JOURNAL OF VIROLOGY, Oct. 1993, p. 6056–6063
0022-538X/93/106056-08$02.00/0

Vol. 67, No. 10

Identification of the Catalytic Sites of a Papain-Like Cysteine Proteinase of Murine Coronavirus

SUSAN C. BAKER,[1]* KYOKO YOKOMORI,[2] SHANGHONG DONG,[1] RICARDA CARLISLE,[1] ALEXANDER E. GORBALENYA,[3] EUGENE V. KOONIN,[4] AND MICHAEL M. C. LAI[2]

Department of Microbiology and Immunology, Loyola University of Chicago, Stritch School of Medicine, Maywood, Illinois 60153[1]; Howard Hughes Medical Institute and Department of Microbiology, University of Southern California School of Medicine, Los Angeles, California 90033[2]; Institute of Poliomyelitis and Viral Encephalitides, Academy of Medical Sciences, Moscow, Russia[3]; and National Center for Biotechnology Information, Bethesda, Maryland 20894[4]

Received 1 March 1993/Accepted 30 June 1993

The murine coronavirus mouse hepatitis virus gene 1 is expressed as a polyprotein, which is cleaved into multiple proteins posttranslationally. One of the proteins is p28, which represents the amino-terminal portion of the polyprotein and is presumably generated by the activity of an autoproteinase domain of the polyprotein (S. C. Baker, C. K. Shieh, L. H. Soe, M.-F. Chang, D. M. Vannier, and M. M. C. Lai, J. Virol. 63:3693–3699, 1989). In this study, the boundaries and the critical amino acid residues of this putative proteinase domain were characterized by deletion analysis and site-directed mutagenesis. Proteinase activity was monitored by examining the generation of p28 during in vitro translation in rabbit reticulocyte lysates. Deletion analysis defined the proteinase domain to be within the sequences encoded from the 3.6- to 4.4-kb region from the 5′ end of the genome. A 0.7-kb region between the substrate (p28) and proteinase domain could be deleted without affecting the proteolytic cleavage. However, a larger deletion (1.6 kb) resulted in the loss of proteinase activity, suggesting the importance of spacing sequences between proteinase and substrate. Computer-assisted analysis of the amino acid sequence of the proteinase domain identified potential catalytic cysteine and histidine residues in a stretch of sequence distantly related to papain-like cysteine proteinases. The role of these putative catalytic residues in the proteinase activity was studied by site-specific mutagenesis. Mutations of Cys-1137 or His-1288 led to a complete loss of proteinase activity, implicating these residues as essential for the catalytic activity. In contrast, most mutations of His-1317 or Cys-1172 had no or only minor effects on proteinase activity. This study establishes that mouse hepatitis virus gene 1 encodes a proteinase domain, in the region from 3.6 to 4.4 kb from the 5′ end of the genome, which resembles members of the papain family of cysteine proteinases and that this proteinase domain is responsible for the cleavage of the N-terminal peptide.

Mouse hepatitis virus (MHV) is an enveloped RNA virus belonging to the coronavirus family. The positive-sense RNA genome is 31 kb long (22, 27), making it the largest viral RNA genome identified so far. In the infected cell, the virus encodes an RNA-dependent RNA polymerase, which is responsible for the replication of the viral genomic RNA and the synthesis of seven subgenomic RNAs. These mRNAs form a 3′-coterminal, nested-set structure and contain an identical 5′-end leader sequence of approximately 72 nucleotides (3, 21, 32), which has been proposed to be the result of a unique discontinuous transcription mechanism (20).

In vitro translation studies indicate that only the 5′ unique region of each of the mRNAs is translated (29). The largest mRNA, which is of genomic length, has a 5′ unique region of 22 kb (22, 27). This unique region constitutes the 5′-most gene, termed gene 1, of the MHV RNA genome and is thought to encode the viral RNA-dependent RNA polymerase. Sequencing studies of MHV and avian coronavirus (infectious bronchitis virus [IBV]) genomic RNAs suggest that this gene is expressed as a single long polyprotein via a ribosomal frameshifting of two overlapping open reading frames (4–6, 22). Analyses of IBV and MHV gene 1 sequences have also suggested that this gene encodes many functional domains, including either one or two papain-like proteinases, a poliovirus 3C protein-like proteinase, helicase, zinc finger-like sequence, polymerase, cysteine-rich domains, and membrane-spanning regions (13, 22). Interestingly, the papain-like proteinase domain is duplicated in the MHV genome, but not in IBV, at approximately 3.5 to 4.3 kb and 5.3 to 6.1 kb from the 5′ end of the genome (22). The large size of the gene 1 open reading frames, which have a combined coding capacity of more than 800 kDa (22), and the presence of proteinase-like sequences suggest that this region may be expressed as a polyprotein that is subsequently processed into different functional subunits.

The first indication of the co- or posttranslational processing of the MHV gene 1 polyprotein product came from the studies of in vitro translation of MHV genomic RNA, which yielded a major product of 220 kDa and a 28-kDa protein (8, 23), the latter being the N-terminal cleavage product of the MHV polyprotein (1, 31). This protein product was subsequently confirmed by in vitro transcription and translation of a plasmid cDNA containing the 5′ portion of MHV gene 1, which yielded a p28 protein and a 160-kDa protein (1, 31). p28 and other cleavage products of the polyprotein have also been detected in MHV-infected cells, indicating that these were normal products of the gene 1 polyprotein processing (8–11). The proteinases responsible for the processing of these proteins are mostly unknown. However, our previous studies showed that the peptide sequences encoded from the region between 3.9 and 5.3 kb from the 5′ end of the genome

* Corresponding author.

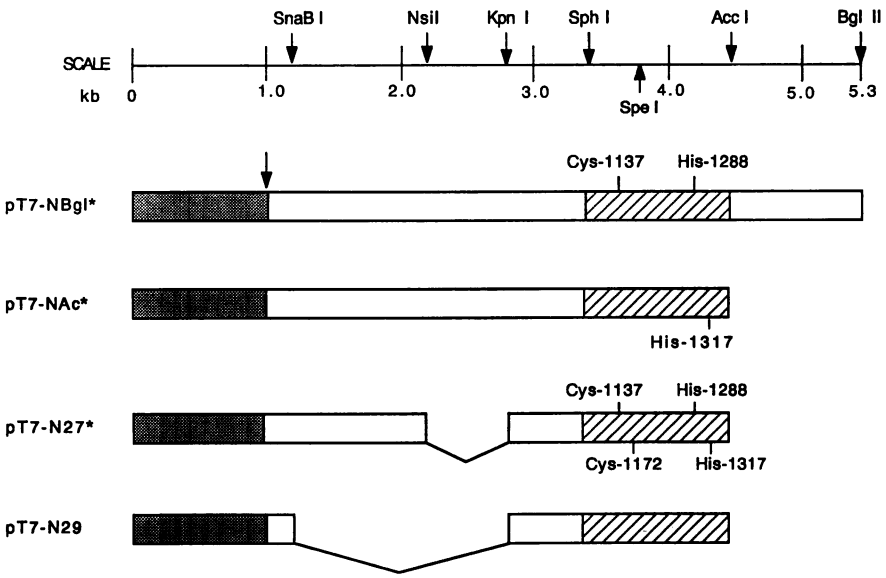

FIG. 1. Schematic diagram of the structure of the pT7-NBgl family of plasmids. A partial restriction map of the 5'-end 5.3 kb of MHV-JHM genomic cDNA is shown to scale. The p28 region, which is released by the autoproteinase activity, is denoted by the stippled box, and the putative cysteine proteinase domain is indicated by the hatched box. Constructs used to generate site-specific mutants are designated by an asterisk (*), with Cys and His residues which were mutated indicated. The amino acids are numbered by using the system of Baker et al. (1).

was required for the proteolytic cleavage of p28 from the primary translation product (1). Significantly, this region overlaps the first papain-like proteinase domain predicted from the sequence analysis (22), suggesting that this is the proteinase which cleaves the p28 protein. However, this study did not rule out the possibility that this region is needed only for maintaining the correct protein conformation as a substrate for cellular proteinases. In the current study, we have used site-directed mutagenesis and deletion analysis to support the presence of a viral proteinase and have identified the Cys and His residues critical for the proteolytic activity, confirming that this MHV proteinase is related to the papain family of thiol proteinases.

MATERIALS AND METHODS

Construction of deletion mutant plasmids. T7 polymerase transcription plasmid pT7-NBgl (1), which contains the MHV-JHM genomic cDNA sequence from the *Nar*I site at nucleotide 187 to the *Bgl*II site at nucleotide 5273 (22), was used as the parent plasmid. Deletion mutant pT7-NAc (Fig. 1) was constructed by digestion of pT7-NBgl with *Acc*I (which cuts at position 4434) and *Sma*I (which cuts in the 3'-polylinker region), removal of the 3'-most 0.9-kb region, blunting with T4 DNA polymerase, and religation of the plasmid DNA. This plasmid was used as the parent plasmid for the construction of internal deletion mutants. Plasmid pT7-NAc was digested with *Nsi*I (which cuts at position 2068) and *Kpn*I (which cuts at position 2811), blunted with T4 DNA polymerase, and religated. The resultant plasmid, with a deletion of 741 nucleotides, maintains the correct open reading frame and was designated pT7-N27. Plasmid pT7-NAc was also digested with *Sna*BI (which cuts at nucleotide 1167) and *Kpn*I (which cuts at nucleotide 2811), blunted with T4 DNA polymerase, and religated. The resultant plasmid, with a deletion of 1,642 nucleotides, maintains the correct open reading frame and was designated pT7-N29. Maintenance of the correct open reading frame in pT7-N27 and pT7-N29 was confirmed by DNA sequencing (Sequenase System; U.S. Biochemicals).

PCR mutagenesis. Site-directed mutagenesis of Cys-1137 and His-1288 in plasmid pT7-NBgl and His-1317 in pT7-NAc was accomplished by the procedure of Higuchi et al., with synthetic oligonucleotides containing mismatches and using the polymerase chain reaction (PCR) to amplify mutated DNA (16). Briefly, oligonucleotide 98 (5'-GTGAGGCATGC GACAGGGAA-3'), encompassing the unique *Sph*I site at nucleotide 3339, and oligonucleotide 99 (5'-CGCAGCCAA GAATTAGTACG-3'), containing a mismatch at position 3624, were used to prime DNA synthesis from plasmid pT7-NBgl in PCR. Likewise, oligonucleotide 101 (5'-CGC TCTTAACTAGTTTGTCC-3'), encompassing a unique *Spe*I site at nucleotide 3743, and oligonucleotide 100 (5'-CGTAC TAATTCTTGGCTGCG-3'), containing a mismatch at position 3624, were used to prime DNA synthesis by PCR from pT7-NBgl. The product DNAs from these two reactions were then mixed, denatured, and reannealed, and oligonucleotides 98 and 101 were added to prime DNA synthesis by PCR. The PCR product DNA consisted of MHV genomic cDNA sequence from nucleotides 3334 to 3756 with a specific mutation (G to C) at nucleotide 3624, resulting in an amino acid change from cysteine to serine at residue 1137. The amplified DNA was digested with *Sph*I and *Spe*I and inserted into the *Sph*I and *Spe*I sites of pT7-NBgl, and the resultant plasmid, which has a mutation at position 1137 from Cys to Ser, was designated pT7-NB.1. The specific mutation was identified by sequencing.

Mutagenesis of histidine residue 1288 in pT7-NBgl and histidine residue 1317 in pT7-NAc was performed in a similar manner by using oligonucleotides and PCR as described above. The region from the *Spe*I site (oligonucleotide 125, 5'-GGACAAACTAGTTAAGAGCG-3') to the *Acc*I site (oligonucleotide 115, 5'-ACCACCGGTAGACTCATAGC-3') was amplified by PCR with oligonucleotides containing mismatches at positions 4067 and 4068 (encoding His-1288) or positions 4164 and 4165 (encoding His-1317), resulting in

TABLE 1. Oligonucleotides used to generate site-specific mutations in the MHV proteinase domain

Wild-type amino acid	Mutagenic oligonucleotide (5′ to 3′)[a]	Mutation	
		Codon	Amino acid
Cys-1137	CGCAGCCAXXXATTAGTACGC	TAT	Tyr
		CGT	Arg
		AGT	Ser
		GGT	Gly
		TGG	Trp
His-1288	GCCATAGAXXXGCAATCATT	ACA	Thr
		CGT	Arg
His-1317	GTCATACCXXXACCCACCAT	CAG	Gln
		GTG	Val
		TCT	Ser
		CCG	Pro
Cys-1172	CCACAAAGXXTTGATCATAGCC	AAC	Asn
		TCC	Ser

[a] X denotes the positions of degenerate nucleotides.

a change from histidine to glycine in each case. The amplified DNA containing the His-1288–to–Gly mutation was digested with *Spe*I and *Acc*I, inserted into the *Spe*I and *Acc*I sites of pT7-NBgl, and designated pT7-NB.3. The amplified DNA containing the His-1317–to–Gly mutation was digested with *Spe*I and *Acc*I, inserted into the *Spe*I and *Acc*I sites of pT7-NAc, and designated pT7-NB.8.

pSELECT mutagenesis. Additional amino acid substitution mutants were constructed by the procedure of Hutchinson et al., using oligonucleotides containing mismatches annealed to single-stranded DNA templates (17). Briefly, the gene 1 sequence in pT7-N27 (from the *Pvu*II site upstream of the T7 promoter to the *Eco*RI site) was inserted into the polylinker region (*Sma*I-*Eco*RI sites) of the pSELECT-1 vector (Promega Biotec) (24), and the resulting plasmid was designated pS-N27. Single-stranded pS-N27 DNA was isolated after infection with R408 helper phage (35). Phosphorylated oligonucleotides containing degenerate sequences at specific sites (see Table 1) were annealed to single-stranded pS-N27 DNA in conjunction with an ampicillin repair oligonucleotide. Subsequent DNA synthesis with T4 DNA polymerase and ligase (Promega Biotec) linked the two annealed, mismatch-containing oligonucleotides into the newly synthesized DNA. The DNA was amplified in a repair-deficient *Escherichia coli* BHM71-18mutS in the presence of ampicillin followed by transformation into *E. coli* JM109. Plasmid DNA isolated from individual colonies was sequenced across the region encompassing the degenerate oligonucleotide binding site by double-stranded DNA sequencing. Specific mutations found are listed in Table 1.

In vitro transcription-translation and immunoprecipitation. Recombinant pT7 plasmids were linearized by *Eco*RI digestion, transcribed in vitro with T7 RNA polymerase, and translated in rabbit reticulocyte lysates as previously described (1) or in a coupled transcription-translation system (TNT lysates; Promega Biotec). Briefly, approximately 0.5 μg of linearized plasmid DNA was incubated for 90 min at 30°C with 12.5 μl of TNT rabbit reticulocyte lysate, 1 μl of TNT reaction buffer, 0.5 μl of T7 RNA polymerase, 20 U of RNasin RNase inhibitor (Promega Biotec), 1 mM methionine-free amino acid mixture, and [^{35}S]methionine (20 μCi per translation reaction; Amersham) in a final volume of 25 μl. Then 2 μl of translation products was removed and precipitated with trichloroacetic acid (25) for measurement of [^{35}S]Met incorporation, and the remainder was mixed with an equal volume of 2× Laemmli sample buffer (19) containing 5% β-mercaptoethanol and 3% sodium dodecyl sulfate (SDS) to stop the translation reaction. Equal trichloroacetic acid-precipitable counts of each reaction (0.5×10^5 to 5×10^5 cpm) were immunoprecipitated by incubation with 3 μl of anti-p28-antibody in 1 ml of RIPA buffer (1). After 12 to 16 h of rotation at 4°C, 30 μl of a 1:1 suspension of protein A-Sepharose beads (Pharmacia) in RIPA buffer was added, and the incubation reaction was continued for an additional 4 h at 4°C. The beads were pelleted by centrifugation, washed three times with 1 ml of RIPA buffer, suspended in 50 μl of 2× Laemmli sample buffer, heated at 100°C for 2 min, and microcentrifuged for 5 min. The supernatant was separated by electrophoresis on a 10% polyacrylamide gel containing 0.1% SDS (SDS-PAGE) (19). Following electrophoresis, the gel was fixed and enhanced with Fluoro-Hance (Research Products International Corp.) for 30 min before being dried and exposed to Kodak X-ray film at −70°C.

RESULTS

Effects of deletion on proteinase activity. We have previously identified in the MHV-JHM gene 1 polyprotein a region responsible for the cleavage of the amino-terminal peptide, p28, from the polyprotein (1). Analysis of the predicted amino acid sequence reveals a putative proteinase domain, which shows some sequence homology to the papain family of thiol proteinases (14, 22), suggesting that this is the autoproteinase that cleaves p28. To define this putative autoproteinase domain more precisely, we constructed a series of deletion mutants from the plasmid pT7-NBgl (Fig. 1). The plasmids were transcribed by T7 RNA polymerase and translated in rabbit reticulocyte lysates in the presence of [^{35}S]methionine, and the protein products were analyzed directly or precipitated with anti-p28 serum and separated by SDS-PAGE (Fig. 2). The pT7-NBgl translation yielded proteins of approximately 188, 160, and 28 kDa (Fig. 2A, lane 2). We have previously shown that the 188-kDa protein is the primary translation product and the 160- and 28-kDa proteins are the products of an autoproteolytic cleavage (1, 31). The p28 is the N-terminal cleavage product, which is immunoprecipitated with anti-p28 antibody (Fig. 2B, lane 2). The cleavage efficiency for pT7-NBgl translation products was approximately 70% as determined by quantitation of the primary and cleavage products with a Betegen scanner. Truncation of the 3′-end 0.9 kb in plasmid pT7-NAc had little effect on proteinase activity, with approximately 70% cleavage efficiency (Fig. 2, lanes 3). Furthermore, deletion of an internal 0.74-kb region had only a minor effect on the ability of the proteinase to cleave p28 from the precursor polyprotein (60 to 65% cleavage efficiency), as shown by translation of pT7-N27 RNA (Fig. 2, lanes 4). This result indicates that the proteinase domain is located in the region from 3.6 to 4.4 kb from the 5′ end and that some spacer sequences between the proteinase and the substrate can be deleted. However, a larger internal deletion of 1.6 kb (pT7-N29 RNA) resulted in an inability to cleave the p28 protein and in the appearance of two protein products of approximately 95 and 90 kDa, which were precipitable by p28 antiserum (Fig. 2B, lane 5). The 95-kDa band was the primary translation product from the pT7-N29, and the 90-kDa protein probably resulted from premature termination of translation or from a proteinase cleavage at an alternative site. This result indicates that the internal region between the proteinase domain and the p28 cleavage site

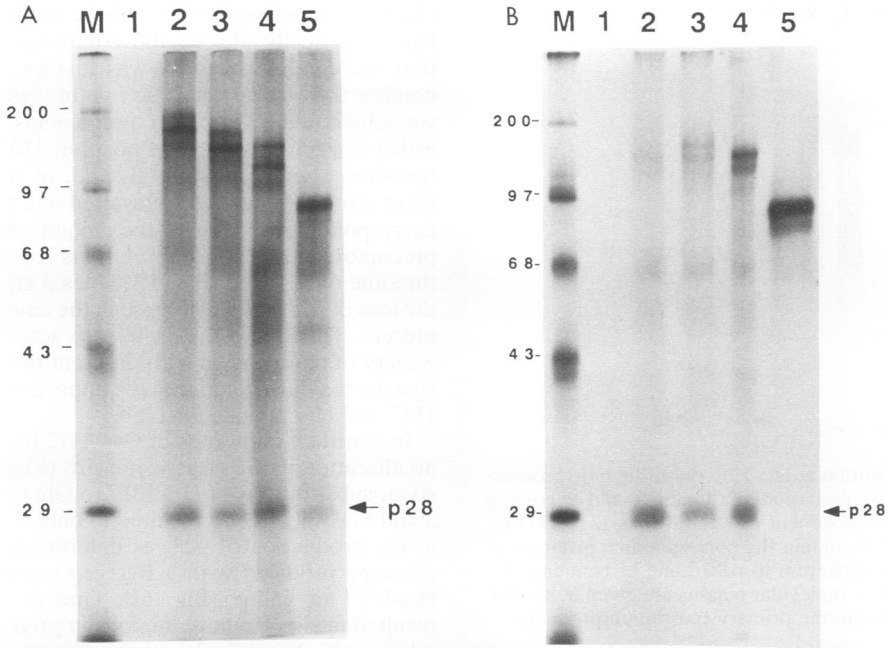

FIG. 2. Effects of deletion on proteolytic processing of p28. Linearized plasmid DNA was transcribed and translated in rabbit reticulocyte lysates in the presence of T7 RNA polymerase and [^{35}S]methionine as described in Materials and Methods. Equal trichloroacetic acid counts of translation products were analyzed by PAGE (10% polyacrylamide) directly (A) and following immunoprecipitated with p28-specific antiserum (B). Lanes: 1, no RNA; 2, pT7-NBgl RNA; 3, pT7-NAc RNA; 4, pT7-N27 RNA; 5, pT7-N29 RNA. Lane M contained ^{14}C-labeled marker polypeptides; molecular masses are given in kilodaltons.

may act as a spacer to provide a correct conformation essential for autoproteolytic activity. Alternatively, the protein product encoded by this region may play a supporting role in the processing of the polyprotein. Overall, these deletion studies supported the prediction of a proteinase domain in the 5′-end 3.6- to 4.4-kb region of gene 1 and also indicated an important conformational role for the spacer region between the proteinase domain and the cleavage site.

Effects of site-specific mutations of Cys and His residues. Extensive amino acid sequence homology searches have suggested that the proteinase responsible for the cleavage of p28 belongs to the papain-like class of thiol proteinases with characteristic catalytic residues of cysteine and histidine (14, 22). Among the Cys and His residues within the predicted proteinase domain, Cys-1137 and His-1288 are the only two conserved among all the cellular and viral papain-like thiol proteinases, and they were therefore predicted to be the catalytic amino acids (22). To determine whether the predicted residues were indeed essential for catalytic activity, we altered the putative catalytic amino acids as well as His-1317 and Cys-1172 in the proximity for comparison of their role in proteolytic activities.

In vitro transcription-translation of plasmid pT7-NB.1, in which Cys-1137 had been mutated to serine, did not yield any p28 (Fig. 3, lane 3), suggesting that Cys-1137 is essential for autoproteolytic cleavage of the polyprotein precursor. A complete loss of autoproteolytic activity was also seen when His-1288 was mutated to a glycine residue (plasmid pT7-NB.3 [Fig. 3, lane 4]). Immunoprecipitations of the in vitro translation products with an anti-p28 antibody (1) further confirmed the absence of p28 and the presence of uncleaved primary translation product when Cys-1137–to–Ser (lane 7) or His-1288–to–Gly (lane 8) mutant RNAs were translated. These results support the hypothesis that Cys-1137 and His-1288 are the catalytic residues essential for the autoproteolytic activity for release of p28. In contrast, the mutation of His-1317 (pT7-NB.8) reduced but did not completely eliminate the production of p28 (Fig. 4, lanes 1 and 2). This conclusion was confirmed by immunoprecipitation with anti-p28 antibody (Fig. 4, lanes 3 and 4). The slight reduction in the amount of p28 produced by the His-1317–to–Gly mutant suggests that the substitution at position 1317 has a minor

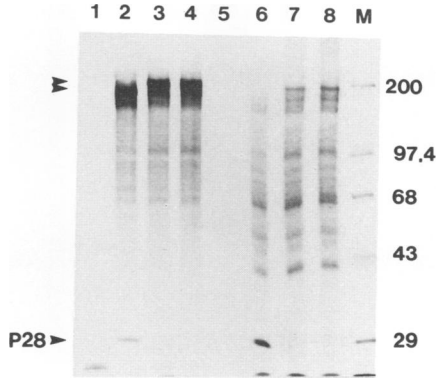

FIG. 3. Effect of amino acid substitution at Cys-1137 and His-1288 on proteolytic processing of pT7-NBgl translation products. Lanes 1 to 4 contain translation products of no RNA, pT7-NBgl, pT7-NB.1 (Cys-1137 to Ser), and pT7-NB.3 (His-1288 to Gly) RNAs, respectively. Lanes 5 to 8 contain the corresponding products of immunoprecipitation with p28-specific antiserum. Lane M contains ^{14}C-labeled marker polypeptides; molecular masses are given in kilodaltons. The arrowheads indicate the primary translation products.

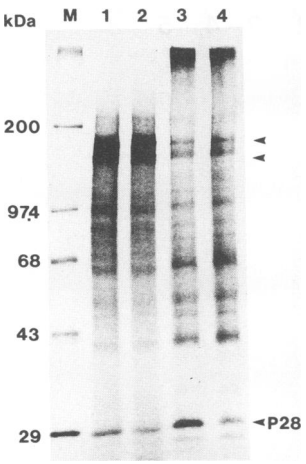

FIG. 4. Effect of substitution at His-1317 on proteolytic processing of the pT7-NAc translation product. Lanes 1 and 2 contain translation products of pT7-NAc and pT7-NB.8 (His-1317 to Gly), respectively. Lanes 3 and 4 contain the corresponding products of immunoprecipitation with antiserum to p28. Lane M contains ^{14}C-labeled marker polypeptides; molecular masses are given in kilodaltons. The arrowheads indicate the primary translation products.

effect on the efficiency of the cleavage of p28 but that His-1317 is not an essential catalytic amino acid of the autoproteinase.

Although the above mutagenesis results strongly support the hypothesis of a papain-like proteinase with Cys-1137 and His-1288 as the catalytic amino acids involved in the p28 cleavage, it is possible that these amino acid substitutions have drastically altered the conformation of proteins and that the catalytic activity was lost as a result. To further confirm the critical catalytic role of Cys-1137 and His-1288, we substituted a variety of amino acids with different properties at each position. At position 1137, substitution with tyrosine, arginine, serine, glycine, or tryptophan (Fig. 5A, lanes 2 to 7, respectively) resulted in a total loss of p28 and corresponding increase in the amount of an uncleaved p188 precursor polypeptide. Mutations of His-1288 to either threoine or arginine (Fig. 5B, lanes 3 and 4) also resulted in the loss of p28 and an increase in the amount of the precursor protein. The loss of the proteinase activity by mutation to a variety of amino acids with different properties at these two specific residues confirmed the importance of Cys at position 1137 and His at position 1288.

In contrast, mutations of Cys-1172 to Asn or Ser revealed no alteration in the cleavage of p28 (Fig. 6A, lanes 3 and 4). Also, substitutions of His-1317 to Glu or Ser (Fig. 6B, lanes 3 and 5, respectively) resulted in only a 10 to 20% reduction in the production of p28, as determined by quantitation of cleavage products with a Betegen scanner. Substitution of His-1317 to valine (Fig. 6B, lane 4) or proline (lane 6) resulted in significant reductions in proteolytic processing to release p28. In particular, the presence of proline at position 1317 completely eliminated the production of p28 and resulted a primary translation product which migrated slightly faster on the gel (lane 6). This is not surprising since the addition of a proline residue may cause significant alteration in protein folding. These data suggest that His-1317 and Cys-1172 are not the catalytic residues of the proteinase but that these residues may play a role in maintaining the

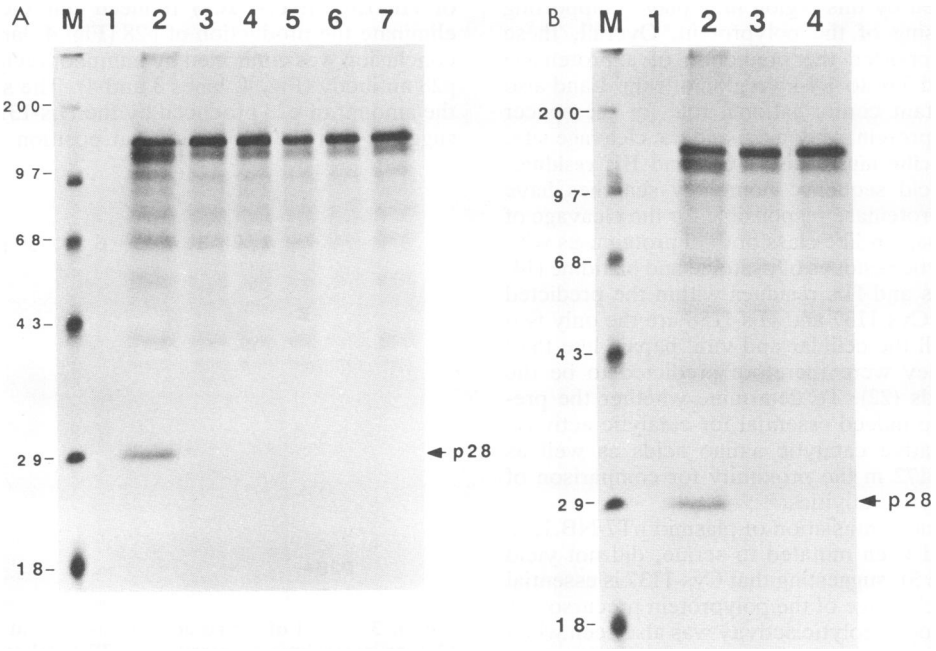

FIG. 5. Effect of substitutions at putative catalytic residues, Cys-1137 and His-1288, on proteolytic processing of pS-N27 translation product. A series of mutants with mutations in positions 1137 and 1288 were generated by oligonucleotide mutagenesis as described in Materials and Methods. Plasmid DNA was linearized by *Eco*RI digestion, and in vitro transcription-translation was performed as in the experiment in Fig. 2. Equal trichloroacetic acid counts of in vitro translation products were immunoprecipitated with p28-specific antiserum and analyzed by SDS-PAGE (10% polyacrylamide). (A) Lanes 1 to 7 contain products with the following amino acids at positions 1137: no RNA, Cys, Tyr, Arg, Ser, Gly, and Trp; (B) lanes 1 to 4 contain products encoding the following amino acids at position 1288: no RNA, His, Thr, and Arg.

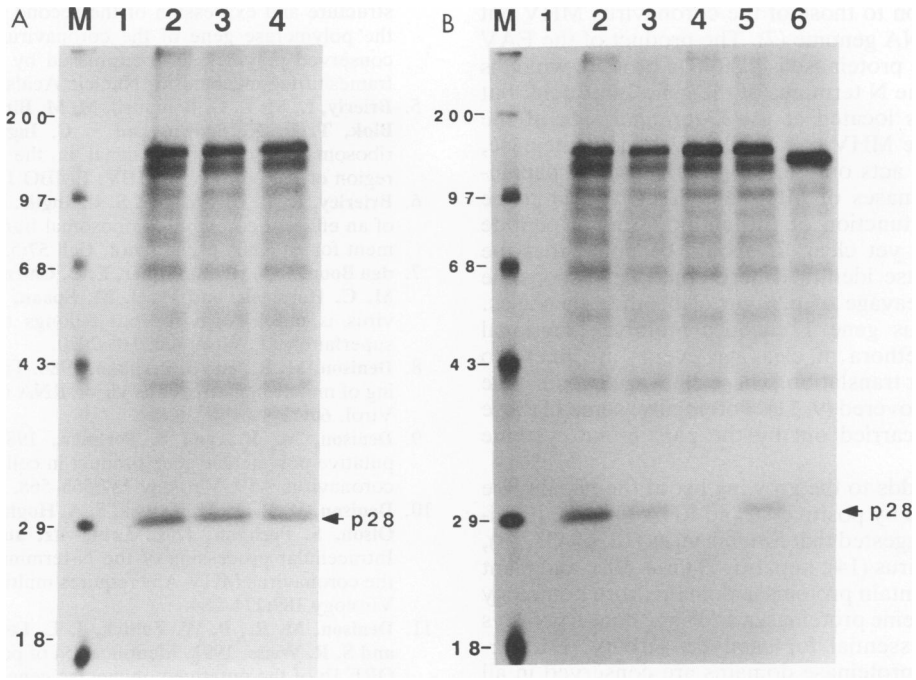

FIG. 6. Effect of substitutions at Cys-1172 and His-1317 residues on proteolytic processing of pS-N27 translation products. Various Cys-1172 or His-1317 mutants with mutations in pS-N27 were analyzed as in the experiment in Fig. 5. (A) Lanes 1 to 4 contain products encoding the following amino acids at position 1172: no RNA, Cys, Asn, Ser; (B) lanes 1 to 6 contain products with the following amino acids at position 1317: no RNA, His, Gln, Val, Ser, Pro.

conformation of the proteinase domain to allow accurate and efficient cleavage of p28.

The above extensive site mutagenesis analysis of Cys-1137, His-1288, Cys-1172, and His-1317 demonstrated that the catalytic residues of this proteinase are Cys-1137 and His-1288 and that it belongs to the papain-like class of thiol proteinases.

DISCUSSION

The mutagenesis studies reported in this paper have identified the Cys-1137 and His-1288 residues as critical for the proteolysis of p28 protein from the N terminus of the MHV gene 1 polyprotein. These two residues have been proposed to be the catalytic residues of the MHV proteinase on the basis of their sequence similarity with cellular and viral papain-like proteinases (22). Although it cannot be unequivocally ruled out that the amino acid substitutions carried out here simply altered the overall protein conformation but did not directly affect the catalytic center of the proteinase, the finding that all the substitutions of amino acids of diverse properties for Cys-1137 and His-1288 resulted in the total loss of proteinase activities suggested strongly that these amino acids are indeed at the catalytic center. In contrast, substitutions of the neighboring His or Cys residues did not inactivate proteinase activities, except for the proline substitution, which is expected to alter global protein conformation. These studies therefore strongly supported the computer prediction that MHV gene 1a contains a papain-like proteinase which is responsible for the cleavage of the N-terminal peptide, p28 (22).

It is interesting that the proteinase domain is separated from the cleavage site (substrate) by approximately 1,000 amino acids (3 kb). Furthermore, some of the intervening sequence between the proteinase domain and the substrate could be deleted without affecting the proteolytic cleavage. Also, the kinetics of protein cleavage during in vitro translation of the MHV genomic RNA suggested that p28 is cleaved as soon as the proteinase domain is synthesized and does not wait until the entire protein is synthesized (1, 8). It is likely that the tertiary conformation of the polyprotein, particularly those formed by sequences neighboring the substrate or the proteinase domain, are critical for the proteolytic cleavage. This conclusion is supported by the failure of pT7-N29, which has an internal deletion of 1.6 kb, to generate p28.

This is the first of three predicted proteinases of MHV to be demonstrated. Interestingly, a second papain-like cysteine proteinase domain is predicted approximately 2 kb downstream from the proteinase domain described here (22). We have not yet investigated the biological activity of this second papain-like proteinase domain. Because of the finding that the first papain-like proteinase works only in *cis* and requires at least some spacer sequences between the substrate and the proteinase domain, it seems unlikely that the downstream papain-like proteinase will also cleave p28. Rather, it is more likely that the second proteinase has other substrates. Significantly, IBV appears to have only one papain-like proteinase, and the N-terminal portion of its gene 1a is substantially diverged from that of MHV. IBV also lacks the corresponding p28 sequence. Therefore, the downstream papain-like proteinase may cleave the region of the gene 1 polyprotein, which is more conserved between IBV and MHV.

A papain-like cysteine proteinase domain similar to the one described here for MHV has been identified in another coronavirus superfamily member, equine arteritis virus (EAV) (30). EAV has a similar replication strategy and

genomic organization to those of the coronavirus MHV but has only a 13-kb RNA genome (7). The product of the EAV papain-like cysteine proteinase is a 30-kDa protein, which is also derived from the N terminus of the gene 1a protein, but the cleavage site is located at the C-terminal side of the proteinase. Like the MHV papain-like cysteine proteinase, this proteinase also acts only in *cis* (30). Thus, the papain-like cysteine proteinases of MHV and EAV appear to be quite related. The function of the amino-terminal peptide (p28 or p30) is not yet clear. It should be noted that the papain-like proteinase identified here could be responsible for not only the cleavage of p28 but also other cleavages. Because coronavirus gene 1 encodes multiple functional domains (22), a plethora of cleavage events are likely to occur during or after translation. These cleavage events have just begun to be uncovered (9, 11). Potentially, some of these cleavages may be carried out by the papain-like cysteine proteinase.

This proteinase adds to the growing list of the papain-like proteinases encoded by positive-strand RNA viruses. Previous studies have suggested that Sindbis virus (12, 15, 33, 34), EAV (30), rubella virus (14), hepatitis E virus (18), and plant potyvirus (26) all contain proteinase domains with homology to the family of cysteine proteinases and have conserved Cys and His residues essential for catalytic activity. Furthermore, the cysteine proteinase domains are conserved in all togaviruses, potyviruses, and coronaviruses sequenced so far (33). The Sindbis virus proteinase nsP2 has been shown to play a pivotal role in the regulation of processing and enzymatic activity of viral nonstructural proteins, especially nsP4, which is the proposed RNA polymerase protein (12, 34). Conceivably, MHV proteinase may play a similar role in the regulation of the coronavirus life cycle. Genetic recombination mapping of temperature-sensitive mutants of MHV suggest that there are at least five distinct functional units encoded by the polymerase gene which are essential for viral RNA transcription (2, 28). These functional units may be involved in the synthesis of various viral RNA species. Viral proteinases may act to release or cleave these functional subunits of the polyprotein. Further study of the defects in the temperature-sensitive mutants of MHV and in the pathway of proteolytic processing will allow us to define the functional subunits of the polymerase polyprotein.

ACKNOWLEDGMENTS

We thank Nicola La Monica for helpful discussions and Dolph Ellefson and Heidi Kjos for technical assistance.

This work was supported by Public Health Service research grants AI19244 and NS18146 (to M.M.C.L.) and AI32065 (to S.C.B.) from the National Institutes of Health. K.Y. was a postdoctoral fellow and M.M.C.L. is an Investigator of the Howard Hughes Medical Institute.

REFERENCES

1. **Baker, S. C., C. K. Shieh, L. H. Soe, M.-F. Chang, D. M. Vannier, and M. M. C. Lai.** 1989. Identification of a domain required for autoproteolytic cleavage of murine coronavirus gene A polyprotein. J. Virol. **63**:3693–3699.
2. **Baric, R. S., K. Fu, M. C. Schaad, and S. A. Stohlman.** 1990. Establishing a genetic recombination map for murine coronavirus strain A59 complementation groups. Virology **177**:646–656.
3. **Baric, R. S., S. A. Stohlman, and M. M. C. Lai.** 1983. Characterization of replicative intermediate RNA of mouse hepatitis virus: presence of leader RNA sequence on nascent chains. J. Virol. **48**:633–640.
4. **Bredenbeek, P. J., C. J. Pachuk, A. F. H. Noten, J. Charite, W. Lyutjes, S. R. Weiss, and W. J. M. Spaan.** 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.
5. **Brierly, I., M. E. G. Boursnell, M. M. Binns, B. Bilimoria, V. C. Blok, T. D. K. Brown, and S. C. Inglis.** 1987. An efficient ribosomal frame-shifting signal in the polymerase-en-coding region of the coronavirus IBV. EMBO J. **6**:3779–3785.
6. **Brierley, I., P. Digard, and S. C. Inglis.** 1989. Characterization of an efficient coronavirus ribosomal frameshift signal: requirement for an RNA pseudoknot. Cell **57**:537–547.
7. **den Boon, J. A., E. J. Snijder, E. D. Chirnside, A. A. F. de Vries, M. C. Horzinek, and W. J. M. Spaan.** 1991. Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. J. Virol. **65**:2910–2920.
8. **Denison, M. R., and S. Perlman.** 1986. Translation and processing of mouse hepatitis virus virion RNA in a cell-free system. J. Virol. **60**:12–18.
9. **Denison, M. R., and S. Perlman.** 1987. Identification of a putative polymerase gene product in cells infected with murine coronavirus A59. Virology **157**:565–568.
10. **Denison, M. R., P. W. Zoltick, S. A. Hughes, B. Giangreco, A. L. Olson, S. Perlman, J. L. Leibowitz, and S. R. Weiss.** 1992. Intracellular processing of the N-terminal ORF 1a proteins of the coronavirus MHV-A59 requires multiple proteolytic events. Virology **189**:274–284.
11. **Denison, M. R., P. W. Zoltick, J. L. Leibowitz, C. J. Pachuk, and S. R. Weiss.** 1991. Identification of polypeptides encoded in ORF 1b of the putative polymerase gene of the murine coronavirus MHV-A59. J. Virol. **65**:3076–3082.
12. **de Groot, R., W. R. Hardy, Y. Shirako, and J. H. Strauss.** 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.
13. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov.** 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.
14. **Gorbalenya, A. E., E. V. Koonin, and M. M. C. Lai.** 1991. Putative papain-related thiol proteases of positive-strand RNA viruses. FEBS Lett. **288**:201–205.
15. **Hardy, W. R., and J. H. Strauss.** 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.
16. **Higuchi, R., B. Krummel, and R. K. Saiki.** 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res. **16**:7351–7367.
17. **Hutchison, C. A., S. Phillips, M. H. Edgell, S. Gilliam, P. Jahnke, and M. Smith.** 1978. Mutagenesis at a specific position in a DNA sequence. J. Biol. Chem. **253**:6551–6560.
18. **Koonin, E. V., A. E. Gorbalenya, M. A. Purdy, M. N. Rozanov, G. R. Reyes, and D. W. Bradley.** 1992. Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: delineation of an additional group of positive-strand RNA plant and animal viruses. Proc. Natl. Acad. Sci. USA **89**:8259–8263.
19. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
20. **Lai, M. M. C.** 1990. Coronavirus: organization, replication, and expression of genome. Annu. Rev. Microbiol. **44**:303–333.
21. **Lai, M. M. C., C. D. Patton, R. S. Baric, and S. A. Stohlman.** 1983. Presence of leader sequence in the mRNA of mouse hepatitis virus. J. Virol. **46**:1027–1033.
22. **Lee, H.-J., C.-K. Shieh, A. E. Gorbalenya, E. V. Koonin, N. La Monica, J. Tuler, A. Bagdzhyadzhyan, and M. M. C. Lai.** 1991. The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerases. Virology **180**:567–582.
23. **Leibowitz, J. L., S. R. Weiss, E. Paavola, and C. W. Bond.** 1982.

Cell-free translation of murine coronavirus RNA. J. Virol. **43:**905–913.
24. **Lewis, M. K., and D. V. Thompson.** 1990. Efficient site directed in vitro mutagenesis using ampicillin selection. Nucleic Acids Res. **18:**3439–3443.
25. **Maniatis, T. E., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Oh, C.-S., and J. C. Carrington.** 1989. Identification of essential residues in potyvirus proteinase HC-Pro by site-directed mutagenesis. Virology **173:**692–699.
27. **Pachuk, C. J., P. J. Bredenbeek, P. W. Zoltick, W. J. M. Spaan, and S. R. Weiss.** 1989. Molecular cloning of the gene encoding the putative polymerase of mouse hepatitis coronavirus, strain A59. Virology **171:**141–148.
28. **Schaad, M. C., S. A. Stohlman, J. Egbert, K. Lum, K. Fu, T. Wei, and R. S. Baric.** 1990. Genetics of mouse hepatitis virus transcription: identification of cistrons which may function in positive and negative strand RNA synthesis. Virology **177:**634–645.
29. **Siddell, S. G.** 1983. Coronavirus JHM: coding assignments of subgenomic mRNAs. J. Gen. Virol. **64:**113–125.
30. **Snijder, E. J., A. L. M. Wassenaar, and W. J. M. Spaan.** 1992. The 5′ end of the equine arteritis virus replicase gene encodes a papainlike cysteine protease. J. Virol. **66:**7040–7048.
31. **Soe, L. H., C.-K. Shieh, S. C. Baker, M.-F. Chang, and M. M. C. Lai.** 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.
32. **Spaan, W. J. M., H. Delius, M. Skinner, J. Armstrong, P. Rottier, S. Smeekens, B. A. M. van der Zeijst, and S. G. Siddell.** 1983. Coronavirus mRNA synthesis involves fusion of noncontiguous sequence. EMBO J. **2:**1839–1844.
33. **Strauss, E. G., R. J. de Groot, R. Levinson, and J. H. Strauss.** 1992. Identification of the active site residues in the nsP2 proteinase of Sindbis virus. Virology **191:**932–940.
34. **Strauss, E. G., and J. H. Strauss.** 1990. Alphavirus proteinases. Semin. Virol. **1:**347–356.
35. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153:**3–11.