Foot-and-Mouth Disease Virus Leader Proteinase: Specificity at the P2 and P3 Positions and Comparison with Other Papain-like Enzymes†

Elisabeth Kuehnel,§ Regina Cencic, Nicole Foeger,‡ and Tim Skern*

Max F. Perutz Laboratories, University Departments at the Vienna Biocenter, Department of Medical Biochemistry, Medical University of Vienna, Dr. Bohr-Gasse 9/3, A-1030 Vienna, Austria

Received April 5, 2004; Revised Manuscript Received July 7, 2004

ABSTRACT: The foot-and-mouth disease virus Leader proteinase (L^{pro}) frees itself from the growing viral polyprotein by self-processing between its own C-terminus and the N-terminus of the subsequent protein VP4. The ArgLysLeuLys*GlyAlaGlyGln sequence is recognized. The proteinase subsequently cleaves the two isoforms of host cell protein eukaryotic initiation factor (eIF) 4G at the AlaAsnLeuGly*ArgThrThrLeu (eIF4GI) and LeuAsnValGly*SerArgArgSer (eIF4GII) sequences. The enzyme does not, however, recognize the sequence on eIF4GII (AlaAspPheGly*ArgGlnThrPro) which is analogous to that recognized on eIF4GI. To investigate the basis for this specificity, we used site-directed mutagenesis to show that the presence of Phe at the P2 position or Asp at the P3 position severely compromises self-processing. Furthermore, these substitutions also give rise to the production of aberrant cleavage products. As Leu is the preferred amino acid at P2, the specificity of L^{pro} is reminiscent of that of cathepsin K. This cellular proteinase can also process collagen through its ability to accept proline at the P2 position. Investigation of the L^{pro} substrate specificity showed, however, that in contrast to cathepsin K, L^{pro} cannot accept Pro at P2 and does not cleave collagen. Subtle variations in the arrangement of the S2 binding pockets on the enzymes are responsible for these differences in specificity.

Foot-and-mouth disease virus, an animal pathogen of economic importance, expresses the information in its genome as a single polyprotein (*1*). Subsequent proteolysis of the polyprotein by proteinases encoded within it gives rise to the mature viral proteins. The initial cleavage of the FMDV[1] polyprotein is performed by L^{pro}, the first protein encoded on the polyprotein, between its own C-terminus and the N-terminus of VP4 (*2*). L^{pro} is a papain-like cysteine proteinase, showing the typical α-helix−β-sheet globular domain found in papain (*3*). In contrast to papain, L^{pro} also possesses a CTE representing the residues between the end of the β-sheet domain and the C-terminus generated by self-processing. In the crystal structure, the final four residues of the CTE of one molecule are bound to the active site of the neighboring molecule, thus defining subsites on the enzyme contacting the substrate. Subsites on the enzyme recognizing the P1−P4 substrate residues [subsite nomenclature from Schechter and Berger (*4*)] have been identified on L^{pro} (*3*).

In the infected cell, L^{pro} also cleaves the two isoforms of the cellular protein eIF4G, eIF4GI and eIF4GII (*5*−*7*). The 18 amino acids of the CTE have been shown to play a role in the recognition of the eIF4G isoforms (*3*, *8*, *9*). The identification of further cellular substrates of FMDV L^{pro} has been hampered, however, by an incomplete definition of L^{pro} specificity, despite the identification through structural work of the subsites binding the substrates. Comparison of the cleavage sites of L^{pro} on the viral polyprotein and the eIF4GI isoform suggested that the enzyme requires a basic residue at either the P1 or P1′ position [Table 1 (*5*, *10*)]. However, the recent elucidation of the L^{pro} cleavage site on eIF4GII [Table 1 (*7*)] indicates that a basic residue adjacent to the cleavage site is not an absolute requirement. In addition, L^{pro} does not cleave eIF4GII at the site analogous to that cleaved on eIF4GI. Examination of the differences in the amino acid sequences between the site cleaved on eIF4GI and the analogous site on eIF4GII (termed eIF4GII*; see Table 1) implies that the enzyme may be able to discriminate at the P2 position between Phe on one hand and Leu or Val on the other. As the P2 position is the major determinant of specificity in papain-like enzymes (*11*), this finding could be considerably important in understanding L^{pro} specificity.

The inability of L^{pro} to cleave the site on eIF4GII which is analogous to that cleaved on eIF4GI provides an opportunity to examine L^{pro} specificity. Here, we investigate which of the amino acids prevent L^{pro} from cleaving the analogous eIF4GII* sequence and compare L^{pro} specificity at the P2 and P3 positions to those of other papain-like

† This research was supported by a grant from the Austrian Science Foundation to T.S. (P-16189).

§ Present address: Divison of Tumour Genetics, German Cancer Research Center, Im Neuheimer Feld 280, D-69120 Heidelberg, Germany.

‡ Present address: Division of Cell Biology, German Cancer Research Center, Im Neuheimer Feld 280, D-69120 Heidelberg, Germany.

* To whom correspondence should be addressed. Telephone: +43 1 4277 61620. Fax: +43 1 4277 9616. E-mail: timothy.skern@univie.ac.at.

[1] Abbreviations: eIF, eukaryotic initiation factor; FMDV, foot-and-mouth disease virus; L^{pro}, leader proteinase, containing amino acids 1−201; Lb^{pro}, shorter form of L^{pro}, containing amino acids 29−201; PAGE, polyacrylamide gel electrophoresis; RRL, rabbit reticulocyte lysate; CTE, C-terminal extension; rms, root-mean-square.

10.1021/bi049340d CCC: $27.50
Published on Web 08/20/2004

Table 1: Lb^{pro} Cleavage Sites on Known Substrates Processed during Viral Replication[a]

Lb^{pro}/VP4	QRK**L**K^{201}	GAGQ
eIF4GI	FAN**L**G^{674}	RTTL
eIF4GII	LLN**V**G^{700}	SRRS
eIF4GII*	FAD**F**G^{685}	RQTP

[a] P2 residues are in bold. The Lb^{pro} cleavage sites in the viral polyprotein and human eIF4G isoforms are taken from Palmenberg (*28*), Kirchweger *et al.* (*5*), and Gradi *et al.* (*7*). The asterisk indicates that this eIF4GII sequence, analogous to that cleaved in eIF4GI, is not recognized by Lb^{pro} (*7*).

enzymes. The observed differences are then interpreted using the available three-dimensional structures.

EXPERIMENTAL PROCEDURES

Plasmids and Site-Directed Mutagenesis. The FMDV L^{pro} is the first protein encoded on the polyprotein. L^{pro} frees itself by cleavage between its own C-terminus and the N-terminus of VP4. As the initiation of protein synthesis on the FMDV RNA can occur at one of two AUG codons lying 84 nucleotides apart, two forms of L^{pro} (designated Lab^{pro} and Lb^{pro}) are synthesized in the infected cell. The reason for this is not clear, as both forms appear to have the same enzymatic properties (*12*). All work described here was carried out with the Lb^{pro} form.

The plasmid pCITE FMDV Lb^{pro}VP4VP2 containing FMDV nucleotides 892−1896 (encoding the 173 amino acids of the mature Lb^{pro}, all 85 amino acids of VP4, and 78 amino acids of VP2) of the FMDV $O1_k$ cDNA (*13*) is described in ref *8*. This plasmid contains the *Bpu*10I and *Sac*I restriction enzyme sites which lie at either side of the Lb^{pro}−VP4 junction and were introduced by site-directed mutagenesis (*8*). The mutations do not change the amino acid sequence. Replacement of the 38 bp *Bpu*10I−*Sac*I fragment with the synthetic oligonucleotide cassettes shown in Table 2 enabled the amino acid substitutions used here to be introduced. Briefly, the plasmid pCITE FMDV Lb^{pro}VP4VP2 was incubated overnight with the restriction enzymes *Bpu*10I and *Sac*I (New England Biolabs). The 4.7 kb fragment was purified by elution from an agarose gel and ligated overnight to the required oligonucleotides which had been previously phosphorylated with T4 polynucleotide kinase (New England Biolabs) and annealed to each other. Annealing was performed by heating the oligonucleotides to 90 °C for 30 s, followed by incubation at 37 °C for 5 min. Standard techniques were used throughout, and all mutations were confirmed by DNA sequencing.

In Vitro Transcription and Translation. Plasmid pCITE FMDV Lb^{pro}VP4VP2 was linearized with *Sal*I. *In vitro* transcription with T7 RNA polymerase and *in vitro* translation were carried out as described in refs *8* and *14*. *In vitro* translation reaction mixtures (typically 50 μL) contained 70% rabbit reticulocyte lysate (RRL, Promega), 20 μCi of [^{35}S]-Met (1000 Ci/mmol, Hartmann Analytic), and amino acids (except methionine) at 20 μM. After preincubation for 2 min at 30 °C, translation was started by addition of RNA to a concentration of 10 ng/μL unless stated otherwise. Aliquots (10 μL) were removed at the designated time points, and the reaction was stopped when the mixture was immediately transferred to ice, unlabeled methionine and cysteine were added to a final concentration of 2 mM, and Laemmli sample buffer was added.

Electrophoresis and Immunoblotting. The polyacrylamide gel electrophoresis system described by Dasso and Jackson (*15*) was used to separate translation products (gels contained 15% acrylamide) and to monitor the state of eIF4GI (gels contained 6% acrylamide). Translation products were detected by fluorography (exposure to film for 15 h); the state of eIF4GI was determined by immunoblotting using the anti-(eIF4GI peptide 7) antiserum as described in refs *8* and *14*. Quantification of cleavage efficiencies was achieved by densitometric scanning of fluorograms using a personal densitometer SI (Amersham Biosciences) and ImageQuant software from Molecular Dynamics.

Proteolysis of Collagen. Bovine collagen type I (Chemicon) was dialyzed overnight against PBS at 4 °C. Five micrograms of this material was digested in 50 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM ETDA, and 5 mM DTT with either 2 μM FMDV Lb^{pro} (prepared as described in ref *5*), 0.3 μM cathepsin K (His-tagged, Calbiochem), or PBS alone for 17 h at 30 °C. Chondroitin 4-sulfate (0.15%) was present in each reaction mixture.

RESULTS AND DISCUSSION

Replacement of the Viral Polyprotein Cleavage Site with Those from the eIF4G Isoforms. The previous definition of the FMDV Lb^{pro} cleavage site was based on the two known cleavage sites, namely, self-processing and eIF4GI (Table 1). The enzyme appears to be able to accept a basic residue at P1 or P1′ with a hydrophobic residue at P2. However, the recent elucidation of the sequence recognized by Lb^{pro} on eIF4GII implied a high degree of specificity at the P2 position [Table 1 (*7*)]. Above all, the inability of Lb^{pro} to cleave eIF4GII at the site analogous to that on eIF4GI (designated eIF4GII*) indicated that the enzyme could discriminate between Leu and Phe at this position. Furthermore, the examination of the cleavage sites also suggested an exclusion of acidic residues at P3 (Table 1).

To investigate Lb^{pro} specificity at P2 and P3, we first examined whether the two Lb^{pro} cleavage sites on the eIF4G isoforms could serve as substrates in the self-processing reaction. To this end, we used a system described previously (*8*) in which an *in vitro*-transcribed RNA encoding the Lb^{pro} followed by the adjacent VP4 and VP2 proteins is translated in RRLs. The production of the viral proteins can be followed by separation of the S^{35}-labeled proteins on PAGE and their detection by fluorography.

The synthesis of wild-type Lb^{pro}VP4VP2 is shown in Figure 1A (top panel). Two major products are visible, corresponding to Lb^{pro} and the VP4VP2 fragment. Lb^{pro} has four methionines, whereas VP4VP2 has only two; thus, the intensity of the Lb^{pro} band is about twice that of VP4VP2. Only a minor amount of uncleaved Lb^{pro}VP4VP2 is observed,

Table 2: Oligonucleotides Used To Construct Plasmids Used in This Study

Plasmid	Oligonucleotides
	Bpu10I Q R K L K↓G A G Q *SacI*
Lb^{pro}VP4VP2wt	5′ GCT AAG GTT CAA CGC AAG CTC AAA GGG GCT GGA CAG AGC TCT 3′
	3′ CGA TTC CAA GTT GCG TTC GAG TTT CCC CGA CCT GTC TCG AGA 5′
	F A N L G↓R T T L
Lb^{pro}VP4VP2eIF4GI	5′ T AAG TCT TTT GCC AAC CTC GGA CGG ACT ACA CTG AGC T 3′
	3′ C AGA AAA CGG TTG GAG CCT GCC TGA TGT GAC 5′
	L L N V G↓S R R S
Lb^{pro}VP4VP2eIF4GII	5′ T AAG CCT TTG TTG AAT GTT GGG TCA CGA AGA TCG AGC T 3′
	3′ C GGA AAC AAC TTA CAA CCC AGT GCT TCT AGC 5′
	F A D F G↓R Q T P
Lb^{pro}VP4VP2eIF4GII*	5′ T AAG GCT TTT GCC GAT TTT GGA CGG CAG ACA CCG AGC T 3′
	3′ C CGA AAA CGG CTA AAA CCT GCC GTC TGT GGC 5′
	F A N F G↓R T T L
Lb^{pro}VP4VP2eIF4GI FANFG RTTL	5′ T AAG TCT TTT GCC AAC TTC GGA CGG ACT ACA CTG AGC T 3′
	3′ C AGA AAA CGG TTG AAG CCT GCC TGA TGT GAC 5′
	F A D L G↓R T T L
Lb^{pro}VP4VP2eIF4GI FADLG RTTL	5′ T AAG TCT TTT GCC GAC CTC GGA CGG ACT ACA CTG AGC T 3′
	3′ C AGA AAA CGG CTG GAG CCT GCC TGA TGT GAC 5′
	F A D F G↓R T T L
Lb^{pro}VP4VP2eIF4GI FADFG RTTL	5′ T AAG TCT TTT GCC GAC TTC GGA CGG ACT ACA CTG AGC T 3′
	3′ C AGA AAA CGG CTG AAG CCT GCC TGA TGT GAC 5′
	F A D F G↓R Q T L
Lb^{pro}VP4VP2eIF4GI FADFG RQTL	5′ T AAG TCT TTT GCC GAC TTC GGA CGG CAG ACA CTG AGC T 3′
	3′ C AGA AAA CGG CTG AAG CCT GCC GTC TGT GAC 5′
	F A N P G↓R T T L
Lb^{pro}VP4VP2eIF4GI FANPG RTTL	5′ T AAG TCT TTT GCC AAC CCC GGA CGG ACT ACA CTG AGC T 3′
	3′ C AGA AAA CGG TTG GGG CCT GCC TGA TGT GAC 5′
	Q R D L K↓G A G Q
Lb^{pro}VP4VP2 K199D	5′ T AAG GTT CAA CGC GAC CTC AAA GGG GCT GGA CAG AGC T 3′
	3′ C CAA GTT GCG CTG GAG TTT CCC CGA CCT GTC 5′
	Q R K F K↓G A G Q
Lb^{pro}VP4VP2 L200F	5′ T AAG GTT CAA CGC AAG TTC AAA GGG GCT GGA CAG AGC T 3′
	3′ C CAA GTT GCG TTC AAG TTT CCC CGA CCT GTC 5′
	Q R K P K↓G A G Q
Lb^{pro}VP4VP2 L200P	5′ T AAG GTT CAA CGC AAG CCC AAA GGG GCT GGA CAG AGC T 3′
	3′ C CAA GTT GCG TTC GGG TTT CCC CGA CCT GTC 5′

indicating that the self-processing reaction is extremely efficient *in vitro*. Quantitation by densitometry of the amounts of the uncleaved Lb^{pro}VP4VP2 precursor and mature Lb^{pro} shows that at 12 min, self-processing had reached a level of 83% (Table 3). This is very similar to values previously reported for *in vitro* translation using RRLs (*8*, *16*). The 12 min time point was chosen as the optimal point for comparing enzymatic activities because earlier time points cannot be measured with any accuracy. In addition, the reaction at this time point is still clearly unimolecular (*8*). For mutants which showed little or no cleavage at the 12 min time point, the time at which self-processing reached 50% is also given (Table 3).

To examine the ability of the two Lb^{pro} cleavage sites from the eIF4G isoforms to serve as substrates in the self-processing reaction, two plasmids were constructed in which the sequence at the Lb^{pro}−VP4 junction was replaced with the amino acids indicated in Tables 1 and 2. RNAs were then synthesized and used to program protein synthesis in RRLs. Figure 1B (top left and middle panels) shows that sequences from the identified cleavage sites on eIF4GI and eIF4GII were recognized as efficient substrates in the self-cleavage reaction of Lb^{pro}. In both cases, the extent of self-processing at 12 min was well above 50% (Table 3). As it is extremely difficult in the RRL system to determine the exact position of a cleavage site by protein sequencing, it should be noted that a slight uncertainity about the position of the cleavage site cannot be eliminated.

Next, we examined whether the introduction of the sequence on eIF4GII which corresponds to that used in eIF4GI (eIF4GII*, Tables 1 and 2) can be used in the self-processing reaction. Figure 1B (top right panel) and Table 3

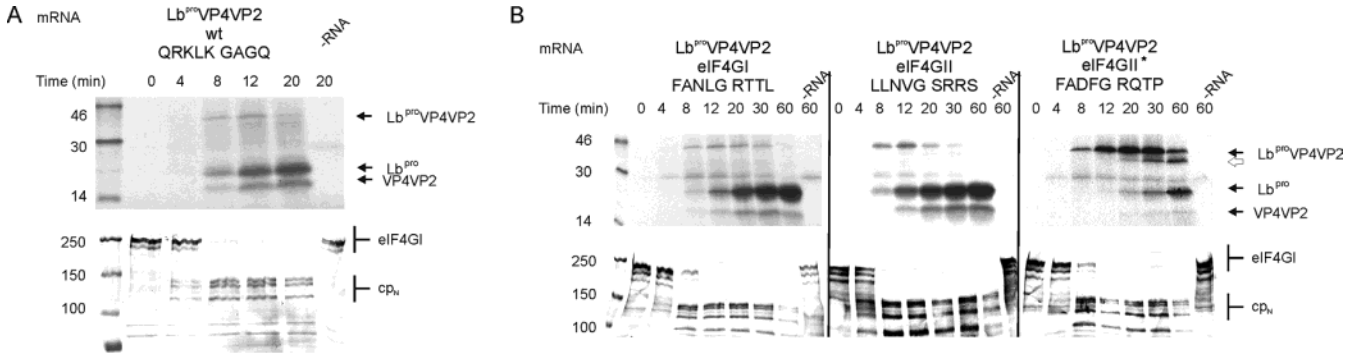

FIGURE 1: Effect of eIF4G isoform sequences at the Lb^{pro}−VP4 junction on self-processing and eIF4GI cleavage. RRLs were incubated with wild-type (A) or the indicated mutated (B) mRNAs (10 ng/μL; transcribed *in vitro* following linearization with *Sal*I) as described in Experimental Procedures; protein synthesis was terminated at the times given by placing the samples on ice followed by the addition of unlabeled methionine and cysteine to final concentrations of 2 mM and Laemmli sample buffer. Aliquots were analyzed on 15% polyacrylamide gels followed by fluorography to detect the synthesis of Lb^{pro}VP4VP2 (top panels) and on 6% polyacrylamide gels followed by immunoblotting for the status of eIF4GI (bottom panels). Fluorographs were exposed for 15 h. The positions of uncleaved Lb^{pro}VP4VP2 and cleavage products Lb^{pro} and VP4VP2 are marked in the top panels; the open arrow in panel B denotes an aberrant cleavage product. In the bottom panels, intact eIF4GI and the N-terminal cleavage products cp$_N$ are marked. Protein standards are given in kilodaltons.

Table 3: Effect on Self-Processing of Amino Acid Substitutions at the Lb^{pro}VP4VP2 Junction

	plasmid	cleavage site	self-processing at 12 min (%)	time at 50% self-processing (min)
Figure 1	Lb^{pro}VP4VP2 wild type	QRKLK GAGQ	83	
	Lb^{pro}VP4VP2 eIF4GI	FANLG RTTL	67	
	Lb^{pro}VP4VP2 eIF4GII	LLNVG SRRS	59	
	Lb^{pro}VP4VP2 eIF4GII*	FADFG RQTP	0	60
Figure 2	Lb^{pro}VP4VP2 eIF4GI	FANFG RTTL	14	20−30
	Lb^{pro}VP4VP2 eIF4GI	FADLG RTTL	16	20
	Lb^{pro}VP4VP2 eIF4GI	FADFG RTTL	14	20
	Lb^{pro}VP4VP2 eIF4GI	FADFG RQTL	19	20
Figure 3	Lb^{pro}VP4VP2 L200F	QRKFK GAGQ	0	30−60
	Lb^{pro}VP4VP2 K199D	QRDLK GAGQ	0	20−30
Figure 4	Lb^{pro}VP4VP2 L200P	QRKPK GAGQ	0	60
	Lb^{pro}VP4VP2eIF4GI	FANPG RTTL	8	20−30

show that self-processing commenced after incubation for at least 12−20 min and had only reached a level of 50% after incubation for 60 min. Furthermore, an aberrant cleavage product, migrating somewhat more rapidly than the uncleaved precursor, was also observed (Figure 1B, open arrow). This indicates that the eIF4GII* sequence is indeed a poor substrate for Lb^{pro} and that some or all of the amino acids present are not accepted by Lb^{pro}.

As any modification at the N-terminal side of the cleavage site also modifies the C-terminus of Lb^{pro}, we sought to ensure that the Lb^{pro}VP4VP2eIF4GII* enzyme was indeed active. For this, we took advantage of the presence of eIF4GI in the RRLs and examined its fate during the synthesis of Lb^{pro}, as reported previously (*14*). Accordingly, aliquots of the translation reactions were subjected to SDS−PAGE and the gels blotted onto PVDF membranes and probed with an antiserum against the N-terminus of eIF4GI (Figure 1, bottom panels). eIF4GI itself migrates as a series of bands with a molecular mass of 220 kDa (*17*). The bands have different N-termini which arise from the use of different AUG codons during synthesis of eIF4GI (*18*, *19*). Cleavage of eIF4GI by Lb^{pro} at its single recognition site at the Gly^{674}Arg sequence (numbering according to ref *18*) generates a series of N-terminal cleavage products which are detected by the N-terminal antiserum used here (cp$_N$, Figure 1A,B, bottom panels). The single C-terminal cleavage fragment is not detected by this antiserum. The cleavage of eIF4GI occurs at wild-type levels (50% cleavage between 4 and 8 min) in all Lb^{pro} variants examined here, indicating that all are synthesized in an active conformation (Figure 1A,B, bottom panels). We have shown previously that eIF4GI cleavage is independent of the extent of self-processing (*8*). Thus, this confirms that the Lb^{pro} in the Lb^{pro}VP4VP2eIF4GII* mutant is catalytically active and that the eIF4GII* sequence cannot serve as a substrate in the self-processing reaction.

Introduction of Mutations into the eIF4GI Cleavage Site. Thus, despite only four rather unspectacular differences between the eIF4GI and eIF4GII* sequences (Table 1), Lb^{pro} discriminates against the eIF4GII* sequence. We set out to identify which of these four residues were responsible for this effect by introducing them one at a time into the eIF4GI cleavage site and examining the effects on the self-processing reaction.

Given the importance of the P2 position for substrate recognition by papain-like enzymes, we began by substituting the Leu at P2 with Phe. As can be seen in Figure 2 (top left panel), the presence of Phe at P2 in the eIF4GI cleavage site reduces the level of self-processing by Lb^{pro}. After 12 min, only 14% self-processing was observed, and 50% cleavage was not achieved until 20−30 min had passed (Table 3). However, a reduction to the level observed with eIF4GII* (see Figure 1B) was not obtained, as cleavage was essentially complete after 60 min. This reduction is also accompanied by the appearance of an aberrant product running slightly faster than the uncleaved material (Figure 2, open arrow). A similar reduction in the level of self-

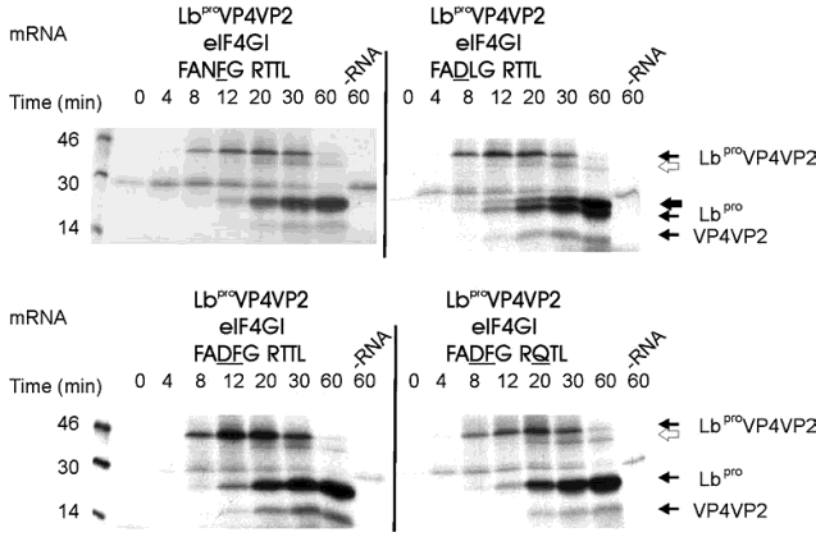

FIGURE 2: Effect of the introduction of substitutions into Lb^proVP4VP2eIF4GI on self-processing. The indicated mRNAs (10 ng/μL; transcribed *in vitro* following linearization with *Sal*I) were used to program RRLs. Analysis of protein synthesis was carried out as described in Figure 1. The open arrows denote an aberrant cleavage product which was also seen with the mRNA Lb^proVP4VP2eIF4GII* (Figure 1B, top right panel). The large filled arrow denotes an aberrant cleavage product only found with the mRNA Lb^proVP4VP2 FADLG RTTL (top right panel).

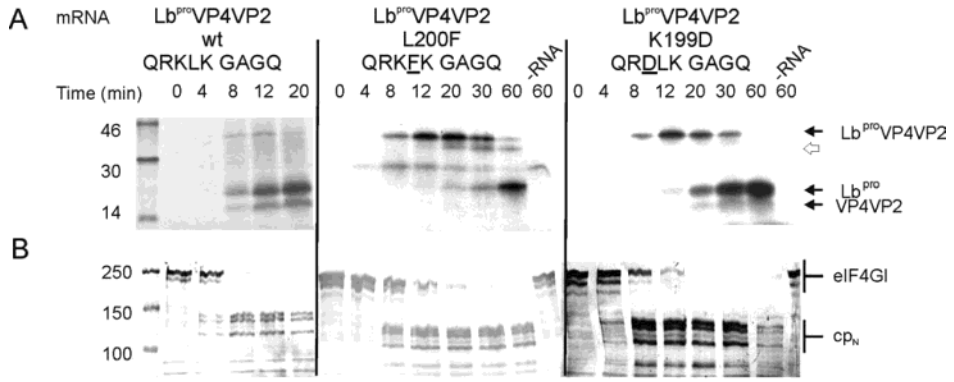

FIGURE 3: Phe at P2 or Asp at P3 is detrimental to Lb^pro self-processing at the wild-type cleavage site. The indicated mRNAs (10 ng/μL; transcribed *in vitro* following linearization with *Sal*I) were used to program RRLs. Analysis of protein synthesis (A) and the state of eIF4GI (B) were as described in the legend of Figure 1. The open arrow denotes an aberrant cleavage product which was also seen with the mRNA Lb^proVP4VP2eIF4GII* (Figure 1B, top right panel).

processing was also observed upon the replacement of Asn at P3 with Asp in the eIF4GI cleavage site (Figure 2, top right panel, and Table 3). Furthermore, a second aberrant cleavage product, in addition to that observed in the left panel, also appeared. This aberrant product migrates slightly slower than wild-type Lb^pro (Figure 2, top right panel, closed arrow). However, once again, the level of self-processing was not reduced to the same extent seen with the eIF4GII* cleavage site.

We therefore created a double mutant containing both the P2 and P3 substitutions and examined Lb^pro processing (Figure 2, bottom left panel, and Table 3). Once again, the level of cleavage was reduced compared to the original level of eIF4GI cleavage, but was still significantly better than that with the eIF4GII* sequence. Interestingly, the aberrant cleavage product found with the P3 Asp mutant was absent from the products of the P3 Asp P2 Phe double mutant. As the subsequent introduction of Gln at P2′ did not reduce the level of cleavage further (Figure 2, bottom right panel, and Table 3), the presence of Pro at P4′ must be required for the final refractory effect seen with the complete eIF4GII* sequence FADFG RQTP (Figure 1B, top right panel) compared to that seen with FADFG RQTL (Figure 2, bottom right panel).

As a control for Lb^pro activity, the state of eIF4GI in the RRLs was monitored (data not shown). Essentially wild-type cleavage rates on eIF4GI were observed in each case, indicating that all Lb^pro variants possessed full catalytic activity.

Introduction of Mutations into the Wild-Type Lb^pro Cleavage Site. The results described above suggested that Lb^pro can discriminate between Leu and Phe at P2 and Asn and Asp at P3. To confirm these results, we introduced each of the amino acids into the wild-type cleavage site and examined Lb^pro processing. In both cases, Lb^pro processing was severely delayed (Figure 3A and Table 3), with self-processing not commencing until 12−20 min after incubation. The P2 Leu to Phe substitution appeared to be the more compromised, showing an inhibition similar to that seen with Lb^pro-VP4VP2eIF4GII* (compare Figure 3A, middle panel, with Figure 1B, right panel; Table 3). Interestingly, the introduction of Phe at P2 or Asp at P3 into the wild-type sequence affected self-processing more significantly than when either was introduced into the eIF4GI sequence. In addition, the

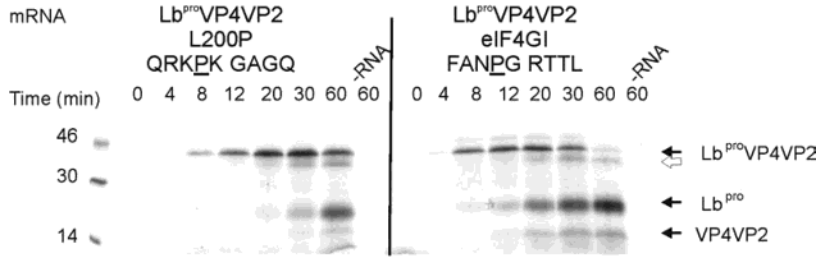

FIGURE 4: Lb^{pro} cannot accept proline at P2 in the self-processing reaction. The indicated mRNAs (10 ng/μL; transcribed *in vitro* following linearization with *Sal*I) were used to program RRLs. Analysis of protein synthesis was as described in the legend of Figure 1. The open arrow denotes an aberrant cleavage product which was also seen with the mRNA Lb^{pro}VP4VP2eIF4GII* (Figure 1B, top right panel).

aberrant cleavage product (Figure 3, open arrow) seen with the Lb^{pro}VP4VP2eIF4GII* construction was also evident. However, the aberrant cleavage product seen with the introduction of Asp at P3 in the eIF4GI site (Figure 2, closed arrow, top panels) was not detected.

Examination of the cleavage of eIF4GI showed a slight delay in the onset of cleavage with both mutant enzymes. However, in both cases, eIF4GI cleavage was essentially complete after 12 min (Figure 3B).

Comparison of the Lb^{pro} Specificity at P2 with Those of Other Papain-like Proteinases. The results described above showed that in the self-processing reaction Lb^{pro} can readily accept Leu or Val but not Phe at the P2 position. Furthermore, the enzyme accepts Asn and Lys at P3, while Asp is less favored. Comparison of the Lb^{pro} specificity with those of other papain-like proteinases revealed that cathepsin K exhibited a similar profile at both the P2 and P3 positions (*20*, *21*). At P2, both cathepsin K and cathepsin S prefer Leu to Phe, whereas cathepsin L accepts Leu and Phe equally well (*20*−*23*). A further ability of cathepsin K is its ability to accept Pro at P2; this has been proposed to explain the ability of this enzyme to cleave collagen (*20*, *24*).

We sought to examine whether the similarities in the substrate specificity between Lb^{pro} and cathepsin K extended to the acceptance of Pro at P2. Figure 4 and Table 3 show that this is not the case and that the substitution of Leu at P2 with Pro has a detrimental effect on Lb^{pro} self-processing. The effect was observed regardless of whether the substitution was carried out in the wild type (Figure 4, left panel) or on the eIF4GI cleavage site background (Figure 4, right panel). Once again, self-processing was more seriously affected in the wild-type background, not beginning until about 20−30 min after initiation of translation and only reaching 50% after 60 min. In contrast, when Phe was present at P2 in the eIF4GI sequence, the level of self-processing reached 8% at 12 min and 50% cleavage was seen after 20−30 min (Table 3). In both cases, no effect was observed on the cleavage of endogenous eIF4GI in the RRLs (data not shown).

Further evidence that Lb^{pro} does not possess a specificity profile similar to that of cathepsin K with respect to the P2 position was obtained using type 1 collagen as a substrate. Lb^{pro} did not show any activity toward type 1 collagen, even when chondroitin 4-sulfate was included in the reaction mixture and the incubation was performed for 17 h (Figure 5). The reaction was performed at pH 8.0 as the Lb^{pro} is inactive below pH 7.0 (*25*). Control reactions that included recombinant cathepsin K confirmed that the collagen could serve as a substrate for proteolytic enzymes (Figure 5).

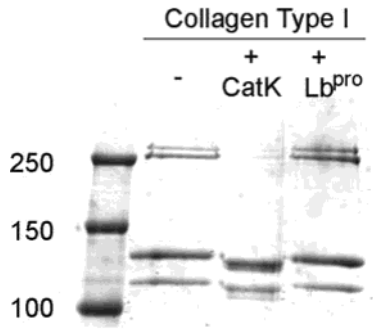

FIGURE 5: Lb^{pro} does not cleave type I collagen. Type I collagen was incubated without enzyme (lane 1) or with cathepsin K (lane 2) or Lb^{pro} (lane 3) as described in Experimental Procedures. The reaction was stopped by the addition of Laemmli sample buffer, and the proteins were separated by SDS−PAGE and detected by staining with Coomassie blue. Molecular mass markers (in kilodaltons) are indicated.

Table 4: Residues Defining the Specificity (in italics) of the S2 Pocket in the Indicated Papain-like Cysteine Proteinases[a]

residue	papain (*Phe/Leu*)	cathepsin L (*Phe/Leu*)	cathepsin S (*Phe/Leu*)	cathepsin K (*Pro/Leu*)	Lb^{pro} (*Leu*)
67	Tyr	Leu	Phe	Tyr	Pro99
68	Pro	Met	Met	Met	Pro100
133	Val	Ala	Gly	Ala	Leu143
157	Val	Met	Val	Leu	
160	Ala	Gly	Gly	Ala	Ala149
205	Ser	Ala	Phe	Leu	Leu178

[a] Residues for papain and the cathepsins are taken from refs *20* and *23* and numbered according to papain numbering (*29*). Residues for Lb^{pro} are taken from and numbered according to ref *3*.

The Arrangement of the S2 Pockets of Lb^{pro} and Cathepsin K Differ. We sought to explain the structural basis of the similarities and differences in P2 specificity between FMDV Lb^{pro} and cathepsin K in terms of their respective three-dimensional structures. Which features of the S2 subsites allow both enzymes to accept Leu but exclude Phe at P2? Why can cathepsin K also accept Pro whereas Lb^{pro} cannot? The residues defining the S2 pocket of papain, the cathepsins L, S, and K, and Lb^{pro} are shown in Table 4; the architecture of the S2 pockets of cathepsin K and Lb^{pro} is illustrated in panels A and B of Figure 6, respectively.

Several reports indicate that the occupancy of residues 67 and 205 is responsible for the differences in specificity between cathepsin K and cathepsin L (*20*, *23*, *26*). Thus, the presence of Ser or Ala at residue 205 correlates with the ability of papain or cathepsin L to accept Leu or Phe at P2. In contrast, the presence of Phe at residue 205 in cathepsin S or Leu at residue 205 in cathepsin K appears to confer a preference for Leu over Phe at P2 in both enzymes (*20*, *21*,

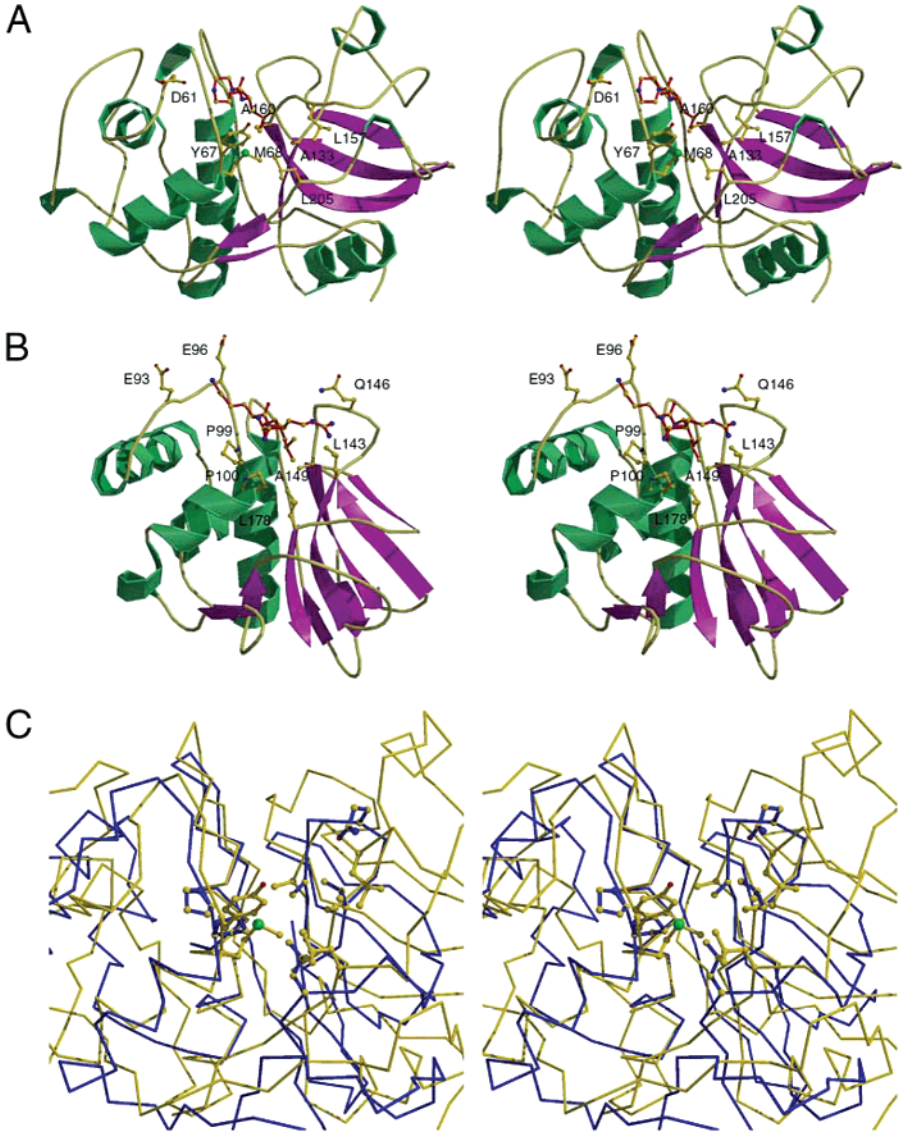

FIGURE 6: Comparison of the structures of cathepsin K and Lb^{pro}. Stereo Molscript drawings of cathepsin K (A) and FMDV Lb^{pro} (B) (α-helices in green, β-sheets in purple, and coils in yellow). Amino acids of the S3 and S2 subsites are shown for cathepsin K and subsites S4, S3, and S2 for Lb^{pro}. Bound inhibitor residues from P3 to P2 (carboxpiperazineLeu) for cathepsin K and substrate residues from P4 to P2 (ArgLysLeu) for Lb^{pro} are shown in red. PDB entries 1MEM for cathepsin K and 1QOL for Lb^{pro} were used. (C) Superimposed Cα traces of Lb^{pro} (blue) and cathepsin K (yellow). Side chains forming the S2 pockets are indicated with the same color scheme. The superimposition of main chain Cα atoms was created using the rigid and rigid2 commands of TURBO-FRODO (*30*), based on the alignment of Lb^{pro} with papain published in ref *3*. An averaged rmsd of 1.2 Å for 76 equivalent residues was obtained, similar to that for papain [1.3 Å for 76 equivalent residues (*3*)]. Drawings were made using Molscript (*31*, *32*) and rendered with Raster3D (*33*).

26). The ability of cathepsin K to accept Pro at P2 has been proposed to result from the presence of Tyr at position 67 (*20*).

The residues in Lb^{pro} equivalent to residues 67 and 205 in cellular papain-like proteinases are Pro99 and Leu178, respectively (Table 4 and Figure 6). Leu178 in Lb^{pro} is at the bottom of the S2 pocket, occupying a space similar to that of Leu205 in cathepsin K (Figure 6C). This similarity would explain the preference of both enzymes for Leu at P2 with the concomitant exclusion of Phe. In contrast, the architecture of Lb^{pro} around Pro99, equivalent to Tyr67 in cathepsin K, is quite different (Figure 6C). The absence of the Tyr side chain found in cathepsin K is presumably responsible for the inability of Lb^{pro} to accept Pro in its S2 binding pocket.

Table 4 also shows that Lb^{pro} uses one less amino acid to form its S2 pocket; there is no Lb^{pro} equivalent to residue 157 of the other papain-like enzymes. Figure 6C shows that the space occupied by Leu157 in cathepsin K is partially occupied by Lb^{pro} residue Leu143, which is equivalent to Ala133 in cathepsin K. This use of fewer residues is all the more remarkable as Lb^{pro} has only 173 amino acids compared to the 212 of papain or the 215 of cathepsin K.

S3 Subsites of FMDV Lb^{pro}. The mutational analysis presented here clearly shows that Lb^{pro} can accept Lys or Asn at P3, while the negatively charged Asp is less favored. In this respect, Lb^{pro} resembles both cathepsin K and cathepsin L (*20*, *21*). The S3 subsite (*3*) of the Lb^{pro} is partially shown in Figure 6B. The P3 Lys residue, through its aliphatic chain, makes van der Waals interactions with the main chain atoms of Gly97 and Gly98 (not shown). The amino group accepts a hydrogen bond from the carbonyl oxygen of Glu93; this interaction is conserved across the eight copies of the complex that were independently refined

in PDB entry 1QOL (*3*). In four of the copies, the P3 Lys residue also makes weak ionic interactions with Glu93 and Glu96.

In contrast, the S3 subsites of cathepsins K and L each show just one acidic amino group, Asp61 and Glu61, respectively. This implies that the viral enzyme has evolved to be more stringent at this position. The results obtained with several of the mutants shown here further indicate the inherent accuracy of the self-processing reaction. Thus, almost all mutants described here which lead to a delay in processing also generate an aberrant product migrating slightly faster than the uncleaved precursor. Presumably, this site is also available in the wild-type substrate; however, the efficiency of cleavage at correct sites is such that cleavage at an aberrant site is negligible.

A second aberrant product of almost the same mobility of the mature Lb^{pro} was obtained on introduction of Asp at P3 into the Lb^{pro}VP4VP2 eIF4GI precursor (Figure 2, top right panel). At the 30 min time point, the amounts of mature Lb^{pro} and the aberrant product were approximately equal. Surprisingly, this product was absent when Phe was introduced into this sequence (Figure 2, bottom left panel) and was not observed with the eIF4GII* mutant which contains Asp and Phe at P3 and P2, respectively (Figure 1B, top right panel). Thus, it appears that the presence of Phe at P2 can suppress the significant aberrant processing observed with Asp at P3 and Leu at P2. Furthermore, this aberrant product was not apparent when Asp at P3 was introduced into the wild-type precursor (Figure 3, top right panel). Thus, the suppression of aberrant cleavage caused by Asp at P3 is not limited to Phe at P2 but can also be achieved by the surrounding residues.

Taken together, these results also indicate that the actual combination of residues in the self-processing substrate is important, rather than the occupancy of each particular position. Finally, given the sensitivity of the self-processing reaction to delay and inaccuracy, it is remarkable that the two known *trans* cleavage sites are efficiently recognized in the *cis* reaction.

CONCLUDING REMARKS

The cleavage specificity of the FMDV L^{pro} has been difficult to define, even though this proteinase has a high degree of specificity. The experiments shown here clarify the nature of the specificity at positions P2 and P3 during the self-processing reaction.

In addition, the experiments confirm that L^{pro} cleaves different sequences on eIF4GI and eIF4GII (*7*). Clearly, the evolutionary advantage to the virus is sufficiently great to drive the development of an enzyme that is able to cleave both eIF4G isoforms at nonanalogous sites. Interestingly, $2A^{pro}$ of HRV2 has also evolved to cleave both isoforms at different sites (*10*, *27*). Nevertheless, for both enzymes, these proteinases still remain very specific, so no other cellular proteins which are required for viral replication are cleaved. The ability of viruses to modulate their enzymes, without breaking the constraints on genome size, is once again amply illustrated.

ACKNOWLEDGMENT

We thank R. Rhoads for the antiserum against eIF4GI and Ignacio Fita for help with Molscript.

REFERENCES

1. Mason, P. W., Grubman, M. J., and Baxt, B. (2003) Molecular basis of pathogenesis of FMDV, *Virus Res. 91*, 9−32.
2. Strebel, K., and Beck, E. (1986) A second protease of foot-and-mouth disease virus, *J. Virol. 58*, 893−899.
3. Guarné, A., Tormo, J., Kirchweger, K., Pfistermueller, D., Fita, I., and Skern, T. (1998) Structure of the foot-and-mouth disease virus leader protease: a papain-like fold adapted for self-processing and eIF4G recognition, *EMBO J. 17*, 7469−7479.
4. Berger, A., and Schechter, L. (1970) Mapping the acttive site of papain with the aid of peptide substrates and inhibitors, *Philos. Trans. R. Soc. London, Ser. B 257*, 249−264.
5. Kirchweger, R., Ziegler, E., Lamphear, B. J., Waters, D., Liebig, H. D., Sommergruber, W., Sobrino, F., Hohenadl, C., Blaas, D., Rhoads, R. E., and Skern, T. (1994) Foot-and-mouth disease virus leader proteinase: Purification of the Lb form and determination of its cleavage site on eIF-4γ, *J. Virol. 68*, 5677−5684.
6. Devaney, M. A., Vakharia, V. N., Lloyd, R. E., Ehrenfeld, E., and Grubman, M. J. (1988) Leader protein of Foot-and-Mouth disease virus is required for cleavage of the p220 component of the cap-binding protein complex, *J. Virol. 62*, 4407−4409.
7. Gradi, A., Foeger, N., Skern, T., and Belsham, G. (2004) Cleavage of eukaryotic translation initiation factor 4GII within foot-and-mouth disease virus-infected cells: identification of the L-protease cleavage site in vitro, *J. Virol. 78*, 3271−3278.
8. Glaser, W., Cencic, R., and Skern, T. (2001) Foot-and-mouth disease Leader proteinase: involvement of C-terminal residues in self-processing and cleavage of eIF4GI, *J. Biol. Chem. 276*, 35473−35481.
9. Foeger, N., Glaser, W., and Skern, T. (2002) Recognition of eukaryotic initiation factor 4G isoforms by picornaviral proteinases, *J. Biol. Chem. 277*, 44300−44309.
10. Lamphear, B. J., Yan, R., Yang, F., Waters, D., Liebig, H. D., Klump, H., Kuechler, E., Skern, T., and Rhoads, R. E. (1993) Mapping the cleavage site in protein synthesis initiation factor eIF-4γ of the 2A proteases from human Coxsackievirus and rhinovirus, *J. Biol. Chem. 268*, 19200−19203.
11. Turk, D., Guncar, G., Podobnik, M., and Turk, B. (1998) Revised definition of substrate binding sites of papain-like cysteine proteases, *Biol. Chem. 379*, 137−147.
12. Cao, X., Bergmann, I. E., Fullkrug, R., and Beck, E. (1995) Functional analysis of the two alternative translation initiation sites of foot-and-mouth disease virus, *J. Virol. 69*, 560−563.
13. Forss, S., Strebel, K., Beck, E., and Schaller, H. (1984) Nucleotide sequence and genome organization of foot-and-mouth disease virus, *Nucleic Acids Res. 12*, 6587−6601.
14. Glaser, W., and Skern, T. (2000) Extremely efficient cleavage of eIF4G by picornaviral proteinases L and 2A in vitro, *FEBS Lett. 480*, 151−155.
15. Dasso, M. C., and Jackson, R. J. (1989) Efficient initiation of mammalian mRNA translation at a CUG codon, *Nucleic Acids Res. 17*, 6485−6497.
16. Schlick, P., Kronovetr, J., Hampoelz, B., and Skern, T. (2002) Modulation of the electrostatic charge at the active site of foot-and-mouth-disease-virus leader proteinase, an unusual papain-like enzyme, *Biochem. J. 363*, 493−501.
17. Etchison, D., Milburn, S. C., Edery, I., Sonenberg, N., and Hershy, J. W. B. (1982) Inhibition of HeLa cell protein synthesis following poliovirus infection correlates with the proteolysis of a 220,000-dalton polypeptide associated with eukaryotic initiation factor 3 and a cap binding protein complex, *J. Biol. Chem. 257*, 14806−14810.
18. Byrd, M. P., Zamora, M., and Lloyd, R. E. (2002) Generation of multiple isoforms of eukaryotic translation initiation factor 4GI by use of alternate translation initiation codons, *Mol. Cell. Biol. 22*, 4499−4511.
19. Bradley, C. A., Padovan, J. C., Thompson, T. L., Benoit, C. A., Chait, B. T., and Rhoads, R. E. (2002) Mass spectrometric analysis of the N terminus of translational initiation factor eIF4G-1 reveals novel isoforms, *J. Biol. Chem. 277*, 12559−12571.
20. Lecaille, F., Choe, Y., Brandt, W., Li, Z., Craik, C. S., and Bromme, D. (2002) Selective inhibition of the collagenolytic activity of human cathepsin K by altering its S2 subsite specificity, *Biochemistry 41*, 8447−8454.
21. Alves, M. F., Puzer, L., Cotrin, S. S., Juliano, M. A., Juliano, L., Bromme, D., and Carmona, A. K. (2003) S3 to S3′ subsite specificity of recombinant human cathepsin K and development

of selective internally quenched fluorescent substrates, *Biochem. J. 373*, 981−986.
22. Harris, J. L., Backes, B. J., Leonetti, F., Mahrus, S., Ellman, J. A., and Craik, C. S. (2000) Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries, *Proc. Natl. Acad. Sci. U.S.A. 97*, 7754−7759.
23. Pauly, T. A., Sulea, T., Ammirati, M., Sivaraman, J., Danley, D. E., Griffor, M. C., Kamath, A. V., Wang, I. K., Laird, E. R., Seddon, A. P., Menard, R., Cygler, M., and Rath, V. L. (2003) Specificity determinants of human cathepsin S revealed by crystal structures of complexes, *Biochemistry 42*, 3203−3213.
24. Lecaille, F., Weidauer, E., Juliano, M. A., Bromme, D., and Lalmanach, G. (2003) Probing cathepsin K activity with a selective substrate spanning its active site, *Biochem. J. 375*, 307−312.
25. Guarné, A., Hampoelz, B., Glaser, W., Carpena, X., Tormo, J., Fita, I., and Skern, T. (2000) Structural and biochemical features distinguish the Foot-and-Mouth Disease Virus leader proteinase from other papain-like enzymes, *J. Mol. Biol. 302*, 1227−1240.
26. Bromme, D., Bonneau, P. R., Lachance, P., and Storer, A. C. (1994) Engineering the S2 subsite specificity of human cathepsin S to a cathepsin L- and cathepsin B-like specificity, *J. Biol. Chem. 269*, 30238−30242.
27. Gradi, A., Svitkin, Y. V., Sommergruber, W., Imataka, H., Morino, S., Skern, T., and Sonenberg, N. (2003) Human rhinovirus 2A proteinase cleavage sites in eukaryotic initiation factors (eIF) 4GI and eIF4GII are different, *J. Virol. 77*, 5026−5029.
28. Palmenberg, A. C. (1989) in *Molecular aspects of picornavirus infection and detection* (Semler, B. L., and Ehrenfeld, E., Eds.) pp 211−241, American Society for Microbiology, Washington, DC.
29. Berti, P. J., and Storer, A. C. (1995) Alignment/phylogeny of the papain superfamily of cysteine proteases, *J. Mol. Biol. 246*, 273−283.
30. Jones, T. A. (1978) A graphics model building and refinement system for macromolecules, *J. Appl. Crystallogr. 11*, 268−272.
31. Esnouf, R. (1997) An extensively modified version of Molscript that includes greatly enhanced colouring capacities, *J. Mol. Graphics 15*, 133−138.
32. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.
33. Merrit, E., and Murphy, M. (1994) Raster3D version 2.0. A program for photorealistic molecular graphics, *Acta Crystallogr. D50*, 869−873.

BI049340D