The Molecular Cloning and Sequence of an Open Reading Frame Encoding for Non-Structural Proteins of Feline Calicivirus F4 Strain Isolated in Japan

Ryoko OSHIKAMO, Yukinobu TOHYA, Yasushi KAWAGUCHI, Keizo TOMONAGA, Ken MAEDA, Naokazu TAKEDA[1)], Etsuko UTAGAWA[1)], Chieko KAI, and Takeshi MIKAMI*

Department of Veterinary Microbiology, Faculty of Agriculture, The University of Tokyo, 1–1–1 Yayoi, Bunkyo-ku, Tokyo 113 and [1)]Department of Epidemiology, National Institute of Health, 1–23–1 Toyama, Shinjuku-ku, Tokyo 162, Japan

(Received 8 June 1994/Accepted 11 August 1994)

ABSTRACT. The nucleotide sequence of the 5'-end of feline calicivirus (FCV) Japanese F4 strain genome was determined. This region had 5311 bases and contained a large open reading frame (ORF1) encoding the non-structural proteins. The nucleotide sequence of the ORF1 region was highly conserved as compared with that of FCV F9 strain. When the deduced amino acid sequence of the ORF1 was compared with those of FCV F9 and CFI strains, the sequence was also highly conserved (88.9% and 88.8%, respectively). Functional motifs of the non-structural proteins were common to these strains. There were 2C polypeptide-, 3C cysteine protease- and 3D RNA-dependent RNA polymerase-like regions. The N-terminal region of 2C-like region continued upstream from the region identified by Neill [Virus Res. 17: 145–160]. Furthermore, the presence of 2B-like region was suggested in the upper stream of the 2C-like region, although the function of the region is unknown. When Kyte and Dolittle hydrophobicity profiles of the predicted amino acid sequences of the ORF1s of FCV F4 and F9 were computed and compared, both the profiles had striking similarities. In the region between residues 950–1000, there was a high rate of basic amino acid residues, suggesting that the polypeptide in this region of FCV may have a nucleic acid-binding function.—KEY WORDS: *Caliciviridae*, feline calicivirus, non-structural protein, nucleotide sequence.

J. Vet. Med. Sci. 56(6): 1093–1099, 1994

The *Caliciviridae* is a family of small non-enveloped viruses that possess a single-stranded, positive-sense RNA genome (7.6 kb) and a single capsid protein [21]. These viruses also produce a major subgenomic mRNA during their replication [21]. The genomic RNA has been shown to have a protein (VPg) covalently attached to the 5'-end [3, 22]. Members belonging to this family are feline calicivirus (FCV), vesicular exanthema of swine virus, San Miguel sea lion virus and rabbit hemorrhagic disease virus [12, 21].

Recently, several studies with regard to the molecular biology of FCV were carried out. Neill and Mengeling [15] first cloned cDNAs from FCV strain CFI/68 FIV (FCV CFI), which was isolated in the United States. Neill [13] determined the nucleotide sequence of the central region of FCV genome and found a large open reading frame (ORF) which terminated at 2.4 kb from the 3'-extreme of the genome. The deduced amino acid sequence of the region was suggested to encode three picornavirus-like non-structural proteins, RNA-dependent RNA polymerase, cysteine protease and 2C polypeptide. Neill *et al.* [16] also performed the nucleotide sequence of the capsid protein gene of FCV CFI. Tohya *et al.* [27] analyzed the nucleotide sequence of the 3'-end of the Japanese F4 strain of FCV (FCV F4) from a cloned cDNA of 3.5 kbp and found three ORFs. They suggested that one of the largest ORFs in the cDNA encodes a capsid precursor protein and have a significant homology with VP3 of picornaviruses. Furthermore, Shin *et al.* [24] carried out the expression of the capsid protein gene of FCV F4 in a transient expression system for antigenic analysis of the protein using a panel of monoclonal antibodies to FCV which were previously produced and characterized [26]. Carter *et al.* [4] reported the complete nucleotide sequence of FCV RNA using F9 vaccine strain (FCV F9) whose origin was the U.S. and confirmed two large ORFs (ORF1 and ORF2) and one small ORF (ORF3). From these studies, it was suggested that ORF1 and ORF2 encode the non-structural proteins and the capsid protein, respectively. The ORF3 encodes a putative small protein which was implied to have a possible nucleic acid-binding function [14]. However, the nucleotide sequence of ORF1 of FCV F4 remains incomplete.

In this communication, we report the nucleotide sequence analysis of cloned cDNAs representing the ORF1 of FCV F4 RNA. Within the sequence of the 5'-terminal 5311 nucleotides (nt), a large ORF encoding the non-structural proteins of FCV was found. The deduced amino acid sequence of this ORF was compared with those of the other FCV strains reported by Carter *et al.* [4] and Neill [13].

MATERIALS AND METHODS

Cloning of the central region of the FCV F4 genome: The plasmid pFCV148 is one of the five clones produced in the previous study [27]. Restriction analysis indicated that pFCV148 contains 5.4 kbp of cDNA insert representing the sequence of the 3'-end of FCV F4 genome (Fig. 1A). In the previous study, the 3.5 kb nucleotide sequence of the 3'-end of FCV F4 genome has been analyzed by the use of pFCV119 [27]. The region (1.9 kbp) in the insert of

* CORRESPONDENCE TO: Dr. MIKAMI, T., Department of Veterinary Microbiology, Faculty of Agriculture, The University of Tokyo, 1–1–1 Yayoi, Bunkyo-ku, Tokyo 113, Japan.

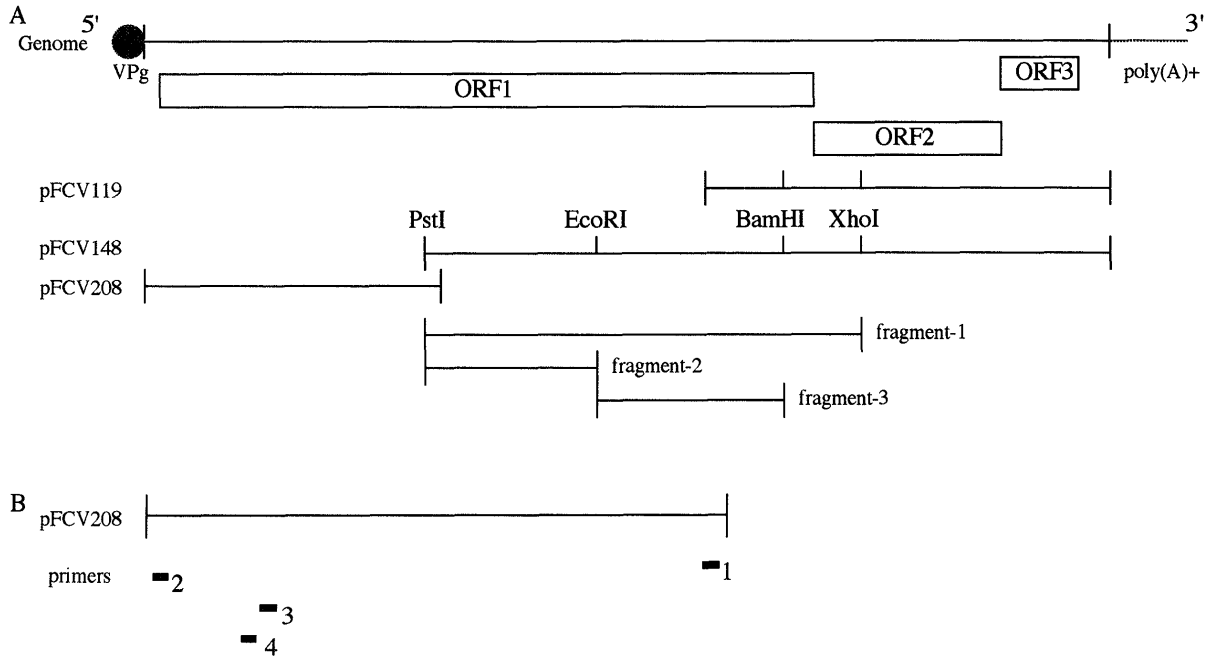

Fig. 1. Physical map and sequencing strategy of the 5'-terminal 5311 bases of the FCV genome. (A) Positions of three ORFs of FCV genome and the cDNA clones pFCV119, pFCV148 and pFCV208 are shown. The restriction sites are indicated. The three fragments were prepared for deletion cloning. (B) Positions of the primers with respect to the genome are shown.

pFCV148 which does not overlap with the insert of pFCV119 was supposed to represent the central region of the FCV genome (Fig. 1A). Three fragments were obtained from pFCV148; fragment-1 (*Pst*I-*Xho*I), fragment-2 (*Pst*I-*Eco*RI) and fragment-3 (*Eco*RI-*Bam*HI) (Fig. 1A), and were further subcloned into the pBluescript (SK(+/−))(Stratagene, Heidellberg, FRG).

Cloning of the 5'-end of the FCV F4 genome: Crandell-Reese feline kidney (CRFK) cell line was grown in Eagle's minimum essential medium containing 10% tryptose phosphate broth, 8% heat-inactivated fetal calf serum(FCS) and antibiotics. The maintenance medium for harvesting the virus RNA was free of FCS. Total RNA was isolated from FCV F4-infected CRFK cells at 4 hr post-infection by the method described previously [5].

The total RNA (1 μg) was reverse-transcribed in the presence of the antisense oligonucleotide primer, primer-1 (5'-ATAAACCTTCTTCCGTTAAA-3'; nt 2271 to 2252 of FCV F4) (Fig. 1B). The RNA was denatured at 70°C for 10 min and cooled on ice. The first-strand synthesis of cDNA was performed in a total volume of 16.8 μl containing 75 mM KCl, 50 mM Tris-HCl (pH 8.3), 5 mM $MgCl_2$, 0.01 M dithiothreitol, 10 mM each deoxynucleoside triphosphate (dNTPmix), 0.5 U RNase inhibitor (Promega Corp., Madison, Wis.), 10 U of Moloney murine leukemia virus reverse transcriptase (GIBCO BRL Life Technologies Inc., Gaithersburg, Md.) and 50 pmol of primer-1. The reaction mixture was incubated at 45°C for 2 hr.

Polymerase chain reaction (PCR) amplification using the synthesized cDNA as a template was carried out after addition of 50 pmol of primer-1, 50 pmol of primer-2 the sense oligonucleotide primer (5'-GTAAAAGAAATTTGAGACAA-3'; nt 1 to 20 of FCVF9) and 1U *Thermus aquaticus* DNA polymerase (*Taq* polymerase; Perkin-Elmer Cetus, Norwalk, Conn.) in 50 μl of buffer (50 mM KCl, 10 mM Tris-HCl [pH 8.3], 1.5 mM $MgCl_2$, 0.1% Triton X-100; supplied by Perkin-Elmer Cetus). The reaction mixture was overlaid with one drop of mineral oil (SIGMA Chemical Co., St. Louis, Mo.) and cycled 25 times for 1 min at 94°C, 1 min at 52°C, and 2 min at 72°C in a DNA thermal cycler (Perkin-Elmer Cetus) [18]. After amplification, only a single band of 2.3 kbp was detected by electrophoresis.

The synthesized cDNA (pFCV208) was ligated into the TA cloning pCR vector according to TA cloning system (Invitrogen, San Diego, CA) and used to transform competent *E. coli* cells supplied in the system. The transformed bacteria containing recombinant plasmids was screened by blue-white color reaction.

To determine the 5'-end sequence of FCVF4 genome, especially 20 bases of the extreme 5'-end of the genome, rapid amplification of cDNA end (RACE) method [7] was performed. Annealing of the antisense oligonucleotide primer-3 (5'-ATCAGGGATGTCACTAGAAG-3', nt 220 to 201 of FCVF4) with FCVF4 RNA and first strand cDNA synthesis were performed as described above. After the cDNA molecules were added homopolymeric tails of dATP at the 3'-ends by terminal deoxynucleotide transferase (GIBCO BRL Life Technologies Inc., Gaithersburg, Md.), PCR amplification was carried out using $(dT)_{17}$-adaptor (5'-GACTCGAGTCGACATCGATTTT-

TTTTTTTTTTTTT-3') and adaptor (5'-GACTCGAG-TCGACATCGA-3') primers, and the other antisense primer-4 (5'-TGCAGTTAGGACAAACGTCG-3'; nt 200 to 181 of FCVF4) in the condition described above. The amplified product was ligated into the pCR vector.

Nucleotide sequencing: The fragments representing the central region and the 5'-end of FCV genome were deleted by exonuclease III (Stratagene, CA) to construct the deletion clones. From these deletion clones, both strands of each insert were sequenced as single stranded DNAs prepared by rescue with R408 helper phage and as double stranded DNAs by dye-primer Taq sequencing method according to the procedure recommended by Applied Biosystems (ABI) (Foster City, CA) [20, 25]. For pFCV208, two clones from separate PCR experiments were sequenced in the reverse direction.

Computer analysis was carried out using the UWGCG software (VAX-11/750) and DNA Strider (Christian

```
                10         20         30         40         50         60         70         80
FCVF4   MSQTLSFVLK THSVRKDFVH SVKVTLARRR DLQYLYNKLA RTMRAEACPS CSSYDVCPNC TSSDIPDNGS STTSIPSWED
FCVF9   ********** ********** ***L****** ****I****S **I******* *A******** **G*V**D** **M*******

                                                                                           160
FCVF4   VTKTSTYSLL LSEDTSDELC PDDLVNVAAH IRKALSTQAH PANTEMCKGQ LTSSLVMAEA MLPQRSWASI PVHHQGLPAR
FCVF9   ***S****** ********** *E******** ********S* ***A****E* **FL****** ******R*** *L*Q*HTA**

                                                                                           240
FCVF4   LEWREKLFFK PLGFLLGKIG VSKDTFQTTA IWKIILEKAC YCKSYGDQWF CAAKQNLREM RTFESDTLKP LVGAFIDGLR
FCVF9   ******F*S* **D***FRV* ****IL**** ********** ******E*** T****K**** KN******** *I*G******

                                          2B?                                              320
FCVF4   FMTVDNPNPM GFLPKLIGLV KPLNLAMIID NHENTLSGWV VTLTAIMELY NITECTIDVI TSLVTGFYDK ISKATKFFSQ
FCVF9   *L******** ********** ********** *****I***I I********* ********I* **VI*A**** *G*****Y*C

                                                                                           400
FCVF4   VKALFTGFRS EDVGNPFWYR AAANPCYLNT GLNPNNGRSP RIKACLSGAT TLVSGIIATQ KLAAMFATWN SESIVNELSA
FCVF9   ********** ***A*S***M ***IL***I* **I*****FS K*****A*** ******V*** ********** **********

                                  2C-like polypeptide region                                #480
FCVF4   RTVAISELNN PTTTSDTDSV ERLLELAKIL HEEIKVHTLN PIMQSYNPIL RDLMSTFDGV ITSCNKRKAI GRKRQVPVCY
FCVF9   ********** ********** ********** *****I**** ********** *N****L*** ********** A*********
CFI                                                                                        *****

                                                                                           560
        +++++++ +                                 #     #    ###                          #  ##
FCVF4   ILTGPPGCGK TTAAQALAKK LSDQEPSVIN LDVDHHDTYT GNEVCIIDEF ASPDKVDYAN FVIGMVNFRP MVLNCDMLEN
FCVF9   ********** ********** ********** ********** ********** D*S******* *******SA* **********
CFI     ********** ********** ********** ********** ********** D*S******* *******SA* **********

        ##   ##       ##   #       #     ##                                                  640
FCVF4   KGKLFTSKYI IMTSNSETPV KPSSKRAGAF YRRVTIIDVT NPLVESHKRA RPGTSVPRSC YKKNFSHLSL AKRGAECWCK
FCVF9   ********** ********** ********** ********** **F******* ********** ********** **********
CFI     ********** ********** ********** *****YH**A -T******** ****A***** ********** **********

                                  #                                                        720
FCVF4   EYVLDPKGLQ HQTIKAPPPT FLNIDSLAQT MKQDFLLKNM ALRTRDGCSE HRYGFICQQS EVETVRRLLN AIRARLNATF
FCVF9   ********** **SM****** ********** ********** *FEAE***A* *****V***E ********** *V***M****
CFI     ********** **ST****** ********** *****A**** *FEAEV**** *****V**** ********** ***M******

                                                                                           800
FCVF4   TVCVGPEASH SIGCTAHVLT PDEPFNGRRF IVSRCNEASL AALEGNCVQS ALGVCMSNKD LTHLCHFIRG KVVNDSVRLD
FCVF9   *******T** ********** *N*T***KK* V********* S*******K* *******D** ********K* *I********
CFI     *****L***N *V********* *******K** V********* S********T ********** ********K* *I********

                                                                                           880
FCVF4   ELPANKHVVT VNSVFDLAWA VRRHLTLTGQ FQAIRAAYDV LTVPDKVPAM LRHWMDETSF SDEHVVTQFV TPGGVVILES
FCVF9   *****Q**** ********** *******A** ********** ******I*** ********** **D******* ****I*****
CFI     *****Q**** ********** L***S***** ********** *H******** ********** *********I **********

                                                                                           960
                                                                                      ! ! !
FCVF4   CGGARIWALG HNVIRAGGVT ATPTGGCVRL LGLSAQTLPW SEIFRELFTL LDRIWSSIKV STLVLTALGM YASRFRPKSE
FCVF9   ********** R********* ********** M****P*M** ********S* *G*****V** *A******** **********
CFI     ********** H********* *I******** M******M** ***LS***S* *G*****V** ***I****S* ********T*

         ! ! ! !                                                                            1040
FCVF4   AKGKTKVKIG PYRGRGVALT DDEYDEWREH NANRKLDLSV EDFLMLRHRA ALGADDADAV KFRSWWNSRT AQEMVLKTVT
FCVF9   ******L*** T********* ********** **S******* ********** ******N*** ********** KMANDYED**
CFI     ******S*** ********** *******K** **A******* ********** ******T*** *********S *LADDFED**
```

```
                                                                    #                                1120
FCVF4   VIGKGGVKHE KIRTSTLRAV DRGYDVSFAE ESGPGTKFHK NAIGSLTDVC GEHKGYCVHM GHGVYASVAH VVKGDSYFLG
FCVF9   ********** ****N**K** ********** ********** *****V**** *******I** ********** ******F***
CFI     ********** ****N***** ********** *****A**** *****V**** ********** ********** ******F***

                                  3C cystein protease like region                              +#  1200
FCVF4   ERIFDVKTNG EFCCFRSTKI LPSGAPFFSG KPTRDPWGSP VATELKPKAY TTTSGKIVGC FATTSTETHP GDCGLPYIDD
FCVF9   *****L**** ********** ***A****** ********** ****W***M* *******L** ********** **********
CFI     *****L**** ********** ***A****** R********* ***DW***P* S********* ********** **********

        #     +    ##                                                                          1280
FCVF4   NGRVTGLHTG SGGPKTPSAK LVVPYIHIDM KNKSVTPQKY DETKPNISYK GLVCKQLGEI RIIPKGTRLH VSPAHVDGYQ
FCVF9   ********** ********** *****V**** *T****A*** *V***D**** *******D** ********** ****TE****
CFI     ********** ********** *****V**** *T****A*** *V***D**** *******D** ********** ****LE*FE*
                                                                                               1360
                                                                                                  #
FCVF4   ECSHQPLFRG QWRPRGSKSL TAIVVDSLKP YCDRVEGPPH DVLHRVHEML IDHLSGFVPM NIFLETFMYP AFHKLNHARS
FCVF9   ******ASL* SGD**CP*** ********** **EN****** ******QK** ********** **SS****LS *******DT*
CFI     ******ASL* SGD**CP*** ********** **VV*N**** *I****QK** ********** **SSD*S*LS *******DT*

        ##   # 3D RNA-dependent RNA polymerase like region### ##   #     # #             1440
FCVF4   CGPYLGGRKK DHMVNGDPDK KFLDLLSSKW KLATQGIALP HEYTIGLKDE LRPIEKMQEG KRRMIWGCDV GVATVCAAAF
FCVF9   ********** ***A**E*** QL*****A** ********** ********** ***V**VS** ********** **********
CFI     ********** ******E*** AL******** ********** ********** ******V*** ********** **********

        #                      #   #                ##   #    #                              1520
FCVF4   KGVSDAITAN HQYGPIQVGI NMDSPSVEVL YQRIKSAAKV FAVDFSNWDS TQSPRVSAAS IDILRYFSDR SPIVDSAANT
FCVF9   ********** ********** ********A* F********* ****Y*K*** ********** ********** *******S**
CFI     ********** ********** ********A* F******R** ****Y*K*** ********** ********** T******T**

                       ## ###   #    # #                    #              ++++          1600
FCVF4   LKSPPIAIFN GVAVKVASGL PSGMPLTSVI NSLNHCMYVG CAILQSLEAR QIPVTWNLFS SFDMMTYGDD GVYMFPTMFA
FCVF9   *****V**** ******S*** ********** ******L*** *********K A********* T**I****** ******I*Y*
CFI     *******V** ******S*** ********** ******L*** ********** NV******** T********* ********Y*

                  #   #                      ####                                              1680
FCVF4   SVSDQIFGNL SAYGLKPTRV DKSVGAIESI DPESVVFLKR TITRTPNGIR GLLDRSSIIR QFFYIKGENS DDWKTPPKTI
FCVF9   *I******** *S******** ********P* **D******* ******Q*** ********** **Y******* ****S***H*
CFI     *I*****A** ********** *****S**P* **N******* ******Q*** ********L* **Y******T *N**E*****

        #                      #                                  #                          1720
FCVF4   DPTSRGQQLW NACLYASQHG VEFYNKVLKL AMRAVEYEGL HLKPPSYSSA LEHYNSQFNG VEARSDQINM SDVTALHCDV
FCVF9   ********** ********** L**F***YR* *E******E* **E**T*A** *D******** ********DS *GM*******
CFI     **M******* ********** ID*****Y** *EK******* **E*****T* ********** ****T***DT *GMA******
                                                                                               1763
FCVF4   FEV
FCVF9   ***
CFI     ***
```

Fig. 2. Comparison of the amino acid sequences for ORF1 of FCV F4, F9 and CFI genomes. The asterisks in the F9 and CFI sequences represent identity with the FCV F4 sequence. The sequences surrounding the functional motif and the active site residues of the 2B, 2C, 3C and 3D regions are lined. The plus signs above the sequences indicate the position of the putative active site of each region; the sharp marks above the sequences show amino acid residues thought to be important in the function of the proteins [13]. The exclamation marks above the segments between 951 and 970 residues denote conserved basic amino acid residues.

Marck).

Primer extension reaction: As a template, virion RNA was extracted from FCVF4 virus particles by the method described previously [10]. Primer-4 was labeled with γ-[^{32}P] ATP using T4 polymerase kinase to prime reverse transcription. After annealing of the labeled primer-4 with FCVF4 RNA, primer extension reaction and polyacrylamide gel electrophoresis were conducted as described previously [19].

RESULTS AND DISCUSSION

The 5'–5311 nt of the FCV genome contained in the cDNA clones, pFCV148 and pFCV208, were subjected to sequence analysis to predict the amino acid sequence of the proteins encoded in this region of the FCV genome. The physical map and sequencing scheme are illustrated in Fig. 1. The sequence analysis of this portion of the FCV genome was facilitated by construction of nested deletions of the cDNA clones. Confirmation of the accuracy of the sequence of pFCV208 came from the sequencing of two clones obtained from the separate PCR amplifications. There was no difference between the sequences of two PCR-derived clones. The cDNA obtained by RACE method was analyzed to determine 20 nt at the 5'-end of FCVF4 genome. The sequence of this area was same as

that of FCVF9. The exact 5'-end of the genome was confirmed by primer extension reaction (data not shown). The computer analysis (DNA Strider) indicated the existence of ORF1 which started at 20 bases from the 5'-end of pFCV208, continued through pFCV148 and terminated at 1142 base from the 5'-end (data not shown).

The total length of ORF1 was 5292 bases (data not shown). The nucleotide sequence of the ORF1 region was compared with that reported by Carter *et al.* [4]. Both sequences were identical in nucleotide length and had ATG and TGA codons for the large ORF at the same positions. In this coding region, there were 1090 nt (20.6%) changes in 5292 nt.

The portion of the ORF1 encodes a polypeptide

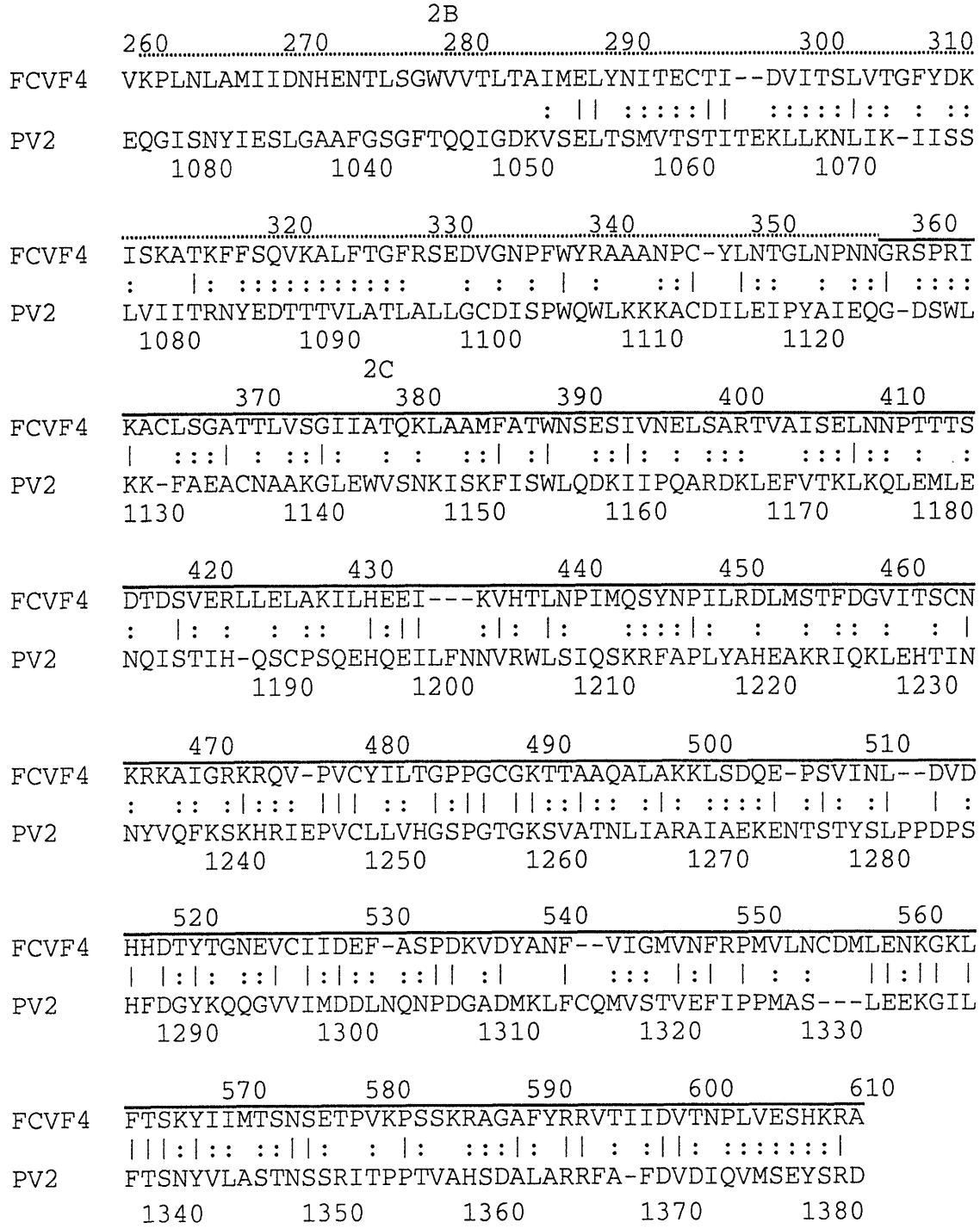

Fig. 3. Amino acid alignment analysis of the deduced 2B- and 2C-like proteins of FCV F4 (amino acid residues 230 through 610 of Fig. 3) with the 2B and 2C polypeptides of poliovirus Sabin 2 (PV2) vaccine strain [28] using the FASTA program of UWCGC. The sequences surrounding the 2B and 2C region are lined.

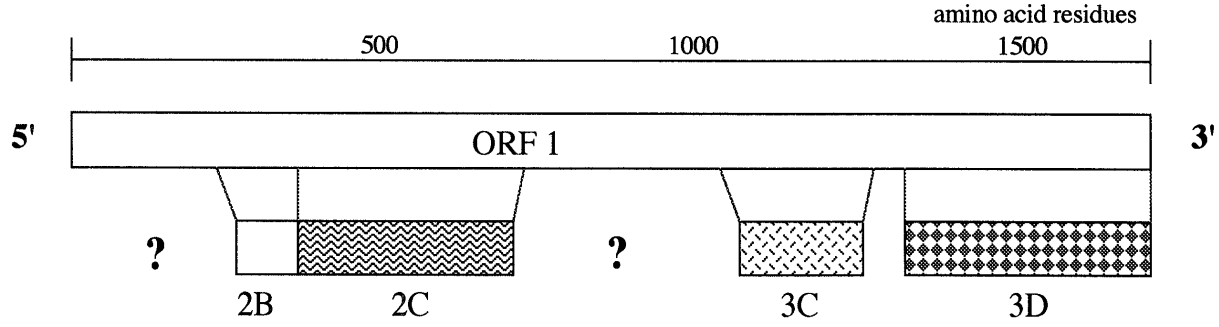

Fig. 4. Proposed ORF1 structures of FCV. The 2B, 2C, 3C and 3D represent coding positions of 2B-, 2C-, 3C cysteine protease-, and 3D RNA-dependent RNA polymerase-like polypeptides, respectively. The question marks represents an area of apparent coding sequences for which the function of the polypeptide product is unknown.

consisting of 1763 amino acids with a molecular weight of 195,229 dalton. When the deduced amino acid sequence of the ORF1 was compared with those reported by Carter *et al.* [4] and Neill [13] (Fig. 2), the sequence was found to be highly conserved (88.9% and 88.8%, respectively). In the ORF1, there is no highly variable region as indicated in ORF2 among FCV strains [11, 23].

Functional motifs and active sites indentified by Neill [13] are common to these strains (Fig. 2). Although N- and C-terminal ends of each protein have not been determined, there are 2C polypeptide-, 3C cysteine protease- and 3D RNA-dependent RNA polymerase-like regions in the deduced amino acid sequences of FCV F4. The computer analysis (UWGCG) indicated that the region of 2C-polypeptide continued upstream from the region indentified by Neill [13]. Furthermore the presence of 2B-like region was also suggested in the upper stream of the 2C-like region.

Figure 3 shows the amino acid aligments between the predicted amino acid sequence of the FCV F4 ORF1 and non-structural proteins encoded by poliovirus Sabin 2 (PV2) vaccine strain [28]. The significant homology was revealed between the amino acid sequences of the 2B and 2C regions of PV2 and the analogous regions of FCV. The positions of the 2B and 2C regions in the ORF1 are also suggested from these results. Although the function of picornavirus 2B polypeptide has not been determined, the polypeptide of FCV in the region might have similar function of 2B polypeptide, if any.

The Kyte and Dolittle hydrophobicity profile [9] of the predicted proteins specified by the ORF1 sequence of FCV F4 were computed and compared with that of FCV F9 [4]. The profiles of the proteins of FCV F4 and F9 had a striking similarity each other (data not shown). This similarity is discernible even in the region 700–1,000 located between 2C-like and 3C-like regions, although the function of this region is still unknown. There are many basic amino acid residues in the hydrophilic region 951–1,000. Especially, in the region 950–970, 7 basic amino acid residues were all conserved among the three strains (Fig. 2). This result suggested that the polypeptide in this region of FCV may have a nucleic acid-binding function [8].

Caliciviruses represent a poorly studied virus family. Elucidation and characterization of the non-structural proteins of the caliciviruses have lagged behind those of other RNA viruses [1, 2, 6, 17]. Comparison of this complete sequence of FCV F4 genome with the two published FCV sequences [4, 13] revealed a considerable degree of similarity. This relationship is evident in the overall sequence homology and the conservation of certain amino acid sequence motifs, indicating the existence of caliciviral 2B, 2C, 3C cysteine protease, and 3D RNA-dependent RNA polymerase. The positions of the sequences encoding these non-structural proteins in the ORF1 are illustrated in Fig. 4. The question marks in Fig. 4 denote areas that apparently encode polypeptides, but the identity and function of these polypeptides remain to be unknown. For 2B-like region, more sequence information and especially data on genome organization and strategy of gene expression will provide better arguments for identification of the function of the polypeptide. This study was done to bring the molecular biology of the caliciviruses up to date and to expand the work that has been done. Further work on these important pathogens will help not only to control and prevent calicivirus-induced diseases but probably also to add new aspects to our knowledge of the molecular biology and evolution of positive-stranded RNA viruses.

Note: The nucleotide sequence data reported in this paper will appear in the GSDB, DDBJ, EMBL and NCBI nucleotide sequence databases with the following accession number D31836.

ACKNOWLEDGEMENTS. We are grateful to Dr. K. Saeki, Department of Molecular Immunology, The University of Tokyo, for providing the primers for RACE cloning and for valuable discussion on the cloning method. This work was supported partly by grants from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. Boeyé, A. and Rombaut, B. 1992. The proteins of Poliovirus. *Prog. Med. Virol.* 39: 139–166.

2. Bredenbeek, P. J. and Rice, C. M. 1992. Animal RNA virus expression systems. *Semin. Virol.* 3: 297–310.
3. Burroughs, J. N. and Brown, F. 1978. Presence of a covalently linked protein on calicivirus RNA. *J. Gen. Virol.* 41: 443–446.
4. Carter, M. J., Milton, I. D., Meanger, J., Bennett, M., Gaskell, R. M., and Turner, P. C. 1992. The complete nucleotide sequence of a feline calicivirus. *Virology* 190: 443–448.
5. Chomczynski, P. and Sacchi, N. 1987. Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. *Anal. Bioch.* 162: 156–159.
6. Dolja, V. V. and Carrington, J. C. 1992. Evolution of positive-strand RNA viruses. *Semin. Virol.* 3: 315–326.
7. Frohman, M. A., Dush, M., K., and Martin, G. R. 1988. Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. U.S.A.* 85: 8998–9002.
8. Koonin, E. V., Boyko, V. P., and Dolja, V. V. 1991. Small cysteine-rich proteins of different groups of plant RNA viruses are related to different families of nucleic acid-binding proteins. *Virology* 181: 395–398.
9. Kyte, J. and Dollitle, R. F. 1982. A simple method for displaying in the hydropathic character of a protein. *J. Mol. Biol.* 157: 105–132.
10. Meyers, G., Wirblich, C., and Thiel, H-J. 1991. Genomic and subgenomic RNAs of rabbit hemorrhagic disease virus are both protein-linked and packaged into particles. *Virology* 184: 677–686.
11. Milton, I. D., Turner, J., Teelan, A., Gaskell, R., Turner, P. C., and Carter, M. J. 1992. Location of monoclonal antibody binding sites in the capsid protein of feline calicivirus. *J. Gen. Virol.* 73: 2435–2439.
12. Moussa, A., Chasey, D., Lavazza, A., Capucci, L., Šmid, B., Meyers, G., Rossi, C., Thiel, H-J., Vlásak, R., Rønsholt, L., Nowotny, N., McCullough, K., and Gavier-Widen, D. 1992. Haemorrhagic disease of lagomorphs: evidence for a calicivirus. *Vet. Microbiol.* 33: 375–381.
13. Neill, J. D. 1990. Nucleotide sequence of a region of the feline calicivirus genome which encodes picornavirus-like RNA-dependent RNA polymerase, cysteine protease and 2C polypeptides. *Virus Res.* 17: 145–160.
14. Neill, J. D. 1992. Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus: Identification of conserved and non-conserved amino acid sequences among calcivirus capsid proteins. *Virus Res.* 24: 211–222.
15. Neill, J. D. and Mengeling, W. L. 1988. Further characterization of the virus-specific RNAs in feline calicivirus infected cells. *Virus Res.* 11: 59–72.
16. Neill, J. D., Reardon, I. M., and Heinrikson, R. L. 1991. Nucleotide sequence and expression of the capsid protein gene of feline calicivirus. *J. Virol.* 65: 5440–5447.
17. Porter, A. G. 1993. Picornavirus nonstructural proteins: Emerging roles in virus replication and inhibition of host cell functions. *J. Virol.* 67: 6917–6921.
18. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239: 487–491.
19. Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular Cloning. A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York.
20. Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74: 5463–5467.
21. Schaffer, F. L. 1991. Caliciviridae. *Arch. Virol. (Suppl.)* 2: 300–302.
22. Schaffer, F. L., Ehresmann, D. W., Fretz, M. K., and Soergel M. E. 1980. A protein, VPg, covalently linked to 36S calicivirus RNA. *J. Gen. Virol.* 47: 215–220.
23. Seal, B. S., Ridpath, J. F., and Mengelling, W. L. 1993. Analysis of feline calicivirus capsid protein genes; Identification of variable antigenic determinant regions of the protein. *J. Gen. Virol.* 74: 2519–2524.
24. Shin, Y-S., Tohya, Y., Oshikamo, R., Kawaguchi, Y., Tomonaga, K., Miyazawa, T., Kai, C., and Mikami, T. 1993. Antigenic analysis of feline calicivirus capsid precursor protein and its deleted polypeptides produced in a mammalian cDNA expression system. *Virus Res.* 30: 17–26.
25. Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H., and Hood, L. E. 1986. Fluorescence detection in automated DNA sequence analysis. *Nature (Lond.)* 321: 674–679.
26. Tohya, Y., Masuoka, K., Takahashi, E., and Mikami, T. 1991. Neutralizing epitopes of feline calicivirus. *Arch. Virol.* 117: 173–181.
27. Tohya, Y., Taniguchi, Y., Takahashi, E., Utagawa, E., Takeda, N., Miyamura, K., Yamazaki, S., and Mikami, T. 1991. Sequence analysis of the 3'-end of feline calicivirus genome. *Virology* 183: 810–814.
28. Toyoda, H., Kohara, M., Kataoka, Y., Suganuma, T., Omata, T., Imura, N., and Nomoto, A. 1984. Complete nucleotide sequences of all three poliovirus serotype genomes. *J. Mol. Biol.* 174: 561–585.