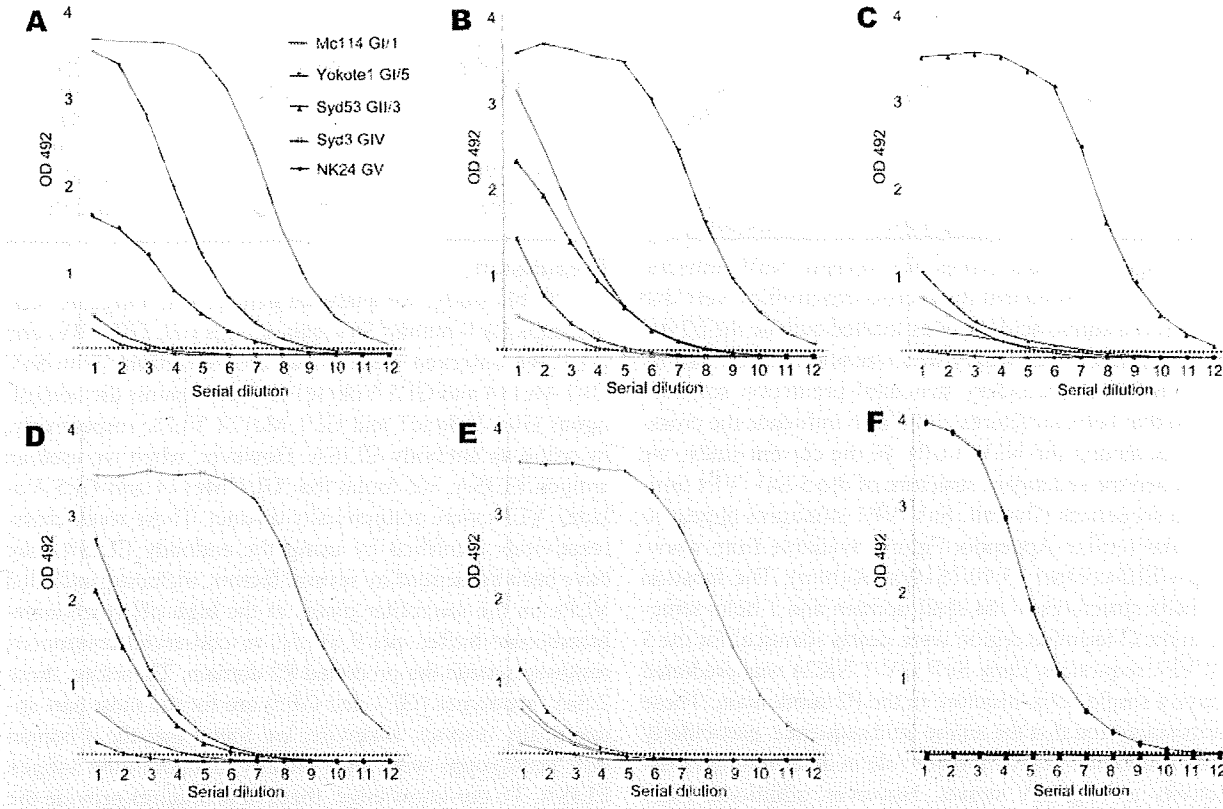

Figure 3. Antibody ELISA absorbances at 492 nm of 6 viruslike particles (VLPs) of sapovirus. A) Mc114 GI/1; B) Yokote1 GI/5; C) Syd53 GII/3; D) Syd3 GIV/1; E) NK24 GV/1; F) NoV Osaka659. Wells were coated with ≈100 ng of purified VLPs. Antiserum was used in 2-fold dilutions from 1:500 to 1,024,000 in phosphate-buffered saline, 0.1% Tween 20, 5% skim milk shown as 1–12 along the x-axis. OD, optical density; NoV, Norwalk virus.

P/N 19.59; and GV/1 NK24 P – N 1.03, P/N 21.58. In the current study, we examined the cross-reactivity among the newly expressed VLPs and those previously prepared by using the antigen ELISA. The antiserum samples reacted only with homologous VLPs, i.e., SaV GII/3 Syd53 P – N 1.02, P/N 21.33; and GIV/1 Syd3 P – N 1.44, P/N 29.74 (Table 3). Several antisera appeared to cross-react with heterologous VLPs, i.e., where the P – N was <0.10, but the P/N ratio was >1.34. SaV GI/5 Yokote1 antisera appeared to cross-react with GII/3 Syd53 VLPs (P/N 1.97); GIV/1 Syd3 antisera appeared to cross-react with GI/5 Yokote1 VLPs (P/N 1.42); and GV/1 NK24 antisera appeared to cross-react with GII/3 Syd53 VLPs (P/N 1.50). However, all of these cross-reactivity results were considered negative because P – N was <0.10.

Amino Acid Alignment and Secondary Structure Prediction

An amino acid alignment of the 5 SaV VP1 sequences (GI/1 Mc114, GI/5 Yokote1, GII/3 Syd53, GIV Syd3, and GV NK24) showed that the shell domain contained more conserved residues than the predicted P domains (Figure 4). However, SaV GI/1 Mc114 and GI/5 Yokote1 shared more conserved continuous residues in the predicted P2 subdomain than other genogroups. The NoV P2 subdomain is thought to contain the determinants of strain specificity, cell binding, and antigenicity. For example, monoclonal antibodies that recognize regions in the P2 subdomain inhibit binding of NoV VLPs to cells (*28*,*29*). In a recent study, we analyzed cross-reactivities among 26 different NoV VLPs (6 GI and 12 GII genotypes) (*30*) and found

Table 2. Titers of antisera to viruslike particles of sapovirus strains in an antibody ELISA

Antisera	Mc114 (GI/1)	Yokote1 (GI/5)	Syd53 (GII/3)	Syd3 (GIV/1)	NK24 (GV/1)
Mc114 (GI/1)	512,000	64,000	32,000	1,000	<1,000
Yokote1 (GI/5)	32,000	512,000	32,000	2,000	4,000
Syd53 (GII/3)	4,000	8,000	512,000	4,000	<1,000
Syd3 (GIV/1)	2,000	16,000	8,000	2,056,000	1,000
NK24 (GV/1)	2,000	2,000	4,000	1,000	2,056,000

Table 3. Antigen ELISA absorbance values of viruslike particles of sapovirus strains

Antisera	Viruslike particles, P – N (P/N)*				
	Mc114 (GI/1)	Yokote1 (GI/5)	Syd53 (GII/3)	Syd3 (GIV/1)	NK24 (GV/1)
Mc114 (GI/1)	0.41 (9.19)	0 (1.00)	0 (1.00)	0 (1.00)	0 (1.00)
Yokote1 (GI/5)	0.01 (1.13)	0.93 (19.59)	0.05 (1.97)	0 (1.00)	0 (1.00)
Syd53 (GII/3)	0 (1.00)	0.02 (1.42)	1.02 (21.33)	0.01 (1.14)	0.03 (1.50)
Syd3 (GIV/1)	0 (1.00)	0 (1.00)	0 (1.00)	1.44 (29.74)	0 (1.00)
NK24 (GV/1)	0 (1.00)	0 (1.00)	0 (1.00)	0.01 (1.11)	1.03 (21.58)

*Measured at 492 nm. P, hyperimmune serum; N, preimmune serum.

broad-range cross-reactivities for several NoV antisera. Our results suggested that these cross-reactivities were due to conserved amino acid residues located outside the P2 domain. Conversely, secondary structure predictions made by using PSIPRED secondary structural prediction software showed that helix structures could also influence the cross-reactivity among the NoV VLPs. In the current study, we determined the secondary structure of the 5 SaV VP1 amino acid sequences. Overall, SaV VP1 structures appear to be similar (online Appendix Figure, available from www.cdc.gov/EID/content/13/10/1519-appG.htm). The location of 3 helix structures in the shell domain and 1 helix structure in the C-terminal region were nearly identical for the 5 SaV VP1 sequences. Only SaV GV/1 NK24 was predicted to have a single helix structure in the P2 subdomain. These results suggested that the amino acid sequence, particularly the P2 subdomain, plays a major role in determining cross-reactivity among SaV strains. However, additional studies, including high-resolution VLP structural analysis, are needed.

Discussion

In this study, we analyzed genetic and antigenic relationships for 4 human SaV genogroups (GI, GII, GIV, and GV). We observed weak 2-way cross-reactivity with SaV GI/1 Mc114 and GI/5 Yokote1 antisera against the heterologous GI/5 Yokote1 and GI/1 Mc114 VLPs, respectively, by using an antibody ELISA. However, when we used an antigen ELISA, we found that GI/1 Mc114 and GI/5 Yokote1 VLPs were antigenically distinct. These weak cross-reactivities identified by using the antibody ELISA may have been influenced by several factors, including unfolded VLPs on the microtiter plates at the high pH (carbonate-bicarbonate buffer, pH 9.6) (*31*) or conserved continuous residues outside the predicted P2 domain. Therefore, these 2 SaV genotypes (GI/1 and GI/5) are for the most part antigenically distinct. Likewise, we found that the 4 human SaV genogroups were antigenically distinct in the antigen ELISA. To our knowledge, these new findings provide the first evidence that SaV antigenicity corresponded well with VP1 genogrouping and genotyping.

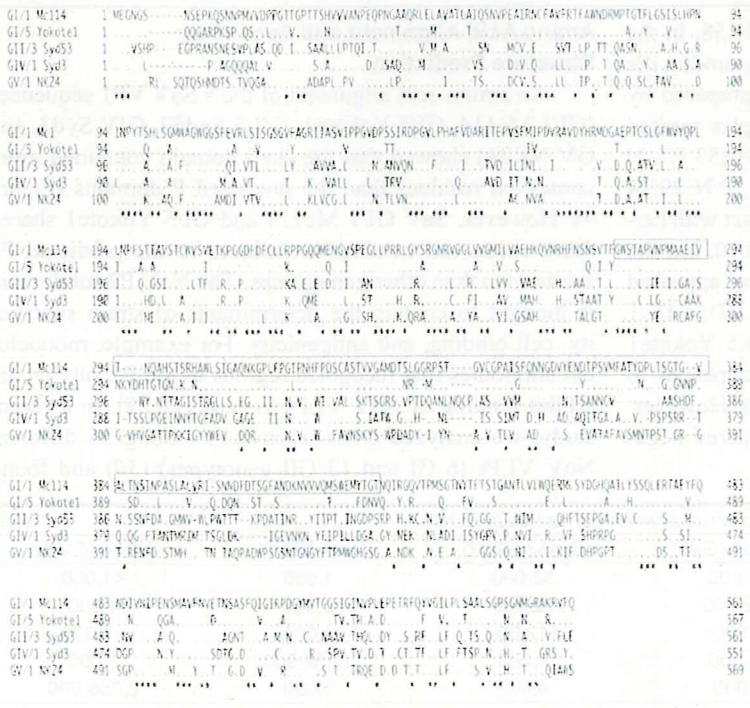

Figure 4. Amino acid alignment of capsid (VP1) sequences of sapoviruses GI/1 Mc114, GI/5 Yokote1, GII/3 Syd53, GIV/1 Syd3, and GV/1 NK24. Sequences in rectangular boxes represent predicted P2 domains (*13*). Asterisks indicate conserved amino acids among these 3 VP1 sequences.

This study was supported in part by a grant for Research on Emerging and Re-emerging Infectious Diseases from the Ministry of Health, Labor and Welfare of Japan, and a grant for Research on Health Science Focusing on Drug Innovation from The Japan Health Science Foundation.

Dr Hansman is a scientist at the National Institute of Infectious Diseases, Tokyo, Japan. His research interests include the epidemiology, expression, and cross-reactivity of viruses that cause gastroenteritis in humans, namely, sapovirus and norovirus.

References

1. Chiba S, Sakuma Y, Kogasaka R, Akihara M, Horino K, Nakao T, et al. An outbreak of gastroenteritis associated with calicivirus in an infant home. J Med Virol 1979;4:249–54.
2. Oka T, Katayama K, Hansman GS, Kageyama T, Ogawa S, Wu FT, et al. Detection of human sapovirus by real-time reverse transcription-polymerase chain reaction. J Med Virol. 2006;78:1347–53.
3. Nakata S, Chiba S, Terashima H, Sakuma Y, Kogasaka R, Nakao T. Microtiter solid-phase radioimmunoassay for detection of human calicivirus in stools. J Clin Microbiol. 1983;17:198–201
4. Hansman GS, Guntapong R, Pongsuwanna Y, Natori K, Katayama K, Takeda N. Development of an antigen ELISA to detect sapovirus in clinical stool specimens. Arch Virol. 2006;151:551–61.
5. Hansman GS, Natori K, Ushijima H, Katayama K, Takeda N. Characterization of polyclonal antibodies raised against sapovirus genogroup five virus-like particles. Arch Virol. 2005;150:1433–7.
6. Okada M, Yamashita Y, Oseto M, Shinozaki K. The detection of human sapoviruses with universal and genogroup-specific primers. Arch Virol. 2006;151:2503–9.
7. Hansman GS, Sano D, Ueki Y, Imai T, Oka T, Katayama K, et al. Sapovirus in water, Japan. Emerg Infect Dis. 2007;13:133–5.
8. Hansman GS, Takeda N, Oka T, Oseto M, Hedlund KO, Katayama K. Intergenogroup recombination in sapoviruses. Emerg Infect Dis. 2005;11:1916–20
9. Katayama K, Miyoshi T, Uchino K, Oka T, Tanaka T, Takeda N, et al. Novel recombinant sapovirus. Emerg Infect Dis. 2004;10:1874–6.
10. Hansman GS, Natori K, Oka T, Ogawa S, Tanaka K, Nagata N, et al. Cross-reactivity among sapovirus recombinant capsid proteins. Arch Virol. 2005;150:21–36.
11. Guo M, Qian Y, Chang KO, Saif LJ. Expression and self-assembly in baculovirus of porcine enteric calicivirus capsids into virus-like particles and their use in an enzyme-linked immunosorbent assay for antibody detection in swine. J Clin Microbiol. 2001;39:1487–93.
12. Hansman GS, Katayama K, Oka T, Natori K, Takeda N. Mutational study of sapovirus expression in insect cells. Virol J. 2005;2:13.
13. Chen R, Neill JD, Noel JS, Hutson AM, Glass RI, Estes MK, et al. Inter- and intragenus structural variations in caliciviruses and their functional implications. J Virol. 2004;78:6469–79.
14. Jiang X, Zhong W, Kaplan M, Pickering LK, Matson DO. Expression and characterization of Sapporo-like human calicivirus capsid proteins in baculovirus. J Virol Methods. 1999;78:81–91.
15. Numata K, Hardy ME, Nakata S, Chiba S, Estes MK. Molecular characterization of morphologically typical human calicivirus Sapporo. Arch Virol. 1997;142:1537–52.
16. Oka T, Hansman GS, Katayama K, Ogawa S, Nagata N, Miyamura T, et al. Expression of sapovirus virus-like particles in mammalian cells. Arch Virol. 2006;151:399–404.
17. Farkas T, Deng X, Ruiz-Palacios G, Morrow A, Jiang X. Development of an enzyme immunoassay for detection of sapovirus-specific antibodies and its application in a study of seroprevalence in children. J Clin Microbiol. 2006;44:3674–9.
18. Prasad BV, Hardy ME, Dokland T, Bella J, Rossmann MG, Estes MK. X-ray crystallographic structure of the Norwalk virus capsid. Science. 1999;286:287–90.
19. Hansman GS, Saito H, Shibata C, Ishizuka S, Oseto M, Oka T, et al. Outbreak of gastroenteritis due to sapovirus. J Clin Microbiol. 2007;45:1347–9.
20. Kageyama T, Shinohara M, Uchida K, Fukushi S, Hoshino FB, Kojima S, et al. Coexistence of multiple genotypes, including newly identified genotypes, in outbreaks of gastroenteritis due to norovirus in Japan. J Clin Microbiol. 2004;42:2988–95.
21. Hansman GS, Katayama K, Maneekarn N, Peerakome S, Khamrin P, Tonusin S, et al. Genetic diversity of norovirus and sapovirus in hospitalized infants with sporadic cases of acute gastroenteritis in Chiang Mai, Thailand. J Clin Microbiol. 2004;42:1305–7.
22. Guntapong R, Hansman GS, Oka T, Ogawa S, Kageyama T, Pongsuwanna Y, et al. Norovirus and sapovirus infections in Thailand. Jpn J Infect Dis. 2004;57:276–8.
23. Hansman GS, Takeda N, Katayama K, Tu ET, McIver CJ, Rawlinson WD, et al. Genetic diversity of sapovirus in children, Australia. Emerg Infect Dis. 2006;12:141–3.
24. Katayama K, Shirato-Horikoshi H, Kojima S, Kageyama T, Oka T, Hoshino F, et al. Phylogenetic analysis of the complete genome of 18 Norwalk-like viruses. Virology. 2002;299:225–39.
25. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics. 2000;16:404–5
26. Farkas T, Zhong WM, Jing Y, Huang PW, Espinosa SM, Martinez N, et al. Genetic diversity among sapoviruses. Arch Virol. 2004;149:1309–23.
27. Okada M, Shinozaki K, Ogawa T, Kaiho I. Molecular epidemiology and phylogenetic analysis of Sapporo-like viruses. Arch Virol. 2002;147:1445–51
28. Lochridge VP, Jutila KL, Graff JW, Hardy ME. Epitopes in the P2 domain of norovirus VP1 recognized by monoclonal antibodies that block cell interactions. J Gen Virol. 2005;86:2799–806.
29. White LJ, Ball JM, Hardy ME, Tanaka TN, Kitamoto N, Estes MK. Attachment and entry of recombinant Norwalk virus capsids to cultured human and animal cell lines. J Virol. 1996;70:6589–97
30. Hansman GS, Natori K, Shirato-Horikoshi H, Ogawa S, Oka T, Katayama K, et al. Genetic and antigenic diversity among noroviruses. J Gen Virol. 2006;87:909–19.
31. White LJ, Hardy ME, Estes MK. Biochemical characterization of a smaller form of recombinant Norwalk virus capsids assembled in insect cells. J Virol. 1997;71:8066–72

Address for correspondence: Grant S. Hansman, Department of Virology II, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashimurayama, Tokyo 208-0011 Japan; email: ghansman@nih.go.jp

WILEY InterScience DISCOVER SOMETHING GREAT
Reviews in Medical Virology

Rev. Med. Virol. 2007; **17**: 133–141.
Published online in Wiley InterScience
(www.interscience.wiley.com)
DOI: 10.1002/rmv.533

REVIEW

Human sapoviruses: genetic diversity, recombination, and classification

Grant S. Hansman*, Tomoichiro Oka, Kazuhiko Katayama and Naokazu Takeda

Department of Virology II, National Institute of Infectious Diseases, Tokyo, Japan

SUMMARY

The family *Caliciviridae* contains four genera *Sapovirus*, *Norovirus*, *Lagovirus* and *Vesivirus*, which include *Sapporo virus* (SaV), *Norwalk virus* (NoV), *Rabbit hemorrhagic disease virus* (RHDV) and *Feline calicivirus* (FCV), respectively. SaV is a causative agent of gastroenteritis in children and adults. SaV can be divided into five genogroups (GI–GV), among which GI, GII, GIV and GV are known to infect humans, whereas SaV GIII infects porcine species. Detection methods include ELISA, RT-PCR and real-time RT-PCR. Since few SaV studies have been conducted, it is difficult to draw correlations between or conclusions about rates of incidence, detection and overall prevalence. Nevertheless, most studies agree that SaV infection is more frequent in young children than adults and that infection in children almost always occurs by 5 years of age. In addition, children at day-care centres and institutions are at greatest risk of SaV-associated infection and transmission. Recently, a number of important findings concerning human SaV were discovered. SaV strains were detected in water samples, which included untreated wastewater specimens, treated wastewater samples and river samples. SaV strains were also detected in shellfish samples destined for human consumption, and recombinant SaV strains were identified in a number of different countries. The purpose of this review was to highlight the current knowledge of human SaV, which appears to be an increasingly important virus causing gastroenteritis in humans. Copyright © 2007 John Wiley & Sons, Ltd.

Received: 20 September 2006; Accepted: 19 December 2006

INTRODUCTION

The family *Caliciviridae* contains four genera *Sapovirus*, *Norovirus*, *Lagovirus* and *Vesivirus*, which include *Sapporo virus* (SaV), *Norwalk virus* (NoV), *Rabbit hemorrhagic disease virus* (RHDV) and *Feline calicivirus* (FCV), respectively. SaV and NoV are etiological agents of human gastroenteritis. SaV is a positive-sense polyadenylated single-stranded RNA virus that infects humans and porcine species [1]. The prototype strain of human SaV, the Sapporo virus, was originally discovered from an outbreak in an orphanage in Sapporo, Japan, in 1977 [2]. In that study, Chiba *et al.* identified viruses with the typical animal calicivirus morphology, the Star-of-David structure, by electron microscopy (EM). Besides having this classical structure, SaV particles are typically 41–46 nm in diameter and have a cup-shaped depression and/or 10 spikes on the outline (see Figure 3). Because only a limited number of SaV studies have been conducted, it has been difficult to check for correlations between or to draw conclusions about rates of incidence, detection and overall prevalence. SaV can be divided into five genogroups (GI–GV) (Figure 1), among which GI, GII, GIV and GV are known to infect humans, whereas SaV GIII infects porcine species. Phylogenetic studies have also designated SaV clusters or genotypes.

*Corresponding author: G. S. Hansman, Department of Virology II, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashimurayama, Tokyo 208-0011, Japan.
E-mail: ghansman@nih.go.jp

DETECTION OF SaV

SaV was first detected by EM [2]. However, this technique is tedious, since virus particles are difficult to correctly identify and the number of particles is generally low [3]. Several groups have used enzyme immunoassays (EIAs) to screen for SaV antibodies [4–8]. Recently, we designed an SaV ELISA based on hyperimmune rabbit and guinea pig antisera raised against SaV virus-like particles (VLPs) [9,10]. The ELISA had 100% specificity and

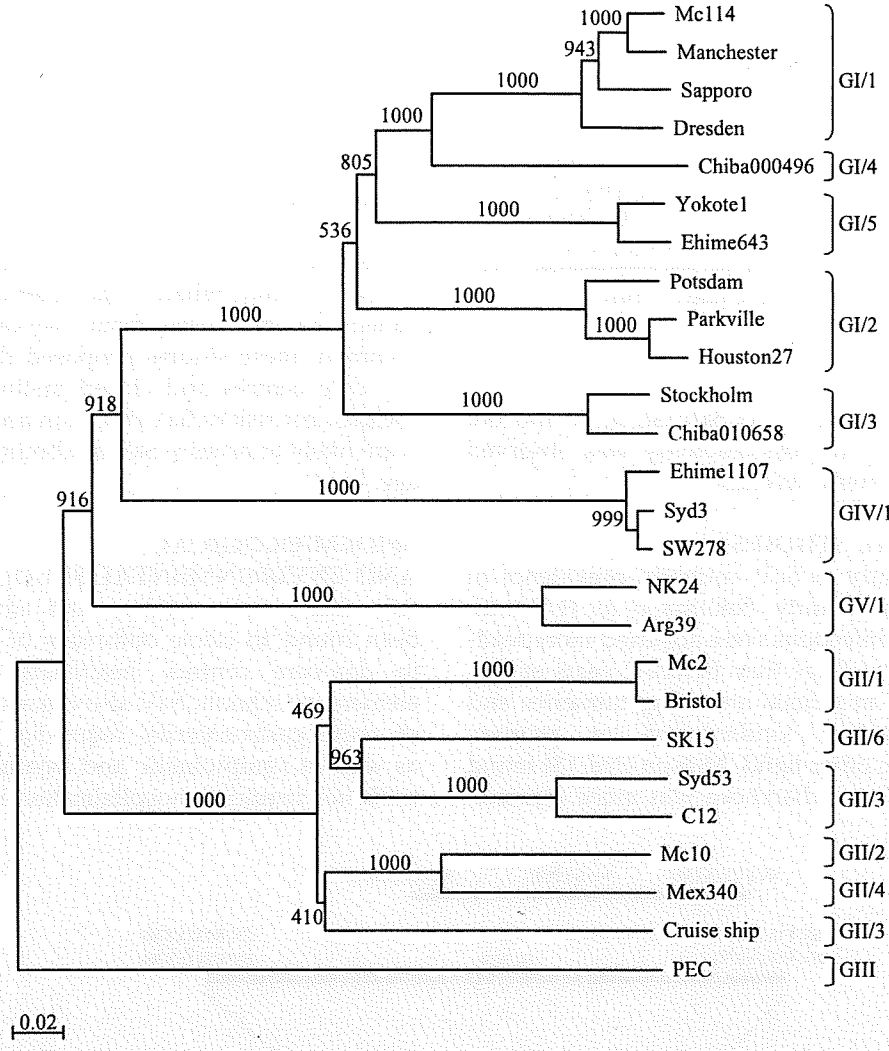

Figure 1. Phylogenetic tree of SaV based upon the entire VP1 nucleotide sequences. Different genogroups and genotypes are indicated. The numbers on each branch indicate the bootstrap values for the genotype. Bootstrap values of 950 or higher were considered statistically significant for the grouping. The scale represents nucleotide substitutions per site. GenBank accession numbers for the reference strains are as follows: Arg39, AY289803; Bristol, HCA249939; C12, AY603425; Chiba000496, AJ412800; Chiba010658, AJ606696; Cruise ship, AY289804; Dresden, AY694184; Ehime643, DQ366345; Ehime1107, DQ058829; Houston27, U95644; Manchester, X86560; Mc2, AY237419; Mc10, AY237420; Mc114, AY237422; Mex340, AF435812; NK24, AY646856; Parkville, U73124; PEC, AF182760; Potsdam, AF294739; Sapporo, U65427; SK15, AY646855; Stockholm, AF194182; SW278, DQ125333; Syd3, DQ104357; Syd53, DQ104360; and Yokote1 AB253740

sensitivities of 60 and 25% when compared to single-round PCR and nested PCR, respectively [9]. Our results showed that ELISA was useful for detecting SaV antigens in clinical stool specimens collected 2 days after the onset of illness. However, the method most widely used these days is reverse transcription-PCR (RT-PCR), which has a high sensitivity and can also be used for genetic analysis. Several groups have designed RT-PCR primers that can detect a broad range of SaV strains (Table 1 and Figure 2a) [11–15]. More recently, we designed a novel TaqMan-based real-time RT-PCR assay that is sensitive and can detect a broad range of genetically diverse human SaV strains [16]. Analysis using clinical stool specimens revealed that the TaqMan-based real-time

DOI: 10.1002/rmv

Table 1. Primers for detection of SaV [15]

Name	Location[a]	Sequence
SV-F13	5074–5094	GAYYWGGCYCTCGCYACCTAC
SV-F14	5074–5094	GAACAAGCTGTGGCATGCTAC
SV-R13	5876–5861	GGTGANAYNCCATTKTCCAT
SV-R14	5876–5861	GGTGAGMMYCCATTCTCCAT
SV-F22	5154–5172	SMWAWTAGTGTTTGARATG
SV-R2	5591–5572	GWGGGRTCAACMCCWGGTGG

[a]Nucleotide numbers of Manchester virus.

RT-PCR assay could detect SaV GI, II, IV and GV sequences, and no cross-reactivity was observed against other enteric viruses.

SEROLOGICAL STUDIES

In an early study of SaV antibody prevalence in the general community, Sakuma *et al.* [17] indicated that SaV infections were acquired more readily after 2 years of age than before 2 years of age, and especially in infants attending nurseries and children attending kindergarten or primary schools. On the other hand, Matson *et al.* [3] found that SaV-associated diarrhoea was more common in infants aged 6 months or less and that the SaV detection rate was higher in a day-care centre population than in hospitalised children (in a concurrent study). Grohmann *et al.* [18] also found that infants under 6 months of age had the highest rate of SaV-associated gastroenteritis. However, a direct comparison between these studies is difficult, as the specimens collected by Sakuma *et al.* were from children without gastroenteritis at an outpatient clinic whereas the other two studies collected the specimens from day-care centres. In common, these studies proposed that children in day-care centres and closed institutions were at the greatest risk of SaV infection and that children were likely to acquire SaV antibodies by 5 years of age.

EPIDEMIOLOGICAL AND ENVIRONMENTAL STUDIES

SaV infects both children and adults and have been found to cause outbreaks of gastroenteritis in day-care centres, healthcare facilities and elementary schools. SaV also causes sporadic cases of acute gastroenteritis requiring hospitalisation as well as symptomatic and asymptomatic infections not requiring hospitalisation in the commu-

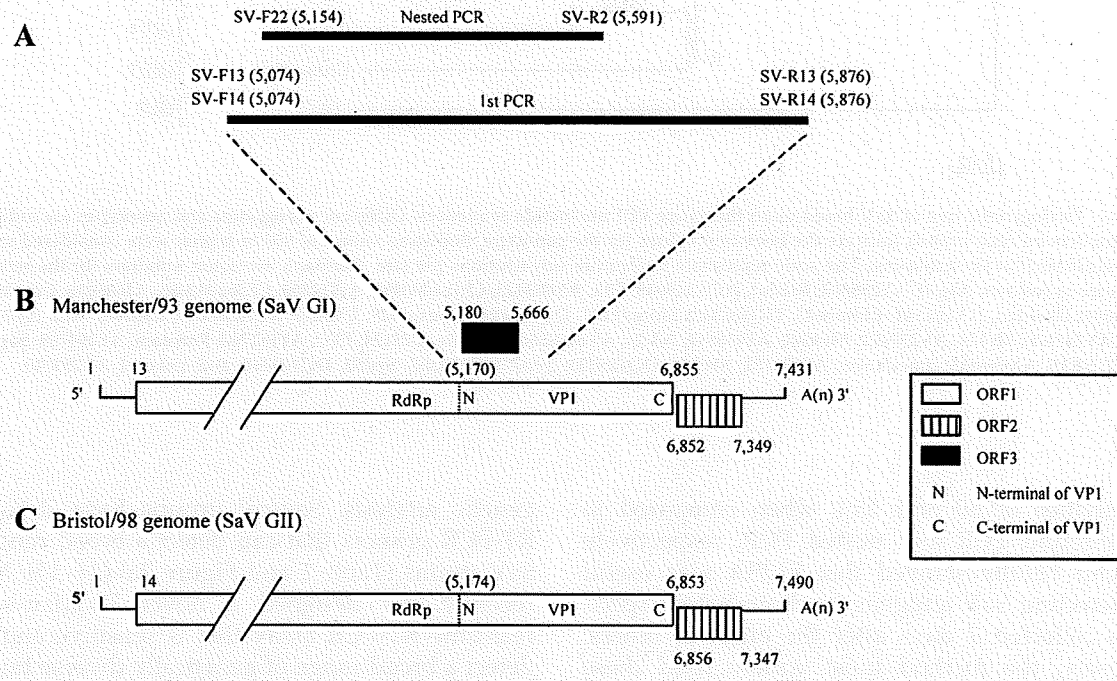

Figure 2. (A) Degenerate primers are directed against a conserved junction sites between the RdRp and the VP1 region and two conserved VP1 sites. (B) The SaV GI, GIV and GV genomes contain three ORFs, whereas (C) the SaV GII genome contains two ORFs

Rev. Med. Virol. 2007; **17**: 133–141.
DOI: 10.1002/rmv

nity [3,5,6,11–14,18–26]. Epidemiological studies have been conducted in many countries, including Australia, Canada, Finland, France, Japan, Mexico, Mongolia, Spain, Sweden, Taiwan, Thailand, UK, US and Vietnam. The rates of incidence, detection and overall prevalence of SaV infections vary in each country and setting and are likely affected by the diagnostic techniques used [27]. A number of reports have noted that SaV detection rates were usually much less frequent than NoV detection rates [24,28–30]. In addition, SaV gastroenteritis appears to produce symptoms milder than those by NoV, so hospitalisation is often unnecessary [28,29,31]. On the other hand, we performed two similar SaV and NoV epidemiological studies among hospitalised infants with gastroenteritis in Thailand and Vietnam, and found SaV in 3.8% (4/105 with single infection) and 0.2% (1/448 with single infection) of stool specimens, respectively [32,33]. This dissimilarity in the SaV detection rates was statistically significant (Fisher's exact $p < 0.005$), whereas the dissimilarity in the NoV detection rates was not statistically significant (in Thailand, 8/105 were NoV-positive with single infection, whereas in Vietnam, 72/448 were NoV-positive with single infection; Fisher's exact $p = 0.1$). The same primers and conditions were used in both studies, thereby suggesting that SaV was an uncommon etiological agent of gastroenteritis in Vietnam. Climatic and environmental conditions as well as cultural differences, including eating habits and hygiene practices, may be important factors in these differences in SaV detection rates between these two countries [34]. NoVs have been detected in oysters, shellfish, drinking fountains, bottled mineral water, ice and community drinking water [35–39]. We recently found SaV in 10% of concentrated water samples, which included untreated wastewater specimens, treated wastewater samples and river samples, but we did not detect SaV in open seawater samples [40]. The SaV samples were isolated in both hot and cold months, that is summer and winter and the sequences matched previously reported SaV sequences. Our results indicated that SaV persists in the natural environment. We also found SaV in shellfish samples destined for human consumption, but we did not detect SaV in oyster samples (manuscript in press). The closely matching SaV sequences detected in the water, shellfish, and patients in Japan suggest that SaV contaminations in the natural environment may lead to food-borne infections in humans. However, further studies are needed to determine if shellfish contaminated with SaV can cause gastroenteritis in humans.

GENOMIC ORGANISATION

The SaV GI, GIV, and GV genomes contain three open reading frames (ORFs), whereas the SaV GII genome contains two ORFs (Figure 2B,C). ORF1 encodes all the non-structural proteins, including the RNA-dependent RNA polymerase (RdRp), and the major capsid protein (VP1). ORF2 encodes a small protein, believed to be similar to VP2 of NoV [1], and ORF3 encodes a protein of unknown function. The SaV ORF1 is organised in the same way as that of RHDV. Matson *et al.* [41] found that the SaV RdRp gene sequence was closer to RHDV and FCV than to those of other human NoVs. Although human SaV is noncultivable, the expression of the recombinant VP1 (rVP1) in a baculovirus expression system or mammalian expression system results in the self-assembly of VLPs that are morphologically similar to the native virus particles, that is approximately 41–46 nm in diametre with cup-shaped depressions and/or 10 spikes on the outline (Figure 3) [42–45]. In a recent study, we expressed constructs containing SaV N- and C-terminal-deleted rVP1 in a baculovirus expression system, which included two N-terminal-deleted rVP1 constructs that began at 49 and 143 nucleotides from the VP1 start and four C-terminal-deleted rVP1 constructs that had an introduced stop codon at nucleotides 687, 1068, 1260 and 1583 (with respect to the VP1 start) [46]. Our results were similar to those reported in a RHDV N- and C-terminal-deleted rVP1 expression study [47], but were distinct from those reported in a NoV N- and C-terminal-deleted rVP1 expression study [48]. For instance, only SaV and RHDV proteins derived from N-terminal-deleted rVP1 constructs, that is constructs that began at 49 and 90 nucleotides from the VP1 start for SaV and RHDV, respectively, assembled into VLPs, whereas both NoV N- and C-terminal-deleted rVP1 constructs assembled into VLPs. This suggested that SaV and RHDV may have similar expression requirements. Furthermore, the cleavage map of SaV ORF1 appeared to be more closely related to that of RHDV than to those of other caliciviruses [49].

DOI: 10.1002/rmv

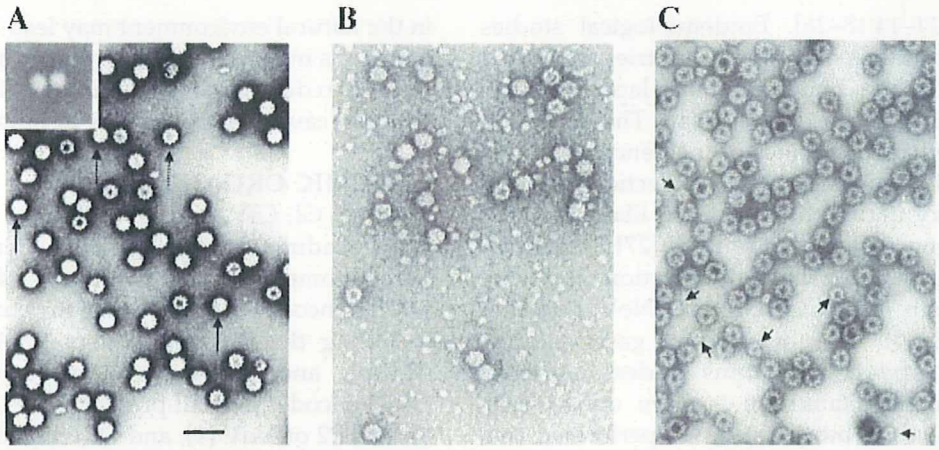

Figure 3. Electron-microscopic images of CsCl purified VLPs negative-stained with 4% uranyl acetate (pH 4) (A) GI Mc114 (Insert, native SaV), (B) GII C12 and (C) GV NK24. The long arrows indicate the Star-of-David structure and the dashed arrows indicate the 10 spikes (A) and the short arrows indicate the small VLPs (C). The bar indicates 100 nm

RECOMBINATION

We recently sequenced the complete genomes of 15 SaV strains (data not shown) and identified the first naturally occurring intragenogroup recombinant SaV strains (Figure 4, i.e. strains Mc10 and C12) [50]. Initially, we grouped Mc10 and C12 into two distinct GII genotypes based on their VP1 sequences (Figure 4B). In addition, the overall genomic nucleotide similarity between Mc10 and C12 was 84.3%, while ORF1 and ORF2 shared 85.5 and 73.3% nucleotide identity, respectively. Our results indicated that they were genetically distinct. However, by comparison of sequence similarity across the length of the genomes, using SimPlot software [51], we discovered a potential recombination site, where the similarity analysis showed a sudden drop in nucleotide identity after the RdRp region (Figure 5). Nucleotide sequence analysis of ORF1 without the downstream VP1 sequence and the VP1 sequence revealed 90.1 and 71.3% nucleotide identity, respectively (Figure 5). These results suggested a single point recombination event occurred at the RdRp–VP1 junction. Interestingly, for Mc10 and C12, there were 44 nucleotides at the RdRp–VP1 junction that matched 100%. This conserved site may represent either the break and rejoin site or the site for copy choice for Mc10 and C12 during viral replication [52], though direct evidence is lacking. More recently, we identified the first intergenogroup recombinant SaV strains (Figure 4, i.e. strains SW278 and Ehime1107) [53]. Based on the classification scheme of either the partial or complete VP1 sequences in our previous studies, we grouped Manchester and Dresden into GI; Bristol, Mc2, Mc10, C12 and SK15 into GII; PEC into GIII; SW278 and Ehime1107 into GIV; and NK24 into GV [11,45,54]. These genogroups were not maintained when we analysed the non-structural region (i.e. between genome start and VP1 start). We found that SW278 and Ehime1107 clustered into GII for the non-structural region-based-grouping (Figure 4A), but clustered into GIV for the structural region-based-grouping (Figure 4B). Comparisons of the complete genome sequences showed that SW278 and Ehime1107 shared over 97% nucleotide identity and likely represented the same strain, although isolated from different countries (Sweden and Japan, respectively). By comparison of sequence similarity across the length of the genomes, we observed a sudden drop in nucleotide similarity after the RdRp region for SW278 and Ehime1107, indicating that a recombination event occurred at the RdRp–VP1 junction (data not shown). The non-structural region of SW278 and Ehime1107, that is a GIV sequence, has not yet been identified, despite the fact that we have conducted a number of molecular epidemiological studies using broad-range primers. Our studies have found that GIV strains (i.e. VP1 sequences) were detected less often than strains from the other genogroups

DOI: 10.1002/rmv

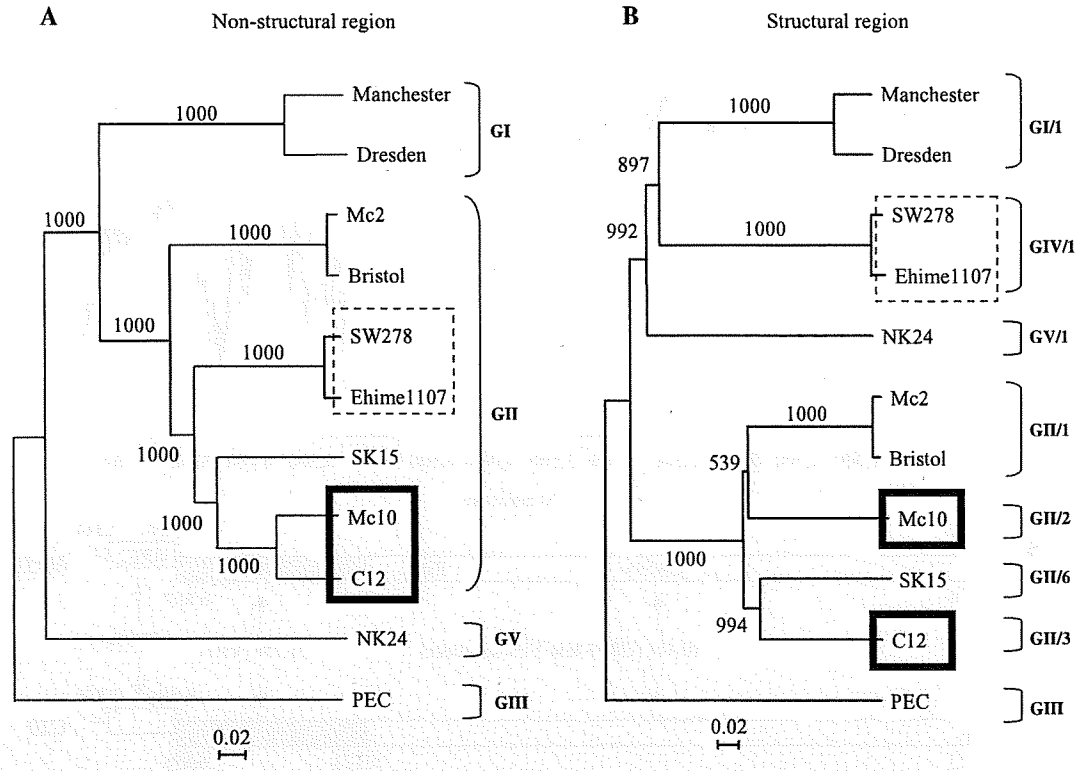

Figure 4. Intra- and intergenogroup recombination among SaV. Phylogenetic tree of (A) non-structural region (i.e. between genome start and VP1 start) and (B) structural region (i.e. between VP1 start and genome end), showing the different genogroups. The numbers on each branch indicate the bootstrap values for the genotype. Bootstrap values of 950 or higher were considered statistically significant for the grouping [62]. The scale represents nucleotide substitutions per site. Mc10 and C12 were the intragenogroup recombinant SaV strains (black box), whereas SW278 and Ehime1107 were intergenogroup recombinant SaV strains (dashed box)

[11,33,40,45,54,55]. Further complete genome analysis of other SaV strains is needed to identify other recombinant strains and determine the extent of recombination in the *Sapovirus* genus.

BINDING TO HUMAN BLOOD GROUP ANTIGENS

In the past several years, increasing evidence has emerged that indicates that human NoVs bind to histo-blood group antigens (HBGAs) [56–60]. These carbohydrate epitopes are present in mucosal secretions and throughout many tissues of the human body, including the small intestine, which certain NoV strains may specifically target. In a recent study, we examined the binding activities of human SaV VLPs to HBGAs present in human saliva and to synthetic carbohydrates [61]. We found that SaV GI and GV VLPs showed no binding activity to HBGAs or synthetic carbohydrates. However, a number of studies have found that different NoV strains exhibit different binding patterns to HBGAs [58–60]. For example Hawaii virus VLPs had no binding to HBGAs, whereas Norwalk virus VLPs had binding activity to HBGAs. Therefore, further studies are needed to examine the possibility that other human SaV genogroups or genotypes have binding activity to HBGAs.

CONCLUDING REMARKS

SaV-associated infection is becoming increasingly important as a result not only of findings by improved detection techniques but also of greater knowledge of the strains' genetic diversity. Diagnostic techniques are now available to properly screen specimens in order to better understand the overall prevalence of SaV and their epidemiologic characteristics. The recent detection of SaV in

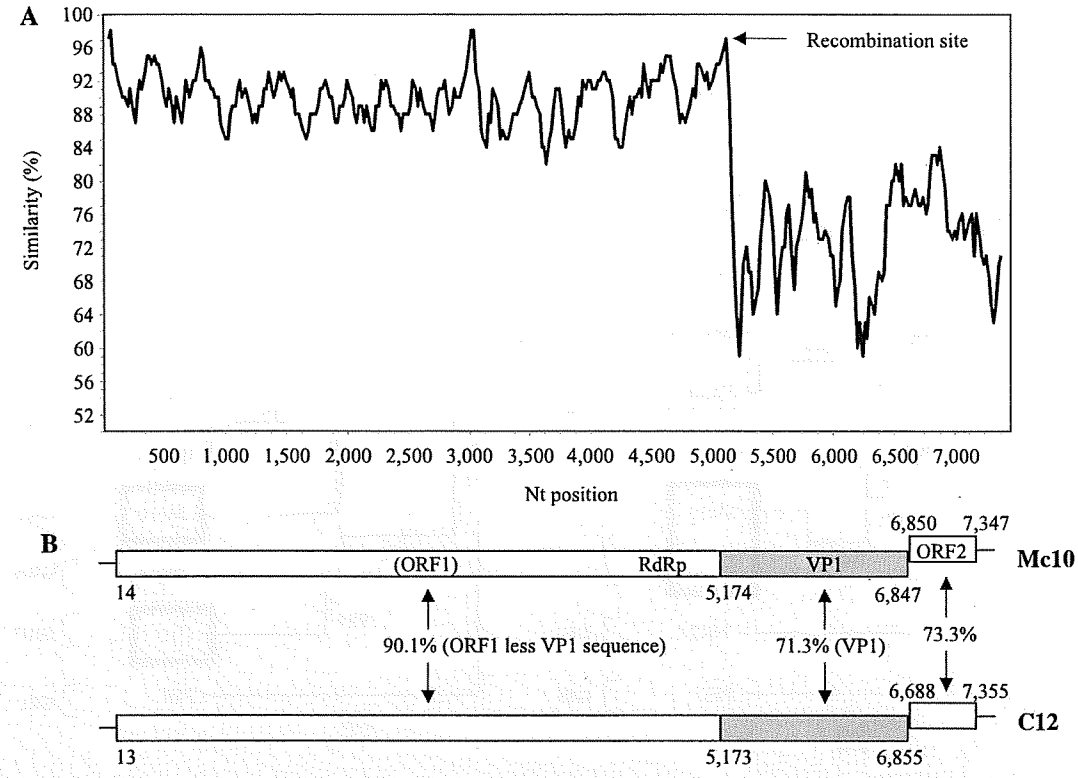

Figure 5. Intragenogroup recombination observed in Mc10 and C12 strains. (A) SimPlot analysis of Mc10 and C12. The Mc10 genome sequence was compared to that of C12 by using a window size of 100 bp with an increment of 20 bp. All gaps were removed. The recombination site is suspected to be located between RdRp and VP1 genes, as the arrow shows. (B) Genome organisation of Mc10 and C12 strains and the nucleotide sequence similarity of different genomic regions

water and shellfish samples creates a greater awareness of these viruses and highlights a possible food-borne transmission route. Recombinant SaV strains are also important and have been identified in a number of countries. And finally, a consensus classification scheme is needed to reduce confusion and properly classify strains while taking recombinant strains and antigenic relatedness into account.

REFERENCES

1. Atmar RL , Estes MK. Diagnosis of noncultivatable gastroenteritis viruses, the human caliciviruses. *Clin Microbiol Rev* 2001; **14**(1): 15–37.
2. Chiba S, Sakuma Y, Kogasaka R, *et al.* An outbreak of gastroenteritis associated with calicivirus in an infant home. *J Med Virol* 1979; **4**(4): 249–254.
3. Matson DO, Estes MK, Glass RI, *et al.* Human calicivirus-associated diarrhea in children attending day care centers. *J Infect Dis* 1989; **159**(1): 71–78.
4. Nakata S, Chiba S, Terashima H, *et al.* Microtiter solid-phase radioimmunoassay for detection of human calicivirus in stools. *J Clin Microbiol* 1983; **17**(2): 198–201.
5. Farkas T, Jiang X, Guerrero ML, *et al.* Prevalence and genetic diversity of human caliciviruses (HuCVs) in Mexican children. *J Med Virol* 2000; **62**(2): 217–223.
6. Wolfaardt M, Taylor MB, Booysen HF, *et al.* Incidence of human calicivirus and rotavirus infection in patients with gastroenteritis in South Africa. *J Med Virol* 1997; **51**(4): 290–296.
7. Chiba S, Nakata S, Numata-Kinoshita K, Honma S. Sapporo virus: history and recent findings. *J Infect Dis* 2000; **181**(Suppl. 2): S303–S308.
8. Nakata S, Estes MK, Chiba S. Detection of human calicivirus antigen and antibody by enzyme-linked immunosorbent assays. *J Clin Microbiol* 1988; **26**(10): 2001–2005.
9. Hansman GS, Guntapong R, Pongsuwanna Y, *et al.* Development of an antigen ELISA to detect sapo-

virus in clinical stool specimens. *Arch Virol* 2006; **151**(3): 551–561.
10. Hansman GS, Natori K, Ushijima H, *et al.* Characterization of polyclonal antibodies raised against sapovirus genogroup five virus-like particles. *Arch Virol* 2005; **150**(7): 1433–1437.
11. Hansman GS, Katayama K, Maneekarn N, *et al.* Genetic diversity of norovirus and sapovirus in hospitalized infants with sporadic cases of acute gastroenteritis in Chiang Mai, Thailand. *J Clin Microbiol* 2004; **42**(3): 1305–1307.
12. Okada M, Shinozaki K, Ogawa T, Kaiho I. Molecular epidemiology and phylogenetic analysis of Sapporo-like viruses. *Arch Virol* 2002; **147**(7): 1445–1451.
13. Vinje J, Deijl H, van der Heide R, *et al.* Molecular detection and epidemiology of Sapporo-like viruses. *J Clin Microbiol* 2000; **38**(2): 530–536.
14. Jiang X, Huang PW, Zhong WM, *et al.* Design and evaluation of a primer pair that detects both Norwalk- and Sapporo-like caliciviruses by RT-PCR. *J Virol Methods* 1999; **83**(1–2): 145–154.
15. Okada M, Yamashita Y, Oseto M, Shinozaki K. The detection of human sapoviruses with universal and genogroup-specific primers. *Arch Virol* 2006; **151**(12): 2503–2509.
16. Oka T, Katayama K, Hansman GS, *et al.* Detection of human sapovirus by real-time reverse transcription-polymerase chain reaction. *J Med Virol* 2006; **78**(10): 1347–1353.
17. Sakuma Y, Chiba S, Kogasaka R, *et al.* Prevalence of antibody to human calicivirus in general population of northern Japan. *J Med Virol* 1981; **7**(3): 221–225.
18. Grohmann G, Glass RI, Gold J, *et al.* Outbreak of human calicivirus gastroenteritis in a day-care center in Sydney, Australia. *J Clin Microbiol* 1991; **29**(3): 544–550.
19. Noel JS, Liu BL, Humphrey CD, *et al.* Parkville virus: a novel genetic variant of human calicivirus in the Sapporo virus clade, associated with an outbreak of gastroenteritis in adults. *J Med Virol* 1997; **52**(2): 173–178.
20. Hedlund KO, Rubilar-Abreu E, Svensson L. Epidemiology of calicivirus infections in Sweden, 1994–1998. *J Infect Dis* 2000; **181**(Suppl. 2): S275–S280.
21. Schuffenecker I, Ando T, Thouvenot D, *et al.* Genetic classification of 'Sapporo-like viruses.' *Arch Virol* 2001; **146**(11): 2115–2132.
22. Nakata S, Honma S, Numata KK, *et al.* Members of the family caliciviridae (Norwalk virus and Sapporo virus) are the most prevalent cause of gastroenteritis outbreaks among infants in Japan. *J Infect Dis* 2000; **181**(6): 2029–2032.
23. Robinson S, Clarke IN, Vipond IB, *et al.* Epidemiology of human Sapporo-like caliciviruses in the South West of England: molecular characterisation of a genetically distinct isolate. *J Med Virol* 2002; **67**(2): 282–288.
24. Buesa J, Collado B, Lopez-Andujar P, *et al.* Molecular epidemiology of caliciviruses causing outbreaks and sporadic cases of acute gastroenteritis in Spain. *J Clin Microbiol 2002;* **40**(8): 2854–2859.
25. Pang XL, Honma S, Nakata S, Vesikari T. Human caliciviruses in acute gastroenteritis of young children in the community. *J Infect Dis* 2000; **181**(Suppl. 2): S288–S2894.
26. Matson DO, Estes MK, Tanaka T, *et al.* Asymptomatic human calicivirus infection in a day care center. *Pediatr Infect Dis J* 1990; **9**(3): 190–196.
27. Lopman BA, Reacher MH, Van Duijnhoven Y, *et al.* Viral gastroenteritis outbreaks in Europe, 1995–2000. *Emerg Infect Dis* 2003; **9**(1): 90–96.
28. Pang XL, Honma S, Nakata S, Vesikari T. Human caliciviruses in acute gastroenteritis of young children in the community. *J Infect Dis* 2000; **181**(Suppl. 2): S288–S294.
29. Kirkwood CD, Bishop RF. Molecular detection of human calicivirus in young children hospitalized with acute gastroenteritis in Melbourne, Australia, during 1999. *J Clin Microbiol* 2001; **39**(7): 2722–2724.
30. Wolfaardt M, Taylor MB, Booysen HF, *et al.* Incidence of human calicivirus and rotavirus infection in patients with gastroenteritis in South Africa. *J Med Virol* 1997; **51**(4): 290–296.
31. Sakai Y, Nakata S, Honma S, *et al.* Clinical severity of Norwalk virus and Sapporo virus gastroenteritis in children in Hokkaido, Japan. *Pediatr Infect Dis J* 2001; **20**(9): 849–853.
32. Hansman GS, Katayama K, Maneekarn N, *et al.* Genetic diversity of norovirus and sapovirus in hospitalized infants with sporadic cases of gastroenteritis in Chiang Mai, Thailand. *J Clin Microbiol* 2004; **42**(3): 1305–1307.
33. Hansman GS, Doan LT, Kguyen TA, *et al.* Detection of norovirus and sapovirus infection among children with gastroenteritis in Ho Chi Minh City, Vietnam. *Arch Virol* 2004; **149**(9): 1673–1688.
34. Matsui SM, Greenberg HB. Immunity to calicivirus infection. *J Infect Dis* 2000; **181**(Suppl. 2): S331–S335.
35. Hoebe CJ, Vennema H, Husman AM, van Duynhoven YT. Norovirus outbreak among primary schoolchildren who had played in a recreational water fountain. *J Infect Dis* 2004; **189**(4): 699–705.
36. Beuret C, Kohler D, Luthi T. Norwalk-like virus sequences detected by reverse transcription-polymerase chain reaction in mineral waters imported into or bottled in Switzerland. *J Food Prot 2000;* **63**(11): 1576–1582.
37. Brugha R, Vipond IB, Evans MR, *et al.* A community outbreak of food-borne small round-structured virus

DOI: 10.1002/rmv

gastroenteritis caused by a contaminated water supply. *Epidemiol Infect* 1999; **122**(1): 145–154.
38. Cannon RO, Poliner JR, Hirschhorn RB, *et al.* A multistate outbreak of Norwalk virus gastroenteritis associated with consumption of commercial ice. *J Infect Dis* 1991; **164**(5): 860–863.
39. Murphy AM, Grohmann GS, Christopher PJ, *et al.* An Australia-wide outbreak of gastroenteritis from oysters caused by Norwalk virus. *Med J Aust* 1979; **2**(7): 329–333.
40. Hansman GS, Sano D, Ueki Y, *et al.* Sapovirus in water, Japan. *Emerg Infect Dis* 2007; **13**(1): 133–135.
41. Matson DO, Zhong WM, Nakata S, *et al.* Molecular characterization of a human calicivirus with sequence relationships closer to animal caliciviruses than other known human caliciviruses. *J Med Virol* 1995; **45**(2): 215–222.
42. Jiang X, Zhong W, Kaplan M, *et al.* Expression and characterization of Sapporo-like human calicivirus capsid proteins in baculovirus. *J Virol Methods* 1999; **78**(1–2): 81–91.
43. Numata K, Hardy ME, Nakata S, *et al.* Molecular characterization of morphologically typical human calicivirus Sapporo. *Arch Virol* 1997; **142**(8): 1537–1552.
44. Oka T, Hansman GS, Katayama K, *et al.* Expression of sapovirus virus-like particles in mammalian cells. *Arch Virol* 2006; **151**(2): 399-404.
45. Hansman GS, Natori K, Oka T, *et al.* Cross-reactivity among sapovirus recombinant capsid proteins. *Arch Virol* 2005; **150**(1): 21–36.
46. Hansman GS, Matsubara N, Oka T, *et al.* Deletion analysis of the sapovirus VP1 gene for the assembly of virus-like particles. *Arch Virol* 2005; **150**(12): 2529–2538.
47. Barcena J, Verdaguer N, Roca R, Morales M, *et al.* The coat protein of Rabbit hemorrhagic disease virus contains a molecular switch at the N-terminal region facing the inner surface of the capsid. *Virology* 2004; **322**(1): 118–134.
48. Bertolotti-Ciarlet A, White LJ, Chen R, *et al.* Structural requirements for the assembly of Norwalk virus-like particles. *J Virol* 2002; **76**(8): 4044–4055.
49. Oka T, Katayama K, Ogawa S, *et al.* Proteolytic processing of sapovirus ORF1 polyprotein. *J Virol* 2005; **79**(12): 7283–7290.
50. Katayama K, Miyoshi T, Uchino K, *et al.* Novel recombinant sapovirus. *Emerg Infect Dis* 2004; **10**(10): 1874–1876.
51. Lole KS, Bollinger RC, Paranjape RS, *et al.* Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 1999; **73**(1): 152–160.
52. Bull RA, Hansman GS, Clancy LE, *et al.* Norovirus recombination in ORF1/ORF2 overlap. *Emerg Infect Dis* 2005; **11**(7): 1079–1085.
53. Hansman GS, Takeda N, Oka T, *et al.* Intergenogroup recombination in sapoviruses. *Emerg Infect Dis* 2005; **11**(12): 1916–1920.
54. Guntapong R, Hansman GS, Oka T, *et al.* Norovirus and sapovirus infections in Thailand. *Jpn J Infect Dis* 2004; **57**(6): 276–278.
55. Hansman GS, Kuramitsu M, Yoshida H, *et al.* Viral gastroenteritis in Mongolian infants. *Emerg Infect Dis* 2005; **11**(1): 180–182.
56. Lindesmith L, Moe C, Marionneau S, *et al.* Human susceptibility and resistance to Norwalk virus infection. *Nat Med* 2003; **9**(5): 548–553.
57. Hutson AM, Atmar RL, Graham DY, Estes MK. Norwalk virus infection and disease is associated with ABO histo-blood group type. *J Infect Dis* 2002; **185**(9): 1335–1337.
58. Huang P, Farkas T, Marionneau S, *et al.* Noroviruses bind to human ABO, Lewis, and secretor histo-blood group antigens: identification of 4 distinct strain-specific patterns. *J Infect Dis* 2003; **188**(1): 19–31.
59. Harrington PR, Lindesmith L, Yount B, *et al.* Binding of Norwalk virus-like particles to ABH histo-blood group antigens is blocked by antisera from infected human volunteers or experimentally vaccinated mice. *J Virol* 2002; **76**(23): 12335–12343.
60. Huang P, Farkas T, Zhong W, *et al.* Norovirus and histo-blood group antigens: demonstration of a wide spectrum of strain specificities and classification of two major binding groups among multiple binding patterns. *J Virol* 2005; **79**(11): 6714–6722.
61. Shirato-Horikoshi H, Ogawa S, Wakita T, *et al.* Binding activity of norovirus and sapovirus to histo-blood group antigens. *Arch Virol* 2006; online version.
62. Katayama K, Shirato-Horikoshi H, Kojima S, *et al.* Phylogenetic analysis of the complete genome of 18 Norwalk-like viruses. *Virology* 2002; **299**(2): 225–239.

JOURNAL OF VIROLOGY, July 2007, p. 6798–6806 Vol. 81, No. 13
0022-538X/07/$08.00+0 doi:10.1128/JVI.02840-06

Highly Conserved Configuration of Catalytic Amino Acid Residues among Calicivirus-Encoded Proteases▽

Tomoichiro Oka,[1]* Mami Yamamoto,[1] Masaru Yokoyama,[2] Satoko Ogawa,[1] Grant S. Hansman,[1] Kazuhiko Katayama,[1] Kana Miyashita,[1] Hirotaka Takagi,[3] Yukinobu Tohya,[4] Hironori Sato,[2] and Naokazu Takeda[1]

Department of Virology II, National Institute of Infectious Diseases, Gakuen 4-7-1, Musashi-murayama, Tokyo 208-0011, Japan[1]; Center for Pathogen Genomics, National Institute of Infectious Diseases, Gakuen 4-7-1, Musashi-murayama, Tokyo 208-0011, Japan[2]; Division of Biosafety Control and Research, National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku-ku, Tokyo 162-8640, Japan[3]; and Department of Veterinary Microbiology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan[4]

Received 22 December 2006/Accepted 11 April 2007

A common feature of caliciviruses is the proteolytic processing of the viral polyprotein catalyzed by the viral 3C-like protease encoded in open reading frame 1 (ORF1). Here we report the identification and structural characterization of the protease domains and amino acid residues in sapovirus (SaV) and feline calicivirus (FCV). The in vitro expression and processing of a panel of truncated ORF1 polyproteins and corresponding mutant forms showed that the functional protease domain is 146 amino acids (aa) in SaV and 154 aa in FCV. Site-directed mutagenesis of the protease domains identified four amino acid residues essential to protease activities: H^{31}, E^{52}, C^{116}, and H^{131} in SaV and H^{39}, E^{60}, C^{122}, and H^{137} in FCV. A computer-assisted structural analysis showed that despite high levels of diversity in the primary structures of the protease domains in the family *Caliciviridae*, the configurations of the H, E, C, and H residues are highly conserved, with these residues positioned closely along the inner surface of the potential binding cleft for the substrate. These results strongly suggest that the H, E, C, and H residues are involved in the formation of a conserved catalytic surface of the SaV and FCV 3C-like proteases.

The family *Caliciviridae* is composed of four genera, *Sapovirus*, *Lagovirus*, *Vesivirus*, and *Norovirus*, which include the species *Sapporo virus* (SaV), *Rabbit hemorrhagic disease virus* (RHDV), *Feline calicivirus* (FCV), and *Norwalk virus* (NoV), respectively (24). Caliciviruses infect a broad range of hosts, including humans and animals, and cause a variety of diseases and disorders, such as gastroenteritis, vesicular lesions, respiratory infections, reproductive failure, and hemorrhagic disease (10).

Calicivirus is a nonenveloped virus, and its genome is a linear, polyadenylated, positive-sense single-stranded RNA of about 7.3 to 8.3 kb with either two or three open reading frames (ORFs) (9). Calicivirus ORF1 encodes a polyprotein that contains amino acid motifs including 2C-like nucleoside triphosphatase (NTPase), VPg, 3C-like protease, and 3D-like RNA-dependent RNA polymerase (polymerase) (11, 25). In *Sapovirus* and *Lagovirus*, the structural protein VP1 is encoded in ORF1, whereas this protein is encoded in a separate ORF (ORF2) in *Vesivirus* and *Norovirus*. Cotranslational proteolytic processing of the ORF1 polyprotein is a common feature in the caliciviruses and is performed with the 3C-like protease encoded in ORF1 (11). The calicivirus 3C-like protease cleaves after the glutamic acid (E) or glutamine (Q) residue of the specific site in the polyprotein (2, 4, 14, 20, 21, 30, 32, 33, 35, 39). The protease itself is released by an autocatalytic cleavage in NoV (1, 2, 8, 15, 21, 32, 39) and RHDV (19, 25, 44), whereas cleavage between the protease and polymerase does not occur in FCV (7, 12, 40, 43) or SaV (7, 28–30).

The critical role of Cys in the calicivirus 3C-like protease motif GDCG in the cleavage activity in NoV and RHDV has been determined previously (4, 15, 20, 32, 34). In addition, the active-site residues of the 3C-like proteases of the NoV Chiba virus strain and the RHDV FRG strain have been identified by site-directed mutagenesis (5, 14, 36, 46). Recently, the X-ray crystal structures of the two 3C-like proteases of NoV (Chiba and Norwalk strains) were determined (27, 46). In contrast, although site-directed mutagenesis of the 3C-like proteases of the SaV Mc10 and FCV Urbana strains showed that C in the GDCG motif is crucial for the proteolytic processing activity (28–30), the remaining amino acids that are important for the activity have not been identified.

The aim of this study was to identify and structurally characterize the functional protease domains and the amino acid residues critical to the activities in SaV and FCV. For this purpose, an in vitro coupled transcription-translation analysis was performed with full-length or C-terminally truncated forms of the ORF1 polyprotein with or without amino acid mutations in the protease domain. In addition, three-dimensional (3-D) structural models of the 3C-like protease domains of the SaV Mc10, FCV F4, and RHDV FRG strains were constructed and compared with the X-ray crystal structure of the 3C-like protease of the NoV Chiba strain (27).

* Corresponding author. Mailing address: Department of Virology II, National Institute of Infectious Diseases, Gakuen 4-7-1, Musashi-murayama, Tokyo 208-0011, Japan. Phone: 81-42-561-0771. Fax: 81-42-561-4729. E-mail: oka-t@nih.go.jp.

▽ Published ahead of print on 25 April 2007.

MATERIALS AND METHODS

Virus strains. The SaV Mc10 strain was isolated from a stool specimen from an infant hospitalized with acute gastroenteritis in Chiang Mai, Thailand (13).

The FCV F4 strain was isolated from a cat with respiratory symptoms in Japan (22).

Preparation of FCV F4. Crandell-Rees feline kidney cells (JCRB9035; Health Science Research Resources Bank, Japan) were grown in a 150-cm^2 flask containing Eagle's minimum essential medium (Sigma-Aldrich, St. Louis, MO) with 5% calf serum (JRH Bioscience Corp., Tokyo, Japan) and virus production serum-free medium with L-glutamine (Invitrogen, Carlsbad, CA). Approximately 10^4 PFU of the virus was added to a monolayer of Crandell-Rees feline kidney cells containing 10 ml of culture medium. The cells were incubated for 3 days at 37°C and were harvested when the cytopathic effect had reached 90%. After three cycles of freezing and thawing, the cell debris were removed by centrifugation and the supernatant was stored at −30°C until use.

Full-length cDNA clones. A plasmid designated pUC19/SaV Mc10 full-length containing a full-length SaV Mc10 genome with the T7 promoter, as well as a plasmid designated as pUC19/SaV Mc10 full-C1171A/ORF1 encoding a ^{1169}GDCG^{1172}-to-GDAG mutation in the protease, were expressed as previously described (29).

The FCV F4 genomic RNA was purified from the culture medium by the QIAamp viral RNA mini kit (QIAGEN, Hilden, Germany). FCV F4 cDNA was synthesized as previously described (16). The 5′ fragment corresponding to nucleotides (nt) 1 to 3785 was amplified with the sense primer 5′-GTAAAAG AAATTTGAGACAATGTCTC-3′ and the antisense primer 5′-GTTTACAAA CTAATCCCTTGTAGC-3′. The middle fragment corresponding to nt 2990 to 6971 was amplified with the sense primer 5′-AATGCCAACAGAAAGCTTG A-3′ and the antisense primer 5′-AGCACGCTAATGCGCACTAC-3′. The 3′ fragment corresponding to nt 6952 to 7681 was amplified with the sense primer 5′-GTAGTGCGCATTAGCGTGC-3′ and the antisense primer 5′-CCCTGGG GTTAGACGCAAATGC-3′. These three fragments were cloned into the pCR-BluntII-Topo vector (Invitrogen), and the resulting constructs were designated FCV F4 5′/Topo, FCV F4 middle/Topo, and FCV F4 3′/Topo, respectively. Several amplification and cloning steps were performed to yield a full-length construct with a T7 RNA polymerase promoter at the 5′ end and both a hepatitis D virus (HDV) ribozyme and a T7 terminator at the 3′ end, as described previously (15). To add the HDV ribozyme and the T7 terminator at the 3′ end of the FCV F4 genome, the region from nt 6952 to 7681 was reamplified from the clone 3′/Topo with a sense primer (5′-GTAGTGCGCATTAGCGTGC-3′) and an antisense primer (5′-GAGGTGGAGATGCCATGCCGACCCT$_{30}$CCCTGG GGTTAGACGCAAATGC-3′). HDV ribozyme and T7 terminator sequences were amplified from the pT7HCV09Luc plasmid (45) with a sense primer (5′-GGGTCGGCATGGCATCTCCACCTC-3′) and an antisense primer (5′-GAA CTAGTGGATCCGAGCTCAGATCTCCTTTCAGCAAAAAACCCCTCAA G-3′) that included a BglII site (underlined) and a SacI site (double underlined). These DNA fragments were joined by a primerless PCR as previously described (15), and the amplified fragment corresponding to nt 6952 to 7681 was purified from the gel by using the QIAGel extraction kit (QIAGEN). This DNA fragment was designated FCV F4 6952-7681 polyA-Rz-T7 term. Following these experiments, the FCV F4 region from nt 2990 to 6971 was reamplified with the sense primer 5′-AATGCCAACAGAAAGCTTGA-3′ and the antisense primer 5′-A GCACGCTAATGCGCACTAC-3′. The amplified DNA fragment was joined with 6952-7681 polyA-Rz-T7 term in a primerless PCR. The joined amplified fragment was purified and cloned into the pCR-BluntII-Topo vector, and the construct was designated FCV F4 2990-7681 polyA-Rz-T7 term/Topo. This plasmid was digested with KpnI and SacI (New England Biolabs, Beverly, MA), and the insert was cloned into a pUC19 vector (Toyobo, Osaka, Japan) which was previously digested with KpnI and SacI. The resultant plasmid was designated FCV F4 middle + 3′ polyA-Rz-T7 term/pUC19. Finally, the FCV F4 5′ region corresponding to nt 1 to 3780 was reamplified from FCV F4 5′/Topo with a sense primer (5′-CAGGGGCCCCGTCGACCTGGTAATACGACTCACTATAGTAAA AGAAATTTGAGACAATGTC-3′) that included a SalI site (underlined) and a T7 RNA polymerase promoter sequence (bold) and an antisense primer (5′-T TGGGCCATGCAGGTGAGCG-3′). The amplified DNA was purified and digested with SalI and KpnI (New England Biolabs) and cloned into SalI- and KpnI-digested middle + 3′ polyA-Rz-T7 term/pUC19. The resultant plasmid was designated FCV F4 T7-GG full-length-Rz-T7 term/pUC19 ver2 (pUC19/FCV F4 full-length). Sequence analysis confirmed that the cloned genome corresponded to the consensus sequence of FCV F4 (GenBank accession number D31836), with the exception of a silent mutation (T to C) at the nucleotide position 4075.

Site-directed mutagenesis. Site-directed mutagenesis was performed using the GeneTailor site-directed mutagenesis system (Invitrogen), with pUC19/SaV Mc10 full-length (29) and pUC19/FCV F4 full-length as the templates. The site-directed mutagenesis primers are listed in Table 1. The resulting nine SaV Mc10 full-length mutant cDNA clones were designated as follows: pUC19/SaV Mc10 full-H1069A/ORF1 (where the H at amino acid [aa] residue 1069 in the ORF1 product is changed to A), pUC19/SaV Mc10 full-H1075A/ORF1, pUC19/SaV Mc10 full-H1086A/ORF1, pUC19/SaV Mc10 full-E1107A/ORF1, pUC19/SaV Mc10 full-H1120A/ORF1, pUC19/SaV Mc10 full-H1136A/ORF1, pUC19/SaV Mc10 full-E1143A/ORF1, pUC19/SaV Mc10 full-E1147A/ORF1, and pUC19/SaV Mc10 full-H1186A/ORF1. Similarly, 12 FCV F4 full-length mutant cDNA clones were constructed: pUC19/FCV F4 full-H1079A/ORF1, pUC19/FCV F4 full-H1093A/ORF1, pUC19/FCV F4 full-H1099A/ORF1, pUC19/FCV F4 full-H1102A/ORF1, pUC19/FCV F4 full-H1110A/ORF1, pUC19/FCV F4 full-E1121A/ORF1, pUC19/FCV F4 full-D1125A/ORF1, pUC19/FCV F4 full-E1131A/ORF1, pUC19/FCV F4 full-D1155A/ORF1, pUC19/FCV F4 full-E1164A/ORF1, pUC19/FCV F4 full-C1193A/ORF1, and pUC19/FCV F4 full-H1208A/ORF1. All of these full-length clones were verified by sequencing analysis to confirm that there were no additional mutations.

TABLE 1. Oligonucleotides used for the site-directed mutagenesis

Amino acid change[a]	Nucleotide sequence (5′–3′)[b]
SaV polyprotein	
H1069A	CACATCTGGGGGTGATgccATTGGGTATGGTTG
H1075A	CATTGGGTATGGTTGTgccATGGGTAATGGGGTG
H1086A	GTTGTCACAGTTACAgccGTGGCCTCTGCGTCTG
E1107A	CAGGAAGACCGAGGGTgcaACCACCTGGGTGAAC
H1120A	CTTGGTCACTTGCCCgccTACCAGATCGGTGATG
H1136A	CTACTACTCGGCGCGCCTAgccCCTGTCACCACGC
E1143A	GTCACCACGCTTGCGgcgGGGACGTATGAG
E1147A	GCGGAGGGGACGTATgcgACACCCAATATCACG
H1186A	CGTTTGGTCGGACTGgccGCAGCCACATCAAC
FCV polyprotein	
C1193A	ACTCACCCTGGAGATgccGGGTTGCCGTATATTGATGAT
H1079A	GGGCCTGGCACTAAATTTgccAAAAATGCCATTGGATCAG
H1093A	GATGTGTGTGGGGAGgccAAGGGTTACTGCGTCCAC
H1099A	CACAAGGGTTACTGCGTCgccATGGGTCATGGGGTCTATG
H1102A	GGTTACTGCGTCCACATGGGTgctGGGGTCTATGCCTCTG
H1110A	TATGCCTCTGTGGCAgccGTGGTGAAAGGTGATTCCTAT
E1121A	GATTCCTATTTCTTGGGTgccAGGATCTTTGATGTAAAAACC
D1125A	GGTGAAAGGATCTTTgccGTAAAAACCAATGGTGAATTCTG
E1131A	GTAAAAACCAATGGTgccTTCTGTTGTTTCCGAAGCACT
D1155A	GGGAAACCAACCCGCgccCCATGGGGGTCGCCAGTTG
E1164A	TGGGGGTCGCCAGTTGCAACAgcgTGGAAACCAAAGGCCTATAC
H1208A	AGGGTTACCGGCTTAgccACTGGATCTGGTGGACCCAAA

[a] Amino acids are shown in the one-letter code. Letters before the numbers indicate the original amino acid residues, and letters after the numbers indicate the mutant amino acid residues.

[b] Only the positive-sense oligonucleotide sequences are shown. The codons corresponding to the changed amino acids are indicated as lowercase letters.

In vitro coupled transcription-translation assay. For an in vitro coupled transcription-translation, linear DNA fragments containing the T7 promoter were generated by PCR with 100 ng of pUC19/SaV Mc10 full-length (29) and nine full-length mutant cDNA clones. DNA fragments corresponding to the entire SaV Mc10 ORF1-encoded region, aa 1 to 2278, were generated with the forward primer SaV Mc10-1F and the antisense primer SaV Mc10-2278R (Table 2). The C-terminally truncated SaV Mc10 ORF1-encoded mutant forms corresponding to the regions from aa 1 to 1246, 1 to 1240, 1 to 1234, 1 to 1229, 1 to 1223, 1 to 1218, 1 to 1212, 1 to 1206, 1 to 1205, 1 to 1204, 1 to 1203, 1 to 1202, 1 to 1201, 1 to 1200, and 1 to 1194 were generated with the forward primer SaV Mc10-1F and the antisense primers SaV Mc10-1246R, SaV Mc10-1240R, SaV Mc10-1234R, SaV Mc10-1229R, SaV Mc10-1223R, SaV Mc10-1218R, SaV Mc10-1212R, SaV Mc10-1206R, SaV Mc10-1205R, SaV Mc10-1204R, SaV Mc10-1203R, SaV Mc10-1202R, SaV Mc10-1201R, SaV Mc10-1200R, and SaV Mc10-1194R (Table 2). Similarly, linear DNA fragments corresponding to the entire and the C-terminally truncated FCV F4 ORF1-encoded regions and containing the T7 promoter were generated by PCR with 100 ng of the pUC19/FCV F4 full-length clone or with 12 full-length mutant cDNA clones. DNA fragments corresponding to the entire FCV ORF1-encoded region, aa 1 to 1763, were generated with the forward primer FCV F4-1F and the antisense primer FCV F4-1763R (Table 2). The truncated FCV F4 ORF1 mutant forms corresponding to the regions from aa 1 to 1419, 1 to 1345, 1 to 1267, 1 to 1245, 1 to 1240, 1 to 1235, 1 to 1230, 1 to 1225, 1 to 1224, and 1 to 1223 were generated with the forward primer FCV F4-1F and the antisense primers FCV F4-1419R, FCV F4-1345R, FCV F4-1267R, FCV F4-1245R, FCV F4-1240R, FCV F4-1235R, FCV F4-1230R, FCV F4-1225R, FCV F4-1224R, and FCV F4-1223R (Table 2).

In vitro T7 polymerase coupled transcription-translation was performed by using the TNT T7 Quick for PCR DNA kit (Promega, Madison, WI) as previ-

TABLE 2. PCR primers to prepare templates for in vitro transcription-translation

Primer name[a]	Sequence (5′–3′)
SaV primers	
SaV Mc10-1F	GCTTCCAAGCCATTCTACCCAATAGAG
SaV Mc10-2278R	TTCTAAGAACCTAACGGCCCGG
SaV Mc10-1246R	TTCTCCTTCCACAGGGTTGGGC
SaV Mc10-1240R	GGGCCATGCAGGTGAGCGGTG
SaV Mc10-1234R	GTGGTAACGAGTCCCCGTGG
SaV Mc10-1229R	CGTGGGCATGCCGCCACAGTC
SaV Mc10-1223R	GTCGGGGCCTCGAACCACTGGTAG
SaV Mc10-1218R	CACTGGTAGACCCTTCCAAGC
SaV Mc10-1212R	AGCAAAAGCATTCTCCAC
SaV Mc10-1206R	CTTGGATGTTTTAGTCAC
SaV Mc10-1205R	GGATGTTTTAGTCACTCGCTGAGCAAGCTTGG
SaV Mc10-1204R	TGTTTTAGTCACTCGCTGAGCAAGCTTGG
SaV Mc10-1203R	TTTAGTCACTCGCTGAGCAAGCTTGG
SaV Mc10-1202R	AGTCACTCGCTGAGCAAGCTTGG
SaV Mc10-1201R	CACTCGCTGAGCAAGCTTGG
SaV Mc10-1200R	TCGCTGAGCAAGCTTGG
SaV Mc10-1194R	CTCACTGAATGACAATTTGGTG
FCV primers	
FCV F4-1F	TCTCAAACTCTGAGCTTCGTG
FCV F4-1763R	TCAAACTTCGAACACATCACAGTG
FCV F4-1419R	CTCTTGCACCTTCTCAATGGG
FCV F4-1345R	TTCAGAGGAGATGTTCATGGG
FCV F4-1267R	AGTGCCTTTTGGTATGATCC
FCV F4-1245R	AGGTTTTGTTTCATCATACTT
FCV F4-1240R	ATACTTTTGAGGGGTTACAGATTTG
FCV F4-1235R	TACAGATTTGTTTTTCATATCAATG
FCV F4-1230R	CATATCAATGTGGATGTAGGGTAC
FCV F4-1225R	GTAGGGTACCACCAATTTTGCGCTTGG
FCV F4-1224R	GGGTACCACCAATTTTGCGCTTGG
FCV F4-1223R	TACCACCAATTTTGCGCTTGG

[a] Primer names ending in F are those of forward primers that include the sequence 5′-GGATCCTAATACGACTCACTATAGGGAACAGCCACCATG-3′ with the T7 promoter (underlined) and an additional start codon (bold) at the 5′ ends. Primer names ending in R are those of reverse primers (5′-T_{30}**TTA**-3′) that include poly(A) with a stop codon (bold) at the 5′ ends.

ously described, except that a 25-μl reaction volume and 3 h of incubation were used. The translation products were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and the radiolabeled proteins were detected as previously described (29, 30).

Nucleotide and amino acid sequence analyses. Nucleotide sequence analysis was performed with the BigDye Terminator (version 3.1) cycle sequencing ready reaction kit (Applied Biosystems, Tokyo, Japan) and an automated sequencer (genetic analyzer 3130; Applied Biosystems). Nucleotide sequences were assembled with the program Sequencher, version 4.2.2 (Gene Codes Corp., Ann Arbor, MI). Nucleotide and amino acid sequences were analyzed with GENETYX Mac software, version 12.2.6 (Genetyx Corp., Tokyo, Japan).

Molecular modeling of SaV, FCV, and RHDV 3C-like proteases. The crystal structure of the NoV 3C-like protease (Protein Data Bank [PDB] code, 1WQS) (27) at a resolution of 2.80 Å was used as the template for the molecular modeling of the SaV, FCV, and RHDV 3C-like proteases. To minimize misalignments of the target and template sequences, multiple-sequence alignments of 3C cysteine proteases, including rhinovirus 3C protease (PDB code, 1CQQ) (23), poliovirus 3C protease (PDB code, 1L1N) (26), and hepatitis A virus 3C protease (PDB code, 1QA7) (3), were used. The alignments were generated using MOE-Align in the Molecular Operating Environment (MOE) package (Chemical Computing Group, Inc., Montreal, Quebec, Canada). 3-D models of SaV Mc10, FCV F4, and RHDV FRG 3C-like proteases were constructed by the homology modeling technique using MOE-Homology in the MOE package as described previously (17, 18). The 3-D structures were thermodynamically optimized by energy minimization with the MOE package and an AMBER99 force field (31). A physically unacceptable local structure of the optimized 3-D model was further refined on the basis of Ramachandran plot evaluation using the MOE package. The quality of the models was assessed using the 3-D-structure evaluation program Verify3D (6, 47).

Strains for amino acid sequence alignments. The 16 SaV strains used for amino acid sequence alignments are as follows, with the nucleotide sequence accession numbers for the corresponding nucleotide regions given in parentheses: Mc114 (AY237422), Manchester (X86560), Dresden (AY694184), N21 (AY237423), Nongkhai50 (AY646853), Chantaburi74 (AY646854), Mc10 (AY237420), Bristol (AJ249939), C12 (AY603425), Mc2 (AY237419), SK15 (AY646855), PEC Cowden (AF182760), PEC LL14 (NC 000940), Ehime1107 (DQ058829), Sw278 (DQ125333), and NK24 (AY646856). The 13 FCV strains used for amino acid sequence alignments are as follows, with the nucleotide sequence accession numbers for the corresponding nucleotide regions given in parentheses: F4 (D31836), F65 (AF109465), 2024 (AF479590), Urbana (L40021), F9 (M86379), CFI/68 (U13992), DD/2006/GE (DQ424892), UTCV-NII1 (AY560113), UTCV-NII2 (AY560114), UTCV-NII3 (AY560115), UTCV-H1 (AY560116), UTCV-H2 (AY560117), and USDA (AY560118).

RESULTS AND DISCUSSION

Identification of the C terminus of the SaV Mc10 3C-like protease. The SaV protease-polymerase (Pro-Pol) is a stable product in both an in vitro translation system and an *Escherichia coli* expression system (28–30). Chang et al. also identified Pro-Pol as a stable product in porcine SaV-infected cells (7). Recently, we identified A^{1056} and E^{1722} as the N and C termini, respectively, of the SaV Mc10 Pro-Pol (Fig. 1A) (28, 30). To define the amino acid residues essential for protease activity, full-length forms and a series of 15 C-terminally truncated forms of the ORF1 polyprotein were expressed in an in vitro transcription-translation system (Fig. 1B; Table 1). The expression was carried out with both wild-type (Pro^{w}) and C1171A mutant (Pro^{mut}) forms of the protease, and the latter were used as a negative control for the proteolytic processing as previously described (29, 30).

At least seven proteins, p28, p32, p35 (NTPase), p46 (p32-VPg), p60 (VP1), p66 (p28-NTPase), and p120 (p32-VPg-Pro-Pol), were detected by direct SDS-PAGE analysis when the ORF1 Pro^{w} form corresponding to aa 1 to 2278 was expressed, and a major 250-kDa product (p250) corresponding to the ORF1 polyprotein appeared when the full-length ORF1 Pro^{mut} form was expressed (29). If the functional protease domain is eliminated, the cleavage patterns of the Pro^{w} forms should be significantly changed. We first analyzed two C-terminally truncated forms corresponding to aa 1 to 1246 and aa 1 to 1194. The proteolytic cleavage occurred in aa 1 to 1246 but not in aa 1 to 1194 (Fig. 1B). The translated product size of the Pro^{w} form of aa 1 to 1194 was identical to that of the Pro^{mut} form of aa 1 to 1194 (Fig. 1B), demonstrating that the C terminus of the functional protease domain is positioned downstream of aa 1194. Next, eight C-terminally truncated forms of the ORF1 polyprotein, aa 1 to 1240, 1 to 1234, 1 to 1229, 1 to 1223, 1 to 1218, 1 to 1212, 1 to 1206, and 1 to 1200, were expressed and analyzed. The proteolytic cleavage occurred efficiently in the Pro^{w} forms of the C-terminally truncated ORF1 polyproteins with aa 1 to 1240, 1 to 1234, 1 to 1229, 1 to 1223, 1 to 1218, 1 to 1212, and 1 to 1206 but not in the form with aa 1 to 1200 (Fig. 1B), demonstrating that the C terminus of the functional protease domain is positioned between aa 1200 and 1206. Finally, we expressed and analyzed five C-terminally truncated forms of the ORF1 polyprotein, aa 1 to 1205, 1 to 1204, 1 to 1203, 1 to 1202, and 1 to 1201. The cleavage occurred efficiently in the Pro^{w} forms of these C-terminally truncated ORF1 polyproteins (Fig. 1B). These results indicated that the C terminus of the functional protease

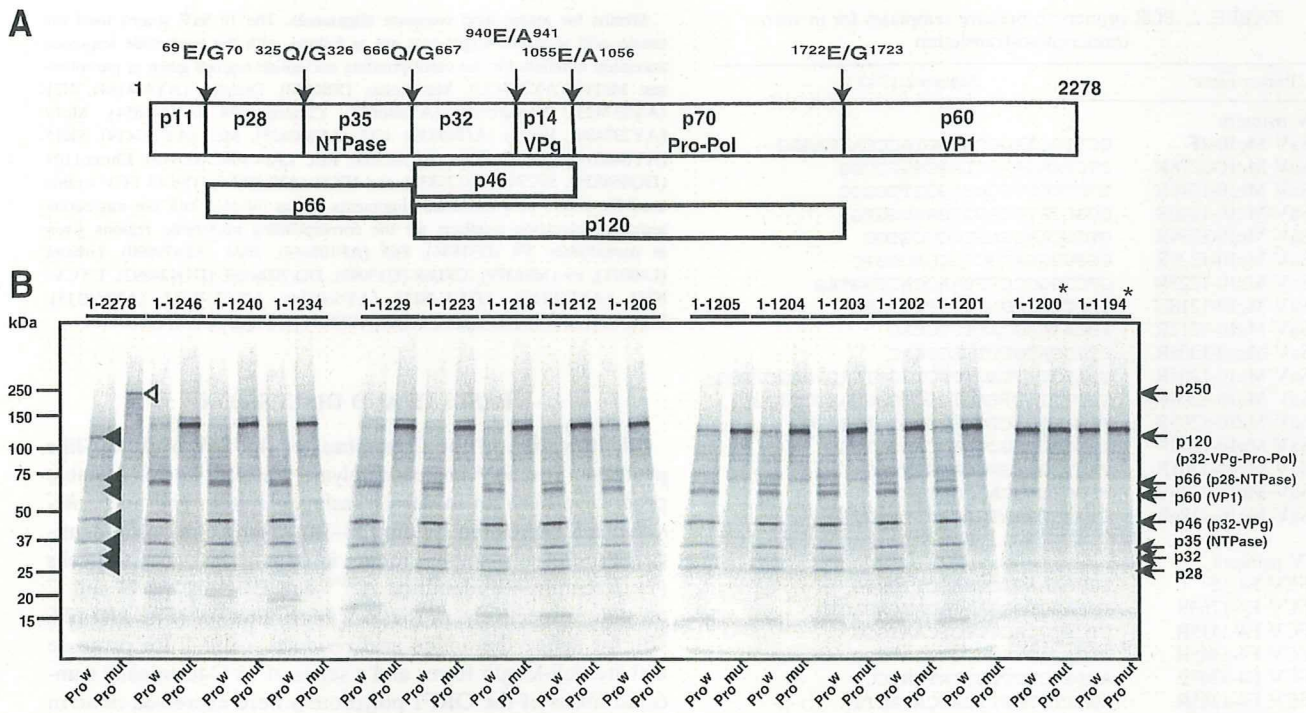

FIG. 1. Identification of the C terminus of the SaV Mc10 protease. (A) Proteolytic cleavage map of the SaV Mc10 ORF1 polyprotein and the processing intermediates (30). (B) SDS-PAGE of in vitro ^{35}S-labeled wild-type (Pro^{w}) and C1171A mutant (Pro^{mut}) forms of the entire ORF1 polyprotein (aa 1 to 2278) and 15 C-terminally truncated polyproteins corresponding to aa 1 to 1246, 1 to 1240, 1 to 1234, 1 to 1229, 1 to 1223, 1 to 1218, 1 to 1212, 1 to 1206, 1 to 1205, 1 to 1204, 1 to 1203, 1 to 1202, 1 to 1201, 1 to 1200, and 1 to 1194. The protein bands corresponding to either the Pro^{w} or Pro^{mut} form of the entire ORF1 polyprotein are indicated by filled and open arrowheads, respectively. The molecular sizes of viral proteins are shown on the right, and size markers are shown on the left. Asterisks indicate two C-terminally truncated forms, corresponding to aa 1 to 1200 and 1 to 1194, which show affected protease activity. Products of approximately 15 to 20 kDa would be the truncated Pro-Pol, released from the truncated ORF1 polyprotein. A 60-kDa product of the C-terminally truncated form corresponding to aa 1 to 1200 was identified as p32-VPg-truncated Pro-Pol.

of SaV Mc10 is V^{1201}, and it seems that the polymerase domain is not important for the protease activity. This amino acid residue was conserved in the 14 human and 2 porcine SaV strains listed in Materials and Methods (data not shown). The Mc10 3C-like protease domain comprises 146 aa (A^{1056} to V^{1201}), similar to the proteases of RHDV (143 aa) and NoV (181 aa) (5, 36).

Identification of the active-site residues of the SaV Mc10 3C-like protease. SaV Mc10 3C-like protease cleaves after the E or Q residue of the specific site in the ORF1 polyprotein (Fig. 1), and the C^{1171} in the GDCG motif was shown to be critical to the protease activity (28–30). Furthermore, the functional SaV Mc10 protease was found to be similar in size to those of RHDV and NoV. These results suggested that the catalytic amino acid residues of the SaV Mc10 protease would also be similar to those of other caliciviruses. In RHDV and NoV, the amino acid residues critical to protease activity have been identified as H^{27}, D^{44}, C^{104}, and H^{119} and H^{30}, E^{54}, C^{139}, and H^{157}, respectively, although it remains controversial whether or not E^{54} has a critical role in NoV protease activity (5, 14, 36, 46). The amino acid essential for the protease activity will definitely be conserved among all SaV strains (at least among human SaV strains) as it is in NoV and RHDV (36) (data not shown). Nine amino acid residues (H^{1069}, H^{1075}, H^{1086}, E^{1107}, H^{1120}, H^{1136}, E^{1143}, E^{1147}, and H^{1186}) within the

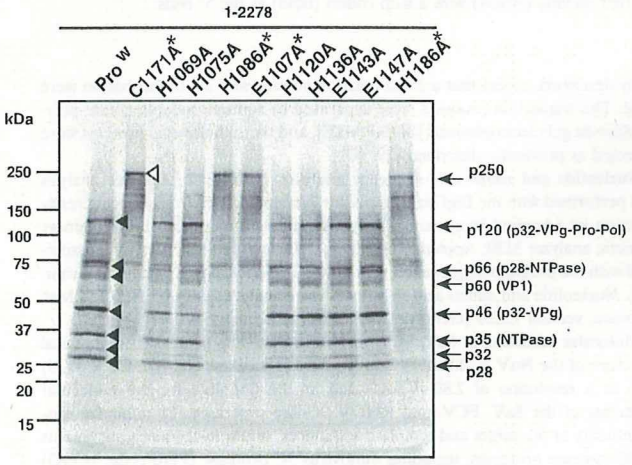

FIG. 2. Identification of the amino acid residues critical to SaV Mc10 protease activity. Shown are the results of SDS-PAGE of in vitro ^{35}S-labeled wild-type (Pro^{w}) and C1171A mutant forms of the entire ORF1 polyprotein (aa 1 to 2278) and nine other mutant forms, the H1069A, H1075A, H1086A, E1107A, H1120A, H1136A, E1143A, E1147A, and H1186A polyproteins. The protein bands corresponding to either the Pro^{w} or C1171A form of the entire ORF1 polyprotein are indicated by filled and open arrowheads, respectively. The molecular sizes of viral proteins are shown on the right, and size markers are shown on the left. Asterisks indicate four mutant forms, the C1171A, H1086A, E1107A, and H1186A polyproteins, which display affected protease activity.

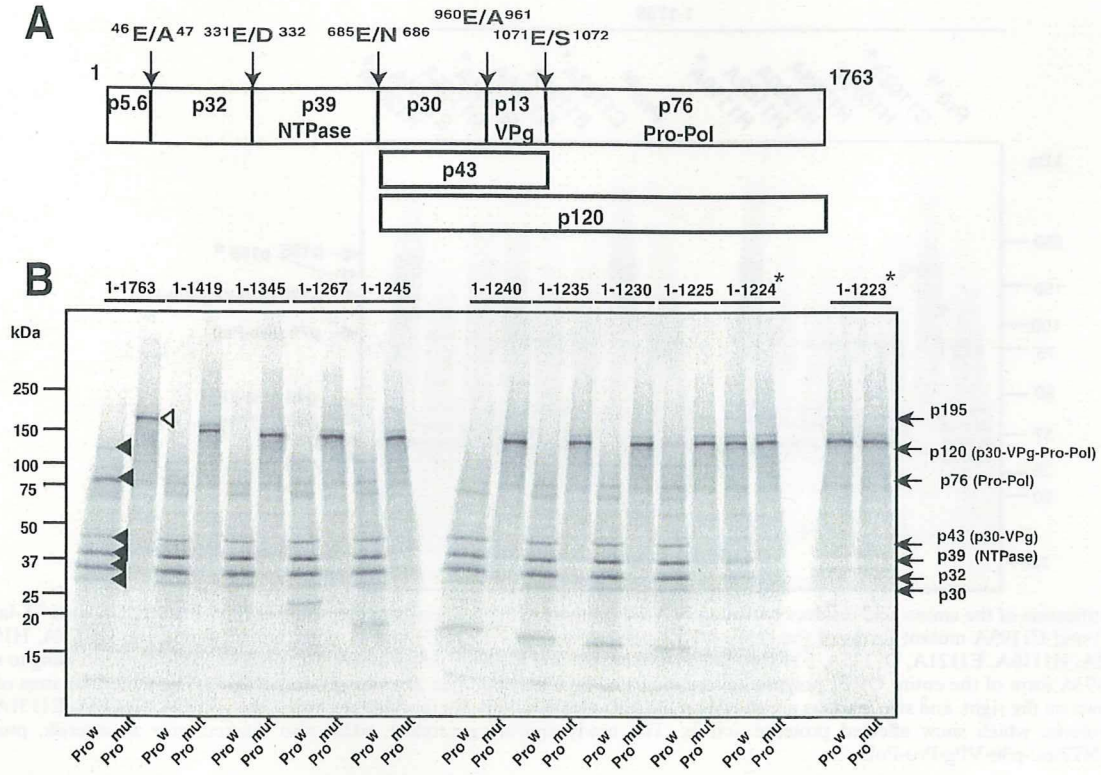

FIG. 3. Identification of the C terminus of the FCV F4 protease. (A) Proteolytic cleavage map of the FCV F4 ORF1 polyprotein and processing intermediates. The locations and designations of the proteins are adopted from the studies by Sosnovtsev et al. (40). (B) SDS-PAGE of in vitro ^{35}S-labeled wild-type (Pro^{w}) and C1193A mutant (Pro^{mut}) forms of the entire ORF1 polyprotein (aa 1 to 1763) and 10 C-terminally truncated polyproteins corresponding to aa 1 to 1419, 1 to 1345, 1 to 1267, 1 to 1245, 1 to 1240, 1 to 1235, 1 to 1230, 1 to 1225, 1 to 1224, and 1 to 1223. The protein bands corresponding to either the Pro^{w} or Pro^{mut} form of the entire ORF1 polyprotein are indicated by filled and open arrowheads, respectively. The molecular sizes of viral proteins are shown on the right, and size markers are shown on the left. Asterisks indicate two C-terminally truncated forms, corresponding to aa 1 to 1224 and 1 to 1223, which display affected protease activity. Products of approximately 15 to 25 kDa would be the truncated Pro-Pol released from the truncated ORF1 polyprotein.

protease domain were selected based on the amino acid alignment of 14 human SaV strains, with one exception: D instead of E^{1147} in a human SaV NK24 strain (data not shown). These nine amino acid residues were changed to A by site-directed mutagenesis (Table 1), and nine mutant forms of the ORF1 polyprotein, the H1069A, H1075A, H1086A, E1107A, H1120A, H1136A, E1143A, E1147A, and H1186A forms, were expressed in an in vitro translation system. Two forms of the ORF1 polyprotein, the full-length Pro^{w} and Pro^{mut} (C1171A) forms, were used as positive and negative controls, respectively (Fig. 2) (29, 30). Three mutant forms of the ORF1 polyprotein, the H1086A, E1107A, and H1186A forms, each produced a major 250-kDa product (Fig. 2), demonstrating that the proteolytic processing of the ORF1 polyprotein was completely blocked with these mutant forms. In contrast, the cleavage products of six mutant forms, the H1069A, H1075A, H1120A, H1136A, E1143A, and E1147A forms, were identical to those of Pro^{w}, although their cleavages were slightly affected (Fig. 2). From these results, we concluded that the amino acid residues critical to SaV Mc10 3C-like protease activity are H$^{1086(31)}$, E$^{1107(52)}$, C$^{1171(116)}$, and H$^{1186(131)}$. The first three of these would form the catalytic triad (general base, anion, and nucleophile, respectively), and the last one would correspond to part of the binding pocket as previously described for other calicivirus proteases (5, 27, 36, 46). Two amino acids, E^{1107} and H^{1186}, essential to the SaV Mc10 protease activity are not conserved in porcine SaV; (i) E^{1107} in Mc10 is D in the PEC LL14 and Cowden strains, and (ii) H^{1186} in Mc10 is Y in the PEC LL14 strain, whereas H^{1186} is conserved in the PEC Cowden strain (data not shown).

Identification of the C terminus of the FCV F4 3C-like protease. The FCV protease-polymerase (Pro-Pol) has been identified as a stable product in both infected cells and an in vitro translation system (12, 40). Sosnovtseva et al. reported that the entire Pro-Pol region is not essential for the autocatalytic polyprotein processing in the Urbana strain in vitro translation system (42). However, the exact C terminus of the functional protease domain has not been determined.

We generated two full-length cDNA clones, pUC19/FCV F4 full-length and pUC19/FCV F4 full-C1193A/ORF1, the latter of which encodes a ^{1191}GDCG^{1194}-to-GDAG mutation in the protease. The expression was carried out with both Pro^{w} and Pro^{mut} (C1193A) forms of the protease, and the latter were used as a negative control for the proteolytic processing. Six major cleavage products, p30, p32, p39 (NTPase), p43 (p30-VPg), p76 (Pro-Pol), and p120 (p30-VPg-Pro-Pol), were detected when the Pro^{w} form corresponding to aa 1 to 1763 was expressed in an in vitro translation system (Fig. 3). These

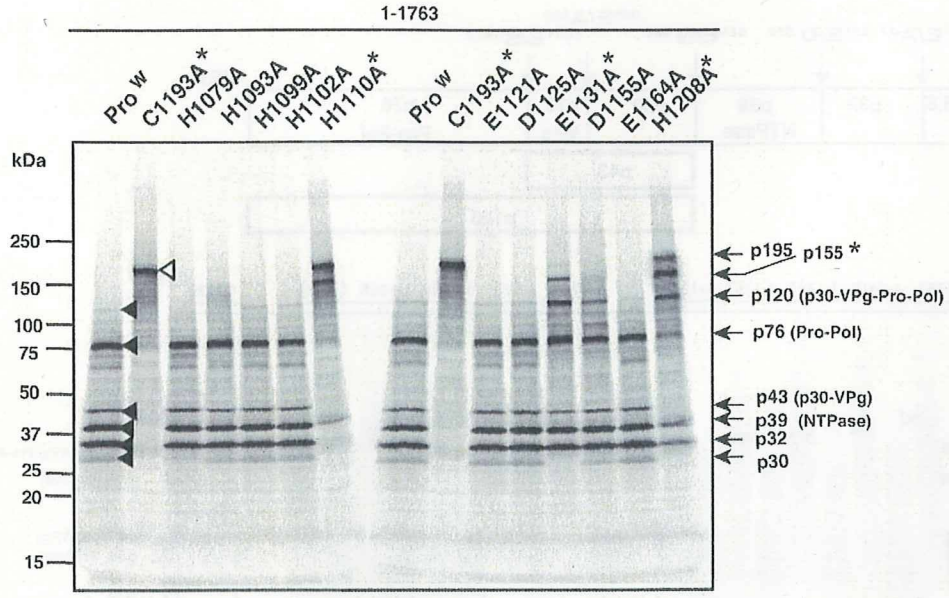

FIG. 4. Identification of the amino acid residues critical to FCV F4 protease activity. Shown are the results of SDS-PAGE of in vitro ^{35}S-labeled wild-type (Pro^{w}) and C1193A mutant forms of the entire ORF1 polyprotein (aa 1 to 1763) and 11 other mutant forms, the H1079A, H1093A, H1099A, H1102A, H1110A, E1121A, D1125A, E1131A, D1155A, E1164A, and H1208A polyproteins. The protein bands corresponding to either the Pro^{w} or C1193A form of the entire ORF1 polyprotein are indicated by filled and open arrowheads, respectively. The molecular sizes of viral proteins are shown on the right, and size markers are shown on the left. Asterisks indicate four mutant forms, the C1193A, H1110A, E1131A, and H1208A polyproteins, which show affected protease activity. The newly appearing product, p155, also marked with an asterisk, probably corresponds to NTPase-p30-VPg-Pro-Pol.

cleavage products were identical to those of the Urbana strain (40). A major 195-kDa product (p195) corresponding to the ORF1 polyprotein appeared when the Pro^{mut} form was expressed (Fig. 3B), demonstrating that the C^{1193} is critical to the FCV F4 3C-like protease activity, consistent with a previous report (42).

To define the C terminus of the functional protease domain of FCV F4, a series of 10 C-terminally truncated ORF1 polyproteins were expressed in an in vitro translation system (Fig. 3B; Table 1). Both Pro^{w} and Pro^{mut} forms of each of these regions were expressed, and the latter were used as a negative control for the proteolytic processing. Eight C-terminally truncated templates corresponding to aa 1 to 1419, 1 to 1345, 1 to 1267, 1 to 1245, 1 to 1240, 1 to 1235, 1 to 1230, and 1 to 1223 were first expressed and analyzed. The proteolytic cleavage occurred in the Pro^{w} forms of the C-terminally truncated ORF1 polyproteins corresponding to aa 1 to 1419, 1 to 1345, 1 to 1267, 1 to 1245, 1 to 1240, 1 to 1235, and 1 to 1230 but not in that corresponding to aa 1 to 1223. The translated product size of the Pro^{w} form of aa 1 to 1223 was identical to that of the Pro^{mut} form of aa 1 to 1223 (Fig. 3B), demonstrating that the C terminus of the functional protease domain is positioned upstream of aa 1223. Next, we expressed and analyzed two additional C-terminally truncated forms, aa 1 to 1225 and 1 to 1224. The proteolytic cleavage occurred efficiently when the Pro^{w} form of aa 1 to 1225 was expressed, whereas it occurred partially when the Pro^{w} form of aa 1 to 1224 was expressed (Fig. 3B). The C terminus of the functional protease domain for FCV F4 was determined to be Y^{1225}. This amino acid is conserved in 13 FCV strains (data not shown). Although we did not identify the cleavage sites of the FCV F4 ORF1 polyprotein, the cleavage pattern and the sizes of the products were consistent with the results of Sosnovtsev et al. (40). Therefore, the size of the functional protease of FCV F4 would be 154 aa (S^{1072} to Y^{1225}) when the cleavage site of the N terminus of the Urbana strain Pro-Pol is considered (40). The FCV F4 functional protease domain is similar in size to those of other caliciviruses, including SaV.

Identification of the active sites of the FCV F4 3C-like protease. The FCV 3C-like protease cleaves after the E residues of the specific site in the ORF1 polyprotein (Fig. 3A), and the C in the GDCG motif is critical to the protease activity (40–42). The functional FCV F4 protease is similar in size to those of SaV, RHDV, and NoV. Thus, the catalytic amino acid residues of the FCV protease would also be similar to those of SaV, RHDV, and NoV. Therefore, 11 amino acid residues (H^{1079}, H^{1093}, H^{1099}, H^{1102}, H^{1110}, E^{1121}, D^{1125}, E^{1131}, D^{1155}, E^{1164}, and H^{1208}) within the protease domain were selected and were changed to A by site-directed mutagenesis (Table 1).

Eleven mutant forms of the ORF1 polyprotein, the H1079A, H1093A, H1099A, H1102A, H1110A, E1121A, D1125A, E1131A, D1155A, E1164A, and H1208A forms, were expressed in an in vitro translation system. Two forms of ORF1 polyprotein, the full-length Pro^{w} and Pro^{mut} (C1193A) forms, were used as positive and negative controls for the proteolytic processing (Fig. 4). H1110A did significantly affect the ORF1 polyprotein processing, whereas E1131A and H1208A affected the processing partially (Fig. 4). That is, (i) p195, the entire ORF1 polyprotein, was detected when the H1110A form was expressed, (ii) p30 disappeared and p155 (likely the stable intermediate of NTPase-p30-VPg-Pro-Pol) appeared when the E1131A form was expressed, and (iii) p195 and p155 (NTPase-

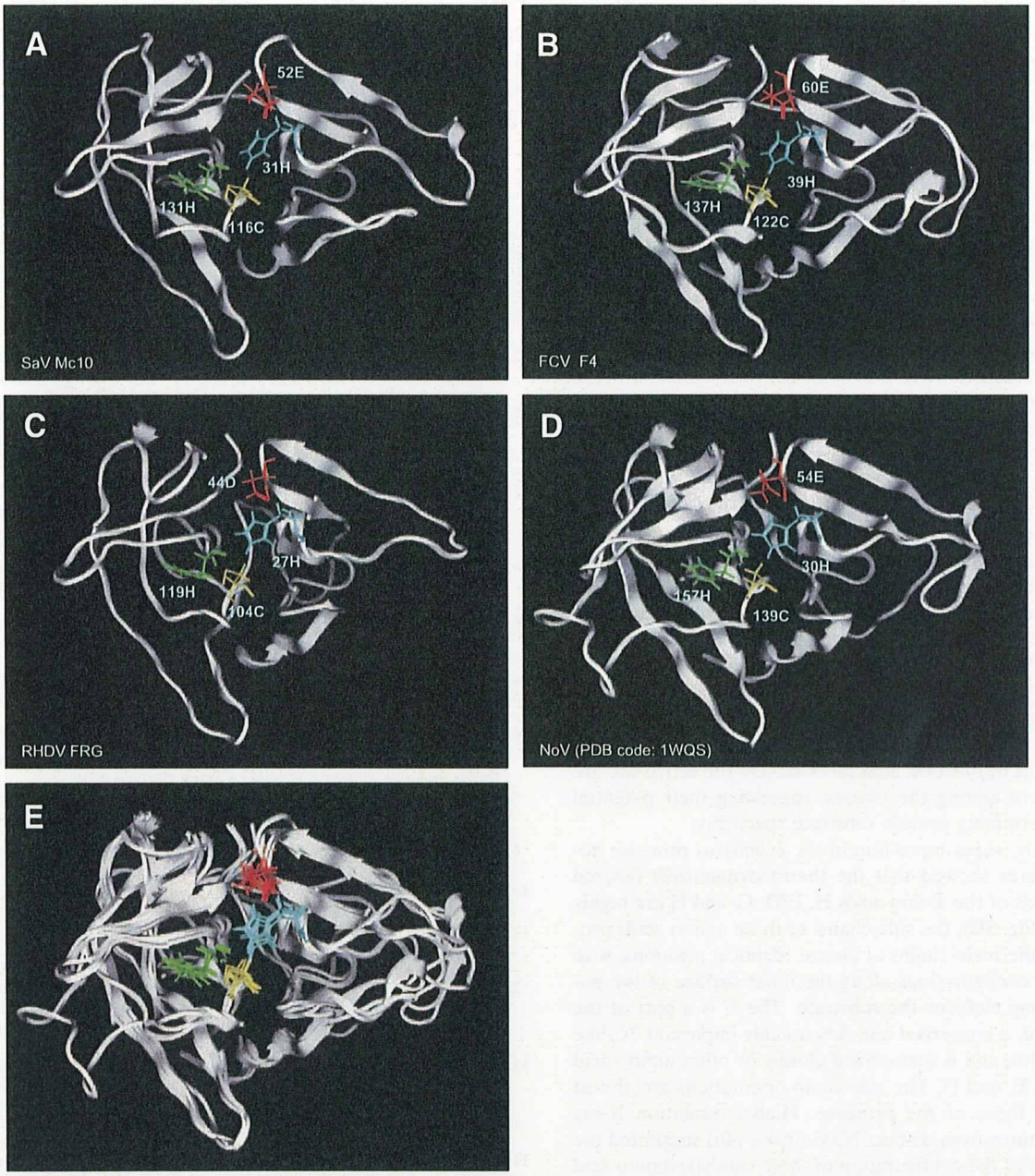

FIG. 5. 3-D models of the calicivirus 3C-like proteases. (A to D) Structural models of calicivirus 3C-like protease domains of the SaV Mc10 (A), FCV F4 (B), and RHDV FRG (C) strains and crystal structure of the NoV Chiba strain protease (27) (D). (E) Superimposition of 3C-like protease structures of the SaV Mc10, FCV F4, RHDV FRG, and NoV Chiba strains. The models were constructed with a homology modeling technique by using programs in the MOE package. Ribbons represent the backbone of the 3C-like protease domain. Side chains of the catalytically important amino acids identified in this study are shown as cyan sticks (H), red sticks (D and E), yellow sticks (C), and green sticks (H).

p30-VPg-Pro-Pol) appeared whereas p43 (p30-VPg) and p30 disappeared when the H1208A form was expressed. Prolonged incubation for up to 16 h did not change the processing patterns of these templates (data not shown). In contrast, the proteolytic processing of the ORF1 polyprotein was barely affected when five mutant forms, the H1099A, H1102A, E1121A, D1125A, and D1155A forms, were expressed (Fig. 4). The p120 (p30-VPg-Pro-Pol) band disappeared when two mutant forms, the H1079A and H1093A forms, were expressed. However, the mutated amino acid residues in these forms are not critical to protease activity, because other cleavage products were produced normally. Therefore, this product would be a precursor intermediate. Combining these results, we concluded that the amino acid residues important to FCV F4 3C-like protease activity are $H^{1110(39)}$, $E^{1131(60)}$, $C^{1193(122)}$, and $H^{1208(137)}$. The former three amino acid residues would form

the catalytic triad, and the last amino acid residue would correspond to a part of the binding pocket as discussed for other calicivirus proteases. A reverse genetics system has been reported for FCV (37, 38, 40), and it would be interesting to evaluate whether the E1131A and H1208A mutant forms are also important in vivo.

Structural modeling of SaV and FCV 3C-like proteases. Processing activities and specificities of the SaV and FCV 3C-like protease were similar when 3D-like RNA-dependent RNA polymerase domains were sequentially deleted (Fig. 1 and 3), strongly suggesting that the structure of the protease active site is self-determined and preserved in the part of the Pro-Pol polypeptide independent of the Pol domain. To obtain structural insights into the roles of the amino acid residues critical to protease activity, 3-D models of the 3C-like protease domains of the SaV Mc10, FCV F4, and RHDV FRG strains were constructed and compared with the X-ray crystal structure of the 3C-like protease of the NoV Chiba strain (27). Despite the low levels of amino acid sequence similarity of proteases among these strains (about 20%), the overall structures were predicted to be similar and to retain structural characteristics seen in the functional protease in general (Fig. 5A to D). The calicivirus proteases consist of N- and C-terminal subdomains, which are separated by a large cleft, probably for substrate binding. H and E in SaV, FCV, and NoV and H and D in RHDV are located along the inner surface of the N-terminal subdomain, whereas C and H in all viruses are located along the inner surface of the C-terminal subdomain. In contrast to the overall similarity, the conformations and configurations of the local structures around the active site are often different among the viruses, suggesting their potential roles in determining protein substrate specificity.

Notably, the superimposition of the calicivirus protease domain structures showed that the thermodynamically favored configurations of the amino acids H, E/D, C, and H are highly conserved (Fig. 5E); the side chains of these amino acids protrude from the main chains at almost identical positions, with very similar configurations along the inner surface of the potential binding cleft for the substrate. The C is a part of the GDCG motif, a conserved and functionally important 3C-like protease motif, and is surrounded closely by other amino acid residues, H, E, and H. The side chain orientations are almost identical to those of the protease. Higher-resolution X-ray crystal structures from distinct NoV strains (46) supported the conservation of the configuration of these catalytic amino acid residues among calicivirus proteases. These findings are consistent with the critical roles of these amino acids in the proteolytic activity and suggest strongly that the H, E, C, and H residues are involved in the formation of a conserved catalytic surface of the SaV and FCV 3C-like proteases.

Conclusions. The functional domains, amino acid residues critical to the proteolytic processing activity, and structural characteristics of SaV and FCV 3C-like proteases were identified. The molecular genetics study showed that the sizes and catalytically important amino acids of the SaV and FCV proteases are similar to those of RHDV and NoV proteases. The computer-assisted structural study strongly suggested that these amino acids are involved in the formation of a conserved catalytic surface of the calicivirus 3C-like protease. In this study, we studied both primary and 3-D structures of SaV and FCV proteases, which are mutually closely related subjects and should be understood in concert. In collaboration with other resources, the data obtained in this study will provide important bases to study the molecular function and inhibitors of calicivirus proteases.

ACKNOWLEDGMENTS

We thank Y. Someya for his critical review of the manuscript.

This work was supported in part by grants for research on emerging and reemerging infectious diseases, as well as food safety, from the Ministry of Health, Labor, and Welfare of Japan and by a grant from the Japan Health Science Foundation for Research on Health Sciences Focusing on Drug Innovation.

REFERENCES

1. **Asanaka, M., R. L. Atmar, V. Ruvolo, S. E. Crawford, F. H. Neill, and M. K. Estes.** 2005. Replication and packaging of Norwalk virus RNA in cultured mammalian cells. Proc. Natl. Acad. Sci. USA **102**:10327–10332.
2. **Belliot, G., S. V. Sosnovtsev, T. Mitra, C. Hammer, M. Garfield, and K. Y. Green.** 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.
3. **Bergmann, E. M., M. M. Cherney, J. McKendrick, S. Frormann, C. Luo, B. A. Malcolm, J. C. Vederas, and M. N. James.** 1999. Crystal structure of an inhibitor complex of the 3C proteinase from hepatitis A virus (HAV) and implications for the polyprotein processing in HAV. Virology **265**:153–163.
4. **Blakeney, S. J., A. Cahill, and P. A. Reilly.** 2003. Processing of Norwalk virus nonstructural proteins by a 3C-like cysteine proteinase. Virology **308**:216–224.
5. **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.
6. **Bowie, J. U., R. Luthy, and D. Eisenberg.** 1991. A method to identify protein sequences that fold into a known three-dimensional structure. Science **253**:164–170.
7. **Chang, K. O., S. S. Sosnovtsev, G. Belliot, Q. Wang, L. J. Saif, and K. Y. Green.** 2005. Reverse genetics system for porcine enteric calicivirus, a prototype sapovirus in the *Caliciviridae*. J. Virol. **79**:1409–1416.
8. **Chang, K. O., S. V. Sosnovtsev, G. Belliot, A. D. King, and K. Y. Green.** 2006. Stable expression of a Norwalk virus RNA replicon in a human hepatoma cell line. Virology **353**:463–473.
9. **Clarke, I. N., and P. R. Lambden.** 2000. Organization and expression of calicivirus genes. J. Infect. Dis. **181**(Suppl. 2):S309–S316.
10. **Clarke, I. N., and P. R. Lambden.** 1997. The molecular biology of caliciviruses. J. Gen. Virol. **78**:291–301.
11. **Green, K. Y., T. Ando, M. S. Balayan, T. Berke, I. N. Clarke, M. K. Estes, D. O. Matson, S. Nakata, J. D. Neill, M. J. Studdert, and H. J. Thiel.** 2000. Taxonomy of the caliciviruses. J. Infect. Dis. **181**(Suppl. 2):S322–S330.
12. **Green, K. Y., A. Mory, M. H. Fogg, A. Weisberg, G. Belliot, M. Wagner, T. Mitra, E. Ehrenfeld, C. E. Cameron, and S. V. Sosnovtsev.** 2002. Isolation of enzymatically active replication complexes from feline calicivirus-infected cells. J. Virol. **76**:8582–8595.
13. **Hansman, G. S., K. Katayama, N. Maneekarn, S. Peerakome, P. Khamrin, S. Tonusin, S. Okitsu, O. Nishio, N. Takeda, and H. Ushijima.** 2004. Genetic diversity of norovirus and sapovirus in hospitalized infants with sporadic cases of acute gastroenteritis in Chiang Mai, Thailand. J. Clin. Microbiol. **42**:1305–1307.
14. **Hardy, M. E., T. J. Crone, J. E. Brower, and K. Ettayebi.** 2002. Substrate specificity of the Norwalk virus 3C-like proteinase. Virus Res. **89**:29–39.
15. **Katayama, K., G. S. Hansman, T. Oka, S. Ogawa, and N. Takeda.** 2006. Investigation of norovirus replication in a human cell line. Arch. Virol. **151**:1291–1308.
16. **Katayama, K., H. Shirato-Horikoshi, S. Kojima, T. Kageyama, T. Oka, F. Hoshino, S. Fukushi, M. Shinohara, K. Uchida, Y. Suzuki, T. Gojobori, and N. Takeda.** 2002. Phylogenetic analysis of the complete genome of 18 Norwalk-like viruses. Virology **299**:225–239.
17. **Kinomoto, M., R. Appiah-Opong, J. A. Brandful, M. Yokoyama, N. Nii-Trebi, E. Ugly-Kwame, H. Sato, D. Ofori-Adjei, T. Kurata, F. Barre-Sinoussi, T. Sata, and K. Tokunaga.** 2005. HIV-1 proteases from drug-naive West African patients are differentially less susceptible to protease inhibitors. Clin. Infect. Dis. **41**:243–251.
18. **Kinomoto, M., M. Yokoyama, H. Sato, A. Kojima, T. Kurata, K. Ikuta, T. Sata, and K. Tokunaga.** 2005. Amino acid 36 in the human immunodeficiency virus type 1 gp41 ectodomain controls fusogenic activity: implications for the molecular mechanism of viral escape from a fusion inhibitor. J. Virol. **79**:5996–6004.
19. **Konig, M., H. J. Thiel, and G. Meyers.** 1998. Detection of viral proteins after infection of cultured hepatocytes with rabbit hemorrhagic disease virus. J. Virol. **72**:4492–4497.