987). Production of single-stranded plasmid DNA. *Methods Enzymol.* **153,** 3–11.