Virology **278,** 490–500 (2000)
doi:10.1006/viro.2000.0672, available online at http://www.idealibrary.com on IDEAL®

metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Complete Nucleotide Sequence of the Chiba Virus Genome and Functional Expression of the 3C-Like Protease in *Escherichia coli*

Yuichi Someya,[1] Naokazu Takeda, and Tatsuo Miyamura

Department of Virology II, National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku, Tokyo 162-8640, Japan

Received July 5, 2000; returned to author for revision September 26, 2000; accepted September 27, 2000; published online November 22, 2000

We cloned the genome RNA of the Chiba virus (ChV; Hu/NLV/Chiba 407/1987/JP) and determined its complete nucleotide sequence. The genome is predicted to be a positive-sense, single-stranded RNA of 7697 bases, excluding a poly(A) tract. Comparison of the nucleotide and amino acid sequences with those of other members of the species *Norwalk virus (NV)* revealed that ChV belongs to genogroup I *NV.* The ChV genome contains three open reading frames (ORFs). A large 5′-terminal ORF (ORF1) encodes a polyprotein with 1785 amino acids that are likely processed into functional proteins, including RNA helicase, VPg, protease, and RNA-dependent RNA polymerase. ORF2 encodes the capsid protein with 544 amino acids, and a small 3′-terminal ORF (ORF3) encodes a basic protein with 208 amino acids. The amino acid sequences of five cleavage sites in ORF1 are highly conserved compared with those of other members of *NV.* When expressed in *Escherichia coli,* the glutathione-*S*-transferase (GST) fusion protein of the ChV protease connected via a short peptide containing a human rhinovirus 3C protease cleavage site was cleaved into GST and the protease; however, this cleavage did not occur when the Cys mutation was introduced into the putative active site of the protease. Moreover, the ChV protease recognized and cleaved the predicted proteolytic sites between VPg and protease and between protease and RNA polymerase. Therefore, the ChV protease expressed in *E. coli* retained an enzymatic activity and a substrate specificity similar to that of the human rhinovirus 3C protease. © 2000 Academic Press

Key Words: Norwalk virus; complete genome sequence; recombinant protein; 3C-like protease; expression in *E. coli.*

INTRODUCTION

Norwalk virus (NV) is the causative agent of nonbacterial acute gastroenteritis that has occurred in various epidemiological settings, including nursing homes (Jiang *et al.,* 1996; Vinje and Koopmans, 1996; Vinje *et al.,* 1997), schools (Kobayashi *et al.,* 1991), hospitals (Vinje *et al.,* 1997; Stevenson *et al.,* 1994), restaurants (Parashar *et al.,* 1998; Fleissner *et al.,* 1989), and cruise ships (Herwaldt *et al.,* 1994; Khan *et al.,* 1994; Ho *et al.,* 1989; Gunn *et al.,* 1980). Transmission of *NV* was carried out by contaminated food, especially oysters (Kohn *et al.,* 1995; Dowell *et al.,* 1995) and water (Beller *et al.,* 1997; Gray *et al.,* 1997), and by person-to-person contact (Kaplan *et al.,* 1982). *NV* was recently assigned as a new species of a genus with the temporary name "Norwalk-like viruses" in the family *Caliciviridae. NV* is usually isolated from stool specimens of the patient with diarrheal illness and is often discovered in bivalves such as oysters and in half-cooked foods (Dowell *et al.,* 1995; Le Guyader *et al.,* 1996). *NV* forms a group of genetically and antigenically diverse caliciviruses. *NV* has been observed as small, round particles with a diameter of 38 nm on electron microscopy, and the term used to describe *NV* was "small round structured viruses" (SRSVs) (Appleton, 1987; Iversen *et al.,* 1987; Brugha *et al.,* 1999).

Since the genome of Norwalk virus (NV/68), a prototype of *NV,* was sequenced in 1993 for the first time (Jiang *et al.,* 1990, 1993; Hardy and Estes, 1996; M87661), determination of the complete nucleotide sequence of *NV* has been carried out in Southampton virus (SAV) (Lambden *et al.,* 1993; L07418), BS5 strain (unpublished, AF093797), Lordsdale virus (LV) (Dingle *et al.,* 1995; X86557), Camberwell virus (CWV) (Seah *et al.,* 1999; AF145896), and Hawaii virus (HWV) (U07611). In addition, a number of partial genomic sequences derived from the RNA-dependent RNA polymerase and capsid protein region of *NV* have been cloned, sequenced, and deposited in databases. Members of *NV* share common features in that their genomes are a positive-sense, single-stranded RNA molecule to which a putative genome-encoded protein called VPg is linked at the 5′ end of the genome (Estes *et al.,* 1997; Clarke *et al.,* 1998). Extensive genetic analyses of *NV* demonstrated that *NV* is further classified into two subgroups, genogroup I (GI), which includes NV/68, SAV, and BS5, and genogroup II (GII), which includes LV, CWV, and HWV (Wang *et al.,* 1994; Green *et al.,* 1994; Lew *et al.,* 1994; Liu *et al.,* 1995). The phylogenetic tree of published *NV* sequences based on part of the RNA-dependent RNA polymerase gene demonstrated that the GI and GII subgroups are composed of at least

[1] To whom correspondence should be addressed. Fax: 81-3-5285-1161. E-mail: someya@nih.go.jp.

0042-6822/00 $35.00
Copyright © 2000 by Academic Press
All rights of reproduction in any form reserved.

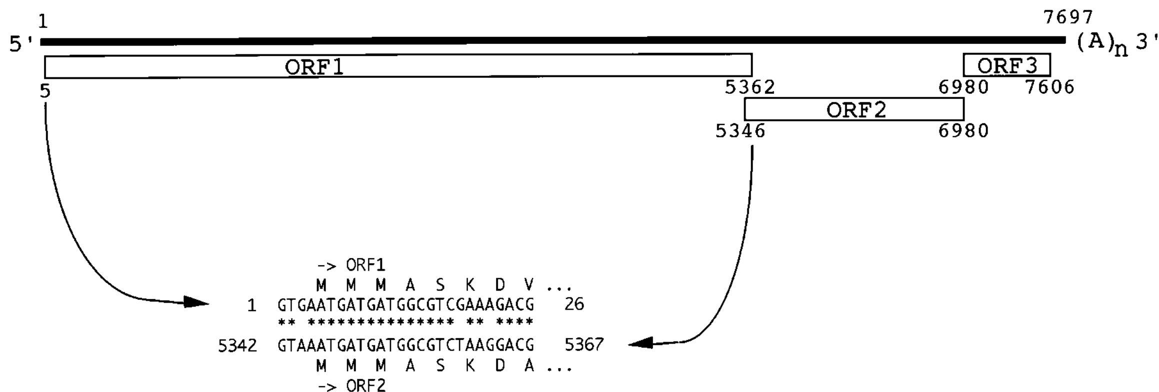

FIG. 1. Genomic organization of ChV. The ChV genome is composed of 7697 bases, excluding the 3′ poly(A) tail. Three ORFs are predicted. ORF1 encodes a polyprotein, ORF2 encodes the capsid protein, and ORF3 encodes a small basic protein with unknown function. The first 26 nucleotides of the genome are aligned with the sequence around the start codon of ORF2, showing 88% identity.

seven and five genetically distinct viruses, respectively (Vinje *et al.,* 2000; Green *et al.,* 2000). Recently, the Jena strain of bovine calicivirus was cloned, and the nucleotide sequence indicated that it is closely related to GI *NV* (Liu *et al.,* 1999a; AJ011099).

The genome of *NV* contains three open reading frames (ORFs). ORF1 encodes a large polyprotein containing amino acid sequence motifs observed in many RNA viruses, such as RNA helicase, 3C-like protease and RNA-dependent RNA polymerase, as well as VPg, originally described in a poliovirus (Lee *et al.,* 1977; Rueckert and Wimmer, 1984). Protease activity is believed to be required for the maturation of functional proteins encoded by ORF1. ORF2 is predicted to encode the capsid protein. ORF3 encodes a small protein abundant in basic amino acids. Although the precise role of the ORF3 protein in the virus replication is unknown, it is likely that ORF3 protein is a minor structural protein that interacts with the genome RNA when the virion formation occurs (Dingle *et al.,* 1995; Jiang *et al.,* 1993; Lambden *et al.,* 1993; Seah *et al.,* 1999; Glass *et al.,* 2000).

Previous studies showed that *in vitro* transcription-translation of a full-length clone of SAV with a T7 polymerase-coupled reticulocyte lysate system yielded a 200-kDa precursor protein corresponding to the ORF1 polyprotein. This large protein was cotranslationally cleaved into three major products with 113, 48, and 41 kDa (Liu *et al.,* 1996). A 3C-like protease encoded in the 113-kDa protein is likely to function in the cleavage of the N-terminal 48-kDa protein and 41-kDa putative 2C-like protein. A recent study demonstrated that the nonstructural precursor polyprotein of SAV is cleaved into at least six smaller products (Liu *et al.,* 1999b).

In this paper, we describe the isolation of the 5′ untranslated region (UTR) by 5′-rapid amplification of cDNA ends (5′-RACE) and amplification of the ORF1-coding region of Chiba virus (ChV) by reverse transcription–polymerase chain reaction (RT–PCR) to complete the entire nucleotide sequence. Based on the deduced amino acid sequence of ORF1 polyprotein, the ChV protease was expressed in *Escherichia coli.* Bacterially expressed protease showed enzymatic activity with specificity similar to that of the human rhinovirus 3C protease.

RESULTS AND DISCUSSION

Sequence analyses of the ChV genome

ChV was first detected in a stool specimen collected from a patient with gastroenteritis in an oyster-associated outbreak that occurred in December 1987 in Chiba Prefecture, Japan (Kasuga *et al.,* 1990). The 3′ half of the genome encoding ORF2, ORF3, and 3′ UTR has been cloned, and the nucleotide sequence has been determined (Utagawa *et al.,* 1994). In this experiment, we obtained 25 different clones covering the entire genome using RT–PCR, as described in Materials and Methods. Moreover, to confirm the 5′ end sequence of the genome, 5′-RACE was also performed using the genomic RNA. The determination of the ORF1 sequence allowed us to complete the entire genome of 7697 bases, excluding the 3′ poly(A) tail. The length and GC content (47.4%) of the genome were comparable with those of other GI *NV*. The ChV entire genome has an approximately 70% nucleotide identity with those of NV/68, SAV, and BS5, whereas it has about 50% identity with those of LV, CWV, and HWV.

The overall genomic organization of ChV is the same as those of known *NV* (Fig. 1), in which three ORFs were identified. The translation of ChV ORF1 was predicted to initiate at nucleotide position 5, the same position as in NV/68, SAV, BS5, LV, CWV, and HWV (Hardy and Estes, 1996; Lambden *et al.,* 1995 and 1994; Seah *et al.,* 1999). ChV ORF1 encodes a nonstructural polyprotein composed of 1785 amino acid residues that is highly homologous to NV/68, SAV, and BS5, with about 85% amino acid identities. Several amino acid sequence motifs are conserved in ChV ORF1, as described later. In addition, the first 26 nucleotides at the 5′ terminus of the genome

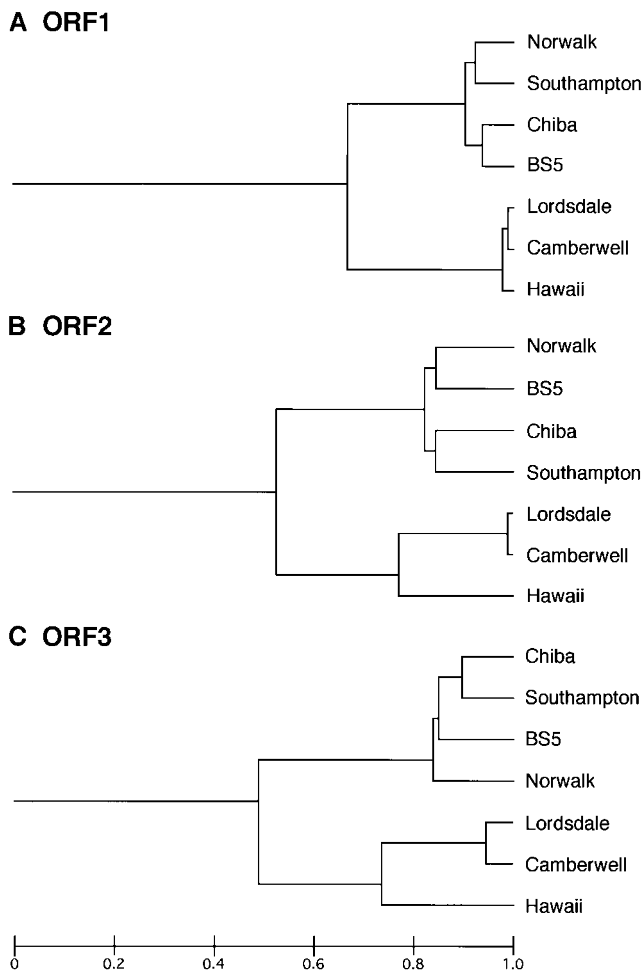

FIG. 2. Phylogenetic trees for ORF1 (A), ORF2 (B), and ORF3 (C) from *NV* whose genomes are thoroughly sequenced. Analyses are based on the UPGMA method (Sneath and Sokal, 1974). The branch length indicates relative evolutionary distances. The relative alignment score is indicated at the bottom. Norwalk virus, M87661; Southampton virus, L07418; BS5 strain, AF093797; Lordsdale virus, X86557; Camberwell virus, AF145896; Hawaii virus, U07611.

were very similar to those of the 5′ end of ORF2, with 23 of the nucleotides being identical, a common feature observed in known *NV* (Hardy and Estes, 1996; Lambden *et al.,* 1995 and 1994; Seah *et al.,* 1999). When N-terminal amino acid sequences of ORF1 and ORF2 of *NV* were compared, we found that both ORF1 and ORF2 of GI *NV* started with MMMASKD, whereas those of GII *NV* started with MKMASN(IS)D.

Phylogenetic trees generated for each ORF1, ORF2, and ORF3 are depicted in Fig. 2. ChV was most closely related to BS5 when the amino acid sequence of ORF1 was compared. The same results were essentially obtained when ORF1 was separated into six putative functional parts. However, ORF2 and ORF3 of ChV were most closely related to those of SAV. Although there is no direct evidence, this suggests that genetic recombination might occur among different *NV,* conferring genetic diversity on *NV* as described previously (Jiang *et al.,* 1999; Vinje *et al.,* 2000).

Putative cleavage sites of ORF1 proteins

Five proteolytic cleavage sites have been identified in the SAV ORF1 polyprotein (Liu *et al.,* 1999b), and the boundary amino acid sequences of the cleavage sites were highly conserved among the members of *NV* when seven *NV* ORF1 sequences were aligned. As shown in SAV, the protease encoded in ORF1 may be involved in the cleavage of the polyprotein at Q/G in the N-terminal protein/helicase and helicase/3A, at E/G in the 3A/VPg and protease/polymerase, and at E/A in the VPg/protease boundaries. Therefore, ORF1 polyprotein of ChV is also expected to be cleaved into six portions at the conserved cleavage sites (Fig. 3A).

The first N-terminal protein was composed of 396 amino acid residues with an estimated molecular mass of 44.6 kDa. This protein is in the equivalent region corresponding to 2B or 2AB of the picornavirus. Although it has 68–80% amino acid identity with other GI *NV,* its function is unknown. The second polypeptide is 2C-like RNA helicase, which is composed of 363 amino acids with an estimated molecular mass of 39.6 kDa. The nucleotide-binding motif GPPGIGKT is found at position 162. The ChV helicase has about 90% amino acid identity with other GI *NV,* having the highest identity with the BS5 strain (94%). The third protein, corresponding to 3A-like protein, is composed of 201 amino acids with an estimated molecular mass of 21.3 kDa and has 67–77% identity with other GI *NV.* The fourth protein, VPg, is believed to bind to the 5′ terminus of the genomic RNA and be highly homologous to other GI strains (82–89% identity). ChV VPg is rich in basic amino acids and migrated as a 20-kDa protein in SDS–PAGE, although it is composed of 136 amino acid residues with an estimated molecular mass of 16.0 kDa (see Fig. 5).

The fifth protein, corresponding to 3C-like protease, is composed of 181 amino acid residues (19.4 kDa). The ChV protease is a serine-like protease in which the 139th cysteine residue within the GDCG motif is predicted to be the active site. The 30th histidine and 54th glutamate are conserved in all *NV* and probably form a catalytic triad. The 157th histidine is also conserved and is thought to be a member of the substrate-binding pocket (Boniotti *et al.,* 1994). The amino acid identity is very high (more than 90%), as high as 95% with the BS5 strain. The C-terminal polypeptide is the 3D RNA-dependent RNA polymerase, containing 508 amino acid residues with an estimated molecular mass of 56.9 kDa. This protein is highly homologous to other GI *NV* (89–93% amino acid identity), and highly conserved GLPSG and YGDD motifs, which are believed to play an important role in the polymerase activity, are found in all *NV.* An endoplasmic reticulum-targeting signal, KDEL, is also conserved in

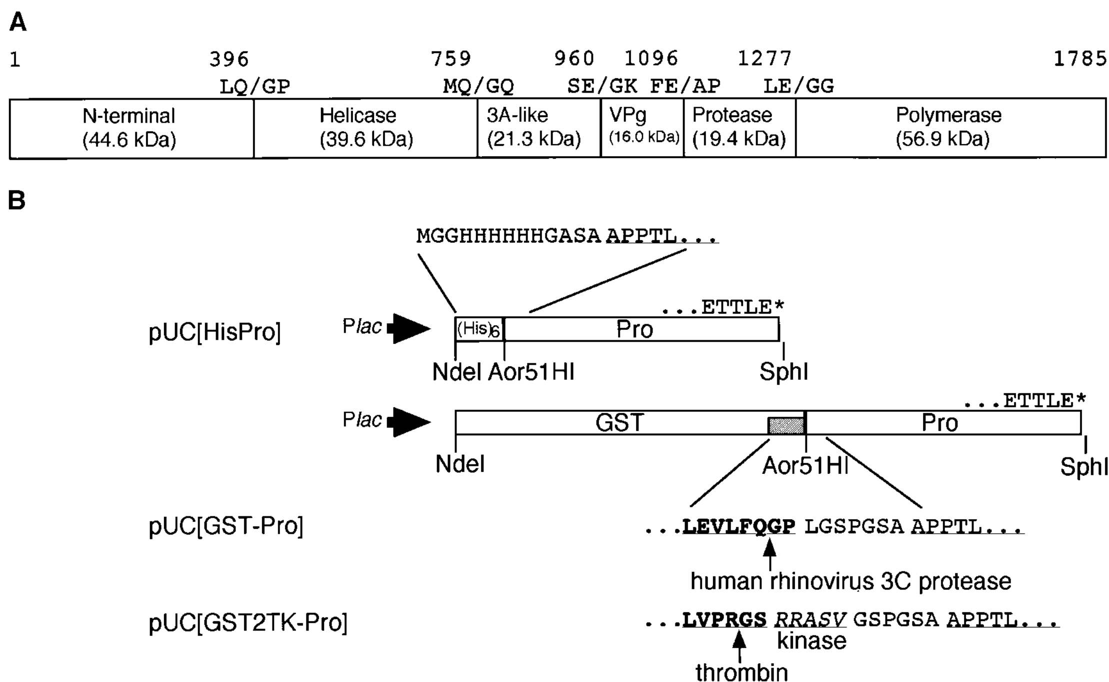

FIG. 3. Schematic representation of the ChV ORF1 polyprotein and construction of the expression plasmids. (A) Map of the proteolytic cleavage sites. Cleavage sites are predicted on the basis of studies of the Southampton virus (Liu *et al.,* 1996 and 1999). The numbers indicate the positions of the amino acids. Slashes indicate the cleavage sites recognized by the protease. The calculated molecular masses of the cleavage products are indicated. (B) Construction of the expression plasmids encoding the ChV protease. An *Aor*51HI site is introduced before the start (APPTL) of the protease, and the stop codon and an *Sph*I site are introduced after the end (ETTLE) of the protease. pUC[His-Pro] encodes the His-tagged protease. pUC[GST-Pro] and pUC[GST2TK-Pro] encode the protease fused to GST derived from pGEX-6P-2 and pGEX-2TK, respectively. The former includes the sequence cleaved by the human rhinovirus 3C protease, which is possibly cleaved by the ChV protease. The latter includes the thrombin site in place of the human rhinovirus 3C protease site. The fusion genes are controlled by the *lac* promoter (P*lac*).

the RNA polymerase region, although direct evidence of the sorting has not been shown.

Functional expression of the ChV 3C-like protease in *E. coli*

In order to characterize the ChV protease, we expressed the protein in *E. coli* using two types of plasmids. The first, pUC[His-Pro], was generated for the production of the ChV protease flanked by a histidine tag (6xHis) at the N terminus, and the second, pUC[GST-Pro], was generated for the production of a GST-protease fusion protein. In both expression vectors, the fusion genes were placed under the control of the *lac* promoter (Fig. 3B).

E. coli BL21-CodonPlus-RIL was transformed with the vector plasmid and grown in a rich medium. The expression was induced with isopropyl-β-D-thiogalactopyranoside (IPTG) for 2 h, followed by disruption of the cells by sonication. After centrifugation, the supernatant was collected, and the proteins were separated in SDS–PAGE and visualized with Coomassie brilliant blue (CBB) R-250 (Fig. 4A). When the lysate from the pUC[His-Pro] was analyzed in the gel, we observed a 23-kDa band (Fig. 4A, lane 3). Although this molecular mass was slightly higher than that calculated from the recombinant ChV protease containing a 6xHis sequence (20.7 kDa), this band was detected by anti-His antibody (Fig. 4B, lane 3) and absent in an *E. coli* extract without the vector plasmid (Fig. 4A, lane 2). Therefore, we concluded that this 23-kDa band was the protease containing the 6xHis sequence. The 23-kDa protein was purified and used to inject rabbits to elicit the antibody against the protease. This anti-Pro antibody reacted with His-Pro protein (Fig. 4D, lane 3). Because no other His-tagged proteins were detected by this anti-Pro antibody (data not shown), we concluded that the antibody reacted with the protease moiety.

When the expression was done in *E. coli* transformed with pUC[GST-Pro], two distinct bands with 26 and 22 kDa were visualized (Fig. 4A, lane 5). If an intact GST-protease fusion protein was generated, a band with 46.5 kDa was expected. The GST gene was derived from pGEX-6P-2 vector, and this GST gene contains an octapeptide (LEVLFQGP) at the C terminus, in which Q/G is the cleavage site recognized by the human rhinovirus 3C protease. When the specificity of the 3C-like proteases of SAV is considered (Liu *et al.,* 1996 and 1999b), it is likely that the Q/G site between the GST and the ChV protease was recognized by the protease and that the 46.5-kDa protein is cleaved into two proteins. Because the 26-kDa protein has mobility similar to that of a control GST-6xHis fusion protein (Fig. 4A, lanes 5 and 9) and has immuno-

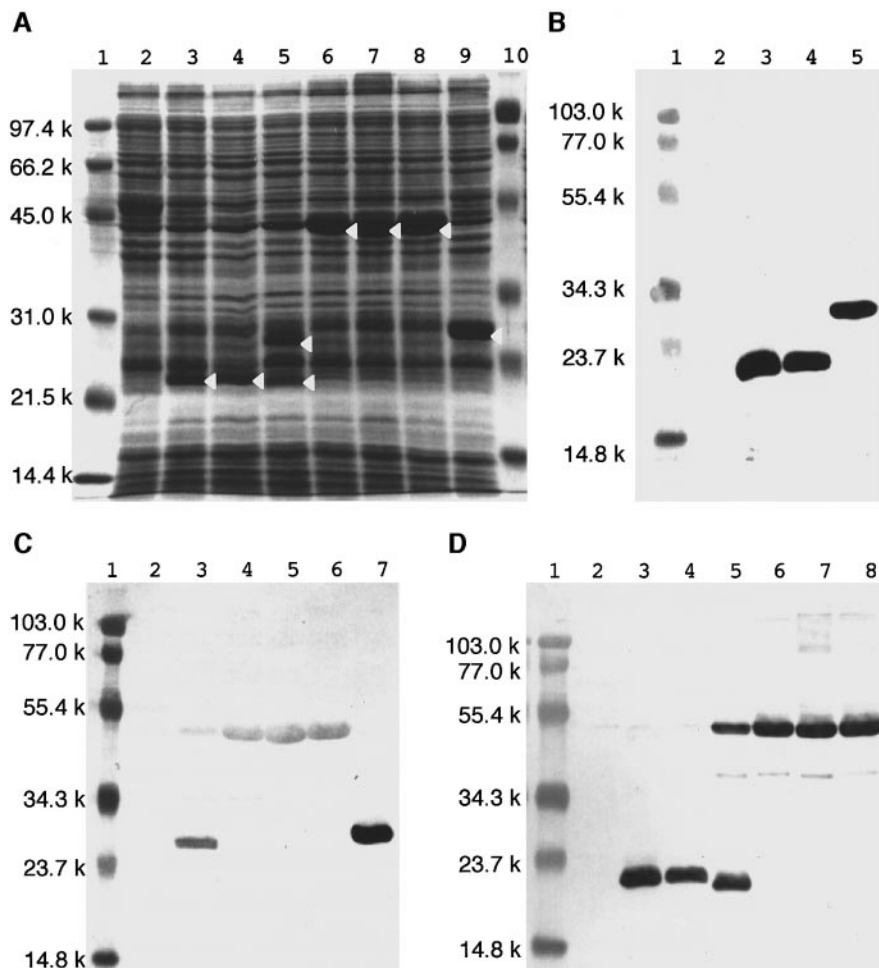

FIG. 4. Expression of the fusion proteins of the ChV protease. Expression plasmids were used to transform *E. coli* BL21-CodonPlus-RIL cells. The gene expression was induced with IPTG. The cell lysates were subjected to SDS–PAGE. (A) CBB staining of the gel. Lanes 1 and 10, molecular weight marker; lane 2, lysates from BL21-CodonPlus-RIL cells without plasmids; lane 3, with pUC[His-Pro]; lane 4, with pUC[His-ProC1235A]; lane 5, with pUC[GST-Pro]; lane 6, with pUC[GST2TK-Pro]; lane 7, with pUC[GST-ProC1235A]; lane 8, with pUC[GST2TK-ProC1235A]; lane 9, with pUC[GST-His]. (B) Western blot analysis with anti-His monoclonal antibody. Lane 1, molecular weight marker; lane 2, without plasmids; lane 3, with pUC[His-Pro]; lane 4, with pUC[His-ProC1235A]; lane 5, with pUC[GST-His]. (C) Western blot analysis with anti-GST antibody. Lane 1, molecular weight marker; lane 2, without plasmids; lane 3, with pUC[GST-Pro]; lane 4, with pUC[GST2TK-Pro]; lane 5, with pUC[GST-ProC1235A]; lane 6, with pUC[GST2TK-ProC1235A]; lane 7, with pUC[GST-His]. (D) Western blot analysis with anti-Pro antibody raised against His-Pro proteins. Lane 1, molecular weight marker; lane 2, without plasmids; lane 3, with pUC[His-Pro]; lane 4, with pUC[His-ProC1235A]; lane 5, with pUC[GST-Pro]; lane 6, with pUC[GST2TK-Pro]; lane 7, with pUC[GST-ProC1235A]; lane 8, with pUC[GST2TK-ProC1235A].

reactivity to anti-GST antibody (Fig. 4C, lanes 3 and 7), this 26-kDa protein should be the GST protein. Therefore, the 22-kDa band represents the ChV protease, because it has a size similar to that containing the 6xHis sequence as described earlier (23 kDa) and it has immunoreactivity to anti-Pro antibody (Fig. 4D, lane 5). These results indicate that ChV protease expressed in *E. coli* is enzymatically active and capable of cleaving the specific site recognized by the human rhinovirus 3C protease.

Identification of the functional element of the ChV protease

As shown previously (Liu *et al.,* 1996 and 1999b), 3C-like protease of SAV is thought to be a serine-like protease, and the active site is predicted to be a cysteine residue within the GDCG motif conserved in the proteases. In the ChV genome, it corresponds to the 1235th cysteine residue in the precursor polyprotein (Cys1235). Therefore, this amino acid was replaced with alanine by using an oligonucleotide-directed site-specific mutagenesis, as described under Materials and Methods. The mutated protease gene was introduced into an expression vector in the same manner as pUC[His-Pro] and pUC[GST-Pro] to generate pUC[His-ProC1235A] and pUC[GST-Pro C1235A], respectively. These plasmids were introduced into *E. coli* and were subjected to the expression. In the extract of *E. coli* harboring pUC[His-ProC1235A], a 23-kDa mutant protease band was de-

tected by CBB staining of the gel (Fig. 4A, lane 4) and Western blot analysis with anti-His antibody (Fig. 4B, lane 4) and anti-Pro antibody (Fig. 4D, lane 4). The size of the band in the gel is identical to that of the wild-type protease (Figs. 4A, lane 3, 4B, lane 3, and 4D, lane 3). In the extract of *E. coli* harboring pUC[GST-ProC1235A], a band with an apparent mass of 44 kDa was visualized (Fig. 4A, lane 7), and Western blot analysis indicated that this band was reactive to anti-GST antibody (Fig. 4C, lane 5) and anti-Pro antibody (Fig. 4D, lane 7). Although the fusion protein contains a sequence recognized by the human rhinovirus 3C protease, it was not cleaved by the mutant protease at all, demonstrating that the C1235A mutation caused a complete loss of cleaving activity. The results confirmed the importance of Cys1235 residue as a functional element.

To further confirm the specificity of the cleavage, we constructed an expression vector containing the active ChV protease designated pUC[GST2TK-Pro]. This vector contains a GST gene derived from an expression vector pGEX-2TK, which is flanked by a thrombin-specific cleavage sequence (LVPR/GS) and a kinase-specific sequence (RRASV) at the C terminus. However, this plasmid lacks the human rhinovirus 3C protease recognition sequence. When the extract from *E. coli* cells harboring this plasmid was analyzed by SDS–PAGE, a fusion protein with an apparent molecular mass of 44 kDa was visualized, a size close to the calculated mass of 46.7 kDa (Fig. 4A, lane 6). This protein was immunoreactive to anti-GST antibody (Fig. 4C, lane 4) and anti-Pro antibody (Fig. 4D, lane 6). The results indicate that the fusion protein was resistant to digestion by ChV 3C-like protease consistent with the lack of a cleavage site. A plasmid containing the C1235A mutant protease gene designated pUC[GST2TK-ProC1235A] resulted in the production of a 44-kDa protein, as with the case of pUC[GST2TK-Pro] (Figs. 4A, lane 8, 4C, lane 6, and 4D, lane 8).

Cleavage at the predicted proteolytic sites by the ChV protease

We were most interested to see whether the ChV protease could cleave the predicted proteolytic sites within the ChV ORF1 polyprotein. We therefore constructed two plasmids, pUC[His-3BCD] and pUC[His-3CD], encoding the VPg (3B)-protease (3C)-RNA polymerase (3D) region and the protease-RNA polymerase region, respectively, and tested whether the sequences between VPg and protease and between protease and RNA polymerase were cleaved by the ChV protease in *E. coli.* When the lysate of *E. coli* cells with pUC[His-VPg] was examined by Western blot analysis, a 23-kDa band was detected with both anti-His antibody (Fig. 5A, lane 3) and anti-VPg antibody (Fig. 5B, lane 3). Although the calculated mass of His-VPg protein is 17.5 kDa, mobility in SDS–PAGE was much slower. This is probably because VPg protein is rich in basic amino acid residues. In the lysate of *E. coli* cells carrying pUC[Pol-His], we detected a 58-kDa band that was reactive with both anti-His antibody (Fig. 5A, lane 7) and anti-Pol peptide antibody (Fig. 5D, lane 3). These results indicated that antibodies raised against VPg and RNA polymerase were highly specific to the respective proteins.

Next, using these antibodies, we examined the expression of His-3BCD and His-3CD proteins. If the His-3BCD protein (93.7 kDa) is cleaved at the predicted proteolytic sites, a His-VPg with 17.5 kDa, a protease with 19.4 kDa, and an RNA polymerase with 56.9 kDa are expected to be produced. When the lysate of *E. coli* cells with pUC[His-3BCD] was separated on SDS–PAGE, a 23-kDa band was detected with both anti-His antibody and anti-VPg antibody (Figs. 5A, lane 4, and 5B, lane 4). This 23-kDa band had the same mobility as the control His-VPg protein, suggesting that the cleavage occurred at the predicted site (FE/AP) between VPg and protease. Two faint bands of 94 and 40 kDa, corresponding to His-3BCD and His-3BC, respectively, were also detected with anti-VPg antibody (Fig. 5B, lane 4). Anti-Pro antibody detected a 20-kDa band corresponding to protease and a 77-kDa band corresponding to 3CD (Fig. 5C, lane 4), and anti-Pol peptide antibody detected a 57-kDa band corresponding to RNA polymerase (Fig. 5D, lane 4). These results suggested that the cleavage between protease and RNA polymerase also occurred at the predicted site (LE/GG).

The 20-kDa band detected with anti-Pro antibody (Fig. 5C, lane 4) was subjected to the N-terminal amino acid sequencing that produced the sequence APITLASRVVRFGSG. The amino acid sequence deduced from the nucleotide sequence is APPTLWSRVVRFGSG. Although the third and sixth amino acid residues were inconsistent, it may be sequencing error due to low yields of sequencing reactions at those positions, because nucleotide sequencing of pUC[His-3BCD] indicated P (proline) at the third position and W (tryptophan) at the sixth position. The data indicated that the cleavage between VPg and protease occurred exactly at the predicted site (FE/AP) and that the ChV protease expressed in *E. coli* retained substrate specificity.

The His-3CD protein is expected to be cleaved into His-Pro with 20.7 kDa and an RNA polymerase with 56.9 kDa. When the lysate of *E. coli* cells harboring pUC[His-3CD] was examined by Western blot analysis, anti-His antibody and anti-Pro antibody detected a 77-kDa band corresponding to His-3CD, as well as a 23-kDa His-Pro band corresponding to His-Pro proteins (Figs. 5A, lane 6, and 5C, lane 5). Much more of the 77-kDa protein remained uncleaved, compared with the case of the 94-kDa His-3BCD protein (see Fig. 5B, lane 4). Anti-Pol peptide antibody detected both a

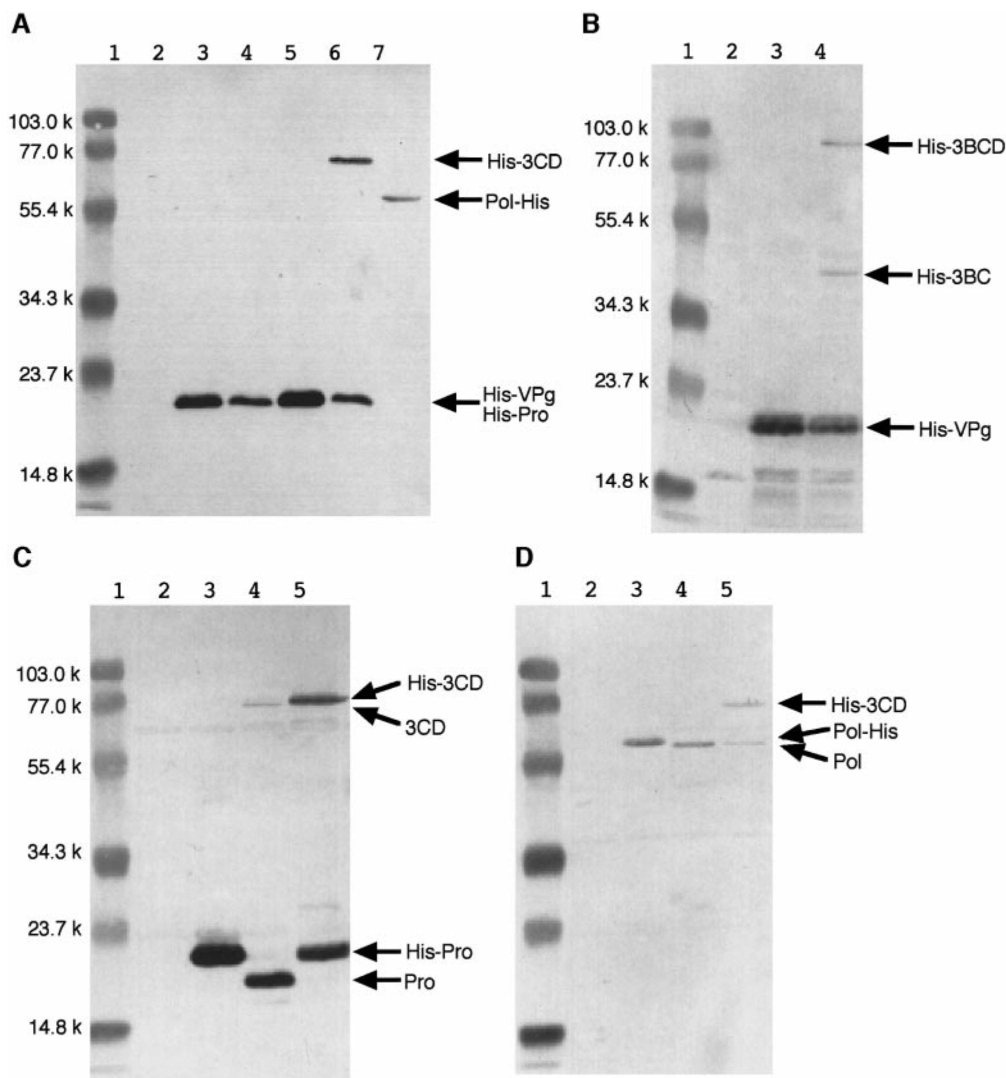

FIG. 5. Cleavage of the predicted intrinsic sequences by the ChV protease. Expression plasmids were used to transform *E. coli* BL21-CodonPlus-RIL cells. The gene expression was induced with IPTG. The cell lysates were subjected to SDS–PAGE. (A) Western blot analysis with anti-His monoclonal antibody. Lane 1, molecular weight marker; lane 2, without plasmids; lane 3, with pUC[His-VPg]; lane 4, with pUC[His-3BCD]; lane 5, with pUC[His-Pro]; lane 6, with pUC[His-3CD]; lane 7, with pUC[Pol-His]. (B) Western blot analysis with anti-VPg antibody raised against His-VPg proteins. Lane 1, molecular weight marker; lane 2, without plasmids; lane 3, with pUC[His-VPg]; lane 4, with pUC[His-3BCD]. (C) Western blot analysis with anti-Pro antibody. Lane 1, molecular weight marker; lane 2, without plasmids; lane 3, with pUC[His-Pro]; lane 4, with pUC[His-3BCD]; lane 5, with pUC[His-3CD]. (D) Western blot analysis with anti-Pol antibody raised against KLH-conjugate of Pol-peptide. Lane 1, molecular weight marker; lane 2, without plasmids; lane 3, with pUC[Pol-His]; lane 4, with pUC[His-3BCD]; lane 5, with pUC[His-3CD].

57-kDa RNA polymerase and a 77-kDa uncut His-3CD (Fig. 5D, lane 5). It is likely that the activity of the protease from His-3CD was lower than that from His-3BCD. An unnatural His-tag sequence introduced prior to protease moiety may affect proteolytic activity. In any case, it was clearly demonstrated that the ChV protease recognized and cleaved the predicted proteolytic site between a protease and an RNA polymerase.

From these results, we concluded that the ChV 3C-like protease expressed in *E. coli* had a proteolytic activity and retained a substrate specificity, which would be exerted in mammalian cells. The molecules expressed in *E. coli* might be useful for structure–function study as well as drug design in diarrheal therapy.

Correction of the nucleotide sequence for the ChV ORF2

Although the sequence data for ORF2 and ORF3 of the ChV have been registered as AB022679 and D38547, two errors were found in ORF2. Two guanine residues at positions 941 and 947 in the sequence from AB022679 should be replaced with cytosine. As a result, amino acids at positions 314 and 316 were changed to Ala in place of Gly.

Nucleotide sequence accession number

The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL, and GenBank nucleotide sequence databases with the accession number AB042808.

MATERIALS AND METHODS

Bacterial strains

E. coli strains TG1, JM109, DH5α (TaKaRa Shuzou Co., Tokyo, Japan), and XL10-Gold (Stratagene, La Jolla, CA) were used for the genetic manipulation. *E. coli* CJ236 strain (TaKaRa Shuzou Co.) was used for the preparation of the deoxyuracil-containing single-stranded DNA. *E. coli* BL21-CodonPlus-RIL strain (Stratagene) was used for the expression of recombinant proteins.

cDNA cloning of the ChV genome

The genomic RNA of ChV was extracted from a 10% stool suspension by using TRIZOL LS (Life Technologies, Grand Island, NY) according to the manufacturer's instruction. First-strand cDNA was synthesized using the RAV-2 reverse transcriptase (TaKaRa Shuzou Co.) with random nanomer (TaKaRa Shuzou Co.) or oligo(dT)$_{18}$ primer (TaKaRa Shuzou Co.) and used as a template for the PCR. Amplification was done by using ExTaq (TaKaRa Shuzou Co.), AmpliTaq Gold (PE Applied Biosystems, Branchburg, NJ), or PfuTurbo (Stratagene) DNA polymerase with a specific primer based on the ChV ORF2, and degenerated primers based on the complete *NV* sequences appeared in the databases. The PCR products were subjected to PAGE, and the separated fragment was electrophoretically eluted from the excised gel, which was then ligated with pCR2.1 T-vector (Invitrogen, San Diego, CA) or pUC118 *Hin*cII/BAP vector (TaKaRa Shuzou Co.). The 5′ UTR of the genomic RNA was amplified with AmpliTaq DNA polymerase by 5′ RACE using cDNA as the template. The nucleotide sequence was determined with the ABI Prism 310 Genetic Analyzer (PE Applied Biosystems). Individual PCR fragments were amplified at least three times independently, and at least three clones from each PCR product were subjected to nucleotide sequence determination. The fragment, including the 5′ UTR and the entire ORF1 coding region of ChV, was reconstructed on pUC118 vector by ligating several restriction fragments and designated pUC-[CVORF1]. The ORF2, ORF3, and 3′ UTR were amplified with the same strategy, cloned into pUC118 vector, and sequenced. The full-length ChV genome was reconstructed on pUC118, which is designated pUC[CVORF1/2/3].

Construction of plasmids used for expression in *E. coli*

pUC118Nd, which contains the *Nde*I site overlapped at the initial codon of the *lacZ* gene, was constructed from pUC118 by site-directed mutagenesis. Mutagenesis was carried out according to the method described by Kunkel *et al.* (1987) using the Mutan-K mutagenesis kit (TaKaRa Shuzou Co.). The pUC118Nd was digested with *Nde*I and *Hin*dIII, and the following four 5′-phosphorylated oligonucleotides were inserted to construct pUCHisNd containing a fragment encoding a 6xHis tag: 5′-TATGGGTGGCCATCATCATCATCACCATGGGGCTA-3′, 5′-GCGCTCCCGGGAGATCTTCGAACGCGTACGCATGCA-3′, 5′-AGCTTGCATGCGTACGCGTTCGAAGATCTCCCGGG-3′, and 5′-AGCGCTAGCCCCATGGTGATGATGATGATGGCCACCCA-3′.

The gene encoding the ChV 3C-like protease was amplified by PCR using pUC[CVORF1] as a template and a set of primers. The forward primer (Pro-5A, 5′-CTCAGTAGCGCTGCCCCACCAACCCTCTGG-3′) contains an *Aor*51HI site (underlined) introduced before the GCC codon for Ala1097, and the reverse primer (Pro-3S, 5′-TTTGCATGCTTACTCTAGGGTGGTTTCACCTTC-3′) contains the stop codon and an *Sph*I site (underlined) after the GAG codon for Glu1277. The PCR product was cloned into pUC118 *Hin*cII/BAP vector and designated pUC-[CVPro-AS]. The *Aor*51HI–*Sph*I fragment encoding the protease was cut out and inserted into the equivalent position on pUCHisNd to generate pUC[His-Pro].

The gene fragment was amplified by PCR with pGEX-6P-2 (Amersham Pharmacia Biotech, Piscataway, NJ) as a template using the forward primer (GST-5N, 5′-ACAGTACATATGTCCCCTATACTAG-3′) containing an *Nde*I site overlapping at the initial codon, as indicated by underlining, and the reverse primer (GST-3A, 5′-AGTCGACCCAGCGCTTCCTGGGGATCCC-3′) containing an *Aor*51HI site, as indicated by underlining. The GST gene was also amplified by PCR with pGEX-2TK (Amersham Pharmacia Biotech) as a template using the same forward primer (GST-5N) and the reverse primer (2TK-3A, 5′-TCAGTCACGAGCGCTTCCCGGGGATCCAAC-3′) containing an *Aor*51HI site, as indicated by underlining. The PCR fragment was cloned into pUC118 *Hin*cII/BAP vector to construct pUC[GST6P-NA] and pUC[GST2TK-NA], and the nucleotide sequence was verified. The *Nde*I–*Aor*51HI fragment containing the GST gene was isolated from the pUC[GST6P-NA] or pUC[GST2TK-NA] and inserted into the *Nde*I–*Sph*I site of pUC118Nd along with the *Aor*51HI–*Sph*I fragment containing ChV 3C-like protease. The plasmid carrying the GST gene derived from pGEX-6P-2 was designated pUC[GST-Pro], and the plasmid carrying the gene from pGEX-2TK was designated pUC[GST2TK-Pro].

The mutation of Cys to Ala at position 1235 in the protease was introduced by a site-directed mutagenesis. The plasmid pUC[CVPro-AS] and a synthetic oligonucleotide (5′-CTACCAGGTGAT**GCC**GGCGCCCCTTATGTG-3′) equivalent to the nucleotide residues 3695–3724 were used as a template and a mutagenic primer, respectively. The introduction of proper mutation was monitored by

the appearance of two unique restriction enzyme sites. One is the *Ehe*I site (GGCGCC), which contains a T-to-C silent mutation at the nucleotide position 3712, and the other is the *Nae*I site (GCCGGC), which contains the Cys-to-Ala change (TGT to GCC, represented in bold letters). The substitutions were verified by DNA sequencing, and no other mutation was confirmed. The resultant expression plasmid to produce mutant proteases was designated pUC[CVPro-ASC1235A]. Plasmids pUC[His-ProC1235A], pUC[GST-ProC1235A], and pUC[GST2TK-ProC1235A] encoding the mutant ChV protease were constructed by the same strategy used to construct pUC[His-Pro], pUC[GST-Pro] and pUC[GST2TK-Pro] encoding the wild-type protease, respectively.

The genes encoding the VPg and the RNA-dependent RNA polymerase were amplified by PCR using pUC-[CVORF1] as a template. For subcloning of VPg, the forward primer (VPg-5A, 5′-ACTAGCGCTGGTAAAAACAAAGGAAAGACCAAG-3′) contains an *Aor*51HI site (underlined) introduced prior to the GGT codon for Gly961, and the reverse primer (VPg-3S, 5′-GGTGCATGCTTATTCAAAACTGAGTTTTTCATTGTA-3′) contains the stop codon and an *Sph*I site (underlined) after the GAA codon for Glu1096. The PCR product was cloned into pUC118 *Hin*cII/BAP vector and designated pUC[CVVPg-AS]. The *Aor*51HI–*Sph*I fragment encoding the VPg was cut out and inserted into the equivalent position on pUCHisNd to generate pUC[His-VPg]. For subcloning of RNA polymerase, the forward primer (Pol-5N, 5′-GGTGAAACCACCCATATGGGTGGTGATAAAGGC-3′) contains an *Nde*I site (underlined) introduced prior to the GGT codon for Gly1278, and the reverse primer (Pol-3A, 5′-GGTGGATCCAGCGCTGACGCCATCATCATTTACGAATTC-3′) contains an *Aor*51HI site (underlined) after the GTC codon for Val1785, the last amino acid of the polyprotein. The PCR product was cloned into pUC118 *Hin*cII/BAP vector and designated pUC[CVPol-NA]. The *Nde*I–*Aor*51HI fragment encoding the RNA polymerase was cut out and inserted into the *Nde*I–*Bam*HI vector fragment of pUC118Nd together with the His-tag linker (the mixture of 5′-GCTGGCCACCATCATCATCACCATGGCTAG-3′ and 5′-GATCCTAGCCATGGTGATGATGATGGTGGCCAGC-3′). The resultant plasmid was designated pUC[Pol-His].

The region encoding 3CD (protease and RNA polymerase) and 3BCD (VPg, protease and RNA polymerase) was also amplified by PCR. For amplification of the 3CD-encoding region, the Pro-5A primer and the reverse primer (Pol-3S, 5′-GTAGCATGCTTAGACGCCATCATCATTTACGAATTC-3′) containing a stop codon and an *Sph*I site after the GTC codon for Val1785 were used. For amplification of the 3BCD-encoding region, we used the VPg-5A primer and the Pol-3S primer. The *Aor*51HI–*Sph*I fragments from amplified products were inserted into the equivalent position on pUCHisNd. The resultant plasmids were designated pUC[His-3CD] and pUC[His-3BCD], respectively.

Expression of wild-type and mutant proteases in *E. coli*

E. coli BL21-CodonPlus-RIL cells were transformed with the expression plasmids. The transformant was grown in 2× YT broth, and protein expression was induced with 0.5 mM IPTG when optical density at 610 nm reached 0.4. After 2 h of incubation, cells were harvested, washed with 20 mM Tris–HCl (pH 7.4), and disrupted by sonication. The cell lysate was centrifuged to remove unbroken cells, and the supernatant was used as a sample. The protein concentration was determined by the Bradford method (1976) using bovine serum albumin as a standard. The supernatant containing 20 μg of proteins was used for SDS–PAGE, followed by staining with CBB R-250 or Western blotting. The His-tagged protein was detected by mouse anti-His monoclonal antibody (Sigma Chemical Co., St. Louis, MO) and alkaline phosphatase-linked goat anti-mouse IgG (Jackson ImmunoResearch Laboratories, West Grove, PA). The GST fusion was detected by goat anti-GST antibody (Amersham Pharmacia Biotech) and alkaline phosphatase-linked rabbit anti-goat IgG (Chemicon International, Temecula, CA).

Purification of His-Pro and His-VPg proteins

E. coli BL21-CodonPlus-RIL cells harboring pUC[His-Pro] were grown in the same manner as described earlier. Cells were suspended in 20 mM Tris–HCl (pH 7.4) to a final concentration of 0.2 g cells/ml and disrupted by sonication. After removal of unbroken cells, the lysate was centrifuged at 200,000 × g for 1 h at 4°C, and the supernatant was taken as the starting material for purification. The supernatant containing 100 mg of proteins was applied to a column (bed volume, 1 ml) packed with TALON Metal Affinity Resin (Clontech, Palo Alto, CA) that had been equilibrated with 20 mM Tris–HCl (pH 8.0) buffer containing 0.1 M NaCl. The column was washed with 20 mM Tris–HCl (pH 8.0) buffer containing 0.1 M NaCl and 10 mM imidazole. His-Pro proteins were eluted with 20 mM Tris–HCl (pH 8.0) buffer containing 0.1 M NaCl and 50 mM imidazole. His-Pro proteins were purified almost to homogeneity as judged by SDS–PAGE. Purified His-Pro proteins were concentrated to 3 mg/ml by using an Ultrafree-15 Centrifugal Filter Device equipped with a Biomax-10 filter (Millipore, Bedford, MA).

His-VPg proteins were purified from *E. coli* BL21-CodonPlus-RIL cells harboring pUC[His-VPg] in the same manner as was used to purify His-Pro proteins, except that pH 9.0 Tris buffer was used for purification with metal affinity chromatography.

Preparation of antisera against VPg, protease and RNA polymerase

Antisera against VPg and protease were prepared using purified His-VPg and His-Pro proteins, respectively. For preparation of antiserum against RNA polymerase, we used the conjugate of the synthetic peptide (N-^{1390}DKTTSSGHPYHKRKNDDWNG^{1409}-C) and KLH (keyhole limpet hemocyanin).

Purified His-tagged proteins (1.5 mg) or conjugate (3.3 mg) was mixed with Freund's complete adjuvant, and rabbits were immunized subcutaneously with each emulsion. At 4 weeks, the rabbits were boosted with the mixture of respective protein and Freund's incomplete adjuvant. At 5 weeks, the rabbits were boosted with the respective protein solution. At 6 weeks after initial immunization, we bled the rabbits to prepare antisera.

Isolation of proteins for amino acid sequencing

The lysate of *E. coli* BL21-CodonPlus-RIL cells harboring pUC[His-3BCD] was separated by SDS–PAGE, followed by electroblotting onto Trans-Blot PVDF membrane (Bio-Rad, Hercules, CA). The membrane was stained by CBB R250, and then the 20-kDa band corresponding to the ChV protease was excised for amino acid sequencing (Sawady Technology, Japan).

ACKNOWLEDGMENTS

We thank Drs. Y. Matsuura and T. Suzuki for helpful discussions and suggestions. We also thank T. Mizoguchi for secretarial work. This work was supported in part by Health Sciences Research Grants, including the Research on Emerging and Re-emerging Infectious Diseases, Research on Environmental Health, Research on Pharmaceutical and Medical Safety, and Research on Health Sciences Focusing on Drug Innovation from the Ministry of Health and Welfare, Japan. We thank Dr. K. Shinozaki for providing fecal specimens.

REFERENCES

Appleton, H. (1987). Small round viruses: Classification and role in food-borne infections. *Ciba Found. Symp.* **128,** 126–143.

Beller, M., Ellis, A., Lee, S. H., Drebot, M. A., Jenkerson, S. A., Funk, E., Sobsey, M. D., Simmons, O. D., 3rd, Monroe, S. S., Ando, T., Noel, J., Petric, M., Middaugh, J. P., and Spika, J. S. (1997). Outbreak of viral gastroenteritis due to a contaminated well: International consequences. *JAMA* **278,** 563–568.

Boniotti, B., Wirblich, C., Sibilia, M., Meyers, G., Thiel, H. J., and Rossi, C. (1994). Identification and characterization of a 3C-like protease from rabbit hemorrhagic disease virus, a calicivirus. *J. Virol.* **68,** 6487–6495.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72,** 248–254.

Brugha, R., Vipond, I. B., Evans, M. R., Sandifer, Q. D., Roberts, R. J., Salmon, R. L., Caul, E. O., and Mukerjee, A. K. (1999). A community outbreak of food-borne small round-structured virus gastroenteritis caused by a contaminated water supply. *Epidemiol. Infect.* **122,** 145–154.

Clarke, I. N., Lambden, P. R., and Caul, E. O. (1998). Human enteric RNA viruses: caliciviruses and astroviruses. *In* "Topley & Wilson's Microbiology and Microbial Infections" (B. W. J. Mahy, and L. Collier, Eds.), Vol. 1, pp. 511–535. Arnold, London.

Dingle, K. E., Lambden, P. R., Caul, E. O., and Clarke, I. N. (1995). Human enteric Caliciviridae: The complete genome sequence and expression of virus-like particles from a genetic group II small round structured virus. *J. Gen. Virol.* **76,** 2349–2355.

Dowell, S. F., Groves, C., Kirkland, K. B., Cicirello, H. G., Ando, T., Jin, Q., Gentsch, J. R., Monroe, S. S., Humphrey, C. D., Slemp, C., *et al.* (1995). A multistate outbreak of oyster-associated gastroenteritis: Implications for interstate tracing of contaminated shellfish. *J. Infect. Dis.* **171,** 1497–1503.

Estes, M. K., Atmar, R. L., and Hardy, M. E. (1997). Norwalk and related diarrhea viruses. *In* "Clinical Virology" (D. D., Richmann, R. J., Whitley, and F. G., Hayden, Eds.), pp. 1073–1095. Churchill Livingstone Inc., New York.

Fleissner, M. L., Herrmann, J. E., Booth, J. W., Blacklow, N. R., and Nowak, N. A. (1989). Role of Norwalk virus in two foodborne outbreaks of gastroenteritis: Definitive virus association. *Am. J. Epidemiol.* **129,** 165–172.

Glass, P. J., White, L. J., Ball, J. M., Leparc-Goffart, I., Hardy, M. E., and Estes, M. K. (2000). Norwalk virus open reading frame 3 encodes a minor structural protein. *J. Virol.* **74,** 6581–6591.

Gray, J. J., Green, J., Cunliffe, C., Gallimore, C., Lee, J. V., Neal, K., and Brown, D. W. (1997). Mixed genogroup SRSV infections among a party of canoeists exposed to contaminated recreational water. *J. Med. Virol.* **52,** 425–429.

Green, S. M., Dingle, K. E., Lambden, P. R., Caul, E. O., Ashley, C. R., and Clarke, I. N. (1994). Human enteric Caliciviridae: A new prevalent small round-structured virus group defined by RNA-dependent RNA polymerase and capsid diversity. *J. Gen. Virol.* **75,** 1883–1888.

Green, J., Vinje, J., Gallimore, C. I., Koopmans, M., Hale, A., and Brown, D. W. G. (2000). Capsid protein diversity among Norwalk-like viruses. *Virus Genes* **20,** 227–236.

Gunn, R. A., Terranova, W. A., Greenberg, H. B., Yashuk, J., Gary, G. W., Wells, J. G., Taylor, P. R., and Feldman, R. A. (1980). Norwalk virus gastroenteritis aboard a cruise ship: An outbreak on five consecutive cruises. *Am. J. Epidemiol.* **112,** 820–827.

Hardy, M. E., and Estes, M. K. (1996). Completion of the Norwalk virus genome sequence. *Virus Genes* **12,** 287–290.

Herwaldt, B. L., Lew, J. F., Moe, C. L., Lewis, D. C., Humphrey, C. D., Monroe, S. S., Pon, E. W., and Glass, R. I. (1994). Characterization of a variant strain of Norwalk virus from a food-borne outbreak of gastroenteritis on a cruise ship in Hawaii. *Virology* **200,** 319–325.

Ho, M. S., Glass, R. I., Monroe, S. S., Madore, H. P., Stine, S., Pinsky, P. F., Cubitt, D., Ashley, C., and Caul, E. O. (1989). Viral gastroenteritis aboard a cruise ship. *Lancet* **2,** 961–965.

Iversen, A. M., Gill, M., Bartlett, C. L., Cubitt, W. D., and McSwiggan, D. A. (1987). Two outbreaks of foodborne gastroenteritis caused by a small round structured virus: Evidence of prolonged infectivity in a food handler. *Lancet* **2,** 556–558.

Jiang, X., Espul, C., Zhong, W. M., Cuello, H., and Matson, D. O. (1999). Characterization of a novel human calicivirus that may be a naturally occurring recombinant. *Arch. Virol.* **144,** 2377–2387.

Jiang, X., Graham, D. Y., Wang, M., and Estes, M. K. (1990). Norwalk virus genome cloning and characterization. *Science* **250,** 1580–1583.

Jiang, X., Turf, E., Hu, J., Barrett, E., Dai, X. M., Monroe, S., Humphrey, C., Pickering, L. K., and Matson, D. O. (1996). Outbreaks of gastroenteritis in elderly nursing homes and retirement facilities associated with human caliciviruses. *J. Med. Virol.* **50,** 335–341.

Jiang, X., Wang, M., Wang, K., and Estes, M. K. (1993). Sequence and genomic organization of Norwalk virus. *Virology* **195,** 51–61.

Kaplan, J. E., Schonberger, L. B., Varano, G., Jackman, N., Bied, J., and Gary, G. W. (1982). An outbreak of acute nonbacterial gastroenteritis in a nursing home: Demonstration of person-to-person transmission by temporal clustering of cases. *Am. J. Epidemiol.* **116,** 940–948.

Kasuga, K., Tokieda, M., Ohtawara, M., Utagawa, E., and Yamazaki, S. (1990). Small round structured virus associated with an outbreak of

acute gastroenteritis in Chiba, Japan. *Jpn. J. Med. Sci. Biol.* **43,** 111–121.

Khan, A. S., Moe, C. L., Glass, R. I., Monroe, S. S., Estes, M. K., Chapman, L. E., Jiang, X., Humphrey, C., Pon, E., Iskander, J. K., *et al.* (1994). Norwalk virus-associated gastroenteritis traced to ice consumption aboard a cruise ship in Hawaii: Comparison and application of molecular method-based assays. *J. Clin. Microbiol.* **32,** 318–322.

Kobayashi, S., Morishita, T., Yamashita, T., Sakae, K., Nishio, O., Miyake, T., Ishihara, Y., and Isomura, S. (1991). A large outbreak of gastroenteritis associated with a small round structured virus among schoolchildren and teachers in Japan. *Epidemiol. Infect.* **107,** 81–86.

Kohn, M. A., Farley, T. A., Ando, T., Curtis, M., Wilson, S. A., Jin, Q., Monroe, S. S., Baron, R. C., McFarland, L. M., and Glass, R. I. (1995). An outbreak of Norwalk virus gastroenteritis associated with eating raw oysters: Implications for maintaining safe oyster beds [published erratum appears in *JAMA* (1995) **273,** 1492]. *JAMA* **273,** 466–471.

Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154,** 367–382.

Lambden, P. R., Caul, E. O., Ashley, C. R., and Clarke, I. N. (1993). Sequence and genome organization of a human small round-structured (Norwalk-like) virus. *Science* **259,** 516–519.

Lambden, P. R., Caul, E. O., Ashley, C. R., and Clarke, I. N. (1994). Human enteric caliciviruses are genetically distinct from small round structured viruses. *J. Hosp. Infect.* **26,** 251–259.

Lambden, P. R., Liu, B., and Clarke, I. N. (1995). A conserved sequence motif at the 5′ terminus of the Southampton virus genome is characteristic of the Caliciviridae. *Virus Genes* **10,** 149–152.

Le Guyader, F., Neill, F. H., Estes, M. K., Monroe, S. S., Ando, T., and Atmar, R. L. (1996). Detection and analysis of a small round-structured virus strain in oysters implicated in an outbreak of acute gastroenteritis. *Appl. Environ. Microbiol.* **62,** 4268–4272.

Lee, Y. F., Nomoto, A., Detjen, B. M., and Wimmer, E. (1977). A protein covalently linked to poliovirus genome RNA. *Proc. Natl. Acad. Sci. USA* **74,** 59–63.

Lew, J. F., Kapikian, A. Z., Valdesuso, J., and Green, K. Y. (1994). Molecular characterization of Hawaii virus and other Norwalk-like viruses: Evidence for genetic polymorphism among human caliciviruses. *J. Infect. Dis.* **170,** 535–542.

Liu, B., Clarke, I. N., and Lambden, P. R. (1996). Polyprotein processing in Southampton virus: Identification of 3C-like protease cleavage sites by in vitro mutagenesis. *J. Virol.* **70,** 2605–2610.

Liu, B. L., Clarke, I. N., Caul, E. O., and Lambden, P. R. (1995). Human enteric caliciviruses have a unique genome structure and are distinct from the Norwalk-like viruses. *Arch. Virol.* **140,** 1345–1356.

Liu, B. L., Lambden, P. R., Gunther, H., Otto, P., Elschner, M., and Clarke, I. N. (1999a). Molecular characterization of a bovine enteric calicivirus: Relationship to the Norwalk-like viruses. *J. Virol.* **73,** 819–825.

Liu, B. L., Viljoen, G. J., Clarke, I. N., and Lambden, P. R. (1999b). Identification of further proteolytic cleavage sites in the Southampton calicivirus polyprotein by expression of the viral protease in *E. coli. J. Gen. Virol.* **80,** 291–296.

Parashar, U. D., Dow, L., Fankhauser, R. L., Humphrey, C. D., Miller, J., Ando, T., Williams, K. S., Eddy, C. R., Noel, J. S., Ingram, T., Bresee, J. S., Monroe, S. S., and Glass, R. I. (1998). An outbreak of viral gastroenteritis associated with consumption of sandwiches: Implications for the control of transmission by food handlers. *Epidemiol. Infect.* **121,** 615–621.

Rueckert, R. R., and Wimmer, E. (1984). Systematic nomenclature of picornavirus proteins. *J. Virol.* **50,** 957–959.

Seah, E. L., Marshall, J. A., and Wright, P. J. (1999). Open reading frame 1 of the Norwalk-like virus Camberwell: Completion of sequence and expression in mammalian cells. *J. Virol.* **73,** 10531–10535.

Sneath, P. H. A., and Sokal, R. R. (1974). “Numerical Taxonomy. The Principles and Practice of Numerical Classification.” W. H. Freeman, San Francisco.

Stevenson, P., McCann, R., Duthie, R., Glew, E., and Ganguli, L. (1994). A hospital outbreak due to Norwalk virus. *J. Clin. Microbiol.* **32,** 861–866.

Utagawa, E. T., Takeda, N., Inouye, S., Kasuga, K., and Yamazaki, S. (1994). 3′-Terminal sequence of a small round structured virus (SRSV) in Japan. *Arch. Virol.* **135,** 185–192.

Vinje, J., Altena, S. A., and Koopmans, M. P. (1997). The incidence and genetic variability of small round-structured viruses in outbreaks of gastroenteritis in The Netherlands. *J. Infect. Dis.* **176,** 1374–1378.

Vinje, J., Green, J., Lewis, D. C., Gallimore, C. I., Brown, D. W., and Koopmans, M. P. (2000). Genetic polymorphism across regions of the three open reading frames of “Norwalk-like viruses.” *Arch. Virol.* **145,** 223–241.

Vinje, J., and Koopmans, M. P. (1996). Molecular detection and epidemiology of small round-structured viruses in outbreaks of gastroenteritis in the Netherlands. *J. Infect. Dis.* **174,** 610–615.

Wang, J., Jiang, X., Madore, H. P., Gray, J., Desselberger, U., Ando, T., Seto, Y., Oishi, I., Lew, J. F., Green, K. Y., *et al.* (1994). Sequence diversity of small, round-structured viruses in the Norwalk virus group. *J. Virol.* **68,** 5982–5990.