KOR. JOUR. MICROBIOL October, 1992 p. 397~402 Vol. 30, No. 5

cDNA Cloning and Nucleotide Sequence Determination for VP7 Coding RNA Segment of Human Rotavirus Isolated in Korea

Kim, Young Bong[1], Kyung Hee Kim[2] and Jai Myung Yang[3]*

[1]*Department of Genetic Engineering, Sung Kyun Kwan University*
[2]*Department of Microbiology, School of Medicine, Hanyang University*
[3]**Biology Department, Sogang University*

The cDNA of RNA segment coding for VP7 of human rotavirus isolated from patient's stool at Seoul area was synthesized, amplified by polymerase chain reaction, filled in with Klenow fragment of DNA polymerase I and cloned into pUC19. The cDNA sequence was determined and compared with that of VP7 coding RNA segments of group A rotaviruses isolated in foreign country. Over 90% sequence homology was found with serotype 1 specific WA1 and RE9 strains. Comparative analysis of the deduced amino acid sequences within the two variable regions (amino acid residue 87 through 101 and 208 through 221) with WA1 and RE9 strains also showed high degree of sequence similarity with each other.

KEY WORDS □ Rotavirus, VP7, PCR, cDNA, Sequence homology

Rotavirus, a member of family *Reoviridae*, is the leading cause of gastroenteritis in young children throughout the world (1). Rotaviral diarrhea is most common in children 6~18 months of age and accounts for up to 50% of such cases during the cooler months of the year in temperate regions (2, 3). In Seoul, Korea, a study using a polyclonal antibody-based enzyme-linked immunosorbent assay (ELISA) revealed that rotavirus was the most common enteropathogen identified in the stools of children with diarrhea, accounting for 47% of cases (4).

The rotavirus genome consists of 11 segments of double-stranded RNA size ranged from 0.2 x 10^6 to 2.2 x 10^6 daltons (5). The genomic RNA is surrounded by two layers of protein capsid. The inner shell consists of a core that contains genomic dsRNA and the structural proteins VP1, VP2, and VP3 surrounded by the major inner shell protein VP6, whereas the outer shell is composed of VP7 and VP4.

The major outer capsid protein, VP7, is a glycoprotein of M.W. 35,500 daltons which has been shown to be an antigen responsible for the induction of major neutralizing antibodies to the virus (6). Serotypes of rotavirus are defined by the neutralizing activity and to date, six serotypes have been identified in group A human rotaviruses (serotypes 1, 2, 3, 4, 8, and 9) (7). VP7 is encoded by the RNA genome segment 7, 8, or 9, depending upon the virus strain (5).

Although electropherotping and serotyping of the human rotavirus (HRV) circulating in Seoul area was investigated (8), molecular genetic studies for the virus has not been reported yet. Comparison of the nucleic acid sequence of the VP7 coding RNA segment would give an information about the degree of diversity among HRV isolated in Korea and foreign country.

We have synthesized cDNA for VP7 coding RNA segment of HRV isolated from patient's stool at Seoul area. Part of the cDNA seqence was determined and compared with that of VP7 coding RNA segment of HRV isolated at foreign country. Over 90% sequence homology was found with serotype 1 specific WA1 and RE9 strains.

MATERIALS AND METHODS

Stool specimen

Stool specimen was kindly supplied from Seoul National University Hospital. The Slidex Rota-Kit 2 (Bio Merieus) positive stool specimen was collected from an infant patient who was admitted for acute diarrhea.

Virus purification

Approximately 1.0 g of stool diluted in 10 m*l* of phosphate buffered saline (PBS, 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4, pH 7.4) were homogenized, and centrifuged at 12,000 x g for 20 min. To the supernatant an equal volume of Freon-113 (1,1,2-trichlorotrifluoroethane, Fluka) was added, homogenized, and centrifuged at 800 x g for 10 min. The virus in aqueous phase

was pelleted through 40% sucrose cushion at 100,000 x g for 2 hr. The virus pellet was resuspended in 200 μl of TE (10 mM Tris-HCl, pH 7.4, 1 mM EDTA) and stored at −20°C until next use.

RNA extraction

10% sodium dodecyl sulfate (SDS) was added to the final concentration of 1% and the mixture was allowed to stand at room temperature for 30 min. After the addition of an equal volume of TE buffer saturated phenol, the mixture was vortexed vigorously and centrifuged at 12,000 x g at room temperature for 2 min. RNA in the upper aqueous phase was precipitated by storing overnight at −20°C after addition of 0.1 volume of 3 M sodium acetate and two volumes of 95% ethanol, and pelleted by centrifugation at 12,000 x g for 20 min. The pelleted RNA was vacuum dried, resuspended in 30 μl of TE (pH 8.0), and stored at −20°C until next use.

Agarose gel electrophoresis

Electrophoresis was carried out in 1.5% agarose gel in TAE buffer (40 mM Tris-acetate, 1 mM EDTA) at 5 V/cm, for 2 hr. The DNA was stained with ethidium bromide (0.5 μg/ml) and visualized by illuminating under U.V.

Polyacrylamide gel electrophoresis (PAGE)

RNA was separated in 7.5% polylacrylamide slab gel polymerized in 89 mM Tris-borate, pH 8.3, 1 mM EDTA (TBE) buffer, at 15 V/cm for 4 hr.

Primers

Oligonucleotides complementary to the 3′ ends of VP7 coding RNA segment 9 (or segment 7 or 8) were synthesized and used as a primer for cDNA synthesis. The sequence of the primers, which are 23 and 18 nucleotides long, is following;

S-856
GGCTTTAAAAAGAGAATTTCCG 1-23
S-857
GGTCACATCATACAACTC 1062-1045

cDNA synthesis

dsRNA was denatured in 90% dimethyl sulfoxide (DMSO, Sigma) at 50°C for 30 min and quickly quenched in a dry ice-ethanol bath to prevent renaturation. Denatured RNA was added to the cDNA synthesis reaction mixture consisting of 10 mM Tris (pH 8.3), 40 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM each dNTPs, and 1 μM of each primers. 4 units of avian myeloblastosis virus (AMV) super reverse transcriptase was added to the reaction tube and incubated at 4°C for 30 min. The RNA template was then digested in 0.5 M KOH at 22°C for 2 hr. After neutralization with 1 N HCl, the reaction mixture was passed through a Sephadex G-50 column in 20 mM Tris-HCl (pH 8.0)/7 M urea and the cDNA was precipitated by adding two volumes of ethanol.

Polymerase chain reaction

The cDNA dissolved in 10 μl of 0.1 X TE was added to the reaction mixture consisted of 10 mM Tris (pH 8.3), 40 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM each dNTPs, 1 μM each primers, and 5 units of Taq DNA polymerase (Pharmacia). The reaction mixture overlaid with mineral oil was heated at 94°C for 5 min. Twenty five cycles (94°C for 1 min, 42°C for 2 min, 72°C for 3 min) of amplification were performed in a DNA thermal cycler (Ericomp Inc.) and incubated at 72°C for 7 min finally.

cDNA cloning

3′ ends of the cDNA were filled in with the Klenow fragment of DNA polymerase I. 5′ ends were phosphorylated by T4 DNA kinase, and the cDNA was ligated into the *Sma*I site of pUC19. The ligation mixture was transformed into *E. coli* strain SURE, and recombinant plasmids with cDNA insert were screened by ampiciline resistance and α-complementation of β-galactosidase in the presence of ampicillin and X-gal.

Sequencing

DNA sequencing was done by Sanger's dideoxy chain-termination method (9). The DNA solution (5 μg in 8 μl of TE) was mixed with 2 μl of 2 N NaOH and kept at 37°C for 30 min and neutralized by adding 8 μl of 5 M ammonium acetate (pH 7.4). The denatured DNA was precipitated by addition of 100 μl of ethanol at −20°C for 1 hr. The precipitated DNA was resuspended in 8 μl of deionized water and added to the reaction mixture containing 0.5 pmol of primer and 2 μl of 5 X annealing buffer (USB) It was then heated at 65°C for 2 min and allowed to cool to room temperature over a period of 30 min and chilled on ice. [α-^{35}S] dATP (5 μCi, Amershame), 100 mM DTT, 2 μl of labeling nucleotide mix and 3 units of sequenase (USB) was added to the annealed template-primer. It was incubated at 16°C for 3 min and immediately divided into four tubes previously filled with 2.5 μl of the appropriate dideoxy termination mixture and labeled as "G", "A", "T", "C". After incubation at 37°C for 3~5 min 4 μl of stop solution was added and samples were kept on ice. The samples were heated at 80 °C for 5 min prior to load on 5% acrylamide-7 M urea gel. At the end of the run, the gel on one glass plate was immersed in 10% methanol-10% acetic acid solution for 15 min, transferred to Whatmann 3 M paper, and dried at 80°C under vacuum. Autoradiography was done for 12~20 hr at room temperature without an intensifying screen.

RESULTS AND DISCUSSION

Preparation of the template dsRNA

HRV genomic dsRNA was extracted from partially purified virus from stool specimen previously screened by Rota clone kit, and separated in 1.5% agarose gel. As shown in Fig. 1, four high-molecular weight segments (segments 1, 2, 3 and 4), five middle-sized segments (segments 5, 6, 7,

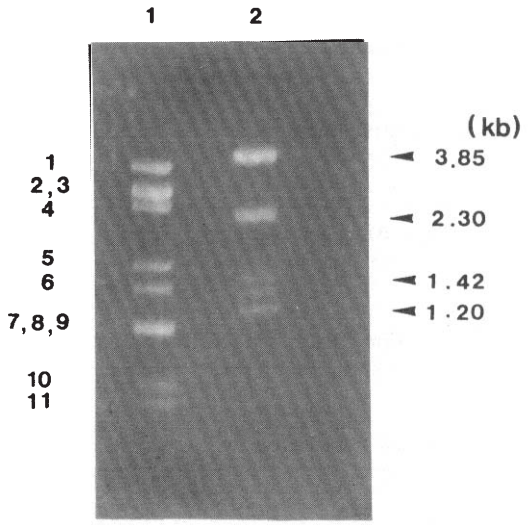

Fig. 1. *Agarose gel electrophoresis of HRV genomic RNA extracted from stool specimen. RNA was separated in 1.5% agarose gel at 5 V/cm, stained with EtBr, and photographed on the U.V. illumination box. Lane 1; HRV genomic RNA, lane 2; reovirus genomic RNA.*

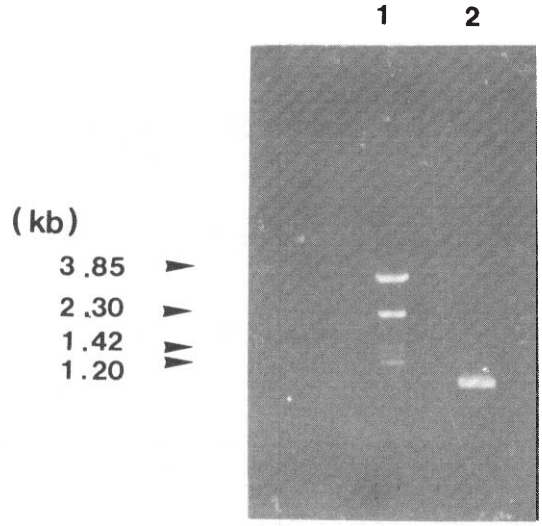

Fig. 2. *Agarose gel electrophoresis of PCR amplified cDNA. Amplified cDNA was electrophoresed in 1.5% agarose gel at 5 V/cm, stained with EtBr, and photographed on the U.V. illumination box. Lane 1; reovirus genomic RNA, lane 2; PCR amplified cDNA.*

8 and 9) including a distinctive triplet of segments (segments 7, 8 and 9), and two smaller segments (segments 10 and 11) were observed. The RNA band containing segments 7, 8 and 9, the probable VP7 coding genome, was cut and RNAs were electroeluted from the gel. The recovered ds RNAs were used as a template for cDNA synthesis.

cDNA synthesis and amplification

dsRNA in 90% DMSO was denatured at 50°C for 30 min and quenched in dry ice-ethanol bath to prevent renaturation. Denatured RNA template was added into the AMV reverse transcriptase reaction mixture, and incubated at 42°C for 30 min. The cDNA was then amplified by twenty five cycles of PCR (94°C for 1 min, 42°C for 2 min, 72°C for 3 min) in a DNA thermal cycler. Small aliquots (5 μl from 100 μl reaction mixture) of the PCR reaction mixture run in 1.5% agarose gel showed the amplified cDNA with the size of 1 Kbp (Fig. 2, lane 2).

Cloning of cDNA into pUC19

3′ end of the amplified cDNA was filled in by Klenow fragment of DNA polymerase I, ligated into the *Sma*I site of the pUC19, and the ligated DNA was transformed into *E. coli* SURE. The recombinant colonies were screened by ampiciline resistance and α-complementation of β-galactosidase from X-gal containing agar plate. Recombinant plasmid DNAs prepared from white colonies were digested with *Pst*I and *Eco*RI, and analyzed in 1% agarose gel. Lane 3 of Fig. 3 showed that the recombinant plasmid generated

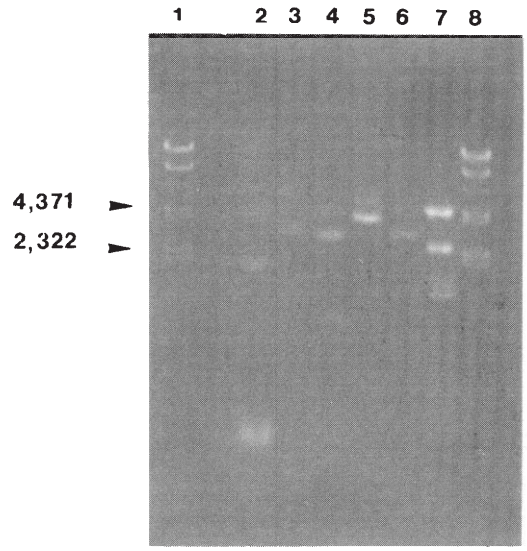

Fig. 3. *Agarose gel electrophoresis of restriction endonuclease digested recombinant plasmid DNA. The digested DNA samples were separated in 1.0% agarose gel. Lane 2; pUC19 uncut, lanes 3; recombinant DNA uncut, lane 4; recombinant DNA EcoRI and Pst1 double cut, lane 5; recombinant DNA EcoRI single cut, lane 6; recombinant DNA SmaI cut, lane 7; reovirus genomic RNA, lanes 1 and 8; λ HindIII marker.*

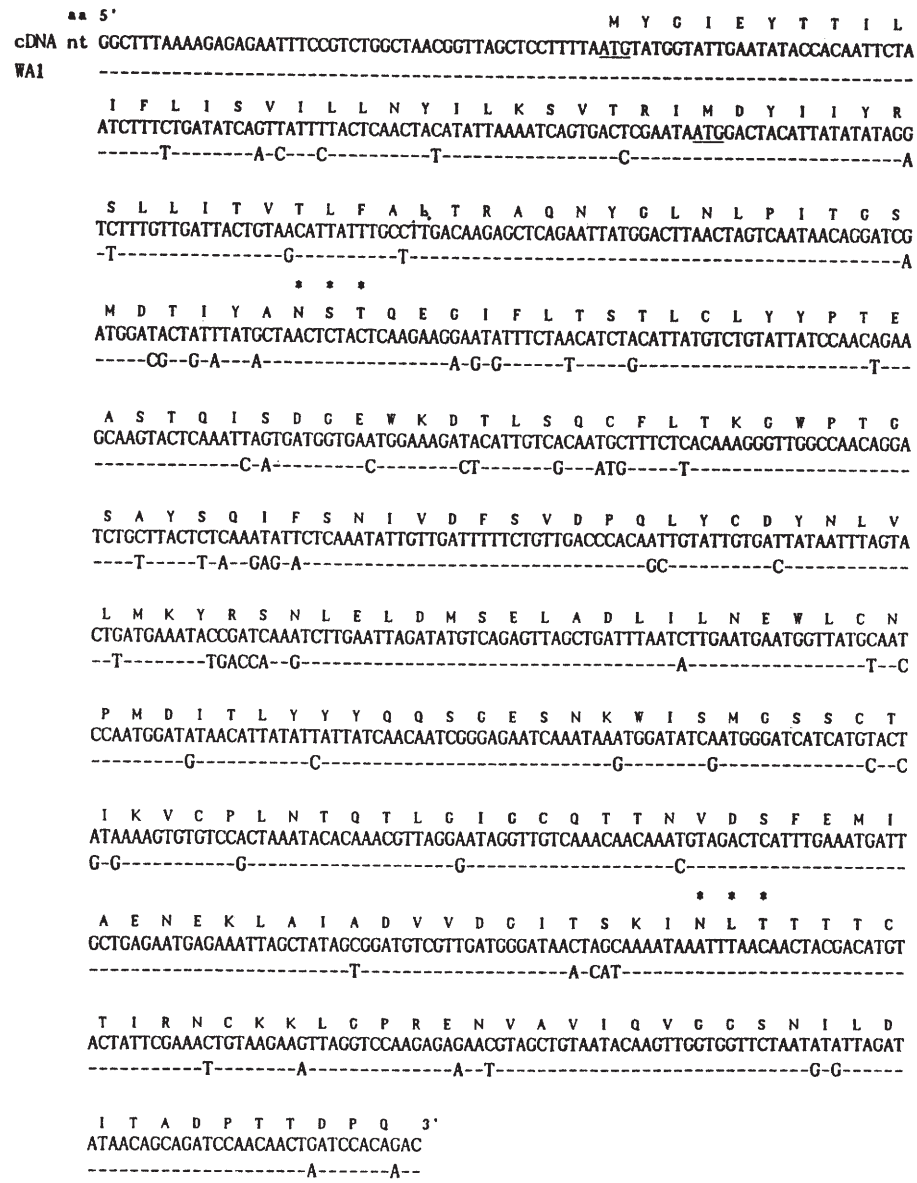

Fig. 4. *Nucleotide sequence of cDNA and it's predicted amino acid sequence. The DNA sequence of the sense stand is shown 5' to 3'. Two in-phase initiation codons are underlined and the potential glycosylation sites are marked with asterisk. For comparison, the changes found in the homologous segment from HRV strain WA1 are shown.*

1 Kbp fragment of insert DNA and 2.7 Kbp fragment of pUC19. Since *Sma*I site was not regenerated after ligation of insert DNA into pUC 19, it did not digest the recombinant DNA (lanes 3 and 6).

DNA sequencing and analysis

Part of the cDNA sequence (890 out of 1062 bases) was determined by Sanger's chain termination method. The nucleotide sequence and it's predicted amino acids sequence is shown in Fig. 4. The base composition of cDNA, like all rotavirus segments sequenced to date, is A/T rich (67%). As found in other HRV genomic RNA, there are two in-phase AUG codons at residues 49 through 51 and 136 through 138. The first potential initiation codon (residues 49 through 51) in cDNA has the sequence TXXATGT, regarded as a weak initiator of protein syntheisis in

VR5 region

```
Strain/
Serotype
                aa: 87  88  89  90  91  92  93  94  95  96  97  98  99  100 101
cDNA        1       ACA GAA GCA AGT ACT CAA ATT AGT GAT GGT GAA TGG AAA GAT TCA
                    Thr Glu Ala Ser Thr Gln Ile Ser Asp Gly Glu Trp Lys Asp Ser
WA1         1       ACT GAA GCA AGT ACT CAA ATC AAT GAT GGT GAC TGC AAA GAC TCA
                    ----------------------------Asn---------Asp Cys -----------
RE9         1       ACT GAA GCA AGT ACT CAA ATC AAT GAT GGT GAC TGG AAA GAC TCA
                    ---------------------------- Asn ------- Asp ---------------
DS1         2       GCA GAA GCT AAA AAT GAG ATT TCA GAT GAT GAA TGG GAA AAT ACT
                    Ala ------- Lys Asn Glu --- --- --- Asp --- --- Glu Asn Thr
P           3       ACT GAA GCA GCA ACA GAA ATA AAT GAT AAT TCA TGG AAG GAT ACA
                    ---------- Ala --- Glu ----Asn---- Asn Ser ---------- Thr
ST3         4       TCA GAA GCT CCA ACT CAA ATT AGT GAC ACT GAA TGG AAA GAT ACA
                    Ser --------Pro ------------------- Thr --------------- Thr
```

VR8 region

```
Strain/
Serotype
                aa: 208 209 210 211 212 213 214 215 216 217 218 219 220 221
cDNA        1       CAA ACA ACA AAT GTA GAC TCA TTT GAA ATG ATT GCT GAG AAT
                    Gln Thr Thr Asn Val Asp Ser Phe Glu Met Ile Ala Glu Asn
WA1         1       CAA ACA ACA AAC GTA GAC TCA TTT GAA ATG ATT GCT GAG AAT
                    ----------------------------------------------------------
RE9         1       CAA ACA ACA AAC GTA GAC TCA TTT GAA ATG ATT GCT GAG AAT

                    ----------------------------------------------------------
DS1         2       AAA ACT ACG GAC GTA AAC ACA TTT GAG ATT GTT GCG TCG TCT
                    lys ------- Asp --- Asn Thr ------- Ile Val --- Ser Ser
P           3       CTA TCT ACT GAT ACA AAC ACG TTT GAA GAA GTT GCA ACA GGT
                    Leu ------- Asp Thr Asn Thr ------- Glu Val --- Thr Ala
ST3         4       CAA ACG ACA AAT ACA GCT ACT TTT GAA ACA GTT GCT GAT AGC
                    --------------- Thr Ala Thr ------- Thr Val --- Asp Ser
```

Fig. 5. *Comparison of the deduced amino acid sequence within two variable regions (VR5, nucleotides 307 through 351 encoding amino acids 87 through 101 and VR8, nucleotides 670 through 711 encoding amino acids 208 through 221) which were divergent in sequence among different serotypes. Serotype 1 specific strains RE9 and WA1, serotype 2 specific strain DS, serotype 3 specific P strain, and serotype 4 specific ST3 were compared with the deduced amino acid sequence from cDNA.*

eucaryotic systems and the second AUG coding triplet (residues 136 through 138) has the consensus sequence AXXATGG, characteristic of strong initiator (10).

The cDNA sequence was compared with VP7 coding RNA segment of HRV strains HU-5, NCDV, RE9, SA-11, UK, and WA1, respectively. 91% sequence homology was observed with segment 9 of serotype 1 specific strains WA1 and RE9 while 71~74% sequence homology was observed with serotype 3 and 6 specific strains HU-5, NCDV, SA-11, and UK. Comparison of the nucleotide sequence of the cDNA with VP7 coding RNA segment of RE9 to which it matched most closely showed that most nucleotide substitutions were transition mutations (Fig. 4).

The amino acid sequence deduced from the cDNA was also analyzed. A potential N-linked glycosylation site (Asn-X-Thr) at amino acid residues 69 through 71 residue was observed. In addition, there is a second potential glycosylation site at residues 238 through 240. The amino acid sequence of two regions within VP7 (VR5, amino acid residues 87~101 and VR8, amino acid residues 208~221) was compared with that of the HRV strains HU-5, NCDV, RE9, SA-11, UK, and WA1. These regions are variable among different serotypes but highly conserved among strains belong to identical serotype. The two regions were also shown to be involved in the formation of a major serotype-specific antigenic site (11).

High degree of sequence similarity was observed in WA1 and RE9 (Fig. 5). Within VR5 region, 94th amino acid Asn was changed into Ser and 97th Asp was changed into Glu. Tryptophan at 98th residue was also changed into Cys in WA1. These changes are conservative, involving the exchange of amino acids with similar chemical properties. Within VR8 region no amino acids substitution was observed. On the other hand considerable substitution was observed in other serotype specific rotavirus strains.

Above results indicated that the HRV used for cDNA synthesis in this study is most probably serotype 1 which is the most dorminant HRV serotype circulating Seoul area (8). The cloned cDNA would be useful for diagnosis of HRV infection by dot blot hybridization and for identifying serotype 1 specific HRV strain by PCR reaction (12).

REFERENCES

1. **Cukor, G. and N.R. Blacklow,** 1984. Human viral gastroenteritis. *Microbiol. Rev.* **48**, 157-179.
2. **Brandt, C.D., H.W. Kim, W.J. Rodrignez, J.O. Arroblo, B.C. Jeffries, E.P. Stallings, C. Lewis, A. J. Miles, R.M. Chanock, A.Z. Kapikian and R.H. Parrott.** 1983. Pediatric viral gastroenteritis during eight years of study. *J. Clin. Mocrobiol.* **18**, 71-78.
3. **Konno, T., H. Suzuki, A. Imai, T. Kutsuzawa, N. Ishida, N. Katsushima, M. Sakamodo, S. Kitnoka, R. Tsuboi and M. Adachi,** 1978. A long-term survey of roavirus infection in Japanese children with acute gastroenteritis. *J. Infect. Dis.* **138**, 569-576.
4. **Kim K.H., Suh I.S., Kim J.M., Kim C.W. and Cho Y.J.,** 1984. Etiology of childhood diarrhea in Korea. *J. Clin. Microbiol.* **27**, 1192-1196.
5. **Ester, M.K. and J. Cohen.,** 1989. Rotavirus gene structure and function. *Microbiol. Rev.* **53**, 410-449.
6. **Kalica, A.Z., H.B. Greenberg, R.G. Wyatt, J. Flores, M.M. Sereno, A.Z. Kapikian, and R.M. Chanock,** 1981. Genes of human (strain Wa)and bovine(strain UK) rotaviruses that code for neutrlization and subgroup antigens. *Virology* **112**, 385-390.
7. **Ester. M.K., E.l. Palmer and J.F. Obijeski,** 1983. Rotaviruses: a review. *Curr. Top. Microbiol.* **19**, 273-278.
8. **Kim K.H, J.M. Yang, S.I. Joo, Y.G. Cho, R.I.**

Glass and Y.J. Cho, 1990. Importance of rotavirus and adenovirus type 40 and 41 in acute gastroenteritis in korean children. *J. Clin. Microbiol.* **28**, 2279-2284.
9. **Masahira Hattori and S. Yoshiyuki,** 1986. Dideoxy Sequencing Method Using Denatured Plasmid Templates. *Anal. Biochem.* **152**, 232-238.
10. **Kozak, M.** 1981. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res.* **9**, 5233-5252.
11. **Green, K.Y., J. Sears, K. Taniguchi, K. Midthun, Y. Hochino, M. Gorziglia, K. Nishikawa, S. Urasawa, A.Z. Kapikian, R.M. Chanock and J. Flores.** 1988. Prediction of human rotavirus serotype by nucleotide sequence analysis of VP7 protein gene. *J. Virol.* **62**, 1819-1823.
12. **Flores, J., K.Y. Green, D. Garcia, J. Sears, I. Perez-Schael, L.F. Avendano, W.B. Rodriguez, K. Taniguchi, S. Urasawa and A.Z. Kapikian.** 1989. Dot hybridization assay for distinction of rotavirus serotypes. *J. Clin. Microbiol.* **27**, 29-34.

(Received August 19, 1992)
(Accepted September 3, 1992)

초 록: 한국에서 분리된 사람 로타바이러스의 VP7 코딩 RNA 분절의 cDNA 합성과 염기서열 결정

김영봉[1], 김경희[2], 양재명[3]*([1]성균관대학교 유전공학과, [2]한양대학교 의과대학 미생물학교실, [3]*서강대학교 생물학과)

서울지역의 소아설사환자 가검물로부터 분리한 로타바이러스의 VP7을 코딩하는 RNA분절 cDNA를 합성한 후 PCR로 증폭하여 pUC19에 클론하였다. 이 cDNA의 염기서열을 결정한 뒤 외국에서 분리한 로타바이러스 혈청형 1인 WA1과 RE9의 아홉번째 RNA 분절과 비교하였더니 90% 이상의 유사성을 보였다. 염기서열로부터 유추된 아미노산 서열중 혈청형간에 변이가 많은 VR5와 VR8 지역을 비교한 결과 역시 혈청형 1인 RE9과 WA1 바이러스주와 매우 높은 유사성을 지님을 알 수 있었다.