Arch Virol (2006) 151: 1319–1327
DOI 10.1007/s00705-005-0717-4

Archives of Virology
Printed in Austria

Genetic diversity of noroviruses in Taiwan between November 2004 and March 2005

F.-T. Wu[1], T. Oka[2], K. Katayama[2], H.-S. Wu[1], D.-S. Donald Jiang[1], T. Miyamura[2], N. Takeda[2], and G. S. Hansman[2]

[1]Center for Disease Control, Taipei, Taiwan
[2]Department of Virology II, National Institute of Infectious Diseases, Tokyo, Japan

Received November 4, 2005; accepted December 20, 2005
Published online February 7, 2006

Summary. Noroviruses are a major health burden and are responsible for the majority of outbreaks of gastroenteritis in the world. Human noroviruses can be genetically divided into two main genogroups (GI and GII) and subdivided into many genotypes. In this study, stool specimens collected from 12 outbreaks of gastroenteritis in Taiwan were screened for viral agents between the 23rd of November 2004 and 9th of March 2005. Noroviruses were detected in all outbreaks. We detected six different norovirus genotypes: GI/11, GI/14, GII/3, GII/4, GII/6, and GII/18. Noroviruses belonging to GII/4 were dominant, 50 of 60 (83%) sequences, and were detected in 10 of 12 outbreaks. Furthermore, the norovirus GII/4 strains were detected throughout Taiwan, demonstrating their widespread distribution. We also found that three outbreaks had noroviruses from multiple genotypes. Our results have shown for the first time that noroviruses are an important cause of gastroenteritis in Taiwan.

Introduction

Viral agents of gastroenteritis affect millions of people worldwide of all ages [25]. The major agents include rotavirus, norovirus, sapovirus, astrovirus, and enteric adenovirus. Rotavirus are an important cause of acute sporadic childhood gastroenteritis [25]. Norovirus strains are a leading cause of outbreaks of gastroenteritis in the world and cause outbreaks in various epidemiological settings including hospitals, cruise ships, schools, and restaurants [3, 8, 10, 12, 17, 18, 22]. Sapovirus, astrovirus, and adenovirus are known to mostly infect infants, occasionally causing outbreaks. Numerous molecular epidemiological studies have revealed a global distribution of these viruses [19, 20, 24]. Transmissions of these viruses can include food-borne transmission, person-to-person contact, and environmental contaminations, including drinking water. Other

agents less commonly screened for include kobuvirus and enteroviruses [1, 26].

Human noroviruses can be divided into two genetically distinct genogroups (GI and GII). Recently, norovirus strains were subdivided into 14 GI and 17 GII genotypes [11]. The norovirus genome contains three open reading frames (ORFs). The first ORF (ORF1) encodes non-structural proteins, including the RNA-dependent RNA polymerase (RdRp), ORF2 encodes VP1, and ORF3 encodes a small capsid protein (VP2) [2]. Norovirus genotypes are generally maintained across the ORFs. However, a number of norovirus strains failed to maintain their sequence identities for RdRp and VP1, and they were shown to be recombinant [13, 15, 23]. Evidence suggested that the recombination site occurred at the conserved polymerase and capsid junction on ORF1. The most widely used method of detection is reverse transcription-PCR (RT-PCR), which has a high sensitivity and can be applied for further genetic analysis. The purpose of this molecular epidemiological study was to describe viral agents of gastroenteritis in Taiwan and then to further characterize the positive specimens.

Materials and methods

Specimens

Stool specimens were collected from 10 counties in Taiwan between the 23rd of November 2004 and 9th of March 2005. These represented 12 separate outbreaks of gastroenteritis based on the geographic location, date of onset, and place of isolation (termed outbreak A–L). Seven other outbreaks were reported to have occurred during this period; unfortunately these specimens were unavailable. Stool specimens were screened for norovirus, sapovirus, astrovirus, rotavirus, enteric adenovirus, enteroviruses, and kobuvirus. Clinical symptoms, when available, were recorded.

Virus detection

A 10% (w/v) stool suspension was prepared with sterilized MilliQ water and centrifuged at 10,000 × g for 10 min. The QIAamp Viral RNA Mini Vacuum Protocol (Qiagen, Hilden, Germany) was used to extract RNA from 140 µl of the clarified supernatant according to the manufacturer's instructions. cDNA synthesis was carried out with 10 µl of the RNA in 20 µl of the reaction mixture containing 50 pmol random hexamer (Takara, Tokyo, Japan), 1 × Superscript III RT buffer (Invitrogen, Carlsbad, Calif.), 10 mM DTT (Invitrogen), 0.4 mM of each dNTP (Roche, Mannheim, Germany), 1 U RNase inhibitor (Toyobo, Tokyo, Japan), and 10 U Superscript RT III (Invitrogen). RT was performed at 37 °C for 15 min, followed by 50 °C for 1 h. For norovirus genogroup I (GI) PCR, G1SKF and G1SKR primers were used and for norovirus GII PCR, G2SKF and G2SKR primers were used [14]. For sapovirus a nested PCR approach was used for all human genogroups [21]. For the first sapovirus PCR, SV-F11 and SV-R1 primers were used, while for the nested PCR, SV-F21 and SV-R2 primers were used. For astrovirus PCR, Mon244 and 82b primers were used [16]. Enterovirus and kobuvirus were screened as previously described [1, 26]. All PCR products were analyzed by 2% agarose gel electrophoresis and visualized by ethidium bromide staining. For screening of rotavirus and adenovirus, a rapid dry-spot latex agglutination test was used, Diarlex Rota-Adeno (Orion Diagnostica, Finland).

Sequencing and phylogenetic analysis

RT-PCR products were excised from the gel and purified by the QIAquick gel extraction kit (Qiagen, Germany). Nucleotide sequences were prepared using the terminator cycle sequence kit (version 3.1) and determined using an ABI 3100 Avant sequencer (PE Biosystems, USA). Nucleotide sequences were aligned with Clustal X and the distances were calculated by Kimura's two-parameter method as mentioned elsewhere [6]. The nucleotide sequence data determined in this study have been deposited in GenBank under accession numbers: DQ263720–DQ263739. Reference strains were adapted from Kageyama et al. [11], and also included: SaitamaT7, AB112127; KU8, AB058547; U1, AB039775; Sakaeo-53; AY646870; KU44, AB058581; Hunter882M, DQ078799; Kawabe, AB234183; and Manchester, X86560.

Results

In this study, stool specimens collected from 10 counties in Taiwan were screened for norovirus, sapovirus, astrovirus, rotavirus, enteric adenovirus, enteroviruses, and kobuvirus, between the 23rd of November 2004 and 9th of March 2005. These represented 12 distinct outbreaks of gastroenteritis (termed outbreak A–L). Clinical symptoms, when available, were recorded, including the number of symptomatic and asymptomatic cases, location, and period of outbreak (Table 1). Noroviruses were detected in all outbreaks. Rotavirus, adenovirus, and astrovirus were detected in a number of outbreaks, including several specimens mixed with norovirus. Sapovirus, kobuvirus, and enterovirus were not detected in any outbreaks.

Five of ten specimens collected from outbreak A were positive for norovirus GII/4 sequences. These five sequences had 100% nucleotide identity and the consensus was termed Taipei-1A (Fig. 1). We found one mixed infection in outbreak A, i.e., norovirus Taipei-1A and astrovirus (serotype 2). Rotavirus and adenovirus were also detected in a single symptomatic case in outbreak A. Three of 13 specimens collected from outbreak B were positive for norovirus GII/4 sequences. These three sequences had 100% nucleotide identity and the consensus was termed Taipei-11B (Fig. 1). The Taipei-11B sequence had almost 100% nucleotide identity (one nucleotide mismatch) to the norovirus sequence isolated from outbreak A (strain Taipei-1A). Adenovirus was also detected in a single symptomatic case in outbreak B. Seven of eight specimens collected from outbreak C were positive for norovirus GII/4 sequences. These seven sequences had 100% nucleotide identity and the consensus was termed Taipei-24C (Fig. 1). We found that the Taipei-24C sequence had 100% nucleotide identity to the sequence isolated from outbreak B (strain Taipei-11B). Fifteen of 20 specimens from outbreak D were positive for norovirus GII/4 sequences. In this outbreak we found that the norovirus sequence from a symptomatic case had 100% nucleotide identity to 11 of 14 sequences from asymptomatic cases and the consensus was termed Taipei-32D (Fig. 1). Three of 15 sequences from outbreak D had an identical single nucleotide mismatch to that of the Taipei-32D sequence. These three sequences were termed Taipei-33D (Fig. 1). Interestingly, the Taipei-32D sequence had 100% nucleotide identity to the sequence isolated from outbreak B (strain Taipei-11B). Four of five specimens collected from outbreak E were positive for norovirus GII/4 sequences; three sequences had 100% nucleotide identity and the consensus was

Table 1. Details of the outbreaks

Outbreak	Number with symptoms	Total persons	Specimens collected symptomatic	Specimens collected asymptomatic	Total norovirus positive	Suspected route	Common symptoms	Period of outbreak	Settings
A	37	108	10	0	5[a]	ND	vomitting, diarrhea, fever	Nov 23, 2004–Dec 1, 2004	Mental nursing center
B	299	1466	13	0	3[b]	ND	diarrhea, URI	Dec 17, 2004	Senior high school
C	27	90	8	0	7	ground water	diarrhea	Dec 26, 2004–Jan 4, 2005	Temple
D	20	96	1	19	15	toilet	diarrhea	Jan 11, 2005	Hospital
E	16	76	5	0	4	nursing care	diarrhea	Jan 4, 2005–Jan 16, 2005	Hospital
F	9	143	6	0	5	ND	diarrhea, abdominal pain, fever, nausea	Jan 19, 2005–Jan 20, 2005	Elementary school
G	13	58	5	0	4	ND	diarrhea	Jan 25, 2005–Jan 27, 2005	Vagrant center
H	8	40	6	0	3[c]	food	vomitting, diarrhea, cough, nausea	Jan 30, 2005–Feb 1, 2005	Education and nursing institute
I	3	39	0	3	1	ND	vomitting, diarrhea, abdominal pain, fever	Feb 4, 2005	Nursing care center
J	>3	398	3	0	1	ND	vomitting, diarrhea, abdominal pain, fever	Mar 11, 2005	Senior high school
K	53	231	3	8	6[d]	cooker	vomitting, diarrhea, abdominal pain, nausea	Mar 9, 2005–Mar 9, 2005	Educaton center
L	>22	5400	0	5	6[e]	food	diarrhea, vomitting	Jan 20, 2005–Jan 21, 2005	Company
Total			49	35	60				

ND not determined; *URI* upper respiratory infection; [a]rotavirus, adenovirus, and astrovirus were also detected; [b]adenovirus was also detected; [c]rotavirus and adenovirus were also detected; [d]adenovirus was also detected; [e]two different norovirus genotypes were detected from one case

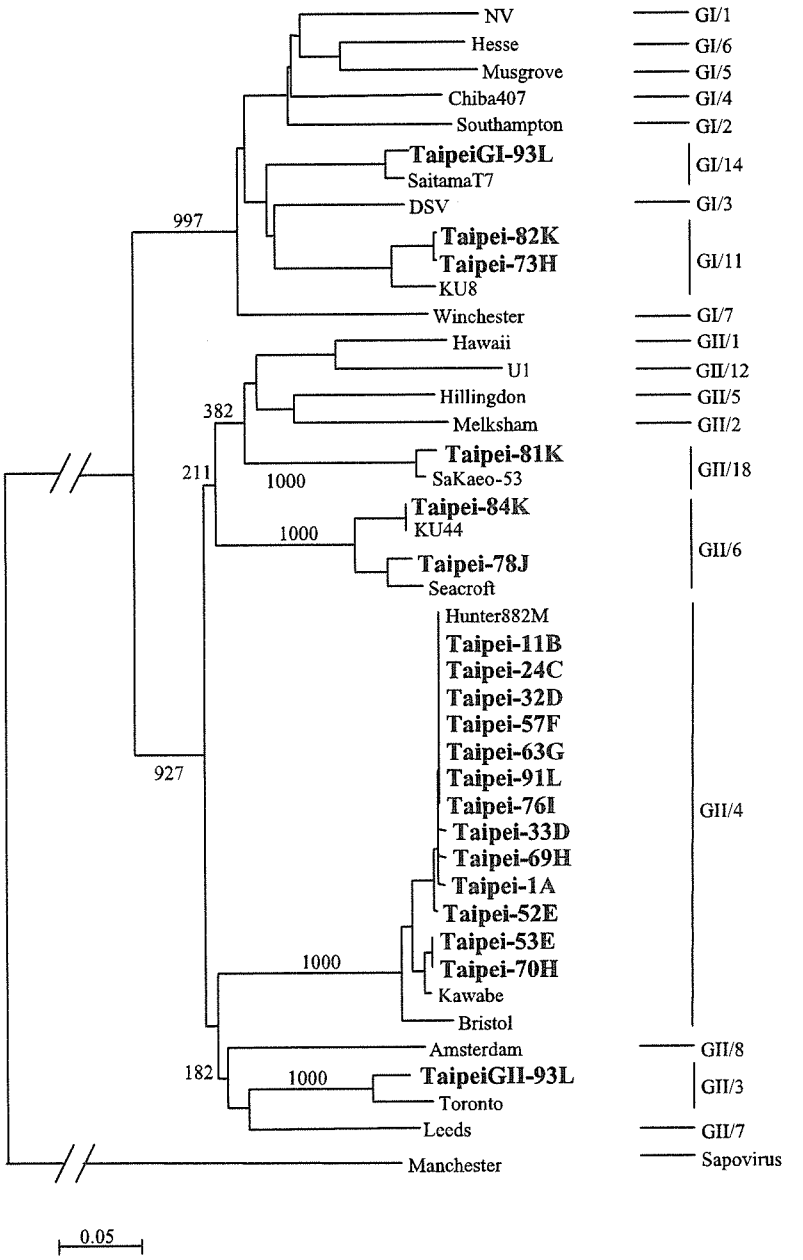

Fig. 1. Phylogenetic tree of the norovirus sequences detected in this study (represented in bold). Norovirus nucleotide sequences were constructed with the partial N-terminal capsid region [11], using sapovirus Manchester strain as an outgroup. We simplified the tree to include only unique sequences, i.e. when two or more sequences had 100% nucleotide identity in the same outbreak we named a single consensus sequence; and if a sequence had one or more nucleotide mismatches in the same outbreak, we gave the sequence a distinct name. The numbers on the branches indicate the bootstrap values for the clusters. Bootstrap values of 950 or higher were considered statistically significant for the grouping [11]. Reference sequences were previously reported [11]

termed Taipei-52E (Fig. 1), whereas one of four sequences had a 10-nucleotide mismatch to that of the Taipei-52E sequence and was termed Taipei-53E (Fig. 1). Five of six specimens collected from outbreak F were positive for norovirus GII/4 sequences. These five sequences had 100% nucleotide identity and the consensus was termed Taipei-57F (Fig. 1). The Taipei-57F sequence had 100% nucleotide identity to the sequence isolated from outbreak B (strain Taipei-11B). Four of five specimens collected from outbreak G were positive for norovirus GII/4 sequences. These four sequences had 100% nucleotide identity and the consensus was termed Taipei-63G (Fig. 1). The Taipei-63G sequence had 100% nucleotide identity to the sequence isolated from outbreak B (strain Taipei-11B). Two of six specimens collected from outbreak H were positive for norovirus GII/4 sequences, while one of six specimens was positive for a norovirus GI/11 sequence. The two norovirus GII/4 sequences isolated in outbreak H were slightly different, having 10 nucleotide differences, and were termed Taipei-69H and Taipei-70H (Fig. 1). Interestingly, the Taipei-69H sequence was more closely related to the sequence isolated from outbreak A (strain Taipei-1A), having only two nucleotide mismatches; whereas the Taipei-70H sequence had 100% nucleotide identity to one of the sequences isolated from outbreak E (strain Taipei-53E). The norovirus GI sequence isolated from outbreak H belonged to GI/11 and was termed Taipei-73H (Fig. 1). Interestingly, the female (15 years of age) infected with the norovirus Taipei-73H strain was also co-infected rotavirus and adenovirus. One of three specimens collected from outbreak I was positive for a norovirus GII/4 sequence, termed Taipei-76I (Fig. 1). The Taipei-76I sequence had 100% nucleotide identity to the Taipei-11B sequence (outbreak B). One of three specimens collected from outbreak J was positive for a norovirus GII/6 sequence, termed Taipei-78J (Fig. 1). Six of eleven specimens collected from outbreak K were positive for norovirus sequences; two sequences had 100% nucleotide identity and belonged to GI/11 (the consensus was termed Taipei-82K) and one of these specimens was also mixed with adenovirus; two sequences had 100% nucleotide identity and belonged to a novel genotype, tentatively named GII/18 (consensus sequences termed Taipei-81K); and two sequences had 100% nucleotide identity and belonged to GII/6 (consensus termed Taipei-84K; Fig. 1). Five specimens were collected from outbreak L and all were positive for noroviruses, one specimen was mixed with both GI and GII sequences, termed TaipeiGI-93L and TaipeiGII-93L (belonging to GI/14 and GII/3, respectively), whereas four sequences had 100% nucleotide identity and belonged to GII/4 (consensus sequences termed Taipei-91L; Fig. 1).

Discussion

Outbreaks of gastroenteritis are dominated by noroviruses worldwide. More than 30 norovirus genotypes have been described, which belong to just two genogroups (GI and GII). Strains have been found to persist in one geographical region, only to suddenly disappear [5, 9]. Shellfish and other contaminated foods, ice, and drinking water are known to harbor noroviruses, leading to outbreaks of gastroenteritis. Seasonal studies have commonly found norovirus outbreaks peaking in the

winter periods, however the incidence rates, detection rates, and overall prevalence rates of infections vary in each country and setting, and are likely affected by the diagnostic techniques. Nevertheless, control measures must be applied swiftly, such as isolating infected cases in order to reduce the burden of this disease.

In this study, we detected six different norovirus genotypes (GI/11, GI/14, GII/3, GII/4, GII/6, and GII/18). We found that 83% (50 of 60) of all isolated norovirus sequences belonged to GII/4, and these were isolated from 10 of 12 outbreaks (outbreaks A, B, C, D, E, F, G, H, I, and L). The GII/4 strains were first detected on the 23rd of November 2004 (outbreak A) and continued to cause outbreaks until the end of the study period (outbreak L). Although only a small part of the norovirus capsid gene was sequenced, we found identical sequences from seven outbreaks; outbreaks B, C, D, F, G, I, and L, corresponding to strains Taipei-11B, Taipei-24C, Taipei-32D, Taipei-57F, Taipei-63G, Taipei-76I, and Taipei-91L, respectively. We also found some variant sequences that had several nucleotide changes (strains Taipei-1A, Taipei-33D, Taipei-52E, and Taipei-69H). Further sequencing of the complete capsid may shed some light onto the evolution of these viruses. We found that sequences from 10 of 12 outbreaks shared between 99 and 100% nucleotide identity to the Hunter882M sequence, which was detected in Australia. Recently, the norovirus Hunter882M strain was found to be a novel variant strain causing the majority of outbreaks in Australia (Dr. Peter White, personal communication). The GII/4 strains were detected throughout Taiwan, demonstrating their widespread distribution.

Norovirus genotypes are generally maintained across the ORFs, however a number of norovirus strains failed to maintain their sequence identities for RdRp and VP1, and they were shown to be recombinant [13, 15, 23]. We found that two outbreaks had norovirus strains belonging to two distinct genotypes and one outbreak had norovirus strains belonging to three distinct genotypes. In outbreaks K and L, norovirus GII strains Taipei-81K/Taipei-84K and Taipei-91L/TaipeiGII-93L were found in the same outbreaks, respectively. We found that these sequences shared short conserved sites of approximately 10 nucleotides at the 5′ end of the capsid gene (data not shown), which may allow for a recombination event to occur, e.g., during copy-choice recombination [4]. Recently, norovirus recombinant strains have received much attention, owing to the fact that they are likely stable and circulating in many countries, including Australia, Japan, Vietnam, Mongolia, and Thailand [4–7]. The result of multiple norovirus strains co-circulating in a single outbreak was recently reported by Kageyama et al. [11] and provides further evidence for norovirus recombination events occurring when two strains come together in physical contact [4]. A number of specimens were mixed with two or more viruses, which included astrovirus/norovirus, rotavirus/adenovirus, rotavirus/adenovirus/norovirus, and adenovirus/norovirus. In these cases the true agent of gastroenteritis could not be determined.

In conclusion, the results of this study provide evidence for the first time of the importance of norovirus infections in Taiwan. We found that norovirus GII/4 strains were the dominant cause of outbreaks of gastroenteritis in Taiwan between November 2004 and March 2005.

Acknowledgements

This work 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. We would like to thank Miss Muo Chin-Hsin for collecting the specimens and epidemiological data.

References

1. Arita M, Zhu SL, Yoshida H, Yoneyama T, Miyamura T, Shimizu H (2005) A Sabin 3-derived poliovirus recombinant contained a sequence homologous with indigenous human enterovirus species C in the viral polymerase coding region. J Virol 79: 12650–12657
2. Bertolotti-Ciarlet A, Crawford SE, Hutson AM, Estes MK (2003) The 3′ end of Norwalk virus mRNA contains determinants that regulate the expression and stability of the viral capsid protein VP1: a novel function for the VP2 protein. J Virol 77: 11603–11615
3. Beuret C, Baumgartner A, Schluep J (2003) Virus-contaminated oysters: a three-month monitoring of oysters imported to Switzerland. Appl Environ Microbiol 69: 2292–2297
4. Bull RA, Hansman GS, Clancy LE, Tanaka MM, Rawlinson WD, White PA (2005) Norovirus recombination in ORF1/ORF2 overlap. Emerg Infect Dis 11: 1079–1085
5. Hansman GS, Doan LT, Kguyen TA, Okitsu S, Katayama K, Ogawa S, Natori K, Takeda N, Kato Y, Nishio O, Noda M, Ushijima H (2004) Detection of norovirus and sapovirus infection among children with gastroenteritis in Ho Chi Minh City, Vietnam. Arch Virol 149: 1673–1688
6. Hansman GS, Katayama K, Maneekarn N, Peerakome S, Khamrin P, Tonusin S, Okitsu S, Nishio O, Takeda N, Ushijima H (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
7. Hansman GS, Kuramitsu M, Yoshida H, Katayama K, Takeda N, Ushijima H, Surenkhand G, Gantolga D, Kuroiwa C (2005) Viral gastroenteritis in Mongolian infants. Emerg Infect Dis 11: 180–182
8. Inouye S, Yamashita K, Yamadera S, Yoshikawa M, Kato N, Okabe N (2000) Surveillance of viral gastroenteritis in Japan: pediatric cases and outbreak incidents. J Infect Dis 181 (Suppl 2): S270–S274
9. Iritani N, Seto Y, Kubo H, Murakami T, Haruki K, Ayata M, Ogura H (2003) Prevalence of Norwalk-like virus infections in cases of viral gastroenteritis among children in Osaka City, Japan. J Clin Microbiol 41: 1756–1759
10. Johansson PJ, Torven M, Hammarlund AC, Bjorne U, Hedlund KO, Svensson L (2002) Food-borne outbreak of gastroenteritis associated with genogroup I calicivirus. J Clin Microbiol 40: 794–798
11. Kageyama T, Shinohara M, Uchida K, Fukushi S, Hoshino FB, Kojima S, Takai R, Oka T, Takeda N, Katayama K (2004) Coexistence of multiple genotypes, including newly identified genotypes, in outbreaks of gastroenteritis due to Norovirus in Japan. J Clin Microbiol 42: 2988–2995
12. Kapikian AZ, Wyatt RG, Dolin R, Thornhill TS, Kalica AR, Chanock RM (1972) Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. J Virol 10: 1075–1081
13. Katayama K, Shirato-Horikoshi H, Kojima S, Kageyama T, Oka T, Hoshino F, Fukushi S, Shinohara M, Uchida K, Suzuki Y, Gojobori T, Takeda N (2002) Phylogenetic analysis of the complete genome of 18 Norwalk-like viruses. Virology 299: 225–239

14. Kojima S, Kageyama T, Fukushi S, Hoshino FB, Shinohara M, Uchida K, Natori K, Takeda N, Katayama K (2002) Genogroup-specific PCR primers for detection of Norwalk-like viruses. J Virol Methods 100: 107–114
15. Lochridge VP, Hardy ME (2003) Snow Mountain virus genome sequence and virus-like particle assembly. Virus Genes 26: 71–82
16. Matsui M, Ushijima H, Hachiya M, Kakizawa J, Wen L, Oseto M, Morooka K, Kurtz J (1998) Determination of serotypes of astroviruses by reverse transcription-polymerase chain reaction and homologies of the types by the sequencing of Japanese isolates. Microbiol Immunol 42: 539–547
17. McEvoy M, Blake W, Brown D, Green J, Cartwright R (1996) An outbreak of viral gastroenteritis on a cruise ship. Commun Dis Rep CDR Rev 6: R188–R192
18. McIntyre L, Vallaster L, Kurzac C, Fung J, McNabb A, Lee MK, Daly P, Petric M, Isaac-Renton J (2002) Gastrointestinal outbreaks associated with Norwalk virus in restaurants in Vancouver, British Columbia. Can Commun Dis Rep 28: 197–203
19. Nakata S, Honma S, Numata K, Kogawa K, Ukae S, Adachi N, Jiang X, Estes MK, Gatheru Z, Tukei PM, Chiba S (1998) Prevalence of human calicivirus infections in Kenya as determined by enzyme immunoassays for three genogroups of the virus. J Clin Microbiol 36: 3160–3163
20. Noel JS, Fankhauser RL, Ando T, Monroe SS, Glass RI (1999) Identification of a distinct common strain of "Norwalk-like viruses" having a global distribution. J Infect Dis 179: 1334–1344
21. Okada M, Shinozaki K, Ogawa T, Kaiho I (2002) Molecular epidemiology and phylogenetic analysis of Sapporo-like viruses. Arch Virol 147: 1445–1451
22. Russo PL, Spelman DW, Harrington GA, Jenney AW, Gunesekere IC, Wright PJ, Doultree JC, Marshall JA (1997) Hospital outbreak of Norwalk-like virus. Infect Control Hosp Epidemiol 18: 576–579
23. Vinje J, Green J, Lewis DC, Gallimore CI, Brown DW, Koopmans MP (2000) Genetic polymorphism across regions of the three open reading frames of "Norwalk-like viruses". Arch Virol 145: 223–241
24. White PA, Hansman GS, Li A, Dable J, Isaacs M, Ferson M, McIver CJ, Rawlinson WD (2002) Norwalk-like virus 95/96-US strain is a major cause of gastroenteritis outbreaks in Australia. J Med Virol 68: 113–118
25. Wilhelmi I, Roman E, Sanchez-Fauquier A (2003) Viruses causing gastroenteritis. Clin Microbiol Infect 9: 247–262
26. Yamashita T, Sugiyama M, Tsuzuki H, Sakae K, Suzuki Y, Miyazaki Y (2000) Application of a reverse transcription-PCR for identification and differentiation of Aichi virus, a new member of the Picornavirus family associated with gastroenteritis in humans. J Clin Microbiol 38: 2955–2961

Author's address: Dr. Grant S. Hansman, Department of Virology II, National Institute of Infectious Diseases, Gakuen 4-7-1, Musashimurayama, Tokyo 208-0011, Japan; e-mail: ghansman@nih.go.jp

Viro-R/883
For Author's Correction Only

Arch Virol (2006) 000: 1–5
DOI 10.1007/s00705-006-0883-z
Printed in The Netherlands

Archives of Virology

Brief Review

Binding activity of norovirus and sapovirus to histo-blood group antigens

H. Shirato-Horikoshi, S. Ogawa, T. Wakita, N. Takeda, and **G. S. Hansman**

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

Received August 22, 2006; accepted October 26, 2006; published online • •, •

Summary

Noroviruses (NoVs) and sapoviruses (SaVs) are causative agents of human gastroenteritis. There is increasing evidence that certain human NoV strains bind to histo-blood group antigens (HBGAs). We found that several NoV virus-like particles (VLPs) showed binding activity to HBGAs, while neither SaV genogroup I (GI) VLP nor SaV GV VLP showed such activity.

*

Human noroviruses (NoVs) and human sapoviruses (SaVs) are etiological agents of human gastroenteritis. Human NoV strains can be grouped into two genogroups (GI and GII), and at least 14 GI and 17 GII genotypes (GI/1–14 and GII/1–17) [11]. SaV strains can be divided into five genogroups (GI–GV), of which the GI, GII, GIV and GV strains infect humans, while the GIII strains infect porcine species [1]. Human NoV and SaV strains are noncultivable, but the expression of the recombinant capsid protein VP1 (rVP1) in insect cells results in the self-assembly of virus-like particles (VLPs) that are antigenically similar to native viruses [2, 9]. In the past several years, increasing evidence has emerged that human NoVs bind to histo-blood group antigens (HBGAs) [8, 12]. These carbohydrate epitopes are present in mucosal secretions and throughout many tissues of the human body, including the small intestine, which may be specifically targeted by certain NoV strains. To the best of our knowledge, the relationship between human SaVs and HBGAs has not yet been reported.

In the present study, we examined the binding activities of human NoV and SaV VLPs to HBGAs present in human saliva and to synthetic carbohydrates. Four NoV strains belonging to different genotypes were examined: the GI/1 124 strain (accession number AB031013), the GI/2 258 strain (AB078335), the GII/4 104 strain (AB078336), and the GII/1 Hawaii strain (U07611). Hawaii VLPs were used as a negative control [5]. Two SaV strains belonging to two different genogroups were also examined: the SaV GI Mc114 strain (AY237422) and the SaV GV NK24 strain (AY646856). Saliva samples were collected from 29 healthy donors. The amounts of Lewis a (Le^a), Lewis b (Le^b), H, A and B antigens in the saliva samples were determined semi-quantitatively by hemagglutination inhibition, and 12 saliva samples with relatively high amounts

Author's address: Dr. Grant Hansman, Department of Virology II, National Institute of Infectious Diseases, Gakuen 4-7-1, Musashi-murayama, Tokyo 208-0011, Japan.
e-mail: ghansman@nih.go.jp

Table 1. Semiquantitation of soluble ABH and Lewis antigens

Donor no.	Interpretation of saliva testing	Grouping	Hemagglutination inhibition titer				
			H	A	B	Lewis-a	Lewis-b
1	Secretor/O/Lewis-positive	H^{high}/Le-b^{high}	>256	0	0	32	32
2	Secretor/O/Lewis-positive		>256	0	0	32	32
3	Secretor/O/Lewis-positive		128	0	0	32	64
4	Secretor/AB/Lewis-positive	A^{high}	16	>256	32	8	2
5	Secretor/AB/Lewis-positive		8	>256	8	8	4
6	Secretor/AB/Lewis-positive		16	128	8	4	4
7	Secretor/B/Lewis-negative	B^{high}	16	0	>256	0	0
8	Secretor/B/Lewis-positive		16	0	128	16	4
9	Secretor/B/Lewis-positive		16	0	128	2	4
10	Nonsecretor	Le-a^{high}	4	8	0	>256	8
11	Nonsecretor		4	1	0	>256	4
12	Nonsecretor		0	1	0	>256	4

of antigens were selected for saliva-VLP binding assay (Table 1). We then used 2 enzyme-linked immunosorbent assay (ELISA)-based assays, a saliva-VLP binding assay and a carbohydrate-VLP binding assay to examine the binding activities of the NoV and SaV VLPs to HBGAs.

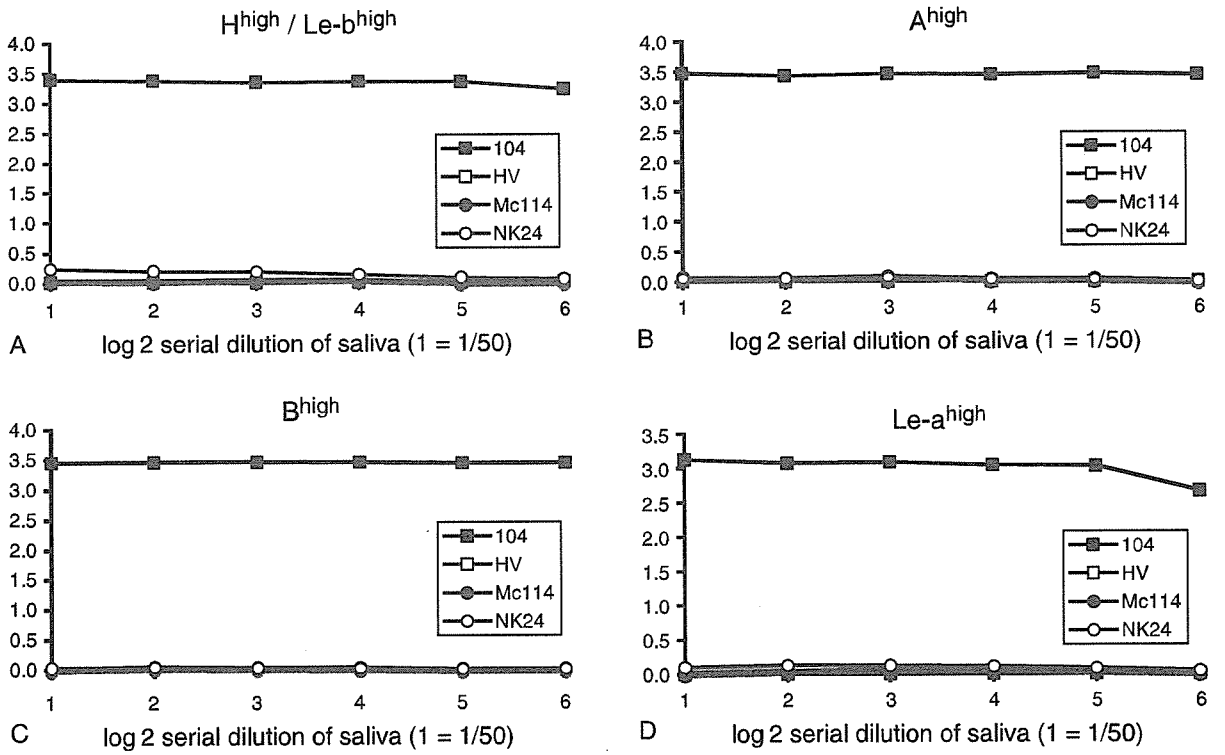

Fig. 1. NoV and SaV VLP binding activity to the saliva samples. Optical densities were measured at 492 nm and were plotted against the serial diluted saliva samples. Each experiment was performed with three donors from each HBGA group (Table 1) and repeated twice. Three samples from the same group produced similar results (data not shown). The results from the donors 1 (**A**), 4 (**B**), 7 (**C**) and 10 (**D**) are shown

In the saliva-VLP binding assay, we examined the possibility that NoV and SaV VLPs may bind to saliva samples. Briefly, 100 µl of serially diluted saliva samples dissolved in carbonate/bicarbonate buffer (50 mmol/L, pH 9.6) were added to wells and incubated at 37 °C overnight. The wells were washed 3–6 times with 300 µl of phosphate-buffered saline containing 0.05% Tween 20 (PBS-T) and washed again after each of the following steps. The wells were blocked with 200 µl of PBS containing 5% skim milk (SM/PBS) for 1 h at room temperature. The purified VLPs, dissolved in 1% SM/PBS-T (final 1 µg/mL), were added (100 µl) to the wells and incubated for 1 h at 37 °C. Next, 100 µl of rabbit anti-rSaV or NoV VLP antiserum (1:2000) in 1% SM/PBS-T was added, and the mixture was incubated for 1 h at 37 °C. Horseradish peroxidase (HRP)-conjugated anti-rabbit IgG (100 µl; Zymed Laboratories Inc., San Francisco, CA, USA) in 1% SM/PBS-T was then added and the mixture was incubated for 1 h at 37 °C. One hundred microliter of O-phenylenediamine (Sigma, St. Louis, MO, USA) was added as substrate, and the mixture was incubated at room temperature for 30 min, at which point 50 µl of 4N H_2SO_4 was added to stop the reaction, and the optical density (OD) at 492 nm was measured. The wells incubated with carbonate/bicarbonate buffer instead of serially diluted saliva samples were used as plate blank (Fig. 1). The Hawaii VLPs have been reported to show no binding to HBGAs [5]; under the present experimental conditions, Hawaii VLPs also showed no binding activity (OD values less than 0.01) at all to HBGAs in saliva. The VLPs of the NoV 104 strain, which resembles Camberwell virus (AF145896) and is classified into GII/4, showed strong binding activity to the diluted saliva at all dilutions of all HBGA samples (OD values greater than 2.7), while SaV Mc114 VLPs showed little binding activity at all dilutions of all samples (OD less than 0.09), as did SaV NK24 VLPs (OD values less than 0.24). These results indicate that the SaV Mc114 and NK24 VLPs have no binding activity with saliva antigens.

In the carbohydrate-VLP binding assay, we examined the possibility that NoV and SaV VLPs may bind to different synthetic carbohydrates, such as H-1 (trisaccharides), A (trisaccharides), B (trisaccharides), Le^a (trisaccharides) and Le^b (tetrasaccharides). Briefly, multivalent carbohydrate-biotin reagents conjugated to polyacrylamide (CHO-PAA-biotin; GlycoTech, Rockville, MD, USA) were resuspended in 0.3 M sodium phosphate buffer at 1 mg/ml, diluted to 20 µg/ml in Tris-buffered saline, and serially diluted twofold, after which 100 µl was added per well to streptavidin-precoated plates (Thermo Labsystems, Basingstoke, United Kingdom) and the mixture was incubated for 2 h at 37 °C. The wells were washed 3–6 times with PBS-T and were washed again after each of the following steps. The plates were blocked with 300 µl of 5% SM/PBS overnight at 4 °C. The VLPs (1 µg/ml in 100 µl of 5% SM/PBS) were added to each well and the mixture was incubated for 4 h at 37 °C. Next, 100 µl of rabbit anti-rNoV VLPs antiserum (1:2000 in 5% SM/PBS) was added and the mixture was incubated for 2 h at 37 °C. One hundred microliter of HRP-conjugated anti-rabbit IgG in 5% SM/PBS was then added and the mixture was incubated for 1 h at 37 °C. The plates were processed as described above. The wells incubated with Tris-buffered saline and 5% SM/PBS instead of serially diluted synthetic carbohydrates and VLPs were used as plate blank.

The NoV 104 VLPs were found to show strong binding activity to three of five synthetic carbohydrates: A, B and Le^b (Fig. 2B, C and E). Although 104 VLPs showed strong binding activity to the saliva samples containing relatively high amounts of H antigen (Fig. 1A) and Le^a antigen (Fig. 1D), they showed only moderate binding activity to H synthetic carbohydrates and no binding activity to Le^a synthetic carbohydrates (Fig. 2A and D). Differences in the reactivity between saliva samples and synthetic carbohydrates may be due to differences between synthetic products and the authentic antigens found in vivo, which are thought to be present on mucin or mucin-like molecules [7]. Therefore, we included two additional NoV VLPs, the 124 and 258 VLPs, as positive controls for the H and Le^a synthetic carbohydrates (Fig. 2A and D), respectively, which showed strong binding activity to the H-high and Le^a-high saliva samples (data not shown). The 124 strain is genetically close to the

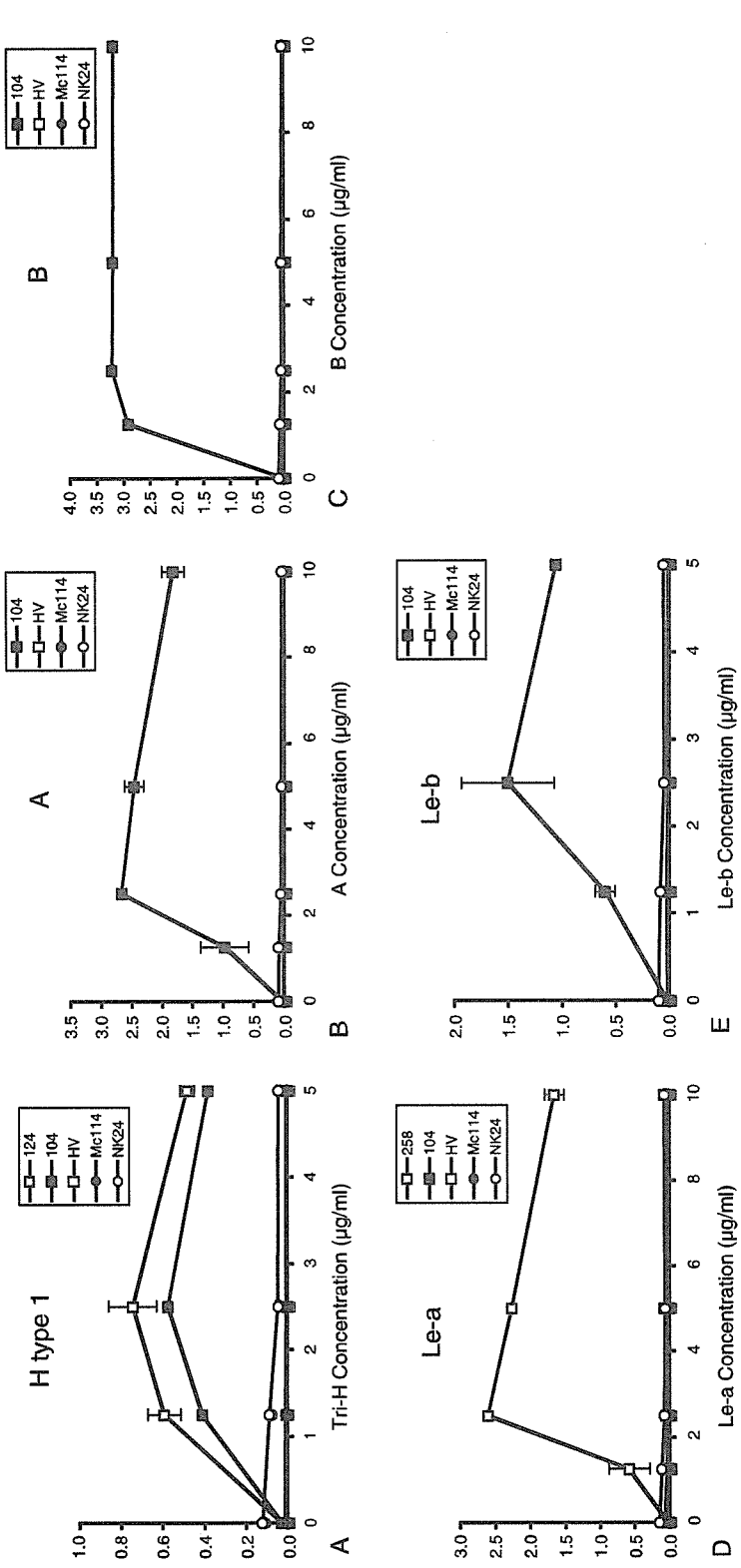

Fig. 2. NoV and SaV VLP binding activity to the synthetic carbohydrates. The optical densities at 492 nm are plotted against the dilutions. Each experiment was performed in duplicate and repeated twice

GI/1 prototype Norwalk virus (NV/68; M87661), and the binding properties of recombinant NV/68 to H antigen have been well characterized [5, 13]. The 124 strain showed 99% amino acid identity with NV/68 in the P2 domain. There is known to be a single amino acid difference at residue 375 which is not related to HBGA binding [14]. The 258 strain resembles Southampton virus (L07418) and is classified into GI/2. The NoV Hawaii VLPs showed no binding activity to any of these synthetic carbohydrates (Fig. 2). The SaV Mc114 and NK24 VLPs also showed no binding activity to any of the synthetic carbohydrates (Fig. 2).

A number studies have found that different NoV strains exhibit different binding patterns to HBGAs [5–7]. In the present study, we found that NoV GII/4 104 VLPs showed binding activity to HBGAs, while SaV GI and GV VLPs showed no such binding activity. Human SaVs are becoming an increasingly important cause of gastroenteritis worldwide [3, 4, 10]. Further studies are needed to examine the possibility that other human SaV genogroups have binding activity.

Acknowledgements

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

References

1. Farkas T, Zhong WM, Jing Y, Huang PW, Espinosa SM, Martinez N, Morrow AL, Ruiz-Palacios GM, Pickering LK, Jiang X (2004) Genetic diversity among sapoviruses. Arch Virol 149: 1309–1323
2. Hansman GS, Natori K, Oka T, Ogawa S, Tanaka K, Nagata N, Ushijima H, Takeda N, Katayama K (2005) Cross-reactivity among sapovirus recombinant capsid proteins. Arch Virol 150: 21–36
3. Hansman GS, Takeda N, Oka T, Oseto M, Hedlund KO, Katayama K (2005) Intergenogroup recombination in sapoviruses. Emerg Infect Dis 11: 1916–1920
4. Hansman GS, Takeda N, Katayama K, Tu ET, McIver CJ, Rawlinson WD, White PA (2006) Genetic diversity of Sapovirus in children, Australia. Emerg Infect Dis 12: 141–143
5. Harrington PR, Lindesmith L, Yount B, Moe CL, Baric RS (2002) 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 76: 12335–12343
6. Huang P, Farkas T, Marionneau S, Zhong W, Ruvoen-Clouet N, Morrow AL, Altaye M, Pickering LK, Newburg DS, LePendu J, Jiang X (2003) Noroviruses bind to human ABO, Lewis, and secretor histo-blood group antigens: identification of 4 distinct strain-specific patterns. J Infect Dis 188: 19–31
7. Huang P, Farkas T, Zhong W, Tan M, Thornton S, Morrow AL, Jiang X (2005) 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 79: 6714–6722
8. Hutson AM, Atmar RL, Graham DY, Estes MK (2002) Norwalk virus infection and disease is associated with ABO histo-blood group type. J Infect Dis 185: 1335–1337
9. Jiang X, Wang M, Graham DY, Estes MK (1992) Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein. J Virol 66: 6527–6532
10. Johansson PJ, Bergentoft K, Larsson PA, Magnusson G, Widell A, Thorhagen M, Hedlund KO (2005) A nosocomial sapovirus-associated outbreak of gastroenteritis in adults. Scand J Infect Dis 37: 200–204
11. Kageyama T, Shinohara M, Uchida K, Fukushi S, Hoshino FB, Kojima S, Takai R, Oka T, Takeda N, Katayama K (2004) Coexistence of multiple genotypes, including newly identified genotypes, in outbreaks of gastroenteritis due to Norovirus in Japan. J Clin Microbiol 42: 2988–2995
12. Lindesmith L, Moe C, Marionneau S, Ruvoen N, Jiang X, Lindblad L, Stewart P, LePendu J, Baric R (2003) Human susceptibility and resistance to Norwalk virus infection. Nat Med 9: 548–553
13. Marionneau S, Ruvoen N, Le Moullac-Vaidye B, Clement M, Cailleau-Thomas A, Ruiz-Palacois G, Huang P, Jiang X, Le Pendu J (2002) Norwalk virus binds to histo-blood group antigens present on gastroduodenal epithelial cells of secretor individuals. Gastroenterology 122: 1967–1977
14. Tan M, Huang P, Meller J, Zhong W, Farkas T, Jiang X (2003) Mutations within the P2 domain of norovirus capsid affect binding to human histo-blood group antigens: evidence for a binding pocket. J Virol 77: 12562–12571

Journal of General Virology (2006), **87**, 3329–3338 DOI 10.1099/vir.0.81799-0

Identification of the cleavage sites of sapovirus open reading frame 1 polyprotein

Tomoichiro Oka, Mami Yamamoto, Kazuhiko Katayama, Grant S. Hansman, Satoko Ogawa, Tatsuo Miyamura and Naokazu Takeda

Correspondence
Tomoichiro Oka
oka-t@nih.go.jp

Department of Virology II, National Institute of Infectious Diseases, Gakuen 4-7-1, Musashi-murayama, Tokyo 208-0011, Japan

Sapovirus (SaV), a member of the family *Caliciviridae*, is a causative agent of acute gastroenteritis in humans and swine and is currently divided into five genogroups, GI–GV. The proteolytic processing of the SaV open reading frame 1 (ORF1) polyprotein with a human GII SaV Mc10 strain has recently been determined and the products are arranged in the following order: NH_2–p11–p28–p35 (NTPase)–p32–p14 (VPg)–p70 (Pro–Pol)–p60 (VP1)–COOH. The cleavage site between p14 (VPg) and p70 (Pro–Pol) was identified as E^{1055}/A^{1056} by N-terminal amino acid sequencing. To identify other cleavage sites, a series of GII SaV Mc10 full-length clones containing disrupted potential cleavage sites in the ORF1 polyprotein were constructed and used to generate linear DNA templates for *in vitro* coupled transcription–translation. The translation products were analysed by SDS-PAGE or by immunoprecipitation with region-specific antibodies. N-terminal amino acid sequencing with *Escherichia coli*-expressed recombinant proteins was also used to identify the cleavage site between p32 and p14. These approaches enabled identification of the six cleavage sites of the Mc10 ORF1 polyprotein as E^{69}/G^{70}, Q^{325}/G^{326}, Q^{666}/G^{667}, E^{940}/A^{941}, E^{1055}/A^{1056} and E^{1722}/G^{1723}. The alignment of the SaV full-length ORF1 amino acid sequences indicated that the dipeptides used for the cleavage sites were either E or Q at the P1 position and A, G or S at the P1′ position, which were conserved in the GI, GII, GIII, GIV and GV SaV ORF1 polyprotein.

Received 28 December 2005
Accepted 26 June 2006

INTRODUCTION

The family *Caliciviridae* is composed of four genera, namely *Sapovirus*, *Lagovirus*, *Vesivirus* and *Norovirus* (Mayo, 2002). Sapovirus (SaV) is a causative agent of gastroenteritis in humans and swine (Guntapong *et al.*, 2004; Guo *et al.*, 2001a; Hansman *et al.*, 2004a, 2006; Johansson *et al.*, 2005; Katayama *et al.*, 2004; Nakata *et al.*, 2000; Noel *et al.*, 1997; Robinson *et al.*, 2002; Vinjé *et al.*, 2000) and is currently divided into five distinct genetic groups, genogroups I (GI)–GV, on the basis of the capsid protein gene (Farkas *et al.*, 2004). The SaV strains that infect pigs belong to genogroup GIII and these strains are able to multiply in cultured cells (Chang *et al.*, 2004; Flynn & Saif, 1988); however, the strains that infect humans, which belong to genogroups GI, GII, GIV and GV, have not yet been grown in cultured cells.

The SaV genome is a linear, polyadenylated, positive-sense, single-stranded RNA of about 7·5 kb with either two or three open reading frames (ORFs) (Clarke & Lambden, 2000). ORF1 encodes a 250 kDa polyprotein that contains amino acid motifs characteristic of caliciviruses, including 2C-like NTPase (NTPase), VPg, 3C-like protease (Pro), 3D-like RNA-dependent RNA polymerase (Pol) and capsid protein (VP1). The functional domains in the ORF1 polyprotein were predicted on the basis of the motifs found in the picornavirus polyprotein, and these domains are highly conserved among members of the family *Caliciviridae* (Green *et al.*, 2000; Meyers *et al.*, 2000). The functions of the proteins encoded by ORF2 and ORF3 have not yet been elucidated.

Proteolytic processing of the ORF1 polyprotein is a common feature of the caliciviruses (Green *et al.*, 2000) and the cleavage sites have been mapped in detail in *Rabbit hemorrhagic disease virus* (RHDV) (König *et al.*, 1998; Martin Alonso *et al.*, 1996; Meyers *et al.*, 2000; Wirblich *et al.*, 1995, 1996), *Feline calicivirus* (FCV) (Sosnovtsev *et al.*, 1998, 2002; Sosnovtseva *et al.*, 1999) and norovirus (NoV) (Belliot *et al.*, 2003; Blakeney *et al.*, 2003; Hardy *et al.*, 2002; Liu *et al.*, 1996, 1999; Seah *et al.*, 1999, 2003; Someya *et al.*, 2000). The 3C-like protease of these viruses cleaves dipeptides containing either glutamic acid (E) or glutamine (Q) at the P1 position (i.e. the amino acid immediately upstream of the scissile bond) and those containing glycine (G), alanine (A), serine (S), threonine (T), aspartic acid (D) or asparagine (N) at the P1′ position (i.e. the amino acid

The GenBank/EMBL/DDBJ accession numbers for the nucleotide sequences reported in this paper are AY237420, AY237422, AY237423, AY646853, AY646854, AY237419, AY646855, DQ058829, AY646856, X86560, AY694184, AJ249939, AY603425, AF182760, NC_000940 and DQ125333.

immediately downstream of the scissile bond). We have recently identified SaV ORF1 polyprotein-cleavage products of the GII Mc10 strain and these products are arranged in the following order: NH_2–p11–p28–p35 (NTPase)–p32–p14 (VPg)–p70 (Pro–Pol)–p60 (VP1)–COOH (Oka *et al.*, 2005b). Site-directed mutagenesis of the GDCG motif in the 3C-like protease fully abolished the proteolytic activity of this enzyme, thus demonstrating that the viral 3C-like protease was responsible for the cleavage (Oka *et al.*, 2005b).

Our recent study using an *Escherichia coli* expression system also revealed that GII Mc10 3C-like protease cleaves the Q/G site in the rhinovirus 3C protease-recognition sequence (Oka *et al.*, 2005a) in a manner similar to that of NoV Chiba virus 3C-like protease (Someya *et al.*, 2000). Although the cleavage site between p14 (VPg) and p70 (Pro–Pol) of GII Mc10 was identified as E^{1055}/A^{1056} by N-terminal amino acid sequencing (Oka *et al.*, 2005a), the other cleavage sites are unknown. Our cleavage-products map indicated that the GII Mc10 ORF1 polyprotein should have at least six cleavage sites between six non-structural proteins and one structural protein (Oka *et al.*, 2005b).

The aim of this study was to identify the remaining five cleavage sites of the Mc10 ORF1 polyprotein. Site-directed mutagenesis, an *in vitro* coupled transcription–translation system and N-terminal amino acid sequencing of *E. coli*-expressed recombinant proteins were used to identify all of the cleavage sites. In addition, the cleavage site between p14 and p70 was confirmed by site-directed mutagenesis. The dipeptide used for the cleavage sites was conserved among 16 SaV strains and was similar to those of other members of the family *Caliciviridae*.

METHODS

Virus strains and their complete nucleotide sequences. The SaV GII Mc10, GI Mc114, GI N21 and GII Mc2 strains were isolated from infants hospitalized with acute gastroenteritis in Chiang Mai, Thailand, in 2000 (Hansman *et al.*, 2004b). The GII C12 strain was isolated from an infant with gastroenteritis in Sakai, Japan, in 2001 (Katayama *et al.*, 2004). The GI NongKhai-50, GI Chanthaburi-74, GV NongKhai-24 (NK24) and GII SaKaeo-15 (SK15) strains were isolated from infants with gastroenteritis in NongKhai, Chanthaburi or SaKaeo, Thailand, between November 2002 and April 2003 (Guntapong *et al.*, 2004). The GIV Ehime1107 and GIV Sw278 strains were isolated from an infant with gastroenteritis in Matsuyama, Japan, in 2002, and from an adult with gastroenteritis in Solna, Sweden, in 2003 (Hansman *et al.*, 2005d). RNA extraction from the stool samples, cDNA synthesis and complete genome sequencing were performed as described previously (Katayama *et al.*, 2002). The SaV sequences were classified phylogenetically according to the method of Farkas *et al.* (2004).

Full-length cDNA clones. Plasmids containing a full-length Mc10 genome sequence with the T7 promoter, designated 'pUC19/SaV Mc10 full-length' and 'pUC19/SaV Mc10 full-C1171A/ORF1', of which the latter contains a ^{1169}GDCG^{1172} to GDAG mutation in the protease, have been described previously (Oka *et al.*, 2005b).

Site-directed mutagenesis. Site-directed mutagenesis was performed by using the GeneTailor site-directed mutagenesis system (Invitrogen) with pUC19/SaV Mc10 full-length as a template. In brief, 100 ng plasmid was methylated in 16 µl reaction mixture according to the manufacturer's instructions and then PCR was performed in 100 µl reaction mixture containing 1 µl methylated DNA, 40 pmol each primer, KOD polymerase buffer, 0·2 mM each dNTP, 1 mM $MgSO_4$ and 2 units KOD-Plus DNA polymerase (TOYOBO). The primers used for site-directed mutagenesis to generate the mutant full-length cDNA clones are represented in Table 1. After initial denaturation at 94 °C for 5 min, 20 cycles of amplification were performed. Each cycle consisted of denaturation at 94 °C for 30 s, primer annealing at 55 °C for 30 s and primer extension at 72 °C for 10 min, followed by a final extension at 72 °C for 15 min. *E. coli* DH5α-T1 cells (Invitrogen) were transformed with 2 µl PCR mixture, and the plasmids containing the mutation(s) in ORF1 were amplified. The resulting 17 full-length mutant cDNA clones were designated as follows: pUC19/SaV Mc10 full-ORF1-E69A (where E at amino acid residue 69 was changed to A), -EE6869AA (where E at amino acid position 68 was changed to A, and E at amino acid position 69 was changed to A), -Q112A, -Q325A, -E385A, -E430A, -EE429430AA, -Q666A, -EQ665666AA, -E940A, -EE939940AA, -EEE938939940AAA, -E1055A, -E10541055AA, -E1679A, -Q1690A and -E1722A. All of the full-length clones were verified by sequencing and no additional mutation was found.

***In vitro* coupled transcription–translation assay and immunoprecipitation.** The template for *in vitro* T7 polymerase coupled transcription–translation, which consisted of linear DNA fragments containing the T7 promoter, was generated by PCR with pUC19/SaV Mc10 full-length, pUC19/SaV Mc10 full-C1171A/ORF1 and 17 full-length mutant cDNA clones. DNA fragments corresponding to the entire ORF1 [Fig. 1(I)] were generated with the forward primer 5′-GGATCCTAATACGACTCACTATAGGGAACAGCCAC-**CATG**gcttccaagccattctacccaatagag-3′, including the T7 promoter sequence (underlined) and a start codon (bold), and antisense primer 5′-T_{30}**TT**Attctaagaacctaacggcccgg, including a stop codon (bold), and designated I-Pro^{w}, -Pro^{mut}, -E69A, -EE6869AA, -Q112A, -Q325A, -E385A, -E430A, -EE429430AA, -Q666A, -EQ665666AA, -E940A, -EE939940AA, -EEE938939940AAA, -E1055A, -E10541055AA, -E1679A, -Q1690A and -E1722A. The truncated ORF1 templates [Fig. 1(II) and (III)] were similarly generated with the forward primers 5′-GGATCCTAATACGACTC-ACTATAGGGAACAGCCACC**ATG**gccaaaggaaagaccaagcatggc-3′ and 5′-GGATCCTAATACGACTCACTATAGGGAACAGCCACC**AT**-**G**gctcccacaccaattgttac-3′, including the T7 promoter-encoding sequence (underlined) and a start codon (bold), and the antisense primer described above. The products were designated II-Pro^{w}, II-Pro^{mut}, -E1055A, -EE10541055AA and III-Pro^{w}, -Pro^{mut}, -E1679A, -Q1690A and -E1722A.

In vitro T7 polymerase coupled transcription–translation was performed by using a TNT T7 Quick for PCR DNA kit (Promega) according to the manufacturer's instructions. The reaction was performed in the presence of ^{35}S-labelled methionine and cysteine (Redivue Pro-mix L-[^{35}S] *in vitro* cell labeling mix; Amersham Biosciences) and 2 µl of the reaction mixture was analysed by SDS-PAGE. Immunoprecipitation was performed with 10 µl reaction mixture and 5 µg region-specific antibodies raised against *E. coli*-expressed recombinant proteins, A (aa 1–231), C (aa 637–812) and D (aa 941–1055), as described previously (Fig. 1) (Oka *et al.*, 2005b). The complex was captured with protein A magnetic beads (New England Biolabs) as described previously (Oka *et al.*, 2005b). The protein in the gel was blotted electrically onto a PVDF membrane (Immobilon-P; Millipore) and radiolabelled proteins were detected by a Bioimage Analyser BAS 2500 (Fujifilm).

Construction of *E. coli* expression plasmids. The DNA fragment corresponding to the amino acid residues 926–1720

Table 1. Oligonucleotides used for the site-directed mutagenesis

Potential cleavage site*	Mutant name†	Nucleotide sequence (5′–3′)‡
FTEE^{69}/G^{70}LLD	E69A	GCCCCCACATTCACGGAGgcgGGCTTGTTAGACTC
	EE68,69AA	CCGCCCCCACATTCACGgcggcgGGCTTGTTAGACTC
AKLQ^{112}/G^{113}ELV	Q112A	CATGATTTGGCAAAATTAgcgGGTGAGCTTGTGGG
FQSQ^{325}/G^{326}PTS	Q325A	GAGCACGTTCCAGTCAgcaGGCCCAACATCTTGCGGC
IIAE^{385}/A^{386}DGK	E385A	GCGAGACATCATAGCAgcaGCTGATGGCAAGGCTAG
LIEE^{430}/A^{431}TEL	E430A	CACACAGCTCATCGAGgcgGGTACTGAGTTGATACAG
	EE429,430AA	GTTTCACACAGCTCATCgcggcgGGTACTGAGTTGATAC
FKEQ^{666}/G^{667}NEH	Q666A	GAAACCAAGTTTAAGGAGgcgGGCAATGAACATCG
	EQ665,666AA	GATGGAAACCAAGTTTAAGgcggcgGGCAATGAAC
REEE^{940}/A^{941}KGK	E940A	GTCTGGTCGGGAGGAGgcgGCCAAAGGAAAGACCAAGC
	EE939,940AA	CCGGTCTGGTCGGGAGgcggcgGCCAAAGGAAAGACC
	EEE938,939,940AAA	CACCCGGTCTGGTCGGgcggcggcgGCCAAAGGAAAGACC
YDEE^{1055}/A^{1056}PTP	E1055A	GTAATCAAGGTTATGATGAAgcaGCTCCCACACCAATTG
	EE1054,1055AA	CGTAATCAAGGTTATGATgcagcaGCTCCCACACCAATTG
AKAE^{1679}/G^{1680}LVL	E1679A	GACACACAGCCAAGGCTgcgGGACTGGTGCTAACC
NYDQ^{1690}/A^{1691}LAT	Q1690A	CAATGTCAACTATGACgcgGCTCTCGCCACCTACG
FEME^{1722}/G^{1723}LGQ	E1722A	CAAATTAGTGTTTGAAATGgcgGGCCTAGGCCAACC

*The potential cleavage sites, including the surrounding amino acids (P4 to P4′), are shown. '/' indicates the potential cleavage site.
†The names of the mutants represent the amino acid change. Amino acids are shown in the one-letter code. Letters before the number indicate the original amino acid residue(s) and letters after the number indicate the mutant amino acid residues.
‡Only the positive-sense oligonucleotide sequence is shown. The codon(s) corresponding to changed amino acid(s) are indicated in lower case.

(nt 2789–5173) was amplified by PCR with 500 ng plasmid pUC19/SaV Mc10 full-ORF1-E1055A, which contains the nucleotide changes in the P1 position of the p14/p70 cleavage site, or UC19/SaV Mc10 full-ORF1-E1055A/C1171A, which contains an additional mutation in the GDCG motif, with a sense primer (5′-CAGGGGCCCCTGGGATCCcacaatgtttcatacctcgcc-3′) including a *Bam*HI site (underlined) and an antisense primer (5′-GCCGCTCGA*GTCGAC*TCAGTGATGGTGATGGTGATGttcaaacact-aatttggtggtctcttcactggggct-3′), including a 6×His tag-encoding sequence (underlined), a stop codon (bold) and a *Sal*I site (italic). The PCR products were purified and digested with *Bam*HI and *Sal*I (New England Biolabs) and cloned into the corresponding sites of

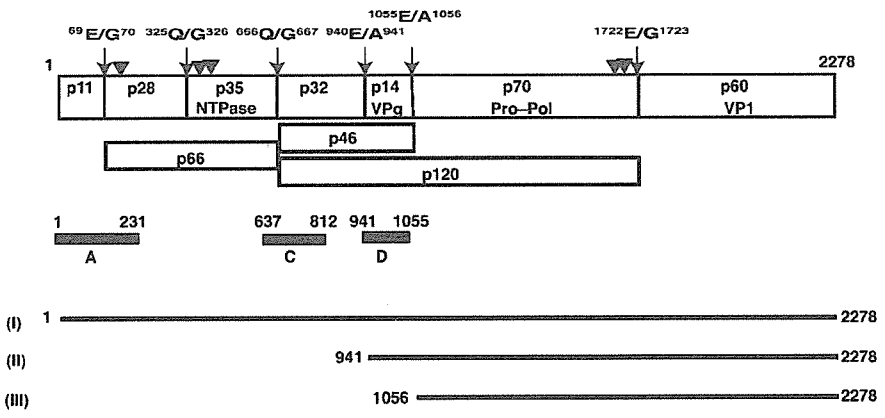

Fig. 1. Diagram of the ORF1 polyprotein-cleavage map of SaV GII Mc10 and linear DNA templates containing the T7 promoter for *in vitro* coupled transcription–translation. The Mc10 ORF1 polyprotein-cleavage map is shown at the top. One previously identified cleavage site (E^{1055}/A^{1056}) and five cleavage sites identified in this study, E^{69}/G^{70}, Q^{325}/G^{326}, Q^{666}/G^{667}, E^{940}/A^{941} and E^{1722}/G^{1723}, are indicated by arrows; the other five potential cleavage sites not utilized for the proteolytic processing, Q^{112}/G^{113}, E^{385}/A^{386}, E^{430}/G^{431}, E^{1679}/G^{1680} and Q^{1690}/A^{1691}, are indicated by arrowheads. The *E. coli*-expressed proteins used to prepare region-specific antibodies are indicated as A, C and D (Oka *et al.*, 2005b). The DNA templates used for the *in vitro* coupled transcription–translation system are shown at the bottom. The template DNA fragments were generated by PCR.

the pGEX-4T-1 vector (Amersham Biosciences). DH5α cells (TOYOBO) were used for the transformation and propagation of the plasmids. The plasmids were designated pGEX-4T-1-Δp32-p14-p70/E1055A and pGEX-4T-1-Δp32-p14-p70/E1055A/C1171A. The truncated ORF1 polyproteins were expressed as fusion proteins with glutathione *S*-transferase (GST) at the N terminus and 6 × His tag products at the C terminus.

Expression of recombinant proteins in *E. coli*. *E. coli* BL21-CodonPlus-RIL cells (Stratagene) were transformed with the expression plasmid and incubated at 37 °C in Luria broth in the presence of 50 μg ampicillin ml^{-1} and 50 μg chloramphenicol ml^{-1} until the OD_{600} value reached 0·6–0·8. Expression was induced by addition of a final concentration of 1 mM IPTG followed by incubation at 37 °C for 3 h. The *E. coli* lysates or purified recombinant proteins were separated by SDS-PAGE and stained with GelCode blue staining reagent (Pierce) (Oka *et al.*, 2005a). The recombinant proteins were purified by using TALON resin (BD Clontech) and subjected to N-terminal amino acid sequencing (APRO Science) (Oka *et al.*, 2005a).

Nucleotide and amino acid sequence analyses. Nucleotide sequence analysis was performed with a BigDye Terminator (version 3.1) cycle sequencing ready reaction kit (Applied Biosystems) and an automated sequencer, the 3100 Avanti genetic analyser (Applied Biosystems). Nucleotide sequences were assembled with the program SEQUENCHER version 4.2.2 (Gene Codes Corporation). Nucleotide and amino acid sequences were analysed with GENETYX Mac software, version 12.2.6 (Genetyx Corporation).

RESULTS

Selection and amino acid substitution of the potential cleavage sites in the ORF1 polyprotein

Our previous study with N-terminal amino acid sequencing revealed that E^{1055}/A^{1056} is the cleavage site between p14 (VPg) and p70 (Pro–Pol) in the GII Mc10 ORF1 polyprotein (Oka *et al.*, 2005a). The other potential cleavage sites were selected on the basis of our findings that SaV 3C-like protease cleaves next to E or Q residues. An amino acid alignment with seven SaV strains, i.e. GI Manchester (GenBank accession no. X86560), GI Dresden (AY694184), GII Bristol (AJ249939), GII Mc10 (AY237420), GII C12 (AY603425), GIII PEC (AF182760) and GIII PEC LL14 (NC_000940), allowed us to identify six potential cleavage sites (E^{69}/G^{70}, Q^{325}/G^{326}, Q^{666}/G^{667}, E^{940}/A^{941}, E^{1679}/G^{1680} and E^{1722}/G^{1723}) (Fig. 1). In addition, we tested four additional dipeptides (Q^{112}/G^{113}, E^{385}/A^{386}, E^{430}/G^{431} and Q^{1690}/A^{1691}) (Table 1), because these sites were conserved among five human SaV strains: GI Manchester, GI Dresden, GII Bristol, GII Mc10 and GII C12. Although our strategy was to alter the E or Q residues at position P1 to an A residue, five of 11 potential cleavage sites, including E^{1055}/A^{1056}, had amino acid(s) that were able to create a potential novel cleavage site immediately upstream of the P1 position (i.e. $^{68}EE^{69}$, $^{429}EE^{430}$, $^{665}EQ^{666}$, $^{938}EEE^{940}$ and $^{1054}EE^{1055}$) (Table 1). Therefore, mutant clones containing either double or triple substitutions were prepared. To this end, we constructed 17 full-length mutant cDNA clones for the 11 potential cleavage sites, as shown in Table 1 and Fig. 1.

Cleavage of the ORF1 polyprotein in an *in vitro* coupled transcription–translation system

To identify the potential cleavage sites in the ORF1 polyprotein, 17 linear template DNAs containing the T7 promoter were amplified by PCR as described in Methods; these sites corresponded to entire ORF1 regions and were designated I-E69A, -EE6869AA, -Q112A, -Q325A, -E385A, -E430A, -EE429430AA, -Q666A, -EQ665666AA, -E940A, -EE939940AA, -EEE938939940AAA, -E1055A, -E10541055AA, -E1679A, -Q1690A and -E1722A. Then, *in vitro* coupled transcription–translation was performed and the expressed proteins were analysed by SDS-PAGE or immunoprecipitation. Two DNA templates, I-Pro^{w}, which encodes the wild-type protease, and I-Pro^{mut}, which encodes the mutant protease, were used as the positive and negative controls for proteolytic processing. I-Pro^{w} produced at least nine proteins, i.e. p11, p14, p28, p32, p35, p46, p60, p66 and p120 (Fig. 2, lanes 10 and 20), whereas I-Pro^{mut} produced a major 250 kDa product in SDS-PAGE (Fig. 2, lanes 11 and 21) (Oka *et al.*, 2005b). To detect p11 and p14, immunoprecipitation with anti-A and anti-D region-specific antibodies was performed (Fig. 1). The detection of p70 (Pro–Pol) was difficult when the entire ORF1 region was expressed, as described previously (Oka *et al.*, 2005b). A ~100 kDa product clearly visible in Fig. 2 was present in all samples analysed, including the I-Pro^{mut} sample. This suggested that it was probably an artefact of the expression system, an internal initiation product or a terminally truncated protein, and is not discussed further in this study. As shown in Fig. 2, 10 constructs – I-E69A, -EE6869AA, -Q325A, -Q666A, -EQ665666AA, -EE939940AA, -EEE938939940AAA, -E1055A, -E10541055AA and -E1722A – demonstrated cleavage patterns different from those of I-Pro^{w}, demonstrating clearly that the proteolytic processing of the ORF1 polyprotein was blocked in these constructs, as described in the following sections.

Cleavage site between p11 and p28

Two potential cleavage sites, E^{69}/G^{70} and Q^{112}/G^{113}, were tested (Table 1; Fig. 1). If the cleavage site was abolished, then these two products (i.e. p11 and p28) would be expected to disappear, whereas p39, corresponding to p11–p28, would appear. *In vitro* coupled transcription–translation with I-E69A as the template resulted in the accumulation of p39 (Fig. 2, lane 1), which was immunoprecipitated with anti-A antibody (Fig. 3a, lane 1), indicating clearly that p39 was indeed p11–p28. The construct I-EE6869AA had a cleavage pattern identical to that of I-E69A (Fig. 2, lanes 1 and 2), demonstrating that the cleavage occurred between E^{69} and G^{70} and that no alternative cleavage occurred between E^{68} and A^{69}. The loss of p66, which corresponds to p28–p35, was also observed in both the I-E69A and I-EE6869AA translation products (Fig. 2, lanes 1 and 2; Fig. 3a, lane 1). I-Q112A had little effect on the proteolytic processing (Fig. 2, lane 3), demonstrating that Q^{112}/G^{113} is not a cleavage site between p11 and p28. Based

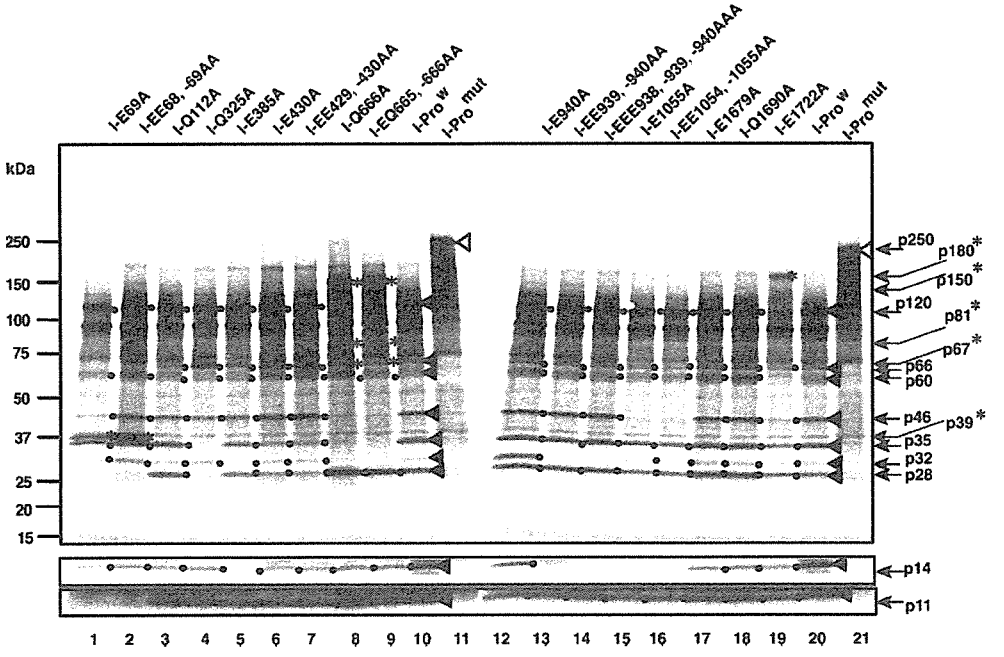

Fig. 2. Analysis of *in vitro* translation products. I-Pro^{w}, I-Pro^{mut} and 17 I- constructs containing mutation(s) in the potential cleavage site were analysed by SDS-PAGE. p11 and p14 were detected by immunoprecipitation with anti-A and anti-D antibodies. The specific products in I-Pro^{w} and I-Pro^{mut} are indicated by filled and open arrowheads, respectively. The products found in I-Pro^{w}, p120, p66, p60, p46, p35, p32, p28, p14 and p11, are indicated by dots and the newly appearing products, p180, p150, p81, p67, p39, are indicated by asterisks. Mc10 ORF1-specific proteins are indicated on the right and molecular size markers are shown on the left.

on these findings, we concluded that the cleavage site between p11 and p28 is E^{69}/G^{70}.

Cleavage site between p28 and p35

Three potential cleavage sites, Q^{325}/G^{326}, E^{385}/A^{386} and E^{430}/G^{431}, were predicted between p28 and p35 (Table 1; Fig. 1). The translation products from the I-E385A, I-E430A and I-EE429430AA constructs were similar to those of I-Pro^{w} (Fig. 2, lanes 5, 6, 7, 10 and 20), indicating that E^{385}/A^{386} and E^{430}/G^{431} are not cleavage sites between p28 and p35. In contrast, the cleavage products from I-Q325A showed a loss of p28 and p35 and an accumulation of p66 (Fig. 2, lane 4; Fig. 3a, lane 2). Therefore, we concluded that the cleavage site between p28 and p35 is Q^{325}/G^{326}.

Cleavage site between p35 and p32

One site, Q^{666}/G^{667}, was tested for the putative cleavage site between p35 and p32 (Table 1; Fig. 1). The translation products from I-Q666A showed a loss of p35 and p32; in addition, a loss of p66, p46 and p120 was observed, as well as an accumulation of p67, p81 and p150 (Fig. 2, lane 8). The antibodies raised against fragment C immunoprecipitated p67, p81 and p150 (Fig. 3b, lane 1), demonstrating that these products corresponded to p35–p32, p35–p32–p14 and p35–p32–p14–p70, respectively. I-EQ665666AA had a cleavage pattern identical to that of I-Q666A (Fig. 2, lanes 8 and 9), indicating that alternative cleavage did not occur between E^{665} and A^{666}. Based on these results, we concluded that the cleavage site between p28 and p35 is Q^{666}/G^{667}.

Cleavage sites between p32 and p14

E^{940}/A^{941} was predicted as the putative cleavage site between p32 and p14 (Table 1; Fig. 1) (Oka *et al.*, 2005b). The cleavage products from I-EE939940AA and I-EEE938939940AAA showed a loss of p32 and p14 (Fig. 2, lanes 13 and 14). In contrast, the translation products from I-E940A were similar to those of I-Pro^{w} (Fig. 2, lanes 12, 10 and 20), indicating that the newly created $^{939}EA^{940}$ was utilized as the alternative cleavage site. These results suggested that the cleavage site between p32 and p14 was E^{940}/A^{941}; however, this interpretation of the results was inconclusive and, therefore, N-terminal amino acid sequencing analysis was carried out. pGEX 4T-1-Δp32-p14-p70/E1055A, a plasmid encoding Δp32–p14–p70 (aa 926–1720) with a mutation at $^{1055}E/A^{1056}$ in the protease, was expressed as an N-terminal GST and C-terminal 6 × His tag fusion protein in *E. coli* and used to analyse the N terminus of p14, because our previous study indicated that the cleavage between p14 and p70 occurred efficiently when p14–p70 (aa 941–1720) was expressed as an N-terminal GST and C-terminal 6 × His tag fusion recombinant protein in *E. coli* (Oka *et al.*, 2005a). Three major products of approximately

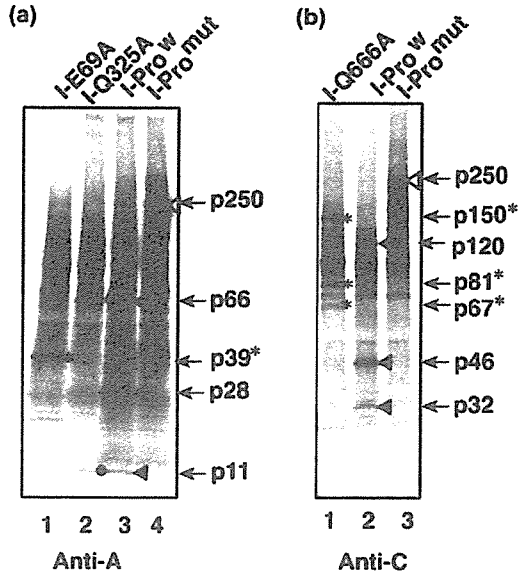

Fig. 3. Identification of the cleavage sites of p11/p28, p28/p35 and p35/p32. (a) Immunoprecipitation of the *in vitro* translation products of I-E69A, I-Q325A, I-Pro^{w} and I-Pro^{mut} with anti-A antibodies. The immunoprecipitated products in I-Pro^{w} and I-Pro^{mut} are indicated by filled and open arrowheads, respectively. The immunoprecipitated products of I-Q325A, p66 and p11, are indicated by dots and the newly appearing product of I-E69A, p39, is indicated by an asterisk. Mc10 ORF1-specific proteins are shown on the right. (b) Immunoprecipitation of the *in vitro* translation products of I-Q666A, I-Pro^{w} and I-Pro^{mut} with anti-C antibodies. The immunoprecipitated products in I-Pro^{w} and I-Pro^{mut} are indicated by filled and open arrowheads. The newly appearing products, p150, p81 and p67, in I-Q666A are indicated by asterisks. Mc10 ORF1-specific proteins are shown on the right and newly appearing products are indicated by asterisks.

110, 84 and 26 kDa were visualized when the total lysate was analysed by SDS-PAGE (Fig. 4a, lane 1). These products were considered to be GST–Δp32–p14–p70–6 × His, p14–p70–6 × His and GST–Δp32, respectively, on the basis of their molecular sizes and their affinity to TALON resin (data not shown). N-terminal amino acid sequencing of the purified 84 kDa protein revealed the sequence AKGKT, which corresponds to aa 941–945 of the Mc10 ORF1 polyprotein. pGEX-4T-1-Δp32-p14-p70/E1055A/C1171A produced a major product of 110 kDa (Fig. 4a, lane 2), demonstrating clearly that the proteolytic processing was dependent on the 3C-like protease, as described previously (Oka *et al.*, 2005a, b). We therefore concluded that the cleavage site between p32 and p14 is E^{940}/A^{941}.

Cleavage site between p14 and p70

Our previous N-terminal amino acid sequencing demonstrated that the cleavage site between p14 and p70 is E^{1055}/A^{1056} (Table 1; Fig. 1) (Oka *et al.*, 2005a). This conclusion was confirmed by site-directed mutagenesis. The translation products from I-E1055A and II-E1055A prevented proteolytic cleavage between p14 and p70 (Fig. 2, lane 15; Fig. 4b, lane 1). The constructs I- and II-EE10541055AA, which contain a double mutation, had a cleavage pattern identical to that of the constructs I- and II-E1055A (Fig. 2, lanes 16 and 17; Fig. 4b, lanes 1 and 2), indicating that no alternative cleavage occurred between E^{1054} and A^{1055}.

Cleavage site between p70 and p60

Three potential cleavage sites, E^{1679}/G^{1680}, Q^{1690}/A^{1691} and E^{1722}/G^{1723}, were predicted between p70 and p60 (Table 1 and Fig. 1). The cleavage products from I-E1679A and I-Q1690A were identical to those of I-Pro^{w} (Fig. 2, lanes 17 and 18), indicating that E^{1679}/G^{1680} and Q^{1690}/A^{1691} are not cleavage sites between p70 and p60. In contrast, the cleavage products from the I-E1722A construct showed a loss of p120 and p60 and an accumulation of p180 (Fig. 2, lane 19). Antibodies raised against the H fragment (aa 1951–2278) (Oka *et al.*, 2005b) immunoprecipitated p180 from the I-E1722A construct (data not shown), demonstrating that p180 corresponds to p120–p60. To further confirm the cleavage between E^{1722} and G^{1723}, III-Pro^{w} was expressed (Fig. 1). We observed two proteins, p70 and p60, as described previously (Fig. 4c, lane 4) (Oka *et al.*, 2005b). In contrast, III-Pro^{mut} produced a single major band, p130 (Fig. 4c, lane 5). III-E1722A produced a p130 band, as did III-Pro^{mut} (Fig. 4c, lanes 3 and 5). In contrast, the cleavage patterns from III-E1679A and III-Q1690A were identical to those of III-Pro^{w} (Fig. 4c, lanes 1, 2 and 4). Therefore, we concluded that the cleavage site between p70 and p60 is E^{1722}/G^{1723}.

Cleavage sites of the ORF1 polyprotein

The cleavage sites of the SaV GII Mc10 ORF1 polyprotein were identified as E^{69}/G^{70}, Q^{325}/G^{326}, Q^{666}/G^{667}, E^{940}/A^{941}, E^{1055}/A^{1056} and E^{1722}/G^{1723} (Table 2). Although seven full-length SaV genome sequences, including Mc10, were available at the beginning of this study, we had recently determined nine human full-length SaV genome sequences: the GI Mc114, N21, Nongkhai50, Chantaburi74, GIIMc2, SK15, GIVEhime1107, Sw278 and GV NK24 strains. The cleavage sites in ORF1 were highly conserved among these 16 SaV strains and either E or Q was found at the P1 position, whereas A, G or S was found at the P1′ position (Table 3). The dipeptide sequences at the cleavage sites were similar to those of other caliciviruses, namely, E or Q at the P1 position and A, G, S, T, D or N at the P1′ position (Table 2).

DISCUSSION

The cleavage sites of SaV ORF1 polyprotein were identified completely by using the GII Mc10 strain, which confirmed our previous predictions (Oka *et al.*, 2005a). Multiple alignment of 16 SaV ORF1 amino acid sequences revealed that the amino acids for the cleavage sites (P1/P1′) are either E or Q at the P1 position and G, A or S at the P1′ position