55. Fluit, A. C., Torensma, R., Visser, M. J. C., Aarsman, C. J. M., Poppelier, M. J. J. G., Keller, B. H. I., Klapwijk, P. and Verhoee, J., *Appl. Land Environ. Microbiol.*, 1993, **59**, 1289–1293.
56. Herman, L., Ridder, H. and DeRidder, H., *Neth. Milk Dairy J.*, 1993, **47**, 23–29.
57. Salzano, G., Coppola, R., Maddonni, M. F., Pallotta, M. L., Villani, F. and Sorlini, C., *Ann. Microbiol. Enzimol.*, 1993, **43**, 159–163.
58. Higuchi, R., Fockler, C., Dollinger, G. and Watson, R., *Biotechnology*, 1993, **11**, 1026–1030.
59. Kitchin, D. A., Szotyori, Z., Fromholc, C. and Almond, N., *Nature*, 1990, **344**, 201.
60. Kwok, S. and Higuchi, R., *Nature*, 1989, **339**, 237–238.
61. Sarkar, G. and Sommer, S. S., *Nature*, 1990, **343**, 27.
62. Furrer, B., Candrian, U., Huefelein, C. and Leuthy, J., *J. Appl. Bacteriol.*, 1991, **70**, 372–379.
63. Johnson, W., Tyler, S., Ewan, E., Ashton, F., Wang, G. and Rozee, K., *Microbiol. Pathogen*, 1992, **12**, 79.
64. Bohnert, M., Dilasser, F., Dalet, C., Mengaud, J. and Cossant, P., *Res. Microbiol.*, 1992, **143**, 271.

Received 4 May 1996; revised accepted 19 August 1996

RESEARCH ARTICLE

Comparative sequence analysis and expression in *E. coli* of the subgroup I-specific antigen VP6 from a G2 serotype human rotavirus IS2

Saima Aijaz* and C. Durga Rao*†

*Department of Microbiology and Cell Biology, †Centre for Genetic Engineering, Indian Institute of Science, Bangalore 560 012, India

VP6, the intermediate capsid protein of the virion, specifies subgroup specificity of rotavirus. It is also the most conserved, both at nucleotide and amino acid levels, among group A rotaviruses and is the target of choice for rotavirus detection. In this study we report the sequence of the subgroup I (SGI)-specific VP6 from the serotype G2 strain IS2 isolated from a child suffering from acute diarrhoea in Bangalore and its comparison with the published VP6 sequences. Interestingly, IS2 gene 6 shared highest homology with that from bovine UK strain and the protein contained substitutions by lysine at amino acid positions 97 and 134. In contrast, the amino acids Met and Glu/Asp at these respective positions are highly conserved in all the other group A rotaviruses sequenced so far. These observations have obvious implications for the evolution of serotype G2 and G2-like strains circulating in India. The SGI VP6, of a human rotavirus, possessing epitopes that are conformationally similar to those found in the native protein in the virion, was successfully expressed in *E. coli* and purified for the first time by single-step affinity chromatography.

DIARRHOEAL diseases are the major cause of morbidity and mortality among infants and young children, especially in developing countries. Of the many diarrhoeal agents, rotavirus is the single most important cause of severe, acute infantile gastroenteritis in humans and a variety of domestic animals[1]. Rotavirus diarrhoea occurs throughout the year and is estimated to account for about a million deaths annually among young children, thus representing an important public health problem[1,2]. In spite of the staggering economic burden, an effective vaccine against rotavirus diarrhoea is yet to be developed.

Rotavirus belongs to the family Reoviridae and consists of a triple-layered capsid[3]. The outer shell consists of two proteins, VP4 and VP7, the intermediate shell comprises of VP6 and the inner shell is composed of VP2 that encloses a genome of 11 segments of double-stranded (ds) RNA[3,4]. The genome encodes 6 structural

and at least 5 nonstructural proteins[5]. The outer capsid proteins VP7 and VP4 specify two independent serotype specificities, the G and P types respectively[6]. VP6, encoded by gene segment 6, contains group as well as subgroup-specific epitopes[6,7]. On the basis of group-specific epitopes, 7 groups A to G have been identified in humans and animals[5]. Group A rotaviruses, the most common pathogens in humans, can be further subdivided into at least 4 subgroups SGI, SGII, SGI + II and non SGI/II on the basis of the SG-specific epitopes[8]. Group A rotaviruses can also be initially characterized as 'short' or 'long' electropherotypes depending on the slower or faster electrophoretic migration, respectively of the dsRNA segment 11 in polyacrylamide gels[1]. In general, majority of the human rotaviruses with SGI specificity belong to G2 serotype and exhibit 'short' RNA electropherotype whereas those with SGII specificity have 'long' RNA pattern and belong to other serotypes[1]. In contrast, the great majority of animal strains appear to possess 'long' RNA pattern but SGI specificity[1,9].

VP6 comprising more than 50% of the virion mass is an important viral antigen and is involved in several viral functions such as replication, transcription and viral morphogenesis[5,8,10,11]. Although majority of the serum antibodies in the infected host are directed against VP6, anti-VP6 IgG antibodies have not been conclusively shown to be capable of virus neutralization[8]. Recent studies, however, indicate that mucosal anti-VP6 secretory IgA antibodies play an important role in protection against the rotavirus disease[12]. VP6 also contains epitopes recognized by cytotoxic T-lymphocytes as well as helper T cells, indicating a role for VP6 in stimulating heterotypic cell-mediated immune response[13–15]. VP6 was shown to function as an excellent immunological carrier for peptides and proteins in vaccine development[16]. Because of its high abundance, stability and high degree of conservation among human and animal group A rotaviruses, majority of the procedures for detection of rotaviruses in clinical samples are based on VP6 (refs 7, 17, 18).

Although the SGI-VP6 from a few animal strains had been expressed[19–12], that from a human rotavirus has not been reported. Because of the importance of VP6 in viral diagnosis and to determine the sequence variation in VP6 from Indian rotaviruses, we report here the cloning, comparative sequence analysis and expression in *E. coli* of the SGI VP6 from an Indian G2 serotype strain IS2 isolated from a child with diarrhoea in Bangalore.

Materials and methods

Extraction of viral genomic dsRNA and purification of the RNA segments

Isolation and serotypic characterization of rotaviruses including the G2 serotype strain IS2 isolated from children suffering from diarrhoea admitted to various hospitals in Bangalore has been described earlier[18]. As IS2 was not adapted to growth in culture, nucleic acids were extracted directly from the clarified supernatants of the 20% suspension of the stool sample as previously described[18]. The dsRNA was electrophoresed on a 1% agarose gel in presence of ethidium bromide and the RNA segments 5, 6, 7, 8 and 9 were electroeluted together onto a dialysis membrane and purified by phenol–chloroform extraction[22]. For identification of gene 6-specific cDNA clones, gene 6 RNA segment was separately purified from the agarose gel.

cDNA synthesis, cloning and identification of VP6 gene-specific clones

In vitro polyadenylation of the denatured dsRNAs, synthesis of cDNA on the oligo(dT)-tailed plasmid (pCDV) primer, construction of cDNA library in *E. coli HB101* were described previously[22–25]. The cDNA library, constructed for segments 5 to 9, was screened with ^{32}[P]-labelled mixed cDNA probe for segments 5, 6, 7, 8 and 9. Gene 6-specific cDNA clones were identified by Southern blot hybridization of the BamHI-digested plasmid DNAs with either the cDNA probe prepared from purified RNA segment 6 or the RNA probe prepared by labelling at the 3′ end using *E. coli* poly(A)-polymerase[25]. Several clones containing inserts ranging in size from 600 to 1500 nucleotides (nt) were identified. Clones OB67 having an insert of 1.0 kb and OB48, OB49 and OB68 containing inserts of 1.5 kb were used for further analysis. Although the reported length of gene 6 from group A rotaviruses is 1356 nt, the observed size of 1.5 kb can be attributed to the presence of poly(dA) and poly(dG)-tails of variable length at the 3′ and 5′ ends of the cDNAs respectively.

Nucleotide sequence analysis

The complete nucleotide sequence of gene 6 was determined from partial and full length cDNA clones. The cDNA inserts from the original clones were subcloned at BamH I site of pBluescript KS$^+$ (pBSKS$^+$) vector (Stratagene, CA, USA). From pBS67, subclones were generated utilizing internal sites for restriction endonucleases Nhe I, Pst I and Xba I. Sequence of both strands of the inserts in the subclones was determined using KS, SK and gene-specific primers. Nucleotide sequence was determined by dideoxynucleotide-mediated chain termination method[26] using sequenase version 2.0. Sequence near the 5′-end of the inserts in the original clones was determined using Okayama and Berg vector-specific primer. Nucleotide and the deduced amino acid sequences were analysed using version 6.1 of the GCG application software at the Distributed Information

Centre, Indian Institute of Science. The sequences of the oligonucleotide primers are: 5′-ATCACAACCAGCT CATGAT-3′ from nt position 527 to 545; 5′-TTAACTACAGCTACAAT-3′ from nt position 675 to 691; 5′-GAAGTGTTACTTCTGCTCT-3′ (5′ primer for Okayama and Berg vector).

Expression in E. coli and purification of VP6

For VP6 expression, the T7-promoter – and the polymerase – based expression system was used[27]. As the complete gene 6 cDNA contained 5′ and 3′ untranslated sequences, the gene 6 ORF was amplified by polymerase chain reaction (PCR)[28] using primers corresponding to the 5′ and 3′ ends of the ORF. The primers contained sites for restriction enzymes of convenience. The sequence of the 5′ primer is 5′-GATATCAAGCTT CCCGGGATGGATGTCCTGTACTC-3′ and that of the 3′ primer is 5′-CTGCAGAAGCTTTTTGACAAGCAT GCTTCT-3′. Gene 6 specific sequence is underlined. The PCR-amplified DNA was digested with Hind III and inserted at the Hind III site of pET20b(+) vector[27] (Novagen, Madison, WI, USA). The resulting construct is named as pETG6. To express VP6 without the pelB leader sequence, pETG6 was digested with Nde I and Sma I and religated after blunting the Nde I end by Klenow fill-in reaction. Deletion of the pelB leader sequence as well as the stretch of 17 aa downstream of it (upstream of the gene 6 AUG codon) brings the gene 6 AUG codon to within the requisite distance (10 nt in this case) from the Shine-Dalgarno (SD) sequence[29]. This construct was represented as pETNDG6. *E. coli HB101* cells were transformed with the ligated DNA and the plasmid DNAs from positive clones were used to transform *E. coli* BL21 (DE3)[27]. Expression of VP6 in pETG6 and pETNDG6 recombinants was examined by analysis of the cell lysates by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) as previously described[30]. The recombinant VP6 was purified by single-step affinity chromatograpy using Ni^{2+}-NTA-agarose as described earlier[24].

Production of polyclonal antiserum and immunoblot analysis

About 500 µg of the recombinant VP6 protein purified from the Ni^{2+}-NTA-agarose column was electrophoresed on a 12% preparatory SDS-polyacrylamide gel. The 46.4 kDa band corresponding to VP6 was excised after staining with Coomassie blue and homogenized in 1 ml of phosphate-buffered saline. Polyclonal antibodies were raised by injecting the gel suspension into a New Zealand white rabbit subcutaneously at multiple sites as previously described[22]. Pre-immune serum collected from the same rabbit before immunization was used as control.

Purity of the column-purified recombinant VP6 as well as the specificity of the polyclonal antiserum was determined by immunoblot analysis of the protein[31]. Western blots containing either the total cell lysate or the affinity-purified protein were first incubated with either the anti-VP6 antiserum or the SGI-specific mAb 255/60 and then with the secondary antibodies, goat anti-rabbit IgG or goat anti-mouse IgG conjugated with the horse radish peroxidase, respectively. For detection of VP6 by the SGI mAb, the gel was soaked in renaturation buffer (50 mM Tris.Cl pH 7.4, 20% glycerol) for 20 min prior to blotting onto nitrocellulose membrane. The antigen antibody interaction was detected by colour development in citrate buffer containing 3,3-diaminobenzidine tetrahydrochloride and H_2O_2. The SGI-specific mAb 255/60 was originally produced using rhesus rotavirus (RRV) as the immunizing virus[7] and provided by Dr Harry B. Greenberg, Stanford University, USA.

Immunoprecipitation

Individual colonies of BL21 (DE3) containing either pETG6 or pETNDG6 were grown overnight in M9 medium. The cultures were then inoculated into fresh sulfate-free M9 medium at 100-fold dilution and grown until the OD_{600} reached 0.4. The cells were then induced with 0.4 mM IPTG in presence of 10 µCi of ^{35}[S]-methionine per ml for 10 min at 37°C. The cells were harvested, lysed and inclusion bodies were prepared as described earlier[24]. The inclusion bodies were dissolved in 0.1 M Tris.Cl pH 8.5 buffer containing 8 M urea by incubating at room temperature for 30 min. Urea was removed from the lysate by centrifugation through a centricon 30 column at 5000 rpm for 30 min. The protein solution remaining in the column was diluted with a buffer containing 10 mM Tris.Cl pH 8.0 and 100 mM NaCl. The process was repeated 3 to 4 times and the final fraction remaining in the column was used for immunoprecipitation. About 50 µl of the protein solution was diluted to 300 µl with RIPA buffer and incubated on ice with 5 µl of mAb 255/60 for 1 h after which 200 µl of protein A-sepharose CL-4B suspension (5 mg of dry gel) was added. The radioactively labelled recombinant VP6 bound to the resin was analysed on a 12% SDS-polyacrylamide gel and the bands were detected by autoradiography as previously described[32].

Results

Nucleotide sequence analysis of gene 6 from the symptomatic G2 serotype Indian strain IS2 revealed that the

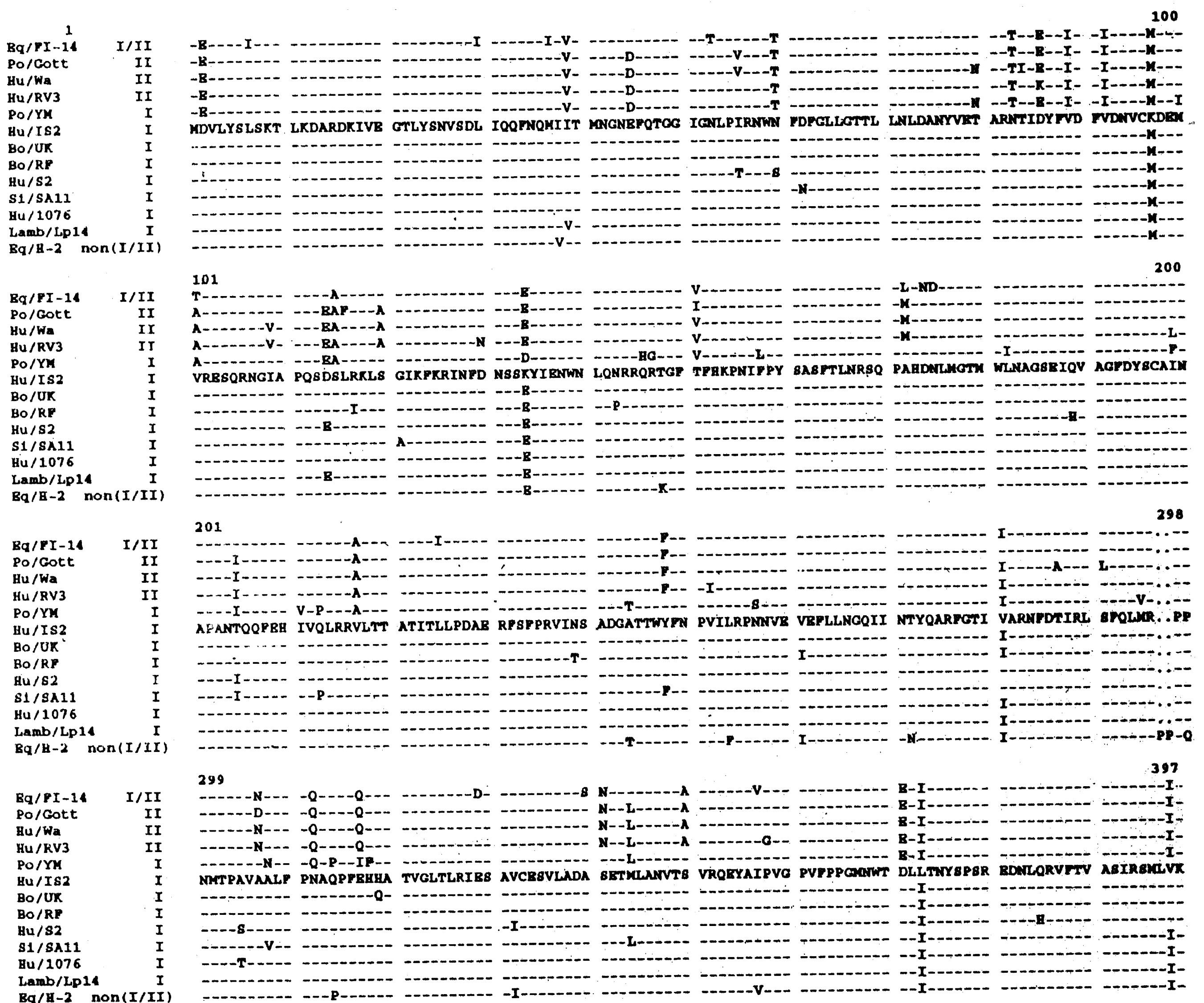

Figure 1. Comparison of the deduced amino acid sequence of IS2 VP6 with VP6 from other group A rotaviruses. Species origin of the rotavirus strains and the SG specificity of each of the strains are indicated on the left side of the corresponding VP6 sequence. Only the aas that differ from IS2 VP6 are shown. The numbering corresponds to the VP6 from majority of the strains that contain 397 aas. VP6 from the equine strain H2 is 399 aas in length and contains a two-aa insertion at position 296. The IS2 gene 6 sequence was submitted to EMBL database with accession number X94617.

Table 1. Per cent nucleotide and amino acid sequence identities of IS2 VP6 with VP6 from other group A rotaviruses

Rotavirus strain	Species origin	SG	nt	aa
IS2	Human	I	–	–
UK	Bovine	I	94.48	98.74
RF	Bovine	I	93.84	97.73
SA11	Simian	I	87.30	97.23
S2	Human	I	86.74	97.23
1076	Human	I	86.60	98.49
LP14	Lamb	I	85.51	98.24
H2	Equine	nonI/II	82.31	96.47
FI-14	Equine	I/II	79.92	91.44
Gottfried	Porcine	II	79.82	92.44
Wa	Human	II	79.10	91.18
YM	Porcine	I	78.78	90.43

gene was 1356 nt in length. The gene contained a 5′ untranslated region (UTR) of 23 nt followed by a long ORF with an AUG codon from position 24 to 26 and a termination codon from position 1215 to 1217. The ORF is followed by a 3′ UTR of 139 nt. The ORF codes for a polypeptide of 397 aa with an apparent molecular weight of 44.87 kDa which is similar to that from other group A rotaviruses[5] (Figure 1). The initiation codon at nt position 24 has the optimal sequence context for a strong initiation codon[33]. There is no polyadenylation signal, AAUAAA, in the 3′ UTR which is a characteristic of the genomes of the members of the family Reoviridae[34].

Comparison of the nt and the deduced aa sequences of gene 6 of IS2 with those of other strains showed a high