JOURNAL OF VIROLOGY, Nov. 1995, p. 7159–7168 Vol. 69, No. 11
0022-538X/95/$04.00+0

3C-Like Protease of Rabbit Hemorrhagic Disease Virus: Identification of Cleavage Sites in the ORF1 Polyprotein and Analysis of Cleavage Specificity

CHRISTOPH WIRBLICH,[1] MARIA SIBILIA,[2]† MARIA BEATRICE BONIOTTI,[2] CESARE ROSSI,[2] HEINZ-JÜRGEN THIEL,[1]‡ AND GREGOR MEYERS[1]*

Federal Research Centre for Virus Diseases of Animals, 72001 Tübingen, Germany,[1] and Istituto Zooprofilattico Sperimentale della Lombardia e dell' Emilia, 25124 Brescia, Italy[2]

Received 25 May 1995/Accepted 31 July 1995

Rabbit hemorrhagic disease virus, a positive-stranded RNA virus of the family *Caliciviridae*, encodes a trypsin-like cysteine protease as part of a large polyprotein. Upon expression in *Escherichia coli*, the protease releases itself from larger precursors by proteolytic cleavages at its N and C termini. Both cleavage sites were determined by N-terminal sequence analysis of the cleavage products. Cleavage at the N terminus of the protease occurred with high efficiency at an EG dipeptide at positions 1108 and 1109. Cleavage at the C terminus of the protease occurred with low efficiency at an ET dipeptide at positions 1251 and 1252. To study the cleavage specificity of the protease, amino acid substitutions were introduced at the P2, P1, and P1′ positions at the cleavage site at the N-terminal boundary of the protease. This analysis showed that the amino acid at the P1 position is the most important determinant for substrate recognition. Only glutamic acid, glutamine, and aspartic acid were tolerated at this position. At the P1′ position, glycine, serine, and alanine were the preferred substrates of the protease, but a number of amino acids with larger side chains were also tolerated. Substitutions at the P2 position had only little effect on the cleavage efficiency. Cell-free expression of the C-terminal half of the ORF1 polyprotein showed that the protease catalyzes cleavage at the junction of the RNA polymerase and the capsid protein. An EG dipeptide at positions 1767 and 1768 was identified as the putative cleavage site. Our data show that rabbit hemorrhagic disease virus encodes a trypsin-like cysteine protease that is similar to 3C proteases with regard to function and specificity but is more similar to 2A proteases with regard to size.

Rabbit hemorrhagic disease virus (RHDV) is a member of the family *Caliciviridae* (39, 44, 51), which includes feline calicivirus, San Miguel sea lion virus, European brown hare syndrome virus, vesicular exanthema virus of swine and Norwalk virus, which causes disease in humans. Caliciviruses are nonenveloped RNA viruses that are characterized by a positive-stranded genomic RNA of 7.5 to 8 kb, a major capsid protein of 59 to 71 kDa, and a subgenomic RNA of 2.2 to 2.4 kb that serves as mRNA for the capsid protein (10, 11, 24, 25, 29, 33–37, 44, 57).

Caliciviruses have some features in common with picornaviruses and related plant viruses (potyviruses, comoviruses, and nepoviruses) and are therefore grouped within the superfamily of picornavirus-like positive-stranded RNA viruses (18). Like other members of the picornavirus superfamily, caliciviruses have a protein (VPg) covalently attached to the 5′ end of their genomic RNA (34). Furthermore, all members of the picornavirus superfamily encode a set of three conserved nonstructural proteins in their genomes: a nucleoside triphosphate-binding protein that presumably functions as an RNA helicase, a trypsin-like cysteine protease, and an RNA-dependent RNA polymerase. The coding sequences of the three conserved nonstructural proteins are contained in a large open reading frame and are arranged in the same order in the genomes of all picornavirus-like viruses. In the case of RHDV, the three proteins are encoded in the 5′ two-thirds of an open reading frame of 7 kb, ORF1, that contains not only the genes for the nonstructural proteins but also the coding sequence of the major capsid protein (Fig. 1). For feline calicivirus and Norwalk virus, two separate open reading frames, ORF1 and ORF2, that apparently replace ORF1 of RHDV have been described. The first open reading frame, ORF1, contains the genes for the nonstructural proteins, while the second open reading frame, ORF2, encodes the major capsid protein (25, 29, 35, 37).

Picornaviruses and their relatives among the plant viruses express their genetic information via synthesis of large polyproteins from which functional proteins are generated by co- and posttranslational proteolytic cleavages (17, 20, 41). The proteolytic cleavage reactions are catalyzed mostly by cysteine proteases that are themselves part of the polyproteins. These cysteine proteases are structurally homologous to the cellular serine proteases of the trypsin family (1, 3, 19, 31). On the basis of their sizes, viral trypsin-like cysteine proteases (TCPs) were divided into two classes, the large and the small TCPs (3). The small TCPs comprise the 2A proteases, which are encoded by entero- and rhinoviruses. These proteases are located immediately downstream of the capsid proteins and catalyze cleavage at their own N termini, which results in the release of the capsid protein precursor from the nascent polyprotein. The large TCPs comprise the 3C and 3C-like proteases, which are encoded by all members of the picornavirus superfamily. 3C proteases and their counterparts in the plant viruses are located immediately upstream of the putative RNA polymerases. These proteases and their precursors are responsible for most,

* Corresponding author. Mailing address: Federal Research Center for Virus Diseases of Animals, P.O. Box 1149, 72001 Tübingen, Germany.

† Present address: Institute of Molecular Pathology, A-1030 Vienna, Austria.

‡ Present address: Institute of Virology, University of Giessen, D-35392 Giessen, Germany.

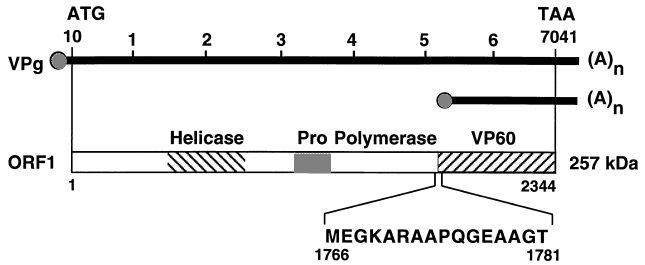

FIG. 1. Schematic diagram of the RHDV genome. The genomic and subgenomic RNAs of RHDV are depicted as black bars with VPg drawn as circles at their 5′ ends. The open reading frame ORF1 is depicted as an open bar, and the limits of ORF1 are indicated by vertical lines and the positions of the first and last nucleotides. The amino acid sequence of the primary translation product of the subgenomic RNA is shown from the N-terminal methionine at position 1766 in ORF1 to the threonine at position 1781.

and in some cases all, of the cleavage reactions that are required to release the different functional entities of the viral polyproteins. 3C and 2A proteases are not only different in their sizes, functions, and locations in the polyprotein but also differ in their cleavage specificities. Cleavage by 3C proteases and most 3C-like proteases occurs preferentially after glutamine and glutamic acid, whereas the 2A proteases of entero- and rhinoviruses cleave after alanine, valine, threonine, phenylalanine, and tyrosine (20, 41).

Recent studies have shown that the TCP of RHDV shares some features with both the 3C and 2A proteases (7). Like 3C proteases, the RHDV protease is located immediately upstream of the putative RNA polymerase. Moreover, sequence comparisons show that amino acids that are thought to be major determinants for the cleavage specificity of 3C proteases are conserved in the TCP of RHDV (1, 3, 7, 19, 31). These findings suggest that the RHDV protease resembles 3C proteases in cleavage specificity. However, the amino acid sequence comparisons also show that the RHDV protease is considerably shorter than 3C proteases and is more similar to 2A proteases with regard to size.

In the present study, various parts of the ORF1 polyprotein of RHDV were synthesized in *Escherichia coli* and in vitro to examine the proteolytic activity and substrate specificity of the TCP of RHDV. These experiments led to the identification of three cleavage sites in the C-terminal half of the polyprotein that are recognized by the TCP. In addition, a mutational analysis was performed at positions P2, P1, and P1′ of the cleavage site at the N-terminal boundary of the protease.

MATERIALS AND METHODS

Plasmid constructions and site directed mutagenesis. (i) Plasmids for expression in *E. coli*. Construction of recombinant plasmids was done by standard cloning procedures (46). Plasmid pEX-P was constructed by ligating the 920-bp *Eco*RI-*Hin*dIII fragment spanning codons 1023 to 1332 of ORF1 between the *Eco*RI and *Hin*dIII sites of plasmid pEX34B, which is identical to plasmid pEX31 except for a deleted *Pst*I site in the ampicillin resistance gene (54). As described previously, plasmid pEX-O contains a 1,039-bp *Eco*RI-*Xba*I fragment ligated between the *Eco*RI and *Xba*I sites of plasmid pEX-34B (7). To construct plasmid pEX-U, a 1,019-bp *Pst*I-*Xba*I cDNA fragment spanning codons 1032 to 1371 of ORF1 was ligated with the 1-kb *Bgl*I-*Pst*I fragment of plasmid pQE50 (Qiagen) and the 2.4-kb *Bgl*II-*Bgl*I fragment of plasmid pQE16 (Qiagen). Prior to ligation, the 5′ overhangs of the *Bgl*II and *Xba*I sites were filled in with Klenow polymerase. Site-directed mutagenesis was performed by the method of Kunkel with a mutagene phagemid in vitro mutagenesis kit (Bio-Rad) as described previously (7). Plasmids pCW55 and pCW247 were used to obtain the single-stranded template DNA for the mutagenesis reaction. Plasmid pCW55 contains a 1,018-bp *Hin*dIII-*Hin*dIII fragment spanning codons 991 to 1331 inserted into the *Hin*dIII site of pTZ18U (Bio-Rad). pCW247 contains a 2,000-bp *Sac*I-*Xba*I fragment (codons 703 to 1371) inserted between the *Sac*I and *Xba*I sites of plasmid pTZ19U (Bio-Rad). Oligonucleotides specifying single or multiple amino acid substitutions were synthesized on a Biosearch 8700 DNA synthesizer (New Brunswick Scientific, Edison, N.J.) by the phosphoamidite method (4). Mutant sequences were identified by sequence analysis (47). For further analysis of the mutants, the *Pst*I-*Kpn*I fragment (nucleotides 3105 to 3521) of the wild-type plasmid pEX-O was replaced by the corresponding fragments bearing the desired mutations. The substituted fragments were sequenced over the entire length to ensure the absence of second-site mutations.

(ii) Plasmids for in vitro expression. As described previously (7), plasmid p672 (in pBS SK+) contains an RHDV cDNA insert of 4,391 bp that extends from nucleotide 3077 to the poly(A) stretch at the 3′ end of the genome and includes the genes for the 3C-like protease, the RNA polymerase, and the VP60 capsid protein. In this construct, the first AUG and the stop codon correspond to amino acids 1023 and 2344, respectively, giving rise to a polyprotein of 143 kDa. Plasmid pM1 was derived from p672, in which the catalytic site of the 3C-like protease domain was mutated by replacement of the cysteine at position 1212 with a glycine. Plasmids pM2 and pM3 were derived from clone p13 and are similar to p672 but start more upstream, at nucleotide 2950 of ORF1, and end at the poly(A) site. Therefore, the first AUG corresponds to amino acid 981, resulting in a slightly larger polyprotein of 150 kDa. In pM2 and pM3 the potential cleavage target sites QG (positions 1775 and 1776) and EG (positions 1767 and 1768) were mutagenized to LG and VG, respectively. Starting from the p13 sequence, sequential PCRs with mutagenic oligonucleotide primers were performed to amplify a short fragment spanning the EG or QG sites. The resulting amplification products were sequenced over the entire length to ensure the absence of unwanted mutations. Subsequently, these fragments were used to replace the equivalent wild-type sequences of p13, giving rise to the plasmids pM2 and pM3.

Expression in *E. coli*. Expression of plasmids pEX-O and pEX-P and the mutant derivatives of pEX-O was carried out as previously described (7). Overnight cultures of transformed *E. coli* cells were induced by a temperature shift from 28 to 42°C. After an induction period of 2.5 h, the cells were lysed by the addition of lysozyme and Triton X-100. The lysate was sonicated, and insoluble material was collected by centrifugation. The pellet was washed with 1 M urea, and protein was extracted with 7 M urea. Aliquots of the 7 M urea extracts were then mixed with an equal volume of sample buffer (10% glycerin, 6 M urea, 2% sodium dodecyl sulfate [SDS], 5% β-mercaptoethanol, 62.5 mM Tris-HCl [pH 6.8], 0.01% bromophenol blue, 0.01% phenol red), boiled for 5 min at 95°C, and analyzed by SDS–12% polyacrylamide gel electrophoresis (SDS–12% PAGE) (27). After electrophoresis, proteins were visualized by staining with Coomassie blue (7). For analysis of the mutant derivatives of plasmid pEX-O, the volume of 7 M urea extract loaded onto the gels was adjusted for the difference in cell density of the *E. coli* cultures at the end of the induction period. For Western blot (immunoblot) analysis of total bacterial lysates, 0.5 ml of the uninduced and 1 ml of the induced bacterial cultures were harvested. The pelleted bacteria were boiled in sample buffer for 3 min at 95°C. Aliquots of the lysates were then run on an SDS–12% polyacrylamide gel. Transfer of the proteins to nitrocellulose membranes and incubation of the membranes with antiserum and substrate were performed as described previously (57). For expression of plasmid pEX-U, an overnight culture of transformed *E. coli* XL1 Blue cells (Stratagene) was diluted 50-fold into fresh Luria-Bertani medium, and growth was continued with selection (100 μg of ampicillin per ml) at 37°C for 2.5 h. The culture was then induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The bacteria were harvested after 1.5 h, and the pellet was resuspended in 6 M guanidinium hydrochloride (pH 8) by stirring at room temperature for 1 h. The lysate was loaded onto an Ni^{2+}-nitrilotriacetic acid-Sepharose column (Quiagen), and unbound protein was removed with 6 M urea (pH 8). Bound protein was desorbed from the column by successive elutions with 6 M urea (pH 6.3), 6 M urea (pH 5.5), and 6 M urea (pH 4.5). Aliquots of the eluted fractions were boiled in sample buffer and analyzed by SDS–12% PAGE.

Protein sequence analysis. For protein sequence analysis, samples were separated on SDS–12% polyacrylamide gels. Following electrophoresis, the gels were soaked in Western transfer buffer (10 mM 3-cyclohexylamino-1-propanesulfonic acid [CAPS] [pH 11], 20% methanol) for 20 min. Electrotransfer to a polyvinylidene difluoride membrane was then performed overnight at 100 V in a Hoefer TE72 Transphor apparatus. After transfer, the membrane was stained with 0.1% Ponceau S in 1% acetic acid for 1 min. The blot was destained with 1% acetic acid and rinsed in water. Sections of the membrane containing the relevant protein were excised and subjected to an automated Edman degradation. Alternatively, the samples were applied in duplicate on the same gel. After transfer, the membrane was cut into two strips. One strip was stained with copper phthalocyanine 3,4′,4″,4‴-tetrasulfonic acid (Aldrich) as described by Bickar and Reid (5) to identify the relevant protein band. The corresponding part of the unstained membrane was excised and used for protein sequence analysis.

In vitro transcription and translation. In vitro transcription with either T3 or T7 RNA polymerase was carried out as described by the manufacturer, using 1 μg of plasmid DNA linearized with the appropriate restriction enzyme. In vitro translation was in a volume of 25 μl containing 12.5 μl of rabbit reticulocyte lysate (Promega) and 2.5 μl of [^{35}S]methionine (Amersham) at 10 μCi/μl. The amount of template RNA added to the translation mixture was normalized to give equimolar quantities. The reaction mixture was incubated for 60 min at 30°C, unlabeled L-methionine was added, and two aliquots were removed and further incubated at 30°C in the presence or absence of extract from *E. coli* cells

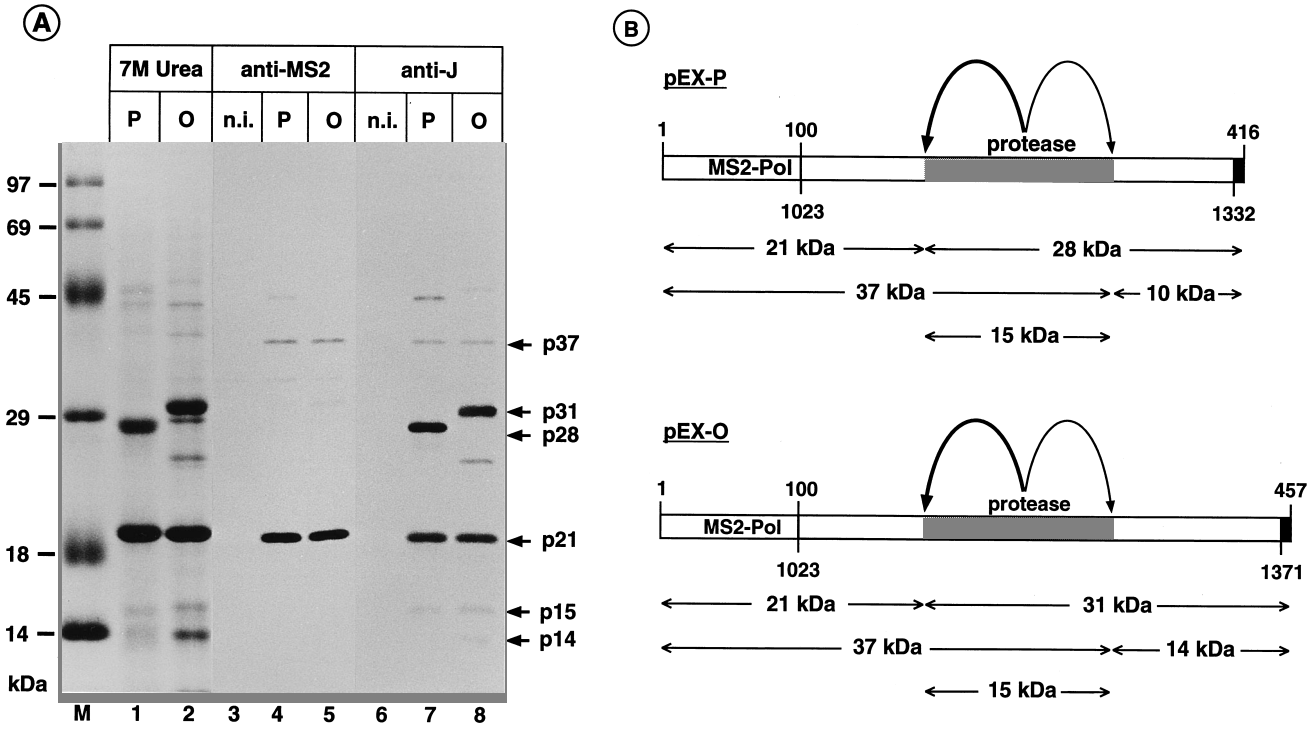

FIG. 2. Expression of the TCP of RHDV in *E. coli*. (A) *E. coli* cells harboring plasmid pEX-O (O) or pEX-P (P) were induced by a temperature shift to 42°C, and inclusion bodies were prepared after 2.5 h of induction. The inclusion bodies were extracted with 7 M urea, and the extracted proteins were separated on SDS–12% polyacrylamide gels. Proteins were visualized by staining with Coomassie blue. In parallel, whole-cell lysates of induced and uninduced (n.i.) bacteria were resolved on SDS–12% polyacrylamide gels and analyzed by Western blotting with antiserum against the MS2 polymerase and antiserum to fusion protein J as described in Materials and Methods. Lane M, marker proteins. (B) Schematic diagram of the cleavage products of fusion proteins O and P. The fusion proteins are depicted as bars, and the protease region of the RHDV part is indicated by stippled boxes. Arrows point to the positions of the cleavage sites at the N and C termini of the protease. The difference between the cleavage efficiencies at these sites is indicated by the thickness of the arrows. Numbers above the bars indicate the size of the MS2 part and the size of the entire fusion protein in amino acids. Numbers below the bars show the positions of the N- and C-terminal amino acids of the RHDV part. The locations of the observed cleavage products are indicated by horizontal arrows below the bars and the apparent molecular masses of the products are given. Pol, polymerase.

expressing the RHDV 3C-like protease. Samples were taken after 60 min and analyzed on SDS-polyacrylamide gels.

RESULTS

Identification of the cleavage site at the N terminus of the protease. We have shown previously that the 3C-like protease of RHDV exhibits proteolytic activity upon expression in *E. coli* and efficiently cleaves the ORF1 polyprotein at a site that has been tentatively mapped to the N-terminal boundary of the protease (7). To identify the exact position of the cleavage site, plasmids pEX-O and pEX-P were constructed. Plasmid pEX-O contains codons 1023 to 1371 of ORF1 fused to the N-terminal 98 codons of the RNA polymerase of bacteriophage MS2. Plasmid pEX-P differs from plasmid pEX-O in lacking codons 1333 to 1371 of ORF1. According to our previous work, the protease domain is located within amino acids 1100 to 1300 of the ORF1 polyprotein and therefore upstream of the deleted part. Both plasmids were expressed in *E. coli*. After 2.5 h of induction of *E. coli* cells harboring plasmid pEX-P, two abundant proteins of 21 and 28 kDa were extracted from inclusion bodies and identified by SDS-PAGE (Fig. 2A, lane 1). A protein of 46 kDa, which probably corresponds to the uncleaved fusion protein, was also detected. This protein was present in very small amounts, indicating almost complete cleavage of the fusion protein. Expression of plasmid pEX-O resulted in two major products of 21 and 31 kDa (Fig. 2A, lane 2). Again, only a small amount of the putative uncleaved fusion protein with an apparent molecular mass of 48 kDa was detectable. The 21-kDa protein was recognized by an antiserum against the MS2 polymerase part of the fusion proteins (Fig. 2A, lanes 4 and 5), showing that this protein represents the N-terminal cleavage product of fusion proteins O and P (Fig. 2B). The 28- and 31-kDa proteins represent the C-terminal cleavage products as shown by their reaction with antiserum against fusion protein J (Fig. 2A, lanes 7 and 8). This fusion protein comprises the C-terminal part of fusion protein P (amino acids 1172 to 1332 of the ORF1 polyprotein) and the N-terminal region of the MS2 polymerase. The anti-J serum therefore recognizes the 21-kDa N-terminal cleavage products and the C-terminal cleavage products of the fusion proteins O and P. Inactivation of the protease by substitution of the putative active-site residues abolished the formation of the 21-, 28-, and 31-kDa cleavage products, indicating that the generation of these products is due to self-cleavage of the protease (data not shown).

The expression products of plasmids pEX-O and pEX-P were separated by SDS-PAGE and transferred to polyvinylidene difluoride membranes to determine the cleavage site at the putative N terminus of the protease. The 28- and 31-kDa C-terminal cleavage products were identified by staining and subjected to an automated Edman degradation. The sequence GLPGF was obtained for both cleavage products. This sequence is found at positions 1109 to 1114 in the ORF1 polyprotein upstream of the putative active-site residues of the protease and is preceded by glutamic acid at position 1108. Cleavage at the N terminus of the protease thus occurs at the EG dipeptide at positions 1108 and 1109.

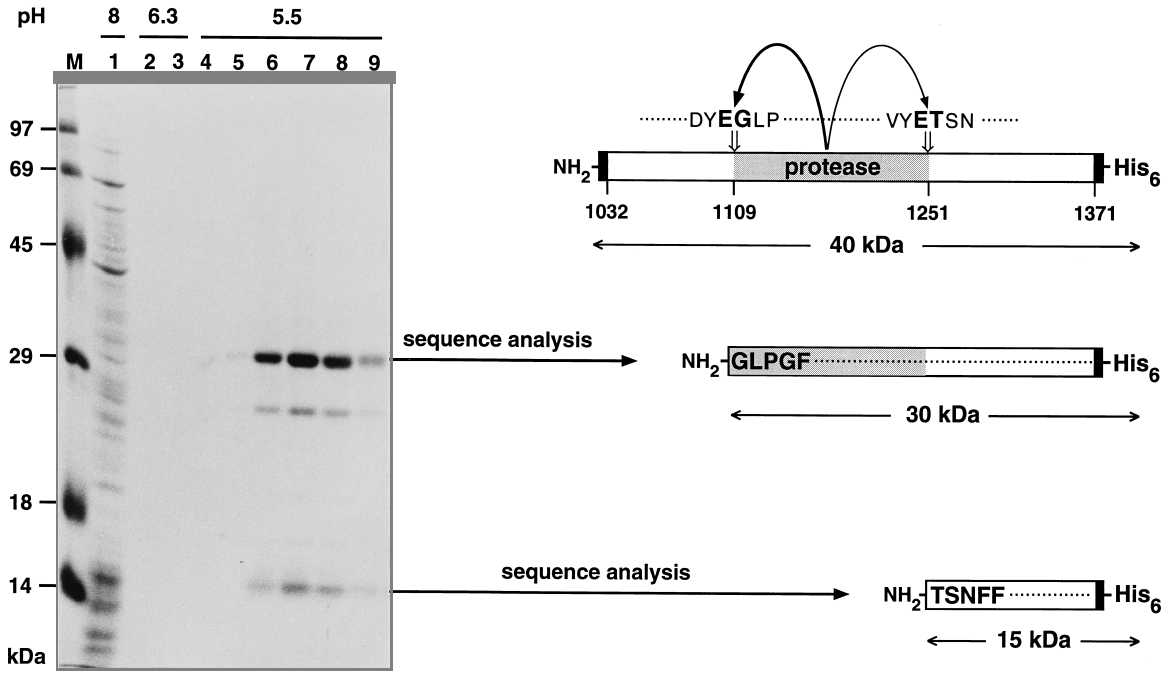

FIG. 3. Identification of the cleavage site at the C-terminal boundary of the TCP. (Left) *E. coli* cells harboring plasmid pEX-U were induced with IPTG for 1.5 h and subsequently lysed with guanidinium-hydrochloride. The lysate was loaded on an Ni^{2+}-nitrilotriacetic acid-Sepharose column, and unbound protein was removed by washing with 6 M urea (pH 8). Bound protein was then eluted by a stepwise reduction of pH. Successive fractions eluted at pH 8, 6.3, and 5.5 were separated by SDS-PAGE and stained with Coomassie blue. Lane M, marker proteins. (Right) Schematic diagram of the uncleaved fusion protein and the C-terminal cleavage products. The uncleaved fusion protein is depicted as a bar in the upper half of the panel. The protease is indicated by stippled boxes, and vector-derived sequences are indicated by black bars. Arrows point to the positions of the cleavage sites at the N and C termini of the protease. The amino acids flanking each cleavage site are given in the one-letter code. The difference between the cleavage efficiencies at these sites is indicated by the thickness of the arrows. Numbers below the bars indicate the positions of the N- and C-terminal amino acids of the RHDV part and the positions of the N- and C-terminal amino acids of the protease. The lower part of the panel shows the location of the C-terminal cleavage products. The N-terminal amino acid sequences of the cleavage products are given in the single-letter code within the bars.

Identification of the cleavage site at the C terminus of the protease. The Western blot analysis of the expression products of plasmids pEX-O and pEX-P revealed some minor bands in addition to the major bands, which arise from cleavage at the N terminus of the protease. A 37-kDa protein was generated from both fusion proteins. This protein was recognized by the MS2 antiserum (Fig. 2A, lanes 4 and 5) and the anti-J serum (lanes 7 and 8), indicating that it represents an N-terminal cleavage product which extends beyond the cleavage site at the N terminus of the protease (Fig. 2B). The 37-kDa band was not observed upon expression of an inactive protease mutant (data not shown). From this finding it could be concluded that the 37-kDa band is generated not by bacterial proteases but by self-cleavage of the protease at its own C terminus. Cleavage of fusion protein O at this site also led to a 14-kDa C-terminal cleavage product (Fig. 2B) which was detected by the anti-J serum (Fig. 2A, lane 8). The corresponding cleavage product of fusion protein P with an expected molecular mass of 10 kDa was not observed, probably because it migrated out of the gel. A minor band of 15 kDa was recognized by the anti-J serum but not by the anti-MS2 serum. This protein was generated from both fusion proteins (Fig. 2A, lanes 7 and 8) and therefore represents an internal cleavage product that is cleaved at its N and C termini. The 15-kDa protein most likely represents the mature protease (Fig. 2B).

In order to determine the position of the C-terminal cleavage site, expression plasmid pEX-U was constructed (Fig. 3). This plasmid contains a cDNA fragment that covers the complete protease gene and extends from codon 1032 to codon 1371. Six histidine codons followed by a stop codon were fused in frame to the 3′ end of this cDNA fragment. The hexahistidine tail was included to allow affinity purification of the expected C-terminal cleavage products of the polyprotein encoded in plasmid pEX-U. The expression products of plasmid pEX-U were extracted from induced *E. coli* cells under denaturing conditions and applied to an Ni^{2+}-metal chelate column. Bound protein was eluted by a pH step gradient. Three proteins with molecular masses of 30, 26, and 15 kDa were specifically retained by the column and eluted at pH 5.5 (Fig. 3, lanes 4 to 9). A microsequence analysis of the 30-kDa protein showed that this protein was generated by cleavage at the EG dipeptide at positions 1108 and 1109. It should be noted that cleavage at this site was complete as indicated by the absence of the uncleaved protein, which has a calculated molecular mass of 40 kDa. N-terminal sequence analysis of the 15-kDa protein led to the amino acid sequence TSNFF. This sequence occurs only once in the ORF1 polyprotein, at positions 1252 to 1256, and is preceded by glutamic acid at position 1251. Provided that no further processing occurs after the primary cleavage, the C terminus of the protease is generated by cleavage at the ET dipeptide at positions 1251 and 1252.

Of the amino acids that constitute the catalytic triad of trypsin-like proteases, the nucleophilic serine or cysteine is located most closely to the C termini of the enzymes. In 2A proteases and 3C proteases of enteroviruses, the distance between the catalytic cysteine and the C-terminal amino acid is in the range of 36 to 40 residues. In the RHDV protease, the glutamic acid at position 1251 is located 39 amino acids downstream of the putative catalytic cysteine, consistent with our

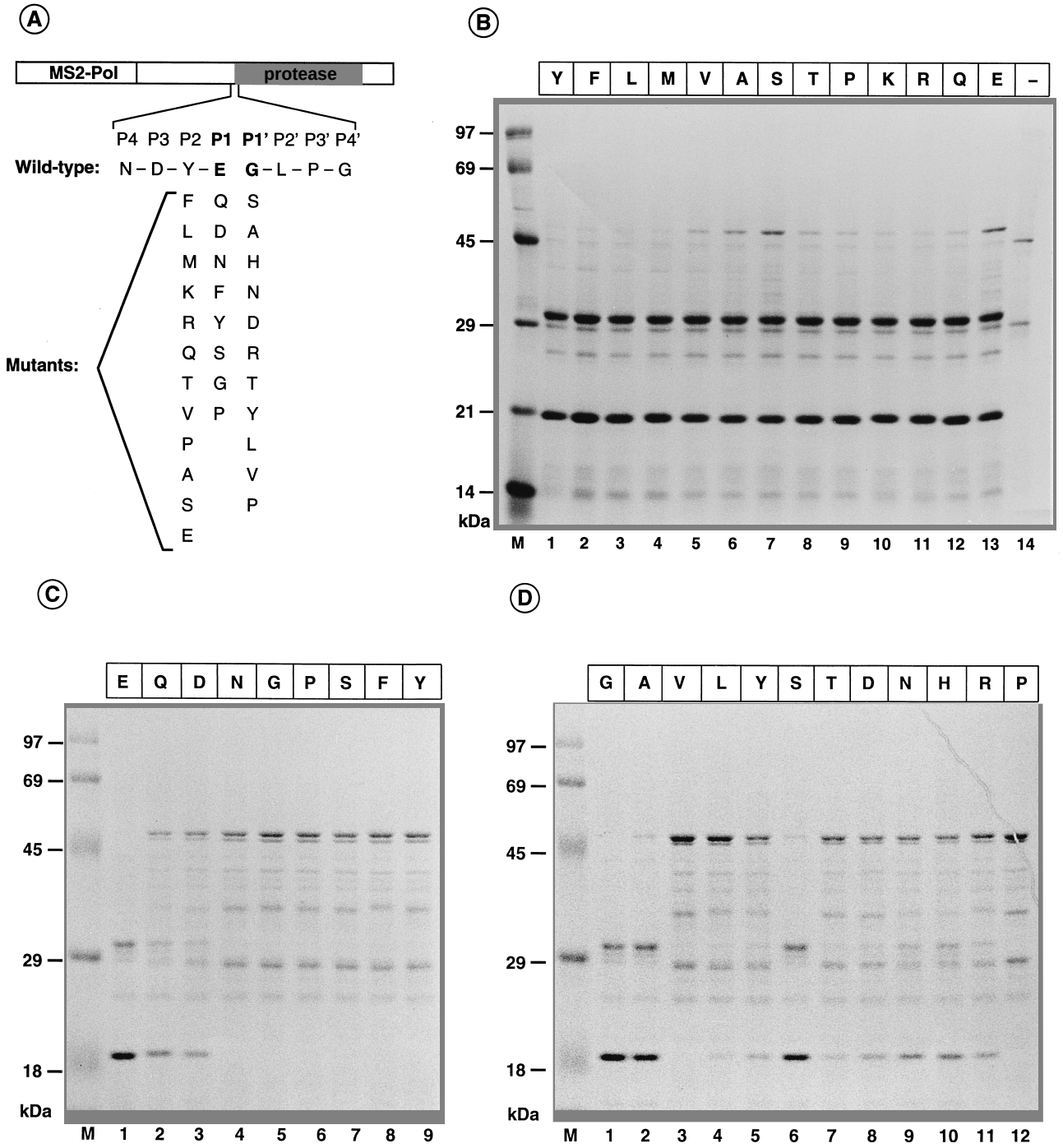

FIG. 4. Site-directed mutagenesis of the cleavage site at the N terminus of the protease and expression of the cleavage site mutants in *E. coli*. (A) Schematic diagram of the cleavage site mutants. Individual amino acid substitutions were introduced at the P2, P1, and P1′ positions of the cleavage site at the N terminus of the protease. The sequence of the wild-type protein is shown from position P4 to P4′, and amino acid replacements are indicated below this sequence. The single-letter code for amino acids is used. Pol, polymerase. (B to D) Expression of cleavage site mutants in *E. coli. E. coli* cells harboring plasmid pEX-O and derivatives of pEX-O encoding single amino acid substitutions at position P2 (B), position P1 (C), and position P1′ (D) of the cleavage site at the N-terminal boundary of the 3C-like protease were induced by a temperature shift to 42°C, and inclusion bodies were prepared after 2.5 h of induction. The inclusion bodies were extracted with 7 M urea, and the extracted proteins were separated by SDS-PAGE and stained with Coomassie blue. Amino acids at the relevant positions are indicated above the lanes. Lane 14 in panel B shows a 7 M urea extract of *E. coli* cells that did not harbor an expression plasmid. Lanes M, marker proteins.

conclusion that the ET dipeptide constitutes the C-terminal boundary of the protease.

In addition to the expected cleavage products, a third protein of 26 kDa was specifically retained by the Ni^{2+}-chelate column (Fig. 3, lanes 6 to 9). Attempts to sequence this protein failed. The origin of this protein therefore remained unclear. It may represent a bacterial contaminant or a cleavage product of the fusion protein that is due to proteolytic degradation by

bacterial proteases or aberrant self-cleavage of the protease. This protein may also correspond to an additional cleavage product with an unknown function, although this is unlikely.

Cleavage specificity of the protease. Cleavage sites that are recognized by viral proteases are partially defined by sequence determinants in the immediate vicinity of the scissile bond. In general, the most stringent sequence requirements are observed at the P1 position. Additional sequence determinants are usually located upstream of the scissile bond (positions P2 to P8) and at position P1′ on the C-terminal side of the scissile bond. On the basis of these observations, we decided to examine the substrate specificity of the protease at the P1 and P1′ positions by introducing amino acid substitutions at the cleavage site at the N terminus of the protease. The P2 position was included in this study because it is the only position upstream of P1 that is strictly conserved for both cleavage sites at the N- and C-terminal boundaries of the protease. The amino acid substitutions that were analyzed in this study are shown in Fig. 4A. Mutated cDNA fragments containing single-codon substitutions were transferred into plasmid pEX-O for further analysis of the mutants. The resulting derivatives of pEX-O were expressed in *E. coli* cells, and the expression products were extracted from inclusion bodies and analyzed by SDS-PAGE. The effect of the mutations on cleavage efficiency was monitored by the accumulation of the uncleaved fusion protein, which has an apparent molecular mass of 48 kDa, and the concomitant reduction in the amounts of the 21- and 31-kDa cleavage products. A 7 M urea extract of induced *E. coli* cells harboring the parental plasmid pEX-O was used for comparison, and an extract of *E. coli* cells lacking an expression plasmid served as a control.

Replacement of the tyrosine at position P2 of the wild-type protein had only a small effect on cleavage efficiency (Fig. 4B). Nearly 100% cleavage was observed for the phenylalanine, leucine, and methionine mutants. The other mutants showed a slight reduction in cleavage efficiency as indicated by the accumulation of the uncleaved fusion protein. In general, inhibition of cleavage was stronger for mutants carrying small and polar amino acids at position P2. The greatest reduction in the extent of cleavage was observed for the serine and glutamic acid mutants (Fig. 4B, lanes 7 and 13). The relatively large decrease in the cleavage efficiency of the mutant carrying glutamic acid at position P2 is probably due to a competition between this amino acid and the following glutamic acid for the P1 binding site of the protease. In summary, the analysis indicated a preference for large and hydrophobic amino acids at position P2.

Amino acid substitutions at positions P1 and P1′ caused a significant reduction in cleavage efficiency up to complete inhibition of cleavage. At position P1, the protease accepted only glutamic acid, glutamine, and aspartic acid as substrates (Fig. 4C, lanes 1 to 3). Of these three amino acids, glutamic acid was the most- and aspartic acid the least-suited substrate for the protease. Cleavage of the glutamine mutant was less efficient than cleavage of the wild-type protein but significantly better than cleavage of the aspartic acid mutant. No cleavage was detected for the mutants with the other amino acid replacements at position P1, including the asparagine mutant. These data showed that efficient cleavage strongly depends on both the size of the side chain and the carboxylate or carboxamide group of the amino acid at position P1. Furthermore, the data indicate a more favorable interaction of the protease with the carboxyl groups of glutamic acid and aspartic acid than with the carboxamide groups of glutamine and asparagine.

At position P1′, glycine allowed the highest cleavage efficiency (Fig. 4D). A slight reduction in the extent of cleavage was observed for the alanine and serine mutants (lanes 2 and 6). All other mutants exhibited a more pronounced inhibition of cleavage. Complete inhibition of cleavage was observed for the proline and valine mutants (lanes 3 and 12). In summary the following order of relative cleavage efficiency was apparent at position P1′: G > A, S > H, N > D, R > T, Y > L > P, V. This order revealed for position P1′ a strong preference for amino acids with small side chains and a weaker preference for polar amino acids.

It is apparent from Fig. 4C and D that inhibition of cleavage at the natural cleavage site by certain amino acid substitutions resulted in the appearance of additional bands not observed for the wild-type protein. In particular, two proteins of 28 and 34 kDa were produced in considerable amounts. Both proteins were recognized by MS2 antiserum (data not shown) indicating that they represent N-terminal cleavage products of the fusion proteins. Neither of the two proteins was produced when the protease was inactivated by replacement of the putative active-site residues (not shown). These findings suggested that the 28- and 34-kDa proteins are aberrant cleavage products that were generated by self-cleavage of the protease at cryptic cleavage sites. Indeed, several potential cleavage sites located within the protease region (ES at positions 1159 and 1160, EG at positions 1169 and 1170, and DS at positions 1218 and 1219) may account for the observed cleavage products. The amount of the aberrant cleavage products increased in parallel to the inhibition of cleavage at the N-terminal boundary of the protease, whereas the extent of cleavage at the C-terminal boundary of the protease was significantly reduced in mutants that exhibited impaired cleavage at the N-terminal boundary (not shown). This observation suggested that cleavage at the N terminus of the protease results in a shift in conformation of the protease that is a prerequisite for cleavage at the ET dipeptide. Alternatively, the impaired cleavage at the protease-polymerase junction may be explained by improper folding of the polypeptide chain due to the presence of the amino acid substitution at the N-terminal boundary of the enzyme. The same argument may also be applied to explain the results of the cleavage site mutations; impaired cleavage at the N-terminal boundary of the protease may be due to improper folding of the protease instead of reduced recognition by the substrate binding pocket. However, this seems unlikely to us, since substitution of the amino acid at position P2 has almost no effect on recognition of the natural cleavage sites.

Taken together, our results indicate that the sequence EG is the optimal substrate of the TCP of RHDV. Furthermore, our data show that the amino acid at position P1 is the most important sequence determinant for substrate recognition. It is also obvious that the requirement for glutamic acid at position P1 and glycine at position P1′ is not absolute. The protease may accept glutamine and aspartic acid at position P1 and various amino acids at position P1′ with a concomitant reduction in cleavage efficiency compared with that for the optimal sequence EG. Finally, our data show that the optimal sequence (YEG) has been selected at the N-terminal boundary of the protease, whereas the suboptimal sequence YET is present at the C-terminal boundary, consistent with the difference in cleavage efficiency at the two sites.

Identification of the cleavage site between the putative polymerase and the capsid protein VP60. We have shown previously that the TCP of RHDV is able to cleave the viral polyprotein at the boundary between the putative polymerase and the capsid protein. The cleavage gives rise to a capsid protein species of 60 kDa that is indistinguishable in size from the translation product of the subgenomic RNA. This finding indicates that the cleavage site is located close to the first

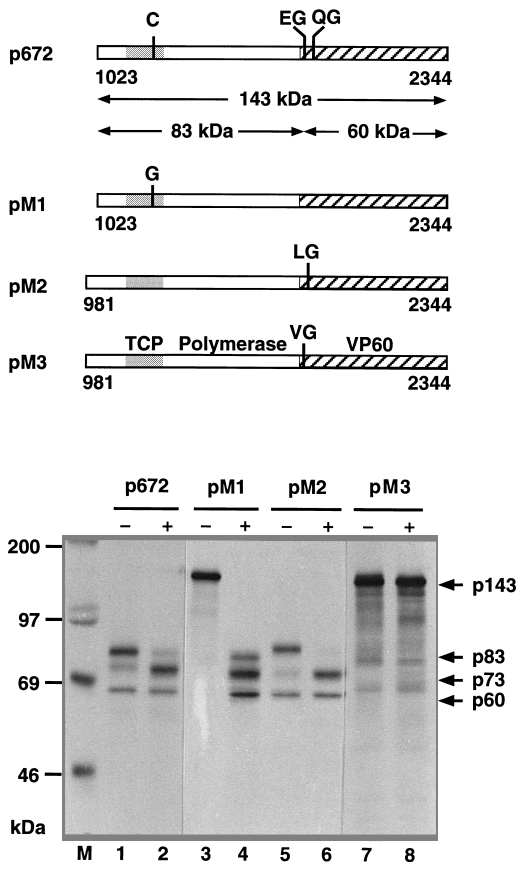

FIG. 5. Identification of the cleavage site between the putative polymerase and the capsid protein VP60. In vitro transcripts were prepared from plasmids p672, pM1, pM2, and pM3 and translated in a rabbit reticulocyte lysate in the presence of [^{35}S]methionine. After 1 h, the translation reaction was chased with unlabeled L-methionine, and two aliquots were removed and further incubated in the absence (−) or presence (+) of exogenous protease expressed in *E. coli*. After 60 min, samples were removed and analyzed on SDS-polyacrylamide gels. Lane M, marker proteins.

methionine encoded by the subgenomic RNA (M-1766), which is the presumed N-terminal amino acid of the major capsid protein species. Two potential cleavage sites are in close proximity to the methionine at position 1766. Both of these sites fulfill the sequence requirements for efficient cleavage. The first site, an EG dipeptide at positions 1767 and 1768, immediately follows the methionine. The second site, a QG dipeptide, is located seven amino acids downstream of the EG dipeptide. To determine if cleavage occurs at one of these sites, two plasmids (pM2 and pM3) in which the glutamine codon and the glutamic acid codon were changed individually by site-directed mutagenesis into a leucine codon and a valine codon, respectively, were constructed (Fig. 5). A third mutant (pM1), in which the protease was inactivated by a change of the cysteine codon at position 1212 into a glycine codon, was also used in this study. All three mutant constructs were derived from the parental plasmid p672 or p13, each of which contains the protease gene, the polymerase gene, and the complete capsid protein gene. Figure 5 shows the in vitro translation products from each plasmid after separation by SDS-PAGE. Expression of the parental plasmid p672 gave rise to two major proteins of 60 and 83 kDa, which represent the fully processed capsid protein VP60 and the remaining part of the polyprotein containing the protease fused to the polymerase, respectively (Fig. 5, lane 1). In addition, a less abundant product of 73 kDa was also visible. This protein most likely originates from further processing of the 83-kDa protein at the EG at positions 1108 and 1109, which is located 90 amino acids downstream of the N terminus of the polyprotein encoded by plasmid p672. Incubation of the p672 in vitro translation reaction mixture with the RHDV 3C-like protease expressed in *E. coli* resulted in an increase in the amount of the 73-kDa band, while the 83-kDa protein decreased in parallel (Fig. 5, lane 2). In contrast, no change in the amount of the cleavage product VP60 was observed. These findings showed that the fused protease-polymerase protein generated in this assay does not cleave the EG at positions 1108 and 1109 as efficiently as it does the cleavage site at the boundary between the polymerase and the capsid protein. Interestingly, in this assay no processing was detectable at the C terminus of the protease. As expected, in vitro expression of plasmid pM1, which encoded the inactive protease, showed no processing of the 144-kDa polyprotein (Fig. 5, lane 3). However, cleavage could be restored completely by addition of the protease expressed in *E. coli* (lane 4). The same cleavage products as for p672 were obtained from plasmid pM2, which encodes the glutamine-to-leucine exchange at the putative QG cleavage target site (Fig. 5, lanes 5 and 6). In contrast, the replacement of the glutamic acid codon at position 1767 with valine (plasmid pM3) (lane 7) completely abolished cleavage at the polymerase-capsid protein boundary. The cleavage could not be restored even after addition of exogenous protease (Fig. 5, lane 8), indicating that the lack of proteolysis was not due to an altered activity of the protease present in the translation product. Taken together, these data strongly suggest that cleavage between the polymerase and the capsid protein occurs at the EG dipeptide at positions 1767 and 1768.

In summary, we have identified three natural cleavage sites in the ORF1 polyprotein of RHDV that constitute the N- and C-terminal boundaries of the protease and the boundary between the putative polymerase and the VP60 capsid protein. It should be kept in mind, however, that bacterial and cell-free expression systems were used in this study, since RHDV cannot yet be propagated in cell culture. We therefore cannot exclude the possibility that in vivo cleavage differs in some aspects from the results reported here, although it is unlikely that cleavage sites other than the ones reported here are recognized in the C-terminal half of the ORF1 polyprotein by the TCP in vivo.

DISCUSSION

In a previous report, we have shown that the ORF1 polyprotein of RHDV contains a TCP that is involved in the proteolytic processing of the polyprotein. In the present study a mutational analysis was performed to study the substrate specificity of the protease. In addition, three natural cleavage sites in the C-terminal half of the ORF1 polyprotein that are recognized by the TCP were identified. The results of these studies show that the RHDV protease, like many other viral proteases, tolerates only a limited number of amino acids at the P1 position (8, 12, 15, 16, 32, 38, 42, 45). Aside from glutamic acid, which is the optimal substrate at the P1 position, cleavage was observed only after glutamine (Gln) and aspartic acid (Asp), whereas no cleavage was detectable when asparagine (Asn) was present at the P1 position. This finding indicates a specific interaction of the substrate binding pocket of the protease with the carboxylic groups of Glu and Asp that is more favorable than the interaction with the carboxamide groups of glutamine (Gln) and asparagine. Furthermore the lack of cleavage after

asparagine strongly suggests that amino acids other than Glu, Gln, and Asp are not accepted by the protease at position P1. Direct experimental proof for this conclusion for several amino acids (G, P, S, F, and Y) was obtained in this study.

Preferred cleavage after glutamic acid and glutamine is a characteristic feature of picornaviral 3C proteases and most 3C-like proteases of the picornavirus-like plant viruses (17, 20, 41). In contrast to the RHDV protease, most 3C and 3C-like proteases appear to prefer glutamine over glutamic acid at the P1 position. The specificity for glutamine is attributed to specific interactions of the amino acid at position P1 with two conserved amino acids in the substrate binding pocket of 3C proteases. These amino acids, a threonine upstream of the nucleophilic cysteine and a histidine downstream of the cysteine, are capable of forming hydrogen bonds with the carboxamide group of the glutamine at position P1, as shown by the X-ray structural analysis of the 3C protease of human rhinovirus (31). In the RHDV protease, both amino acids are conserved at the corresponding positions in the primary sequence. It has already been shown by site-directed mutagenesis that the histidine is important for proteolytic activity. Thus, our data strongly support the notion that the histidine is important for substrate recognition and specificity. However, for final proof of this hypothesis, X-ray structural analysis of the protease will be required. Such an analysis should also reveal the structural basis for the preferred cleavage after glutamic acid instead of glutamine.

At the P1′ position, 10 amino acids were selected to replace the glycine in the wild-type sequence. Two of the mutations, namely, the changes to alanine and serine, had almost no effect on cleavage efficiency, whereas all other amino acid replacements inhibited cleavage considerably. The extent of cleavage was comparable for acidic (D), basic (R), aromatic (Y), and aliphatic (L) amino acids. Nearly complete inhibition of cleavage was observed for the proline and valine mutants. Taken together, these results suggest that the substrate specificity of the RHDV protease at the P1′ position, in contrast to that at the P1 position, is not due to specific interactions of the P1′ amino acid with the substrate binding pocket but is dictated mainly by steric hindrance and conformational constraints.

Several viral proteases exhibit a strong specificity for certain amino acids at the P2 position of a cleavage site. For example, the 2A protease of poliovirus prefers amino acids with small side chains at the P2 position, and the Sindbis virus Nsp2 protease exhibits a pronounced specificity for glycine (22, 50). Our results indicate that the amino acid at the P2 position is only a minor determinant for cleavage specificity of the RHDV protease. This finding was surprising, because the P1 and P2 positions show the strongest conservation of all cleavage sites (Fig. 6). However, we cannot exclude the possibility that the protease exhibits a particularly relaxed sequence specificity in the assay system used. It is also unclear whether any other position further upstream of the scissile bond is an important determinant for efficient recognition by the TCP, as described for a large number of viral proteases (6, 9, 15, 16, 28, 40, 45, 49, 53, 56). Further studies with other assay systems including additional cleavage sites will be required to completely define the features of the primary sequence that define a cleavage site. Nonetheless, the available data should allow the identification of potential cleavage sites in the N-terminal half of the ORF1 polyprotein that might be recognized by the RHDV protease.

Our data show that cleavage of the EG dipeptide bond at the N-terminal boundary of the protease is much more efficient than cleavage of the ET dipeptide bond at the protease-polymerase junction. Since the P1′ position is clearly an important determinant for substrate recognition, it is likely that the exchange of glycine for threonine is mainly responsible for the slow cleavage at the C-terminal boundary of the protease. However, the presence of the threonine at the P1′ position is probably not the sole factor that contributes to the slow cleavage at the protease-polymerase junction, since replacement of the EG dipeptide at the N terminus of the protease by an ET dipeptide did not reduce cleavage efficiency at the altered site to the low level of cleavage that was observed at the C-terminal cleavage site. It remains to be determined if other divergent positions in the vicinity of the ET dipeptide contribute to the inefficient cleavage at the protease-polymerase junction.

	P7	P6	P5	P4	P3	P2	P1	P1'	P2'	P3'	P4'	P5'
X - TCP	F	Y	D	N	D	**Y**	**E**	**G**	L	P	G	F
TCP - Pol	I	T	K	G	V	**Y**	**E**	**T**	S	N	F	F
Pol - VP60	E	F	V	N	V	**M**	**E**	**G**	K	A	R	A

FIG. 6. Comparison of the amino acid sequences flanking the three cleavage sites identified in this study. The sequences are shown in the one-letter code from position P7 to position P5′. The location of each cleavage site within the RHDV polyprotein is indicated to the left of the sequences. Pol, polymerase; VP60, capsid protein; X, protein that precedes the protease in the ORF1 polyprotein.

The varying of rates of cleavage is an important means by which viruses utilizing polyproteins for gene expression regulate the levels of individual gene products. This regulation is particularly obvious in cases in which long-lived precursors are produced that serve essential functions in the replication cycle distinct from those of their constituent proteins that are eventually produced by proteolytic cleavage of the precursor. In the poliovirus system, for example, it is well documented that the 3CD protease-polymerase precursor is much more active in proteolytic processing of the capsid protein precursor than the mature protease (26, 58). On the other hand, cleavage of the 3CD protein is required to activate the polymerase portion 3D of the 3CD protein (21). Similar to the case for the RHDV system, the 3CD precursor of poliovirus accumulates, because cleavage at the N terminus of the 3C protease is more efficient than cleavage at its C terminus. Although the exact functions of the corresponding RHDV proteins remain to be fully elucidated, it seems likely that the slow cleavage of the protease-polymerase precursor of RHDV might be used in a way similar to that of the corresponding 3CD poliovirus protein to activate or inactivate some function that is associated with the precursor protein or its cleavage products.

In our previous studies we have demonstrated that the TCP of RHDV cleaves at the boundary between the polymerase and the capsid protein. It is shown here that this cleavage occurs at an EG dipeptide. This dipeptide immediately follows the N-terminal methionine of the primary translation product of the subgenomic RNA. Parra et al. (43) concluded from a microsequence analysis that the major capsid protein species from mature virions carries the initiator methionine at its N terminus. From that observation together with the data reported here, it seems likely that the primary translation product of the subgenomic RNA is not cleaved at the EG dipeptide and that the major capsid protein species is actually derived from the subgenomic RNA. Many viral proteases require a minimum of four amino acids on the N-terminal side of a scissile bond to cleave efficiently at this site (13, 14, 23, 30, 48, 52, 55). In the primary translation product of the subgenomic RNA, only one amino acid is located upstream of the EG dipeptide. The failure of the protease to cleave the translation product of the subgenomic RNA is therefore probably due to the short dis-

tance between the scissile bond and the amino terminus of the protein.

The data presented here imply that two different capsid protein species that have different N termini are generated. One capsid protein species is generated by translation of the genomic RNA and subsequent cleavage by the 3C-like protease at the junction of the putative polymerase and the capsid protein. This capsid protein species is two amino acids shorter than the other capsid protein, which is generated by translation of the subgenomic RNA. At present it is not known if the shorter capsid protein species is incorporated into virions. The smaller protein has not been demonstrated so far, but it may have been overlooked because of its low abundance and the fact that it has nearly the same size as the translation product of the subgenomic RNA.

Despite all of the similarities between other 3C-like proteases and the RHDV TCP with regard to function and substrate specificity, one important difference has to be mentioned. The results presented here show that the TCP of RHDV is considerably smaller than 3C and 3C-like proteases. From the N-terminal glycine at position 1109 to the C-terminal glutamic acid at position 1251, the RHDV protease comprises 143 amino acids, whereas known 3C proteases are composed of more than 180 amino acids. The TCP of RHDV is even smaller than 2A proteases, which are composed of 146 to 149 amino acids. To our knowledge, the TCP of RHDV is the smallest known trypsin-like protease.

Recent studies indicate that the 3C protease of poliovirus is able to bind to the 5′ end of the viral RNA. This interaction appears to be required for initiation of minus-strand RNA synthesis (2). It is apparent from our previous amino acid comparisons (7) that many of the residues that have been implicated in the RNA binding activity of the poliovirus and rhinovirus 3C proteases (2, 31) are absent from the TCP of RHDV and from 2A proteases. The absence of these residues accounts at least in part for the smaller size of the RHDV protease and 2A proteases compared with 3C proteases. It therefore could be that the RHDV TCP fulfills only part of the functions identified for picornaviral 3C proteases. Certainly, it will be interesting to elucidate the three-dimensional structure of the RHDV protease and to determine if the protease exhibits RNA binding activity similar to that of the poliovirus 3C protease.

ACKNOWLEDGMENTS

This study was supported by grant Th 298/3-1 from the Deutsche Forschungsgemeinschaft and by the Consiglio Nazionale delle Ricerche (Biotecnologie e Biostrumentazione programme). M.B.B. was the recipient of a fellowship from the Fondazione Iniziative Zooprofilattiche, Brescia, and also received support from Telethon, Italy (project 89).

REFERENCES

1. **Allaire, M., M. M. Chernaia, B. A. Malcolm, and M. N. G. James.** 1994. Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteinases. Nature (London) **369:**72–76.
2. **Andino, R., G. E. Rieckhoff, P. L. Achacoso, and D. Baltimore.** 1993. Poliovirus RNA synthesis utilizes an RNP complex formed around the 5′-end of viral RNA. EMBO J. **12:**3587–3598.
3. **Bazan, J. F., and R. J. Fletterick.** 1988. Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: structural and functional implications. Proc. Natl. Acad. Sci. USA **85:**7872–7876.
4. **Beaucage, S. L., and M. H. Caruthers.** 1981. Deoxynucleoside phosporamidites: a new class of key intermediates for deoxynucleotide synthesis. Tetrahedron Lett. **22:**1859–1862.
5. **Bickar, D., and P. D. Reid.** 1992. A high-affinity protein stain for Western blots, tissue prints and electrophoretic gels. Anal. Biochem. **203:**109–115.
6. **Blair, W. S., and B. L. Semler.** 1991. Role for the P4 amino acid residue in substrate utilization by the poliovirus 3CD proteinase. J. Virol. **65:**6111–6123.
7. **Boniotti, B., C. Wirblich, M. Sibilia, G. Meyers, H.-J. Thiel, and C. Rossi.** 1994. Identification and characterization of a 3C-like protease from rabbit hemorrhagic disease virus, a calicivirus. J. Virol. **68:**6487–6495.
8. **Carrington, J. C., and W. G. Dougherty.** 1988. A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. Proc. Natl. Acad. Sci. USA **85:**3391–3395.
9. **Carrington, J. C., and K. L. Herndon.** 1992. Characterization of the potyviral HC-Pro autoproteolytic cleavage site. Virology **187:**308–315.
10. **Carter, M. J.** 1994. Genomic organization and expression of astroviruses and caliciviruses. Arch. Virol. (Suppl.) **9:**429–439.
11. **Carter, M. J., I. D. Milton, J. Meanger, M. Bennett, R. M. Gaskell, and P. C. Turner.** 1992. The complete nucleotide sequence of a feline calicivirus. Virology **190:**443–448.
12. **Chambers, T. J., A. Nestorowicz, and C. M. Rice.** 1995. Mutagenesis of the yellow fever virus NS2B/3 cleavage site: determinants of cleavage site specificity and effects on polyprotein processing and viral replication. J. Virol. **69:**1600–1605.
13. **Cordingley, M. G., P. L. Callahan, V. V. Sardana, V. M. Garsky, and R. J. Colonno.** 1990. Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro. J. Biol. Chem. **265:**9062–9065.
14. **DiIanni, C. L., C. Mapelli, D. A. Drier, J. Tsao, S. Natarajan, D. Riexinger, S. M. Gestion, M. Bolgar, G. Yamanaka, and S. P. Weinheimer.** 1993. In vitro activity of the herpes simplex virus type 1 protease with peptide substrates. J. Biol. Chem. **269:**25440–25454.
15. **Dong, S., and S. C. Baker.** 1994. Determinants of the p28 cleavage site recognized by the first papain-like cysteine proteinase of murine coronavirus. Virology **204:**541–549.
16. **Dougherty, W. G., J. C. Carrington, S. M. Cary, and T. D. Parks.** 1988. Biochemical and mutational analysis of a plant virus polyprotein cleavage site. EMBO J. **7:**1281–1287.
17. **Goldbach, R.** 1990. Plant viral proteinases. Semin. Virol. **1:**335–346.
18. **Goldbach, R., and J. Wellink.** 1988. Evolution of plus-strand RNA viruses. Intervirology **29:**260–267.
19. **Gorbalenya, A. E., A. P. Donchenko, V. M. Blinov, and E. V. Koonin.** 1989. Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. FEBS Lett. **243:**103–114.
20. **Harris, K. S., C. U. T. Hellen, and E. Wimmer.** 1990. Proteolytic processing in the replication of picornaviruses. Semin. Virol. **1:**323–333.
21. **Harris, K. S., S. R. Reddigari, M. J. Nicklin, T. Hämmerle, and E. Wimmer.** 1992. Purification and characterization of poliovirus polypeptide 3CD, a proteinase and a precursor for RNA polymerase. J. Virol. **66:**7481–7489.
22. **Hellen, C. U., C.-K. Lee, and E. Wimmer.** 1992. Determinants of substrate recognition by poliovirus 2A proteinase. J. Virol. **66:**3330–3338.
23. **Jewell, D. A., W. Swietnicki, B. M. Dunn, and B. A. Malcolm.** 1992. Hepatitis A virus 3C protease substrate specificity. Biochemistry **31:**7862–7869.
24. **Jiang, X., M. Wang, D. Y. Graham, and M. K. Estes.** 1992. Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein. J. Virol. **66:**6527–6532.
25. **Jiang, X., M. Wang, K. Wang, and M. K. Estes.** 1993. Sequence and genomic organization of Norwalk virus. Virology **195:**51–61.
26. **Jore, J., B. De Geus, R. J. Jackson, P. H. Pouwels, and B. E. Enger-Valk.** 1988. Poliovirus protein 3CD is the active protease for processing of the precursor protein P1 in vitro. J. Gen. Virol. **59:**1627–1636.
27. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
28. **Lai, C.-J., M. Pethel, L. R. Jan, H. Kawano, A. Cahour, and B. Falgout.** 1994. Processing of dengue type 4 and other flavivirus nonstructural proteins. Arch. Virol. (Suppl.) **9:**359–368.
29. **Lambden, P. R., E. O. Caul, C. R. Ashley, and I. N. Clarke.** 1993. Sequence and genome organization of a human small round-structured (Norwalk-like) virus. Science **259:**516–519.
30. **Long, A. C., D. C. Orr, J. M. Cameron, B. M. Dunn, and J. Kay.** 1989. A consensus sequence for substrate hydrolysis by rhinovirus 3C proteinase. FEBS Lett. **258:**75–78.
31. **Matthews, D. A., W. W. Smith, R. A. Ferre, B. Condon, G. Budahazi, W. Sisson, J. E. Villafranca, C. A. Janson, H. E. McElroy, C. L. Gribskov, and S. Worland.** 1994. Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site and means for cleaving precursor polyprotein. Cell **77:**761–777.
32. **McCann, P. J., D. R. O'Boyle, and I. C. Deckman.** 1994. Investigation of the specificity of the herpes simplex virus type 1 protease by point mutagenesis of the autoproteolysis sites. J. Virol. **68:**526–529.
33. **Meyers, G., C. Wirblich, and H.-J. Thiel.** 1991. Rabbit hemorrhagic disease virus. Molecular cloning and nucleotide sequencing of a calicivirus genome. Virology **184:**664–676.
34. **Meyers, G., C. Wirblich, and H.-J. Thiel.** 1991. Genomic and subgenomic RNAs of rabbit hemorrhagic disease are both protein-linked and packaged into particles. Virology **184:**677–689.
35. **Neill, J. D.** 1990. Nucleotide sequence of a region of the feline calicivirus genome that encodes picornavirus-like RNA-dependent RNA polymerase, cysteine protease and 2C polypeptides. Virus Res. **17:**145–160.

36. **Neill, J. D.** 1992. Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus: identification of conserved and nonconserved amino acid sequences among calicivirus capsid proteins. Virus Res. **24:**211–222.
37. **Neill, J. D., I. M. Reardon, and R. L. Heinrikson.** 1991. Nucleotide sequence and expression of the capsid protein gene of feline calicivirus. J. Virol. **65:**5440–5447.
38. **Nestorowicz, A., T. J. Chambers, and C. M. Rice.** 1994. Mutagenesis of the yellow fever virus NS2A/2B cleavage site: effects on proteolytic processing, viral replication, and evidence for alternative processing of the NS2A protein. Virology **199:**114–123.
39. **Ohlinger, V. F., B. Haas, G. Meyers, F. Weiland, and H.-J. Thiel.** 1990. Identification and characterization of the virus causing rabbit hemorrhagic disease. J. Virol. **64:**3331–3336.
40. **Pallai, P. V., J. Burkhardt, M. Skoog, K. Schreiner, P. Bax, K. A. Cohen, G. Hansen, D. E. Palladino, K. S. Harris, and M. J. Nicklin.** 1989. Cleavage of synthetic peptides by purified poliovirus 3C proteinase. J. Biol. Chem. **264:** 9738–9741.
41. **Palmenberg, A. C.** 1990. Proteolytic processing of picornaviral polyprotein. Annu. Rev. Microbiol. **44:**603–623.
42. **Parks, G. D., J. C. Baker, and A. C. Palmenberg.** 1989. Proteolytic cleavage of encephalomyocarditis virus capsid region substrates by precursors to the 3C enzyme. J. Virol. **63:**1054–1058.
43. **Parra, F., J. A. Boga, M. S. Marin, and R. Casais.** 1993. The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supports its putative subgenomic origin. Virus Res. **27:**219–228.
44. **Parra, F., and M. Prieto.** 1990. Purification and characterization of a calicivirus as the causative agent of a lethal hemorrhagic disease in rabbits. J. Virol. **64:**4013–4015.
45. **Pethel, M., B. Falgout, and C.-J. Lai.** 1992. Mutational analysis of the octapeptide sequence motif at the NS1-NS2A cleavage junction of dengue type 4 virus. J. Virol. **66:**7225–7231.
46. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
47. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
48. **Sardana, V. V., J. A. Wolfgang, C. A. Veloski, W. J. Long, K. LeGrow, B. Wolanski, E. A. Emini, and R. L. LaFemina.** 1994. Peptide substrate cleavage specificity of the human cytomegalovirus protease. J. Biol. Chem. **269:** 14337–14340.
49. **Shapira, R., and D. L. Nuss.** 1991. Gene expression by a hypovirulence-associated virus of the chestnut blight fungus involves two papain-like protease activities. J. Biol. Chem. **266:**19419–19425.
50. **Shirako, Y., and J. H. Strauss.** 1990. Cleavage between nsP1 and nsP2 initiates the processing pathway of Sindbis virus nonstructural polyprotein P123. Virology **177:**54–64.
51. **Smid, B., L. Valicek, J. Stepanek, E. Jurak, and L. Rodak.** 1989. Experimental transmission and electron microscopic demonstration of the virus of haemorrhagic disease of rabbits in Czechoslovakia. J. Vet. Med. B **36:**237–240.
52. **Sommergruber, W., H. Ahorn, H. Klump, J. Seipelt, A. Zoephel, F. Fessl, W. Krystek, D. Blaas, E. Kuechler, H.-D. Liebig, and T. Skern.** 1994. 2A proteinases of coxsackie- and rhinovirus cleave peptides derived from eIF-4γ via a common recognition motif. Virology **198:**741–745.
53. **Sommergruber, W., H. Ahorn, A. Zöphel, I. Maurer-Fogy, F. Fessl, G. Schnorrenberg, H.-D. Liebig, D. Blaas, E. Kuechler, and T. Skern.** 1992. Cleavage specificity on synthetic peptide substrates of human rhinovirus 2 protease 2A. J. Biol. Chem. **267:**22639–22644.
54. **Strebel, K., E. Beck, K. Strohmaier, and H. Schaller.** 1986. Characterization of foot-and-mouth disease virus gene products with antisera against bacterially synthesized fusion proteins. J. Virol. **57:**983–991.
55. **Webster, A., S. Russell, P. Talbot, W. C. Russell, and G. D. Kemp.** 1989. Characterization of the adenovirus proteinase: substrate specificity. J. Gen. Virol. **70:**3225–3234.
56. **Wellink, J., and A. van Kammen.** 1988. Proteases involved in the processing of viral polyproteins. Arch. Virol. **98:**1–26.
57. **Wirblich, C., G. Meyers, V. F. Ohlinger, L. Capucci, U. Eskens, B. Haas, and H.-J. Thiel.** 1994. European brown hare syndrome virus: relationship to rabbit hemorrhagic disease virus and other caliciviruses. J. Virol. **68:**5164–5173.
58. **Ypma-Wong, M. F., P. G. Dewalt, V. H. Johnson, J. G. Lamb, and B. L. Semler.** 1988. Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. Virology **166:**265–270.