Sequence Analysis of VP4 and VP7 Genes of Nontypeable Strains Identifies a New Pair of Outer Capsid Proteins Representing Novel P and G Genotypes in Bovine Rotaviruses

C. Durga Rao, Krishe Gowda, and B. S. Yugandhar Reddy

Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560 012, India

Received February 25, 2000; accepted June 1, 2000

During a limited epidemiological study, the serotype specificities of several isolates of bovine rotavirus, exhibiting identical electropherotypes, from a single cattle farm near Bangalore, India, could not be determined using a panel of serotyping monoclonal antibodies (MAbs) specific for G serotypes 1–6 and 10. To determine the genotypes of these isolates, the nucleotide sequences of the genes encoding the outer capsid proteins VP4 and VP7 of two representative isolates, Hg18 and Hg23, were determined. The corresponding gene sequences from the two isolates were identical, indicating that these isolates represented a single strain of bovine rotavirus. Comparison of the VP4 nucleotide (nt) and the deduced amino acid (aa) sequences with those of several human and animal rotavirus strains representing all of the currently recognized 20 different VP4 (P) genotypes revealed low nt and aa sequence identities of 61.0 to 74.2% and 57.9 to 78.2% for VP4. The percentages of amino acid homology for the VP8* and VP5* regions of VP4 were 37.7 to 67.9 and 68.1 to 84.2%, respectively. The nt and aa sequences of the VP7 gene were also distinct from those of human and animal strains belonging to the previously established 14 VP7(G) serotypes (65.9 to 75.5% nt and 59.5 to 77.6% aa identities). These findings suggest the classification of the VP4 and VP7 genes of the bovine isolates represented by Hg18 as new P and G genotypes and provide further evidence for the vast genetic/antigenic diversity of group A rotaviruses.

INTRODUCTION

Rotaviruses, members of the genus Rotavirus within the family Reoviridae, are the major etiologic agents of severe, acute dehydrating diarrhea in the young of many mammalian species, including humans and calves (Kapikian and Chanock, 1996). Group A rotaviruses are the major pathogens of humans and domestic animals and account for about 800,000 deaths in children under the age of 5 years, besides inflicting staggering economic burden worldwide (Glass et al., 1997). Rotavirus diarrhea is also a serious problem in animal farms, causing severe economic loss due to morbidity and mortality in young calves (Woode and Bridger, 1975; House, 1978; deLeeuw et al., 1980).

The rotavirus genome consists of 11 segments of double-stranded (ds) RNA which are enclosed in a triple-layered protein capsid (Estes, 1996). VP4 and VP7, the two proteins comprising the outer capsid are encoded by gene segments 4 and 7, 8, or 9 (depending on the strain), respectively (Estes and Cohen, 1989; Kapikian and Chanock, 1996). VP6, encoded by gene segment 6, constitutes the intermediate capsid and represents the subgroup (SG)-specific antigen of the virus and VP2, encoded by RNA segment 2, forming the inner capsid (Greenberg et al., 1981; Estes, 1996).

VP4 and VP7 are considered most important for rotavirus vaccine development since they elicit serotype-specific neutralizing antibody response in the infected host (Hoshino et al., 1985; Offit and Blavat, 1986). Rotaviruses have been classified into G and P serotypes based on the VP7 (glycoprotein) and VP4 (protease-sensitive) proteins (Estes and Cohen, 1989). So far, 14 G serotypes have been recognized and of these, several serotypes are shared between humans and animals (Kapikian and Chanock, 1996).

Serotypes G1–G4 are most widespread in humans; serotypes G6, G8, and G10 are major pathogens in cattle (Taniguchi et al., 1991; Hardy et al., 1992; Sato et al., 1997), but strains belonging to these serotypes are frequently detected in humans in recent years (Gerna et al., 1992; Beards et al., 1992; Browning et al., 1992; Das et al., 1993; Gentsch et al., 1993; Kapikian and Chanock, 1994; Santos et al., 1998; Cunliffe et al., 1999; Holmes et al., 1999; Jagannath et al., 2000). Serotypes other than G6, G8, or G10 are rarely observed in cattle (Hussein et al., 1993), with only a single strain each belonging to G1 and G7 serotypes characterized to date (Blackhall et al., 1992; Isegawa et al., 1994). Serotype G3 strains appear to have the broadest host range and were observed in humans and many animal species (Kapikian, 1996). Three widely separated regions, A (aa 87–101), B (aa 143–152), and C (aa 208–223), have been identified as major antigenic domains on VP7 and were suggested to form complex and functionally related and operationally
overlapping conformational epitopes that determine the serotype and neutralization specificities of rotaviruses (Dyall-Smith et al., 1986; Mackow et al., 1988; Taniguchi et al., 1988; Estes and Cohen, 1989; Hum et al., 1989; Kobayashi et al., 1991; Dormitzer et al., 1992; Dunn et al., 1993b).

Due to lack of appropriate antibody reagents, a dual P typing system (P serotypes and P genotypes) is being used to characterize rotaviruses (Estes and Cohen, 1989). Whereas only 11 P serotypes have been characterized, to date, at least 20 P genotypes have been identified (Estes, 1996; Kapikian and Chanock, 1996). Strains sharing >89% aa sequence identities are considered to belong to the same genotype (Gorziglia et al., 1990; Estes, 1996). Although a general correlation between P genotypes and P serotypes has not been clearly established, the presently characterized P serotypes have been observed to correlate to certain genotypes (Estes, 1996). In this typing system, the P serotype is represented by a number immediately after the letter P and the genotype is denoted by a number in square brackets (Estes, 1996). Thus the VP4 genes present in human strains belonging to the four major human G serotypes are represented as P1A[8] and P1B[4].

VP4 from animal strains consists of 776 amino acids, but that from the majority of the human strains contains 775 amino acids due to the absence of a residue corresponding to aa position 136 of the VP4 of animal strains (Estes, 1996). Proteolytic cleavage of VP4 by trypsin at aa positions 247 and/or 241 into the amino-terminal VP8* and carboxy-terminal VP5* subunits results in enhancement of viral infectivity (Estes, 1996). The VP8* region of VP4 exhibits the greatest sequence divergence between amino acids 71 and 204 among different VP4 types and the sequence variation between amino acids 84 and 180 has been suggested to correlate with VP4 serotype and subtype specificities (Larralde et al., 1991; Larralde and Gorziglia, 1992; Estes, 1996).

Development of an effective vaccine against rotavirus disease is partly hampered by the presence in nature of multiple serotypes, many of which are shared between humans and animals. For more than 6 years, in spite of extensive epidemiological studies in several countries, no new rotavirus G serotype has been reported, suggesting that the genetic/antigenic diversity of group A rotaviruses is limited to 14 G serotypes.

In this study, we describe characterization of nontypeable bovine strains isolated in Bangalore, India, which possessed both outer capsid proteins VP4 and VP7 representing new P and G genotypes and thus contribute to the expanding genetic diversity of group A rotaviruses.

RESULTS AND DISCUSSION

All the isolates of bovine rotavirus from the Hesaraghatta Livestock Farm showed “long” RNA elec-
nontypeable isolates represent a single strain of bovine rotavirus. Hence Hg18 was taken as the representative prototype strain of these bovine isolates for discussion in this study.

The VP7 gene of the bovine strain Hg18 is 1062 nt long and contains an open reading frame (ORF) of 981 nt stretching from nt position 49 to 1040 that encodes a protein of 326 amino acids. All the cysteines as well as prolines present in the mature VP7 in the virion of all other serotypes are conserved in Hg18, suggesting that intramolecular disulfide bonds are important for formation of the conformational antigenic epitopes on VP7. The glycosylation site found in the majority of the VP7 types between aa positions 69 and 71 is also conserved in the VP7 of strain Hg18 (Fig. 2).

**FIG. 2.** Comparison of the deduced amino acid sequence of the VP7 protein of the Indian bovine strain Hg18 with that of group A rotavirus strains representing the established 14G serotypes. Only the amino acids that differ from those of Hg18 are shown. Cysteines (●) and prolines (□) that are conserved in all the serotypes are indicated. Variable regions (VR1 to VR9), which include the major antigenic determinant regions A (aa 87–101), B (aa 143–152), and C (aa 208–223), with those of all other serotypes showed extreme divergence. On the basis of earlier predictions that strains belonging to different serotypes share 85% aa sequence conservation in VP7 as well as in the antigenic regions A and C combined (Green et al., 1988, 1989; Kapikian and Chanock, 1996) is also indicated. Note that the variable regions including the three major antigenic regions of Hg18 are highly divergent from those of the strains belonging to other G serotypes, but the numbering corresponds to that of majority of the serotypes. The Accession numbers for the VP7 gene sequences are Wa, K02033; HU5, A01028; SA11, K02028; ST3, X13603; OSU, X04613; NCDV, M12394; Ch2, M23194; L26, M58290; L338, D13549; FI23, M61876. The VP7 sequences of strains 69M and WI61 were taken from reference Green et al. (1989).

Comparison of the nt and the deduced aa sequences of VP7 gene of the bovine strain Hg18 with the VP7 sequences of strains representing the previously established 14G serotypes revealed low sequence identities with all the VP7 types (Table 1, Fig. 2). The percentages of nt and aa identity with all the 14G serotypes ranged from 65.9 to 75.5 and 59.5 to 77.6%, respectively (Table 1).
Table 1: Comparison of the nt and Deduced aa Sequence Identities of the VP7 Gene of the Bovine Strain Hg18 with Strains Representing the Established 14G Serotypes

<table>
<thead>
<tr>
<th>Strain</th>
<th>G serotype</th>
<th>nt</th>
<th>aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wa</td>
<td>1</td>
<td>74.5</td>
<td>75.2</td>
</tr>
<tr>
<td>HU5</td>
<td>2</td>
<td>72.1</td>
<td>69.6</td>
</tr>
<tr>
<td>SA11</td>
<td>3</td>
<td>75.3</td>
<td>77.3</td>
</tr>
<tr>
<td>ST3</td>
<td>4</td>
<td>72.6</td>
<td>72.7</td>
</tr>
<tr>
<td>OSU</td>
<td>5</td>
<td>74.5</td>
<td>77.0</td>
</tr>
<tr>
<td>NCDV</td>
<td>6</td>
<td>74.4</td>
<td>76.1</td>
</tr>
<tr>
<td>Ch2</td>
<td>7</td>
<td>65.9</td>
<td>59.5</td>
</tr>
<tr>
<td>69M</td>
<td>8</td>
<td>75.5</td>
<td>76.4</td>
</tr>
<tr>
<td>WI61</td>
<td>9</td>
<td>74.2</td>
<td>77.3</td>
</tr>
<tr>
<td>B223</td>
<td>10</td>
<td>72.5</td>
<td>76.4</td>
</tr>
<tr>
<td>YM</td>
<td>11</td>
<td>74.6</td>
<td>77.6</td>
</tr>
<tr>
<td>L26</td>
<td>12</td>
<td>71.9</td>
<td>73.9</td>
</tr>
<tr>
<td>L338</td>
<td>13</td>
<td>73.7</td>
<td>72.1</td>
</tr>
<tr>
<td>FI23</td>
<td>14</td>
<td>73.7</td>
<td>74.2</td>
</tr>
</tbody>
</table>

Note. For Accession numbers of VP7 gene sequences analyzed in this study, refer to the legend to Fig. 2.


The present observation that the Indian bovine strains represented by Hg18 possessed both VP4 and VP7 of new P and G type specificities signifies the increasing genetic diversity of group A rotaviruses and possible cosegregation, in nature, of yet another pair of outer capsid proteins bearing novel genotype specificities. Characterization of more nontypeable strains isolated throughout the world from humans and animals would reveal the actual extent of the genetic/antigenic diversity of group A rotaviruses.

The present findings are of epidemiological significance considering the age-old traditions and close association of the majority of the population with cattle under rural settings in India. Prevalence of this bovine rotavirus with new P and G type specificities in cattle as well as its possible transmission to humans in the context of the frequent detection of rotaviruses belonging to bovine serotypes in humans (Dunn et al., 1993a; Das et al., 1993; Gentsch et al., 1993; Jagannath et al., 2000) remains to be determined. Though the significantly low aa sequence identities of VP4 and VP7 (including that of antigenic regions A and C combined) of Hg18 with those of all other established serotypes/genotypes meet the previously predicted criteria (Green et al., 1988; Kapikian and Chanock, 1996) for assigning new P and G serotypes, classification of the Indian bovine strains represented by Hg18 as a new G serotype awaits confirmation by reciprocal neutralization assays.
FIG. 3. Comparison of the deduced amino acid sequence of the VP4 protein of the new bovine strain Hg18 with the VP4 sequences of strains representing all the currently identified 20 different P genotypes (Estes, 1996). The highly conserved cysteines (F) and prolines (n) are indicated. The putative fusion region between aa positions 385 and 404 is indicated. For optimal alignment, VP4-specific gaps were introduced in the sequences. Note the high degree of sequence variation in the P serotype determining VP8* region from aa positions 71 to 204, suggesting to a new P type specificity of the VP4 of the bovine strains represented by Hg18. The Accession numbers of the VP4 sequences analyzed in this study are NCDV, P17 465; SA11, X14204; HCR3, P17 465; RV5, M32559; UK, M22306; 1076, M88480; Gottfried, M33516; OSU, X13190; KU, M21014; AU1, D10970; 69M, M60600; KK3, D14367; EDIM, AF039219; 993/83, D16352; L338, D13399.
MATERIALS AND METHODS

Viruses, virus isolation, and adaptation to cell culture

Fifty fecal samples from young calves below the age of 1 month, suffering from severe diarrhea, were collected from the Composite Livestock Farm and Research Institute, Hesaraghatta, near Bangalore, India. The stool specimens were processed as described earlier (Sukumar et al., 1992) and the supernatants were stored at −20°C. Ten specimens, representing 20% of the samples (Hg2, Hg3, Hg6, Hg8, Hg16, Hg18, Hg21, Hg23, Hg27, and Hg28), were found to be positive for rotavirus by FIG. 3—Continued
polyacrylamide gel electrophoresis of the viral genomic dsRNA. Of these, two isolates, Hg18 and Hg23, were adapted to growth in MA104 cells in presence of trypsin (Ward et al., 1984). The established rotavirus strains Wa (SGI, P1A[8], G1), S2 (SGI, P1B[4], G2), 1040 (SGI, P[4], G2 subtype), RRV (SGI, P5[3], G3), ST3 (SGII, P2A[6], G4), NCDV (SGI, P6[1], G6), MP409 (SGI, P6[1], G8), WI61 (SGI, P1A[8], G9) and I321 (SGI, P8[11], G10) were also grown in MA104 cells and were used as controls in subgrouping and serotyping ELISAs. Strain 1040 belongs to the G2 genotype with no specific reactivity to serotype G2-specific mAbs (unpublished data).
Extraction and electrophoresis of viral genomic dsRNA

Viral genomic dsRNA from stool specimens and cell culture-grown viruses was extracted with phenol–chloroform and precipitated with isopropanol as described previously (Sukumaran et al., 1992). Viral dsRNA was purified by binding to CC41 cellulose matrix (Whatman) as described earlier (Das et al., 1993). Methods for electrophoresis of the viral dsRNA in polyacrylamide gels and detection by staining with silver nitrate were previously described (Herring et al., 1982; Sukumaran et al., 1992).

Subgroup and serotype analysis

Subgrouping ELISA and serotyping ELISA were carried out as described earlier (Greenberg et al., 1983; Noriega et al., 1990; Sukumaran et al., 1992). In subgrouping ELISA, hyperimmune anti-RRV antiserum R2, the SGI-specific monoclonal antibody (MAb) 255/60 and SGII-specific MAb 631/9 were used (Greenberg et al., 1983). MAbs specific for G1–G6 and G10 used in serotyping ELISA were reported earlier (Noriega et al., 1990; Aijaz et al., 1996). All the antibodies were generously provided by Dr. H. B. Greenberg (Stanford University School of Medicine, Stanford, CA).

Cloning of cDNAs of the VP4 and VP7 genes of the bovine isolates Hg18 and Hg23

Viral genomic dsRNAs purified using CC41 cellulose were used for first-strand cDNA synthesis by AMV-reverse transcriptase and ds cDNA was synthesized by polymerase chain reaction (PCR) using Taq DNA polymerase as described previously (Rao et al., 1995). Gene-specific primers were used in cDNA synthesis and PCR amplification. The VP4 gene-specific 5’ and 3’ primers were 5’-CTAAGCTTCCCGGGCTATAAAATGC/GC/GTTC-3’ and 5’-CTAAGCTTCCCGGGTCACATCC/TT-3’, respectively. The respective sequences for VP7 gene primers were 5’-CTTCCGGGGCTTAAAAGA/CGAGAAT-3 and 5’-CTTCCGGGGTCACA/GT/GCTT-3’, respectively.

### TABLE 2

<table>
<thead>
<tr>
<th>Strain</th>
<th>P serotype/genotype/G serotype</th>
<th>Percentage of identity with Hg18</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCDV</td>
<td>P8[1], G6</td>
<td>70.0/76.2/66.1/81.2</td>
</tr>
<tr>
<td>SA11</td>
<td>P8B[2], G3</td>
<td>73.5/76.8/65.7/82.5</td>
</tr>
<tr>
<td>HCR3</td>
<td>P8[3], G3</td>
<td>74.0/78.2/67.9/83.5</td>
</tr>
<tr>
<td>RV5</td>
<td>P1B[4], G2</td>
<td>70.4/69.8/55.0/77.0</td>
</tr>
<tr>
<td>UK</td>
<td>P7[5], G6</td>
<td>70.0/72.8/57.8/80.0</td>
</tr>
<tr>
<td>1076</td>
<td>P2A[6], G2</td>
<td>68.4/70.9/55.7/77.9</td>
</tr>
<tr>
<td>Gottfried</td>
<td>P2B[8], G4</td>
<td>70.2/71.4/57.7/77.8</td>
</tr>
<tr>
<td>OSU</td>
<td>P9[7], G5</td>
<td>71.0/75.1/65.0/80.2</td>
</tr>
<tr>
<td>KU</td>
<td>P1A[8], G1</td>
<td>70.6/70.5/55.3/77.6</td>
</tr>
<tr>
<td>AU1</td>
<td>P3[9], G3</td>
<td>65.9/65.5/54.1/70.9</td>
</tr>
<tr>
<td>69M</td>
<td>P4[10], G8</td>
<td>73.3/78.0/65.0/84.2</td>
</tr>
<tr>
<td>KK3</td>
<td>P8[11], G10</td>
<td>61.0/57.9/37.7/68.1</td>
</tr>
<tr>
<td>H2</td>
<td>P4[12], G3</td>
<td>72.7/75.1/60.9/82.1</td>
</tr>
<tr>
<td>MDR13</td>
<td>P7[13], G3/G5</td>
<td>71.5/74.6/59.2/82.7</td>
</tr>
<tr>
<td>Mc35</td>
<td>P11[14], G10</td>
<td>66.9/66.0/54.9/71.1</td>
</tr>
<tr>
<td>LP14</td>
<td>P7[15], G10</td>
<td>74.2/78.0/67.1/83.7</td>
</tr>
<tr>
<td>EDIM</td>
<td>P10[16], G3</td>
<td>66.7/70.1/56.5/76.7</td>
</tr>
<tr>
<td>993/83</td>
<td>P7[17], G7</td>
<td>64.8/60.0/41.4/69.3</td>
</tr>
<tr>
<td>L338</td>
<td>P12[18], G13</td>
<td>72.6/73.7/61.6/79.5</td>
</tr>
<tr>
<td>4F</td>
<td>P7[19], G5</td>
<td>71.0/72.3/57.3/79.3</td>
</tr>
<tr>
<td>EHF</td>
<td>P7[20], G3</td>
<td>69.2/74.5/64.6/78.7</td>
</tr>
</tbody>
</table>

Note. Accession numbers for gene 4 sequences used in this study are given in the legend to Fig. 3.

Extraction and electrophoresis of viral genomic dsRNA

Viral genomic dsRNA from stool specimens and cell culture-grown viruses was extracted with phenol–chloroform and precipitated with isopropanol as described previously (Sukumaran et al., 1992). Viral dsRNA was purified by binding to CC41 cellulose matrix (Whatman) as described earlier (Das et al., 1993). Methods for electrophoresis of the viral dsRNA in polyacrylamide gels and detection by staining with silver nitrate were previously described (Herring et al., 1982; Sukumaran et al., 1992).

Subgrouping and serotype analysis

Subgrouping ELISA and serotyping ELISA were carried out as described earlier (Greenberg et al., 1983; Noriega et al., 1990; Sukumaran et al., 1992). In subgrouping ELISA, hyperimmune anti-RRV antiserum R2, the SGI-specific monoclonal antibody (MAb) 255/60 and SGII-specific MAb 631/9 were used (Greenberg et al., 1983). MAbs specific for G1–G6 and G10 used in serotyping ELISA were reported earlier (Noriega et al., 1990; Aijaz et al., 1996). All the antibodies were generously provided by Dr. H. B. Greenberg (Stanford University School of Medicine, Stanford, CA).

Cloning of cDNAs of the VP4 and VP7 genes of the bovine isolates Hg18 and Hg23

Viral genomic dsRNAs purified using CC41 cellulose were used for first-strand cDNA synthesis by AMV-reverse transcriptase and ds cDNA was synthesized by polymerase chain reaction (PCR) using Taq DNA polymerase as described previously (Rao et al., 1995). Gene-specific primers were used in cDNA synthesis and PCR amplification. The VP4 gene-specific 5’ and 3’ primers were 5’-CTAAGCTTCCCGGGCTATAAAATGC/GC/GTTC-3’ and 5’-CTAAGCTTCCCGGGTCACATCC/TT-3’, respectively. The respective sequences for VP7 gene primers were 5’-CTTCCGGGGCTTAAAAGA/CGAGAAT-3’ and 5’-CTTCCGGGGTCACA/GT/GCTT-3’.
AC/CATACA-3′. All the primers had sites for restriction enzymes of convenience at the 5′ end. Rotavirus genespecific sequences in the primers are underlined. The PCR-amplified DNAs were digested with appropriate restriction endonucleases and cloned into either pUC18 or pBluescript (KS+) vectors which were also digested with the corresponding enzymes.

Nucleotide sequencing and comparative sequence analysis

Nucleotide sequences of the cloned VP4 and VP7 genes were determined by the dideoxy nucleotide-mediated chain termination method (Sanger et al., 1977). At least two clones for each gene from both Hg18 and Hg23 strains obtained from two independent PCR products were selected for nucleotide sequence determination. Vector-specific primers as well as gene-specific internal primers were used for determining the complete nucleotide sequences of the two genes. Sequences of both the strands were determined. The nucleotide and deduced amino acid sequences of the VP4 and VP7 genes of both strains Hg18 and Hg23 were analyzed and compared with the previously published corresponding rotavirus gene sequences representing all the currently established serotypes/genotypes. Phylogenetic analysis was carried out using the Distances and Growtree programs of the GCG-package, Wisconsin. Since the sequences of the corresponding genes from both isolates of both strains Hg18 and Hg23 were identical, the VP4 and VP7 gene sequences of strain Hg18 only were deposited in the GenBank and the respective Accession numbers are AF237665 and AF237666.

ACKNOWLEDGMENT

This work was supported by a grant from the Department of Biotechnology, Government of India, under the Indo–U.S. Vaccine Action Programme.

REFERENCES


Hardy, M. E., Gorzigilia, M., and Woode, G. N. (1992). Amino acid sequence analysis of bovine rotavirus B233 reveals a unique outer...


