J. Biochem. 2008;144(6)771–780
doi:10.1093/jb/mvn130

Saturation Mutagenesis reveals that GLU54 of Norovirus 3C-like Protease is not Essential for the Proteolytic Activity

Yuichi Someya*, Naokazu Takeda and Takaji Wakita

Department of Virology II, National Institute of Infectious Diseases, Musashi-Murayama, Tokyo 208-0011, Japan

Received May 28, 2008; accepted September 22, 2008; published online October 6, 2008

The norovirus 3C-like protease is a member of the chymotrypsin-like serine protease superfamily. Previous characterization of its crystal structure has implicated the Glu54–His30–Cys139 triad in the catalysis. In the present study, the Glu54 residue of the protease was subjected to site-saturation mutagenesis, with the result that nearly half of the mutants retained the significant proteolytic activity. It was suggested that a carboxylate at position 54 was not essential for the activity. The *in vitro* assays of the proteolysis revealed that most of Glu54 mutants retained relatively high proteolytic activity. When the Glu54 mutation was combined with the Ser mutation of the Cys139 residue, a nucleophile, only the Asp54 and Gln54 mutations showed proteolytic activity comparable to that of the Ser139 single mutant, suggesting that a hydrogen bond between Glu54 and His30 was critical in the Ser139 background. These results suggested that the mechanism of the proteolysis by the wild-type norovirus 3C-like protease was different from that of typical chymotrypsin-like serine proteases.

Key words: 3C-like protease, catalytic triad, norovirus, serine-like cysteine protease, serine protease.

Abbreviations: DNP, 2,4-dinitrophenyl; GST, glutathione *S*-transferase; NMA, 2-(*N*-methyl amino)benzoyl; NTPase, nucleotide triphosphatase; VPg, genome-linked viral protein.

Norovirus is a major cause of acute non-bacterial gastroenteritis in humans, and genetically and antigenically diverse strains have been isolated worldwide (*1–3*). Norovirus is one of the positive-sense single-stranded RNA viruses and belongs to the family *Caliciviridae*. The norovirus genome is ~7.7 kb in length with a poly(A) tail at its 3′-end. VPg, a genome-linked viral protein is believed to be bound to the 5′-end in place of the cap structure (*4*).

The genome encodes three open-reading frames (ORFs). The ORF1 product is a polyprotein and is cleaved by its viral 3C-like protease activity into six non-structural proteins, which include 2C-like nucleotide triphosphatase (NTPase), 3B VPg and 3D RNA-dependent RNA polymerase, in addition to 3C-like protease (*5*, *6*). The nomenclature is based on that for picornaviruses. The ORF2 and ORF3 products are a major and a minor structural protein (VP1 and VP2), respectively. The norovirus virion consists of 180 VP1 molecules. The sequence diversity of the P2 domain of VP1 correlates with a wide variety of the antigenicity of noroviruses (*7–9*) and potentially provides different patterns of binding to histo-blood group antigens (*10*).

Norovirus 3C-like protease is a central enzyme that is solely responsible for the maturation of norovirus ORF1 polyprotein (*5*, *6*). The 3C-like proteases of various strains isolated from humans share high amino acid identity and have significant similarity with those of other caliciviruses. Norovirus 3C-like protease is thought to be a potent target of medication for the treatment of norovirus diarrhoea.

Our previous extensive mutagenesis study for a 3C-like protease of the Chiba strain revealed that His30 and Cys139 were responsible for the proteolysis, the former being estimated as the general base and the latter the nucleophile in the catalytic centre (*11*). This was clearly revealed by the X-ray crystal structure at 2.8 Å resolution, which showed that the norovirus 3C-like protease had the chymotrypsin-like fold (*12*). The crystal structure also suggested that Glu54, as the third member carboxylate that interacted with His30, was involved in the catalysis just like typical chymotrypsin-like serine proteases (*12*). It is suggested that the mechanism of the proteolysis by the norovirus 3C-like protease was identical to that of chymotrypsin; that is, the thiol of Cys139 attacks against a carbonyl carbon of the substrate, followed by protonation of the imidazole ring of His30 by the proton released from the thiol of Cys139. A negatively charged carboxylate of Glu54 interacts with and stabilizes the imidazole ring of His30. However, we had concluded that Glu54 was not essential for the activity because the Glu54-to-Ala mutant protease retained protease activity comparable to that of the wild-type enzyme (*11*). Independently, Zeitler *et al*. (*13*) resolved the crystal structure of Norwalk virus 3C-like protease at 1.5 Å resolution and insisted that Glu54 was a necessary component for the proteolysis. These controversial conclusions may have arisen from the difference in the methods used to assay the protease activity.

*To whom correspondence should be addressed. Tel: +81-42-561-0771, Fax: +81-42-561-4729, E-mail: someya@nih.go.jp

Namely, we observed the cleavage of fusion proteins including the protease moiety expressed in *Escherichia coli* cells (*11*), while Zeitler *et al*. (*13*) used a fluorogenic peptide substrate to assess the activity. The former method may have reflected the proteolytic activity at the steady state, while the latter reflected the initial rate of the cleavage reaction. Therefore, it is possible that the mutation of Glu54 affects the turnover rate of the protease.

Here we focused on biochemistry of the norovirus 3C-like protease and performed site-saturation mutagenesis of the Glu54 residue of the 3C-like protease to determine what kind of side chain at position 54 is required for the proteolytic activity. For this purpose, we first observed the level of the proteolysis exerted by the 3C-like protease in *E. coli* cells. In addition, the proteolytic activities of some mutant 3C-like proteases were assayed using a fluorogenic peptide including the sequence between 2C NTPase and 3A-like protein of Chiba virus. From the results of these two assays, we concluded that the Glu54 residue was not essential for the protease activity, although it was important for the efficient proteolysis.

MATERIALS AND METHODS

Escherichia coli Strains—The JM109 and HST02 strains (Takara Bio, Inc., Tokyo, Japan) were used for plasmid construction. The BL21-CodonPlus(DE3)-RIPL strain (Stratagene, La Jolla, CA, USA) was used for the expression of the recombinant proteins.

Construction of Bacterial Expression Plasmids—We constructed an expression plasmid, pGEX-2TK-3aBC, encoding a glutathione *S*-transferase (GST) fusion protein to part of the ORF1 polyprotein from the norovirus Chiba strain (*14*) (Fig. 1). DNA fragments were amplified by Vent DNA polymerase (NEB, Ipswich, MA, USA) with 3A2813-5Bam and Pro-3X primers (Table 1) using pUCCVORF1 (*15*) encoding the entire ORF1 as a template. Amplified DNA fragments, which encoded the C-terminal 22 amino acids of the 3A-like region plus the entire 3B VPg and 3C region, were digested with *Bam*HI and *Sma*I restriction enzymes and cloned into the corresponding sites of pGEX-2TK (GE Healthcare, Piscataway, NJ, USA). Since the GST fusion protein contains the 3A/3B and 3B/3C cleavage sites, it should be cleaved into three proteins, GST-3a, 3B VPg and 3C-like protease, when *E. coli* cells were transformed with the wild-type plasmid (Fig. 1).

Mutations of Glu54 were introduced into pGEX-2TK-3aBC, and the resultant plasmid was designated pGEX-2TK-3aBC-E54X (X is an introduced amino acid, which is written by one-letter code). The mutagenic primers are listed in Table 1. For the E54A and E54D mutations, the *Bam*HI–*Nru*I digests of DNA fragments that were amplified with 3A2813-5Bam and Pro3454-3Nr primers, and the *Nru*I–*Sma*I digests of DNA fragments that were amplified with E54A-5Nr or E54D-5Nr and Pro-3X primers were cloned into pGEX-2TK. The E54Q mutation was introduced into pGEX-2TK-3aBC with E54Q-1 and E54Q-2 primers using a QuikChange II Site-Directed Mutagenesis kit (Stratagene). For the other 16 kinds of mutation, PrimeSTAR Max DNA polymerase (Takara Bio, Inc.) was used. All mutations were first detected by restriction analysis and verified by DNA sequencing. No unexpected mutations were introduced. As for pGEX-2TK-3aBC-C139A, the *Eco*RI–*Sma*I fragments from pUCHis3BC-C139A (*11*) were exchanged with the corresponding region of pGEX-2TK-3aBC.

pGEX-2TK-3aBC-C139S was constructed by transferring the *Eco*RI–*Sma*I fragments from pUCHis3BC-C139S (*11*) to the corresponding region of pGEX-2TK-3aBC. The Glu54 mutations were introduced to pGEX-2TK-3aBC-C139S by transferring the *Sac*I–*Kpn*I fragments from the respective pGEX-2TK-3aBC-E54X plasmid, resulting

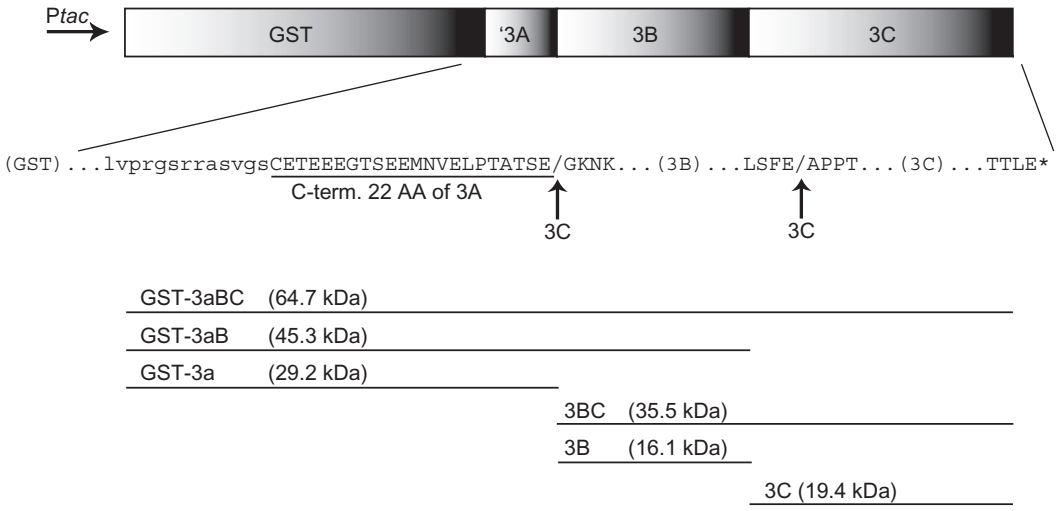

Fig. 1. **Construction of an expression plasmid, pGEX-2TK-3aBC.** The gene fragments encoding the C-terminal 22 amino acid residues of 3A and the entire 3B VPg and 3C-like protease were amplified by PCR as described in the MATERIALS AND METHODS section and fused in-frame to the GST gene of the pGEX-2TK vector. The fusion protein, GST-3aBC, has two 3C-like protease recognition sites (SE/GK and FE/AP). The calculated molecular weight of the parental protein and possible intermediates along with the final products are shown.

in the construction of a series of pGEX-2TK-3aBC-E54X,C139S plasmids.

For the *in vitro* assays of the protease activity, pGEX-2TK-Pro was constructed by ligating the *Nsp*V–*Ehe*I fragment from pUCGST2TK-Pro (*15*) to the *Nsp*V–*Sma*I fragment of pGEX-6P-2 (GE Healthcare). Although a backbone of pGEX-2TK-Pro constructed in this study is that of pGEX-6P-2 vector, there is a thrombin site between GST and the 3C-like protease, but not a Pre-Scission protease site. The *Eco*RI–*Sal*I fragments from mutant pGEX-2TK-3aBC plasmids which contained mutations of Glu54 and/or Cys139 were transferred to pGEX-2TK-Pro, resulting in mutant pGEX-2TK-Pro plasmids.

Expression of the Recombinant Proteins in E. coli Cells and Preparation of Cell Lysate—Escherichia coli BL21-CodonPlus(DE3)-RIPL cells were transformed with each of the pGEX-2TK-3aBC plasmids and grown on 4 ml of MagicMedia *E. coli* Expression Medium (Invitrogen, Carlsbad, CA, USA) at 37°C for 24 h. Cells were harvested and resuspended in 600 µl of PBS containing 1.0% Triton X-100. A one-half volume of glass beads (0.1 mm diameter)

Table 1. **Oligonucleotide primers used for PCR amplification and site-directed mutagenesis.**

Primer	Sequence[a]	Polarity	Related codon change	Restriction site
3A2813-5Bam	5′-AGTATGGATCCTGTGAGACTGAGGAGGAA GGCACCAGCGA-3′	Sense	-	*Bam*HI
Pro-3X	5′-GGCCCCCGGGCTATTACTCTAGGGTGGTTT CACCTTCCCCAGCC-3′	Anti-sense	Stop codons (TAA and TAG) were added after GAG codon for Glu181.	*Sma*I
Pro3454-3Nr	5′-TCACCAGCACGATGGATCGCGATAGATTC AATGGGCTCCCCAAAG-3′	Anti-sense	-	*Nru*I
E54A-5Nr	5′-ATCTATCGCGATCCATCGTGCTGGTGCCTT TACACAATTCAGGTT-3′	Sense	GAA to GCC	*Nru*I
E54D-5Nr	5′-ATCTATCGCGATCCATCGTGCTGGTGATTT TACACAATTCAGGTT-3′	Sense	GAA to GAT	*Nru*I
E54Q-1	5′-GCAATCCATCGTGCTGGCCAATTTACACAAT TCAGG-3′	Sense	GAA to CAA	*Msc*I
E54Q-2	5′-CCTGAATTGTGTAAATTGGCCAGCACGATGG ATTGC-3′	Anti-sense		
E54C-PSF	5′-GTGCCGGCTGTTTTTACACAATTCAGGTTTT-3′	Sense	GAA to TGT	*Nae*I
E54C-PSR	5′-TAAAACAGCCGGCACGATGGATTGCTATAG-3′	Anti-sense		
E54F-PSF	5′-GTGCCGGCTTTTTTACACAATTCAGGTTTT-3′	Sense	GAA to TTT	*Nae*I
E54F-PSR	5′-TAAAAAAGCCGGCACGATGGATTGCTATAG-3′	Anti-sense		
E54G-PSF	5′-GTGCCGGCGGCTTTACACAATTCAGGTTTT-3′	Sense	GAA to GGC	*Nae*I
E54G-PSR	5′-TAAAGCCGCCGGCACGATGGATTGCTATAG-3′	Anti-sense		
E54H-PSF	5′-GTGCTGGCCATTTTACACAATTCAGGTTTT-3′	Sense	GAA to CAT	*Msc*I
E54H-PSR	5′-TAAAATGGCCAGCACGATGGATTGCTATAG-3′	Anti-sense		
E54I-PSF	5′-GTGCCGGCATTTTTACACAATTCAGGTTTT-3′	Sense	GAA to ATT	*Nae*I
E54I-PSR	5′-TAAAAATGCCGGCACGATGGATTGCTATAG-3′	Antisense		
E54K-PSF	5′-GTGCCGGCAAATTTACACAATTCAGGTTTT-3′	Sense	GAA to AAA	*Nae*I
E54K-PSR	5′-TAAATTTGCCGGCACGATGGATTGCTATAG-3′	Anti-sense		
E54L-PSF	5′-GTGCAGGCCTCTTTACACAATTCAGGTTTT-3′	Sense	GAA to CTC	*Aat*I
E54L-PSR	5′-TAAAGAGGCCTGCACGATGGATTGCTATAG-3′	Anti-sense		
E54M-PSF	5′-GTGCCGGCATGTTTACACAATTCAGGTTTT-3′	Sense	GAA to ATG	*Nae*I
E54M-PSR	5′-TAAACATGCCGGCACGATGGATTGCTATAG-3′	Antisense		
E54N-PSF	5′-GTGCCGGCAACTTTACACAATTCAGGTTTT-3′	Sense	GAA to AAC	*Nae*I
E54N-PSR	5′-TAAAGTTGCCGGCACGATGGATTGCTATAG-3′	Anti-sense		
E54P-PSF	5′-GTGCGGGCCCGTTTACACAATTCAGGTTTT-3′	Sense	GAA to CCG	*Apa*I
E54P-PSR	5′-TAAACGGGCCCGCACGATGGATTGCTATAG-3′	Anti-sense		
E54R-PSF	5′-GTGCCGGCCGCTTTACACAATTCAGGTTTT-3′	Sense	GAA to CGC	*Nae*I
E54R-PSR	5′-TAAAGCGGCCGGCACGATGGATTGCTATAG-3′	Anti-sense		
E54S-PSF	5′-GTGCTGGAAGCTTTACACAATTCAGGTTTT-3′	Sense	GAA to AGC	*Hin*dIII
E54S-PSR	5′-TAAAGCTTCCAGCACGATGGATTGCTATAG-3′	Anti-sense		
E54T-PSF	5′-GTGCTGGTACCTTTACACAATTCAGGTTTT-3′	Sense	GAA to ACC	*Kpn*I
E54T-PSR	5′-TAAAGGTACCAGCACGATGGATTGCTATAG-3′	Anti-sense		
E54V-PSF	5′-GTGCCGGCGTATTTACACAATTCAGGTTTT-3′	Sense	GAA to GTA	*Nae*I
E54V-PSR	5′-TAAATACGCCGGCACGATGGATTGCTATAG-3′	Antisense		
E54W-PSF	5′-GTGCCGGCTGGTTTACACAATTCAGGTTTT-3′	Sense	GAA to TGG	*Nae*I
E54W-PSR	5′-TAAACCAGCCGGCACGATGGATTGCTATAG-3′	Anti-sense		
E54Y-PSF	5′-GTGCCGGCTATTTTACACAATTCAGGTTTT-3′	Sense	GAA to TAT	*Nae*I
E54Y-PSR	5′-TAAAATAGCCGGCACGATGGATTGCTATAG-3′	Anti-sense		

[a]Underlines indicate the restriction sites.

was added to the cell suspension, followed by vigorous vortexing for 3 min. After centrifugation at 15,000*g* for 10 min, the supernatant was transferred to a new tube. Proteins were determined with a BCA Protein Assay Kit (Pierce Biotechnology, Rockford, IL, USA) using BSA as a standard.

SDS–PAGE Analysis of Extracted Proteins—Extracted proteins were separated by SDS–PAGE, followed by staining with Coomassie Brilliant Blue R-250 or western blotting with rabbit polyclonal anti-GST (Abcam, Cambridge, MA, USA), anti-VPg (*15*) or anti-protease (*15*) antibodies. Alkaline phosphatase- or horseradish peroxidase-conjugated anti-rabbit IgG (Rockland, Gilbertsville, PA, USA) was used as a secondary antibody, and chemiluminescence developed with Immobilon Western reagents (Millipore, Billerica, MA, USA) was measured by an LAS-3000 Lumino-Image Analyser (Fujifilm, Tokyo, Japan).

GST Binding Assay—Cell extract containing 1 mg of protein was applied to a Glutathione Sepharose 4B MicroSpin column (GE Healthcare), followed by low-speed centrifugation (735*g*) for 1 min. Glutathion-resin-bound proteins were eluted with 100 μl of 50 mM Tris–HCl (pH 8.0) buffer containing 10 mM reduced glutathione. A portion (5 μl) of the eluates was analysed by SDS–PAGE.

Purification of GST-2TK-Pro Fusion Proteins and Assay of the Protease Activity—Escherichia coli BL21-CodonPlus(DE3)-RIPL cells harboring each of the pGEX-2TK-Pro plasmids were grown on MagicMedia *E. coli* Expression Medium at 37°C for 24 h. Proteins were extracted as described above. The GST-2TK-Pro fusion proteins were purified by passing the cell extracts through the Glutathione-Sepharose 4 Fast Flow resin (GE Healthcare). For assays of the protease activity, fluorogenic peptides were synthesized by Peptide Institute, Inc. (Osaka, Japan). The sequence of the peptides is NMA-Glu-Phe-Gln-Met-Gln-Gly-Gln-Val-Lys(DNP)-D-Arg-D-Arg, which includes the sequence between 2C and 3A proteins of Chiba virus. The fluorescence of NMA (2-(*N*-methylamino)benzoyl) moiety, which is quenched by DNP (2,4-dinitrophenyl) moiety when undigested, increases when the peptide bond between Gln and Gly is cleaved by the 3C-like protease. The fluorescence was monitored with the Hitachi F-7000 Fluorescence Spectrophotometer (Tokyo, Japan) (excitation wavelength, 340 nm; emission wavelength, 440 nm). The protease activity was assayed in 300 μl of 20 mM Bis-Tris-Propane–HCl (pH 8.0) buffer containing 5 μM purified GST-2TK-Pro proteins and 10 μM fluorogenic peptides at 37°C. For determination of the kinetic constants, assays were performed in 300 μl of 20 mM Bis-Tris-Propane–HCl (pH 8.0) buffer containing 2 μM purified GST-2TK-Pro proteins and various concentrations (5, 10, 20, 40 or 70 μM) of fluorogenic peptides at 37°C. The peptides were dissolved in DMSO and stored at −30°C. The stock solution of the peptides was added in the assay mixture so that the final concentration of DMSO was below 1%. To calculate the amount of the cleaved substrates, the fluorescence intensity of the reference compound mixture [NMA-labelled FRETS-25-STD1 and DNP-labelled FRETS-25-STD2 (Peptide Institute, Inc.)] at each concentration (5, 10, 20, 40 or 70 μM) was measured and compared with the amount of change in the fluorescence accompanied by enzyme reaction. The kinetic constants were calculated by plotting the initial rate of the proteolysis versus the concentration of peptides.

RESULTS AND DISCUSSION

Effect of the Mutation of Glu54—As shown in Fig. 2A, the GST-3aBC fusion protein including the wild-type 3C-like protease derived from pGEX-2TK-3aBC was cleaved into GST-3a, 3B VPg and the 3C-like protease in *E. coli* cells without any trace of non-cleaved protein or intermediates, indicating that the 3C-like protease was fully active in *E. coli* cells. On the other hand, the C139A mutant protein was not cleaved at all, indicating that the Ala mutation of the Cys139 nucleophile inactivated the enzyme completely, as shown previously (*11*, *15*).

We had previously concluded that Glu54 was not essential for the protease activity since the E54A mutation did not affect the activity (*11*). This identical mutation was introduced into pGEX-2TK-3aBC and the E54A mutant proteins were expressed. Again, as shown in Fig. 2A, the E54A mutation did not affect the cleavage. Furthermore, it was surprising that most of the Glu54 mutants retained the significant proteolytic activity as judged by the expression in *E. coli* cells (Fig. 2A). Based mainly on the results of the GST binding assay, the effects of the mutations could be roughly grouped into five categories (Fig. 2B): (i) mutations that gave a level of activity comparable to that of the wild-type, which consisted of Ala, Cys, Asp, Gly, Gln and Ser; (ii) mutations that affected the activity slightly and left a small amount of the GST-3aB intermediate, which consisted of His, Met and Asn; (iii) mutation that markedly affected the activity and resulted in very small amounts of the GST-3aB intermediate and GST-3a, which consisted of Trp; (iv) mutations that affected the processing of the GST-3aB intermediate to GST-3a and 3B VPg, which consisted of Ile, Leu and Pro; and (v) mutations that inactivated the enzyme completely, which consisted of Phe, Lys, Arg, Thr, Val and Tyr. The results of the GST binding assay were essentially the same as those of western blotting with anti-GST antibody (data not shown).

The E54D mutant, which had a shorter carboxylate than a glutamate, and the E54Q mutant, which had the neutral amide version of a glutamate, exhibited the wild-type protease activity, although a faint band of the GST-3aB intermediate was observed for the E54Q mutant (Fig. 3A). The E54N mutant, which had the neutral amide version of an aspartate, also showed relatively high activity with the GST-3aB intermediate remaining (Fig. 3A). These results indicate that the ability to form a hydrogen bond with His30 is important for the amino acid at position 54, rather than the negative charge. In contrast, as shown in the previous paper (*11*), the E54A mutant also retained high activity comparable to that of the wild-type, although only small amounts of the GST-3aB, 3BC intermediate and parental GST-3aBC protein were detected (Fig. 3A). Moreover, the E54C, E54G and

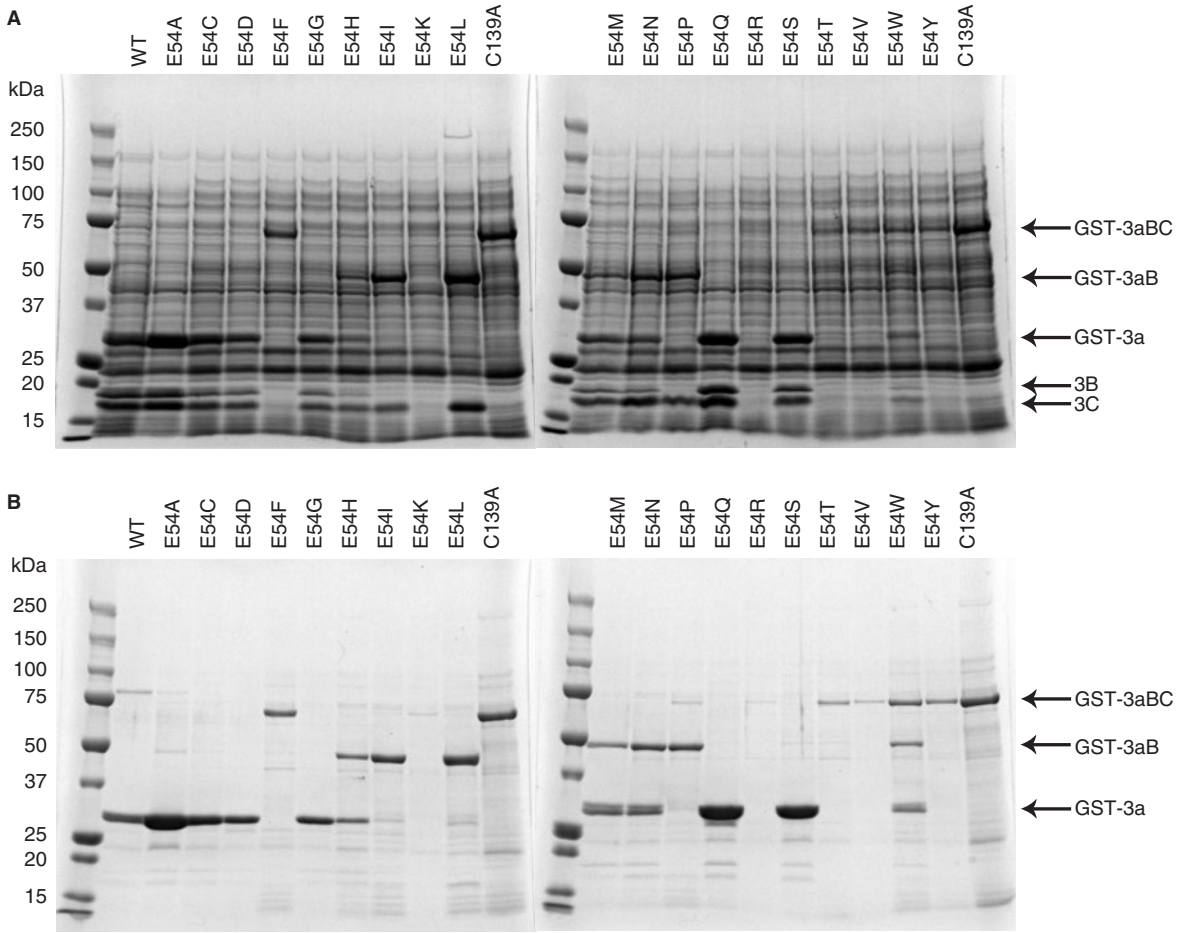

Fig. 2. **Effects of the Glu54 mutations.** (A) Cell lysate (15 μg of proteins) from *E. coli* cells harboring each of the mutant plasmids was subjected to SDS–PAGE, followed by Coomassie Brilliant Blue staining. The bands corresponding to proteins derived from the expression plasmid are indicated. In the SDS–PAGE system used, the 3B VPg protein and the 3C-like protease migrated to around 19 kDa and 17 kDa, respectively. These migrations were confirmed by western blot analysis with the respective antibody (Fig. 3). (B) The GST binding assay was carried out as described in the MATERIALS AND METHODS section. The GST-tagged proteins were also analysed by western blotting with anti-GST antibody (data not shown).

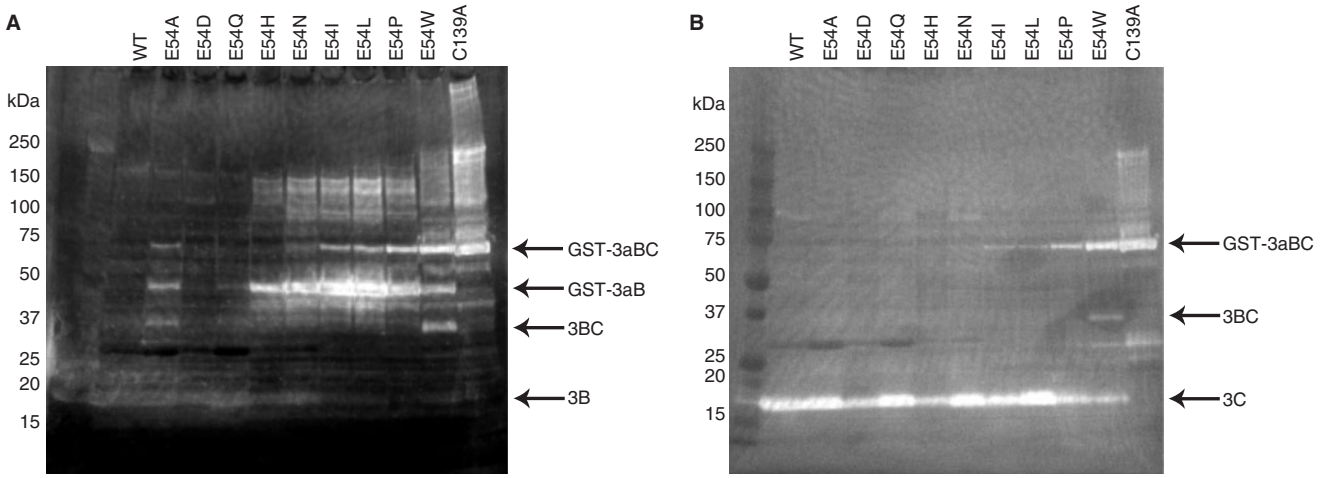

Fig. 3. **Western blot analysis of the Glu54 mutants.** Cell lysate (10 μg of protein) was separated by SDS–PAGE, followed by electroblotting. Proteins were detected by anti-VPg (A) or anti-protease (B) antibodies (*15*). Faint bands which are not indicated with arrows are non-specific. As described in the text, the 3B VPg proteins (~19 kDa) from the wild-type (WT), E54A, E54D, E54Q, E54H and E54N fusions were clearly detected, whereas the amount of the 3B VPg proteins produced from the E54I, E54L and E54P fusions were extremely low (A). This is in marked contrast with the production of the 3C-like protease (B).

E54S mutants, having a small amino acid residue at position 54, exhibited the wild-type activity (western blot data not shown). In these small amino acid mutants, it is possible that a water molecule is located between His30 and the side chain of the amino acid at position 54 and interacts with the imidazole ring of His30. This idea would be supported by the fact that a water molecule interacts with His41 corresponding to His30 of norovirus proteases since the coronavirus main protease (3C-like protease) possesses Val84 at a position equivalent to Glu54 of norovirus proteases (*16*).

On the other hand, to our surprise, the E54T mutant showed no activity. A threonine residue has only an additional methyl residue at the Cβ atom, compared to a serine residue. The E54V mutant also had no protease activity. The side chain of a valine residue is similar to that to a threonine residue, except that a methyl group is substituted for a hydroxyl group. As long as we consider the results of these two mutants, we could assume that the amino acid that had two substituent groups at the Cβ atom was inhibitory for exerting the proteolysis.

However, the E54I mutant showed a significant protease activity, although an isoleucine had two substituent groups (methyl and ethyl) at the Cβ atom and was bigger than a valine. Also, both the E54L mutant (a leucine has a side chain with the same volume as an isoleucine, but a methyl is attached to the Cγ atom.) and the E54P mutant showed the activity similar to that of the E54I mutant. It is notable that the activity of these three mutants was limited because the fused parental GST-3aBC was cleaved into GST-3aB and 3C-like protease, but the GST-3aB intermediate was very rarely cleaved into GST-3a and 3B VPg (Figs 2 and 3). These results indicate that the 3C-like protease having E54I, E54L or E54P mutation is active in the precursor form, but is almost inactive in the mature form. It is interesting that no 3BC intermediate was produced for these three mutants (Fig. 3). These results suggest that the 3B/3C cleavage precedes the 3A/3B cleavage, and that the precursor form might be subtly different from the mature form of the 3C-like protease, at least at its catalytic centre. After cleavage, the conformation of the E54I, E54L or E54P protease might be locked, so that the proteases no longer exert the catalysis, although the precise mechanism is unknown. There is another possibility, namely, that the residue at position 54 is involved in or affects substrate recognition. However, this seems rather unlikely, since Glu54 is far from the substrate according to the model of substrate binding to the 3C-like protease (*12*).

A novel phenomenon was observed for E54I, E54L and E54P mutants, and was not seen in the E54A mutant (as described above) or the E54W mutant (as described below). That is, the E54A and E54W mutants resulted in production of the 3BC intermediate as well as the GST-3aB intermediate, although the amount of 3BC produced appeared to be slightly less than the amount of GST-3aB (Fig. 3). This obviously indicates that both the E54A and E54W proteases (possibly both in the matured protease form and in the precursor form) are able to cleave both the 3A/3B and 3B/3C sites, suggesting that there is virtually no difference in the way of cleaving these two sites, hence there is no difference in the structures of the precursor and mature forms of the protease.

It was likely that an introduction of positively charged residues (a lysine and an arginine) to position 54 was detrimental for the protease and made the proteins unstable, since the mutant proteins were hardly expressed (Fig. 2B). In contrast, the E54H mutant showed significant protease activity comparable to that of the E54N mutant (Figs 2 and 3), although an imidazole ring of the His residue was capable of having a positive charge when protonated. This result is not surprising in light of the fact that human cytomegalovirus serine protease possesses a His–His–Ser catalytic triad (*17*, *18*). In the wild-type norovirus 3C-like protease, a deprotonated, negatively charged side chain of Glu54 interacts with an imidazole ring of His30, which makes the enzyme favourable for the catalysis (*12*, *13*). In the E54H mutant, an imidazole ring of His54 might be a neutral, deprotonated form in order to interact with an imidazole ring of His30.

The E54F and E54Y mutants had no activity (Fig. 2), whereas the E54W mutant, which had an even bigger side chain than a phenylalanine residue and a tyrosine residue, unexpectedly showed low but significant protease activity compared to that of the wild type (Figs 2 and 3). This may have been because a pair of electrons on the N atom of an indole ring was successfully located so as to be able to interact with His30. In the same context, a pair of electrons on the S atom of the Met54 residue in the E54M mutant may have contributed to the hydrogen bond with His30.

Taken together, these results indicate that a glutamate residue itself is not required for an amino acid at position 54, and the Glu54 residue can be replaced with hydrogen bond-forming amino acids such as Asp, Gln and Asn, or small amino acids such as Ala, Cys, Gly and Ser, without a significant loss of activity. In the latter mutant group, it is an intriguing possibility that a water molecule may act in the manner of the side chain of a glutamate, although further investigation into this possibility is needed. It is very interesting that the His, Met and Trp mutants showed a significant activity. In these mutants, the residue at position 54 might donate a pair of electrons for an interaction with His30.

Although norovirus 3C-like proteases have a chymotrypsin-like fold (*12*, *13*), they have a cysteine residue, instead of a serine residue, as a nucleophile. Therefore, they are often called serine-like cysteine proteases. This fact easily leads us to consider the possibility that norovirus 3C-like proteases adopt the mechanism of the proteolysis including a thiolate–imidazolium ion-pair, as indicated by studies on papain, a typical cysteine protease (*19*, *20*). It has been suggested that the rhinovirus 2A protease, one of the chymotrypsin-like cysteine proteases, has a thiolate–imidazolium ion-pair as a catalytic component (*21*). However, Asp38, a third member of the catalytic triad, of the poliovirus 2A protease, which is closely related to the rhinovirus 2A protease, could be replaced with Glu (*22*), but not with Ala (*23*). Thus, norovirus 3C-like proteases are more tolerant than picornavirus 2A proteases to the mutations of the third member carboxylate. On the other hand, some reports have contradicted the possibility that the

viral 3C and 3C-like proteases involve a thiolate–imidazolium ion-pair catalytic mechanism (*24*, *25*). Further kinetics studies will be needed to determine as to which mechanism of proteolysis is adoped by norovirus 3C-like proteases, a general base catalytic mechanism or a thiolate–imidazolium ion-pair catalytic mechanism. In any case, it should be noted that norovirus 3C-like proteases can afford to accommodate various amino acids at position 54 for exerting the proteolysis.

Effect of the Mutation of Glu54 in the Cys139-to-Ser Mutant Background—In order to investigate the mechanism of the proteolysis, we thought that it could be a hint to know the effect of the mutation of the active-site nucleophile on the protease activity. Then we attempted a site-saturation mutagenesis of the Glu54 residue in the background of the Cys139-to-Ser mutation, which mimicked a catalytic triad of typical chymotrypsin-like serine proteases.

As shown in Figs 4 and 5, the C139S single mutant exhibited slightly lower protease activity than the wild-type, and thus could be grouped into category (ii) above. The effects of the Glu54 mutations on the C139S mutant were in contrast to those on the wild-type protease. That is, only the Asp54 and Gln54 mutations supported the proteolysis in the Ser139 background, and the 17 other kinds of mutations did not (Fig. 4). As judged by the amount of the GST-3aB intermediate left behind (Fig. 4B), the activity of the E54Q,C139S mutant was slightly less than that of the E54D,C139S mutant, which in turn was slightly less than that of the parental C139S single mutant, suggesting that a carboxylate was preferable for position 54 in the C139S background. The Gln54

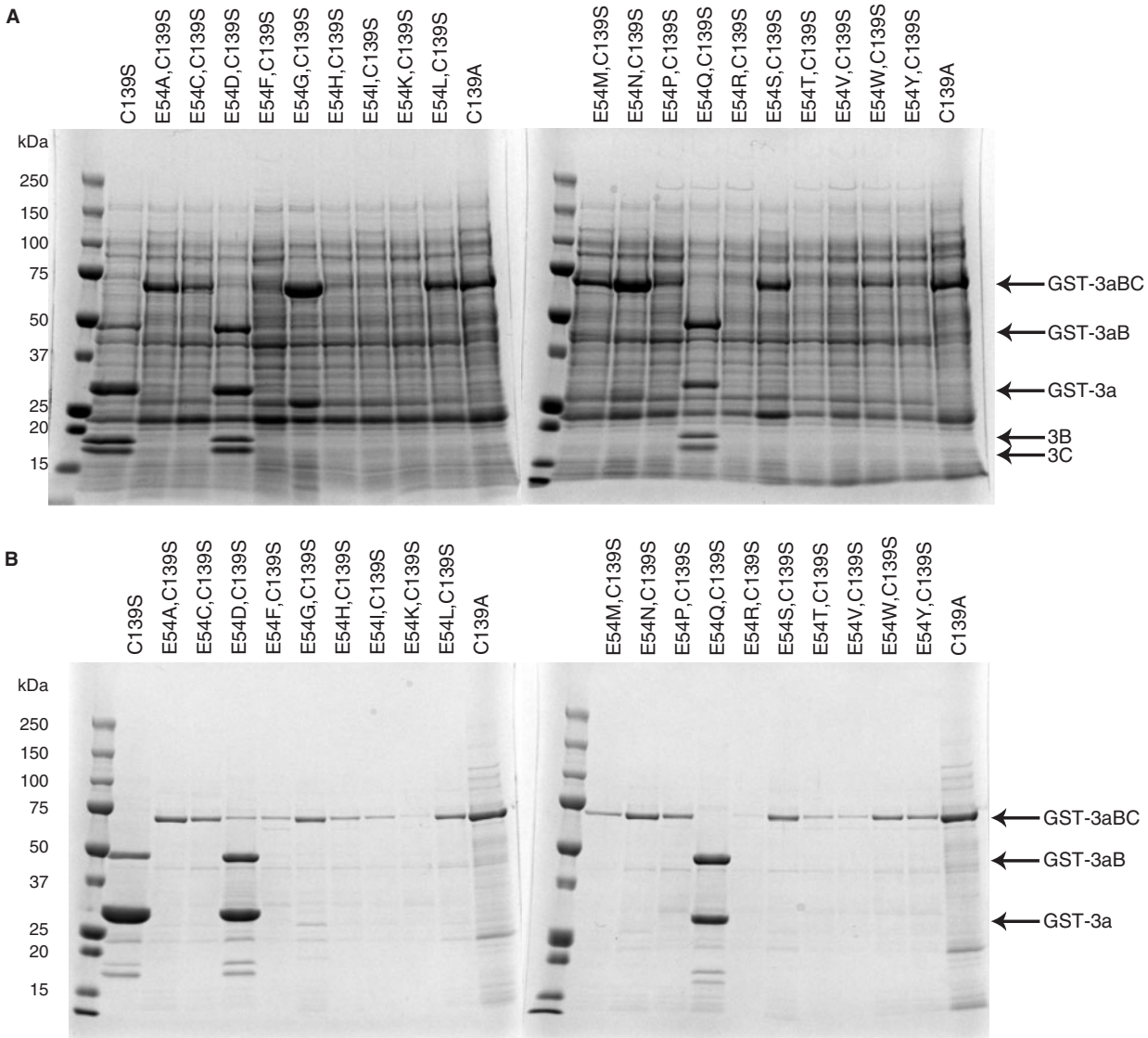

Fig. 4. **Effects of the Glu54 mutations on the C139S mutant.** (A) Cell lysate (15 μg of proteins) from *E. coli* cells harboring each of the mutant plasmids was subjected to SDS–PAGE, followed by Coomassie Brilliant Blue staining. The bands corresponding to proteins derived from the expression plasmid are indicated. (B) The GST binding assay was carried out as described in the MATERIALS AND METHODS section.

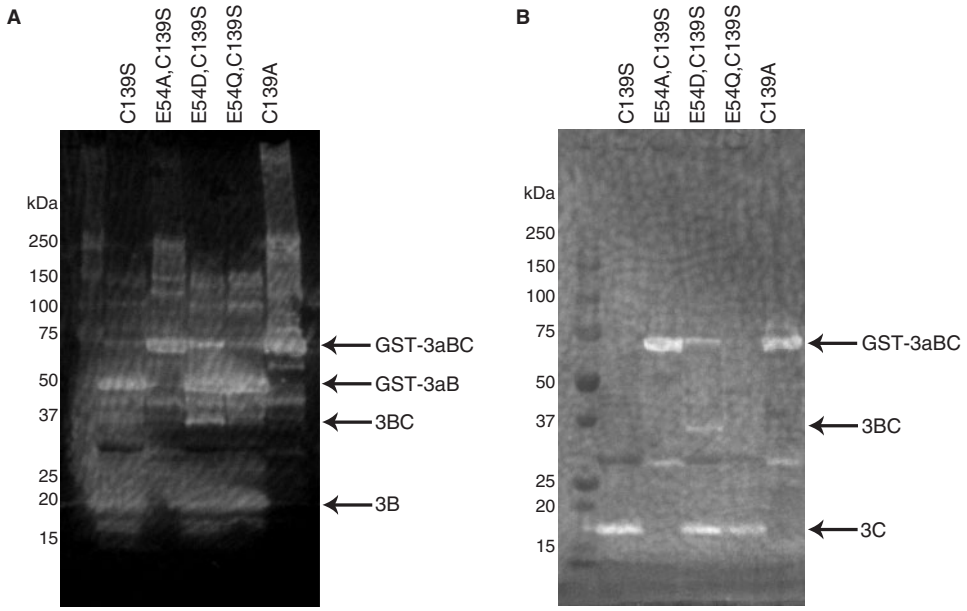

Fig. 5. **Western blot analysis of the Glu54 mutants with the C139S mutation.** Cell lysate (10 μg of protein) was separated by SDS–PAGE, followed by electroblotting. Proteins were detected by anti-VPg (A) or anti-protease (B) antibodies (*15*). Faint bands which are not indicated with arrows were considered non-specific. Both the 3B VPg and 3C-like protease from the C139S, E54D/C139S and E54Q/C139S mutant fusion proteins were clearly detected by the respective antibodies.

residue would provide a pair of electrons that is able to orient the imidazole ring of His30. In contrast, the E54N,C139S mutant showed no proteolytic activity. This was probably because the asparagine side chain was shorter than the glutamine side chain by a carbon atom, resulting in a failure in the operative interaction with His30. As for the E54D,C139S mutant, the Asp54 residue, although an aspartate has a side chain of a size similar to that of an asparagine, would be able to provide a strong negative charge so as to interact with His30. This kind of amino acid requirement for the third member is similar to that in chymotrypsin, in which the Asn mutant of Asp102 was severely defective in the proteolysis (*26*).

We also want to emphasize that the effect of the mutations of Glu54 was quite different between the wild-type and the C139S 3C-like protease. In particular, it was remarkable in the Ala, Cys, Gly and Ser mutations. These mutations were negligible in the wild-type protease, while they were inactivating mutations in the C139S mutant protease. If the Cys139 SH and the Ser139 OH were functionally equivalent during the proteolysis, the effect of the Glu54 mutations would be identical. This was not the case. These experimental results suggest that the Cys139 SH is not functionally equivalent to the Ser139 OH.

The pKa value of the hydroxyl group of a serine is over 10, whereas that of the thiol group of a cysteine is around 8.3. This is why the Cys SH is more ionizable than the Ser OH and is easily deprotonated to be a thiolate. A thiolate is a much better nucleophile than the Ser OH. There is room for argument with regard to the state of the SH group of Cys139 in norovirus 3C-like proteases, but even if the Cys SH is not deprotonated and the thiol is a nucleophile, a proton will be readily withdrawn by His30 upon a nucleophilic attack to a carbonyl carbon of the substrate. It is therefore possible that the proteolysis occurs without a large contribution of an amino acid at position 54 in the wild-type Cys139 background, while in the Ser139 background, the ability for an amino acid at position 54 to interact effectively with His30 is required, so that the residue at position 54 enhances the nucleophilicity of a hydroxyl group of the Ser139 residue.

The In Vitro Assays of the Proteolysis—To further characterize the 3C-like protease, the mutant proteases as well as the wild-type were purified as the GST fusion and subjected to assaying the protease activity using the fluorogenic 2C/3A peptide described under the MATERIALS AND METHODS section. Since the 2C/3A junction was rapidly cleaved when the ORF1 polyproteins were produced in mammalian cells (*5*), we chose the peptide including the 2C/3A junction for the *in vitro* assay. The wild-type GST-2TK-Pro protein cleaved the 2C/3A peptide effectively, and the k_{cat} value of the cleavage by the wild-type protease was $0.63\,s^{-1}$ with the K_m value of 0.53 mM (Table 2). The GST protein derived from the pGEX-2TK vector had no proteolytic activity (data not shown). Therefore, the increase in the fluorescence observed in the assay for the GST-2TK-Pro protein can be attributed to the protease moiety of the fusion. On the other hand, the purified His-tagged 3C-like protease showed only 1% proteolytic activity of the GST-2TK-Pro fusion protein (data not shown). Since at least the GST moiety does not seem to interfere with the protease activity, the GST-2TK-Pro fusion protein was used for *in vitro* assays, and the effects of mutations of Glu54 and Cys139 were analysed.

As for the Glu54 single mutants, the E54A, E54D and E54Q mutants from category (i) and the E54I and E54L mutants from category (iv) were chosen because it was

Table 2. **The kinetic constants for the proteolysis.**

Protease	K_m (mM)	k_{cat} (sec^{-1})	k_{cat}/K_m ($sec^{-1}mM^{-1}$)
Wild-type	0.53 ± 0.37	0.63 ± 0.40	1.19
C139S	0.11 ± 0.02	$1.7 \times 10^{-2} \pm 0.3 \times 10^{-2}$	0.15

The rates of the increase in the fluorescence accompanied by the cleavage of the fluorogenic peptides were measured in the presence of various concentrations (5, 10, 20, 40 or 70 μM) of fluorogenic peptide substrates and 2 μM purified proteases. The K_m and V_{max} values (± SE) were calculated from non-linear regression curves. The k_{cat} value was calculated from the equation, $V_{max} = k_{cat}[E]_0$, where $[E]_0$ is the total concentration of the protease.

Table 3. **The proteolytic activity of the 3C-like proteases.**

Protease	Protease activity (pmol/s)	Relative activity (%)
Wild-type	5.22 ± 0.32	100
E54A	0.45 ± 0.04	8.6
E54D	1.23 ± 0.07	23.6
E54Q	1.06 ± 0.08	20.3
E54I	0.34 ± 0.05	6.5
E54L	0.26 ± 0.04	5.0
C139S	1.02 ± 0.19	19.5
E54D,C139S	0.32 ± 0.06	6.1
E54Q,C139S	0.29 ± 0.01	5.6

The protease activity was measured in the presence of 10 μM fluorogenic peptide substrates and 5 μM purified proteases and expressed as the amount of the cleaved substrate per minute. Data are presented as means ± SE.

expected that the former would show the high activity comparable to that of the wild-type and the latter would have the high activity as the GST fusion that mimicked the precursor form, as judged by the results of Figs 2 and 3. Among these five Glu54 single mutant proteases, the E54D and E54Q mutant proteases had the highest protease activity, being 20–25% activity of the wild-type protease (Table 3). This suggests that the capability of forming a hydrogen bond with His30 is important for the activity. However, we were unable to determine the kinetic constants for these five mutants, even the E54D and E54Q mutants, because the rate of the proteolysis was already saturated at the lowest substrate concentration (5 μM) that was tested and because the activity of mutant proteases tended to be lowered in the higher concentration of the substrate (data not shown). These results indicate that a glutamate is the most appropriate residue at position 54, but clearly several kinds of amino acid residues are permissible at position 54 in order to exert the protease activity.

On the other hand, the C139S mutant protease retained the moderate level of the proteolytic activity and the kinetic constants could be determined successfully (Table 2). Although the k_{cat} value for the C139S mutant is about 3% of that for the wild-type, the C139S mutant showed more than 10% activity of the wild-type when the k_{cat}/K_m values are used for comparison (Table 2). In the presence of 10 μM substrate and 5 μM protease, the C139S mutant showed about 20% activity of the wild-type (Table 3), indicating that the C139S mutant retained relatively high proteolytic activity. Addition of the E54D or E54Q mutation to the C139S mutant decreased the activity (Table 3), being about 30% of the parental C139S single mutant. The degree of the decrease in the activity was almost the same as the effect of the E54D or E54Q mutation on the wild-type protease. This suggests that the side chain that can form a hydrogen bond with His30 is favourable for position 54 regardless of the nucleophile.

It is notable that other mutations, such as E54A, E54I and E54L affected the proteolytic activity of the wild-type protease and the C139S mutant protease in the completely different manner (Figs 2 and 4).

CONCLUSION

The Glu54 residue is not essential for the proteolysis by norovirus 3C-like proteases, but clearly required for the effective proteolysis as indicated by the *in vitro* assays of the proteolysis. Although the mechanism of the proteolysis remains unclear and further *in vitro* analyses would be required, there is an intriguing possibility that norovirus 3C-like proteases adopt the papain-like mechanism involving a thiolate–imidazolium ion pair, despite the fact that they are related structurally to chymotrypsin-like serine proteases. It is also noteworthy that the 3B/3C cleavage appeared to precede the 3A/3B cleavage in the proteolysis, as judged by the cleavage events by the E54I, E54L and E54P mutant fusion proteins. If the cleavage at both the 3A/3B and 3B/3C sites occurs in *trans* (intermolecularly), these two sites are thought to be equivalent against the attack of the 3C-like protease moiety. However, this idea should be denied by the results of the E54I, E54L and E54P mutants. Although these three mutant proteases are active at least in the precursor form, only the GST-3aB intermediate and the 3C-like protease were produced. This indicates that the 3B/3C cleavage event occurs in the manner different from the 3A/3B cleavage event, suggesting the *cis* (intramolecular) event at the 3B/3C site. Released E54I, E54L and E54P mutant 3C-like proteases were inactive probably because the side chain at position 54 sterically hindered the further proteolytic reaction. These results led us to assume that the spatial location of position 54 might vary depending on the 3B/3C cleavage event. The elucidation of the crystal structure of the 3BC intermediate or GST-3aBC fusion protein will be one of the ways of knowing how the 3B/3C site is cleaved by the protease.

FUNDING

Ministry of Health, Labour and Welfare, Japan (partly supported by a Grant-in-aid).

CONFLICT OF INTEREST

None declared.

REFERENCES

1. Goodgame, R. (2007) Norovirus gastroenteritis. *Curr. Infect. Dis. Rep.* **8**, 401–408
2. Estes, M.K., Prasad, B.V., and Atmar, R.L. (2006) Noroviruses everywhere: has something changed? *Curr. Opin. Infect. Dis.* **19**, 467–474

3. Hansman, G.S., Natori, K., Shirato-Horikoshi, H., Ogawa, S., Oka, T., Katayama, K., Tanaka, T., Miyoshi, T., Sakae, K., Kobayashi, S., Shinohara, M., Uchida, K., Sakurai, N., Shinozaki, K., Okada, M., Seto, Y., Kamata, K., Nagata, N., Tanaka, K., Miyamura, T., and Takeda, N. (2006) Genetic and antigenic diversity among noroviruses. *J. Gen. Virol.* **87**, 909–919
4. Daughenbaugh, K.F., Fraser, C.S., Hershey, J.W., and Hardy, M.E. (2003) The genome-linked protein VPg of the Norwalk virus binds eIF3, suggesting its role in translation initiation complex recruitment. *EMBO J.* **22**, 2852–2859
5. Belliot, G., Sosnovtsev, S.V., Mitra, T., Hammer, C., Garfield, M., and Green, K.Y. (2003) In vitro proteolytic processing of the MD145 norovirus ORF1 nonstructural polyprotein yields stable precursors and products similar to those detected in calicivirus-infected cells. *J. Virol.* **77**, 10957–10974
6. Sosnovtsev, S.V., Belliot, G., Chang, K.O., Prikhodko, V.G., Thackray, L.B., Wobus, C.E., Karst, S.M., Virgin, H.W., and Green, K.Y. (2006) Cleavage map and proteolytic processing of the murine norovirus nonstructural polyprotein in infected cells. *J. Virol.* **80**, 7816–7831
7. Prasad, B.V., Hardy, M.E., Dokland, T., Bella, J., Rossmann, M.G., and Estes, M.K. (1999) X-ray crystallographic structure of the Norwalk virus capsid. *Science* **286**, 287–290
8. Cao, S., Lou, Z., Tan, M., Chen, Y., Liu, Y., Zhang, Z., Zhang, X.C., Jiang, X., Li, X., and Rao, Z. (2007) Structural basis for the recognition of blood group trisaccharides by norovirus. *J. Virol.* **81**, 5949–5957
9. Lochridge, V.P. and Hardy, M.E. (2007) A single-amino-acid substitution in the P2 domain of VP1 of murine norovirus is sufficient for escape from antibody neutralization. *J. Virol.* **81**, 12316–12322
10. Shirato-Horikoshi, H., Ogawa, S., Wakita, T., Takeda, N., and Hansman, G.S. (2007) Binding activity of norovirus and sapovirus to histo-blood group antigens. *Arch. Virol.* **152**, 457–461
11. Someya, Y., Takeda, N., and Miyamura, T. (2002) Identification of active-site amino acid residues in the Chiba virus 3C-like protease. *J. Virol.* **76**, 5949–5958
12. Nakamura, K., Someya, Y., Kumasaka, T., Ueno, G., Yamamoto, M., Sato, T., Takeda, N., Miyamura, T., and Tanaka, N. (2005) A norovirus protease structure provides insights into active and substrate binding site integrity. *J. Virol.* **79**, 13685–13693
13. Zeitler, C.E., Estes, M.K., and Prasad, B.V.V. (2006) X-ray crystallographic structure of the Norwalk virus protease at 1.5-Å resolution. *J. Virol.* **80**, 5050–5058
14. Kasuga, K., Tokieda, M., Ohtawara, M., Utagawa, E., and Yamazaki, S. (1990) Small round structured virus associated with an outbreak of acute gastroenteritis in Chiba, Japan. *Jpn J. Med. Sci. Biol.* **43**, 111–121
15. Someya, Y., Takeda, N., and Miyamura, T. (2000) Complete nucleotide sequence of the Chiba virus genome and functional expression of the 3C-like protease in Escherichia coli. *Virology* **278**, 490–500
16. Anand, K., Palm, G.J., Mesters, J.R., Siddell, S.G., Ziebuhr, J., and Hilgenfield, R. (2002) Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J.* **21**, 3213–3224
17. Qiu, X., Culp, J.S., DiLella, A.G., Hellmig, B., Hoog, S.S., Janson, C.A., Smith, W.W., and Abdel-Meguid, S.S. (1996) Unique fold and active site in cytomegalovirus protease. *Nature* **383**, 275–279
18. Chen, P., Tsuge, H., Almassy, R.J., Gribskov, C.L., Katoh, S., Vanderpool, D.L., Margosiak, S.A., Pinko, C., Matthews, D.A., and Kan, C.C. (1996) Structure of the human cytomegalovirus protease catalytic domain reveals a novel serine protease fold and catalytic triad. *Cell* **86**, 835–843
19. Polgár, L. (1974) Mercaptide-imidazolium ion-pair: the reactive nucleophile in papain catalysis. *FEBS Lett.* **47**, 15–18
20. Creighton, D.J., Gessouroun, M.S., and Heapes, J.M. (1980) Is the thiolate–imidazolium ion pair the catalytically important form of papain? *FEBS Lett.* **110**, 319–322
21. Sárkány, Z., Skern, T., and Polgár, L. (2000) Characterization of the active site thiol group of rhinovirus 2A proteinase. *FEBS Lett.* **481**, 289–292
22. Yu, S.F. and Lloyd, R.E. (1991) Identification of essential amino acid residues in the functional activity of poliovirus 2A protease. *Virology* **182**, 615–625
23. Hellen, C.U., Fäcke, M., Kräusslich, H.-G., Lee, C.-K., and Wimmer, E. (1991) Characterization of poliovirus 2A proteinase by mutational analysis: residues required for autocatalytic activity are essential for induction of cleavage of eukaryotic initiation factor 4F polypeptide p220. *J. Virol.* **65**, 4226–4231
24. Sárkány, Z., Szeltner, Z., and Polgár, L. (2001) Thiolate-imidazolium ion pair is not an obligatory catalytic entity of cysteine peptidases: the active site of picornain 3C. *Biochemistry* **40**, 10601–10606
25. Huang, C., Wei, P., Fan, K., Liu, Y., and Lai, L. (2004) 3C-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism. *Biochemistry* **43**, 4568–4574
26. Sprang, S., Standing, T., Fletterick, R.J., Stroud, R.M., Finer-Moore, J., Xuong, N.-H., Hamlin, R., Rutter, W.J., and Craik, C.S. (1987) The three-dimensional structure of Asn102 mutant of trypsin: role of Asp102 in serine protease catalysis. *Science* **237**, 905–909