Sequence analysis and \textit{in vitro} expression of genes 6 and 11 of an ovine group B rotavirus isolate, KB63: evidence for a non-defective, C-terminally truncated NSP1 and a phosphorylated NSP5

S. Shen,$^{1,2}$ T. A. McKee,$^2$ Z. D. Wang,$^3$ U. Desselberger$^2$ and D. X. Liu$^1$

$^1$ Institute of Molecular Agrobiology, 1 Research Link, The National University of Singapore, Singapore 117604
$^2$ Clinical Microbiology and Public Health Laboratory and Division of Virology, Department of Pathology, University of Cambridge, Cambridge CB2 2QW, UK
$^3$ Xinjiang August 1st Agricultural University, Xinjiang, People's Republic of China

An ovine group B rotavirus (GBR) isolate, KB63, was isolated from faeces of a young goat with diarrhoea in Xinjiang, People’s Republic of China. Sequence determination and comparison of genes 6 and 11 with the corresponding sequences of GBR strains ADRV and IDIR showed that they were the cognate genes encoding NSP1 and NSP5, respectively. While the overall identities of nucleotide sequences between these two genes and the corresponding genes of strains ADRV and IDIR were in the range 52\%–57\%, the identities of deduced amino acid sequences were only 34\%–46\%. These results demonstrate that the substantial diversity of NSP1 observed among group A rotaviruses (GAR) also exists within GBRs and that a high degree of diversity also exists among NSP5 of GBRs, in contrast to GAR NSP5. The NSP1 gene of KB63 contains three ORFs, whereas the NSP1 genes of other GBR strains contain only two. ORFs 2 and 3 of the KB63 gene may be derived from a single ORF corresponding to ORF2 of other GBR strains by the usage of a stop codon created by an upstream single base deletion and single point mutations. \textit{In vitro} expression studies showed that ORFs 1 and 2, but not 3, of gene 6 can be translated, suggesting that ORF2 may encode a C-terminally truncated, potentially functional product. It may play a role, together with the product of ORF1, in virus replication, as the virus can be passaged further in kids. Similarly, gene 11 can be translated \textit{in vitro}. Like its counterpart in GARs, the protein encoded by gene 11 was shown to be phosphorylated \textit{in vitro}.

Introduction

Rotaviruses are one of the major aetiological agents causing acute gastroenteritis in children and young animals (Estes, 1996). They constitute a genus within the family \textit{Reoviridae} and contain a segmented, double-stranded (ds) RNA genome. Six of the 11 segments encode six structural proteins (VP1–VP4, VP6 and VP7), and the remainder encode at least five non-structural proteins (NSP1–NSP5). The unique structure of rotaviruses consists of two icosahedral outer layers and a core. On the basis of genetic and antigenic studies, rotaviruses are classified at present into six groups (A to F). Groups A, B and C have been found in both humans and animals, whereas groups D, E and F have been found only in animals. The group A rotaviruses (GARs) are further divided into serotypes, genotypes and genogroups determined by epitopes on the structural protein VP7 (G type) and by sequence diversity of the structural protein VP4 (P type) and the non-structural protein NSP1 (genogroups) (Dunn \textit{et al.}, 1994; Palombo & Bishop, 1994; Kojima \textit{et al.}, 1996). The genetic and antigenic diversity among non-A rotavirus groups has not been clearly defined due to limited data.

Recently, the properties and functions of the non-structural proteins of GARs have begun to be elucidated. The NSP1
sequence is highly variable among GAR strains, especially those from different host species. However, NSP1 contains a cysteine-rich region in its N terminus that is conserved among GARs and group C rotaviruses (GCRs), suggesting the presence of a zinc finger motif and its essential role for the function of this protein. There is evidence suggesting that part or almost all of NSP1 of GAR is not essential for virus replication (Hua & Patton, 1994; Taniguchi et al., 1996), as mutant viruses carrying genes encoding C-terminally truncated NSP1 proteins can grow in cell cultures. To date, non-defective C-terminal deletion of a large region has only been detected in NSP1 from GARS. NSP5 is a serine- and threonine-rich protein and has been shown to be an O-glycosylated (Gonzalez & Burrone, 1991) and phosphorylated protein (Welch et al., 1989; Afrikanova et al., 1996; Poncet et al., 1997). It is highly conserved among different GAR strains. NSP1, as well as NSP2 and NSP3, are RNA-binding proteins associated with replicative intermediates (Hua & Patton, 1994). NSP5 was shown to be a kinase that autophosphorylates (Blackhall et al., 1997). However, their roles in virus replication remain to be clarified.

Group B rotavirus (GBR) infections in humans have been reported (Hung et al., 1984; Bai & Shen, 1985; Eiden et al., 1985) and the viruses have been isolated from a variety of animal species. To our knowledge, only one strain was adapted to grow in cell culture (Sanekata et al., 1996) and a few can be passaged in young animals (Vonderfecht et al., 1984; Wang et al., 1989). This is an obstacle in studying the molecular biology of GBRs. Currently, identification of genes of GBRs depends largely on sequence comparison with GAR strains and between GBR strains, RNA–RNA hybridization assays and analysis of proteins translated in vitro. The assignment of gene 6 (or 7) to NSP1 and gene 11 to NSP5 was based on the similarity of deduced amino acid sequences of GBRs to those of GARs; its validity needs to be confirmed. In contrast to the NSP1 genes of GARs and GCRs, the equivalent gene of GBRs contains two ORFs instead of one, although ORF1 has not been shown to be translatable even in in vitro translation systems (Eiden & Allen, 1992). The ORF2-encoded protein shares the highest similarity with NSP1 among viral proteins of GBRs. The deduced amino acid sequences encoded by ORFs 1 and 2 appear not to contain a zinc finger motif(s), commonly present in NSP1 of GARs and GCRs and assumed to be essential for the function of the protein. This raises an immediate question about the properties and function of the NSP1 gene of GBRs. NSP5 of GBRs has not been shown to be a phosphoprotein. However, its deduced amino acid sequence shows that it is a serine/threonine-rich protein, typical of NSP5 of GARs, despite it being only distantly related.

In this report, we have determined the complete sequences of genes 6 and 11 of an ovine GBR isolate, KB63. Sequence comparisons showed that they are cognate genes of GAR and GBR genes encoding NSP1 and NSP5; however, they share low similarity even with the equivalent genes of GBRs. Sequence analysis showed that gene 6 contains three overlapping ORFs. Of them, only ORFs 1 and 2 were translatable in a cell-free translation system. As ORFs 2 and 3 may be derived from a single ORF corresponding to ORF2 of the NSP1 gene of other GBR strains, the product encoded by ORF2 represents a C-terminally truncated form of the GBR NSP1 protein. Furthermore, as KB63 can be passaged further in young goats, this truncated protein may be non-defective. We also demonstrate that in vitro expression of gene 11 resulted in the phosphorylation of the in vitro-translated polypeptides. In view of the difficulty of adapting GBRs in a cell culture system and the relative rarity of the GBR isolates reported so far, this study has added useful information to our current understanding of the sequence diversity and functions of the non-structural proteins encoded by GBR genes 6 and 11.

Methods

■ Viruses. An ovine GBR, KB63, was isolated from a young goat with diarrhoea in Xinjiang, People's Republic of China, in 1986 (Wang et al., 1989). The virus was amplified by serial passage in newborn kids. Faecal specimens were suspended in PBS and pre-clarified. Partially purified virus was obtained by repeated centrifugation through 30% sucrose cushions at 130000 g for 2 h at 4 °C. The SA11 strain of GAR was grown in MA-104 cells as described previously (Shen et al., 1994).

■ Extraction and PAGE of genomic dsRNA. A partially purified virus suspension (0.5 ml) was used for extraction of dsRNA (Rodger & Holmes, 1979). The genomic segments were separated on 10% polyacrylamide gels and visualized by silver staining (Follett et al., 1984). Viral RNA was further purified by using silica particles (GeneClean II) as described previously (Boom et al., 1990) for amplification of the viral genome.

■ Amplification of viral genes by RT–PCR. A sequence-independent strategy developed by Lambden et al. (1992) was used to amplify the full-length dsRNA fragments of KB63. Briefly, 3’-blocked (ami-}
and extension at 72 °C for 4 min (the extension time on the final cycle was increased to 10 min). After the sequences were obtained for both genes, RT–PCR was also performed, as described above except that the extension time was decreased to 2 min, with primers specific for genes 6 and 11 of KB63. The location and orientation of the specific oligonucleotides used are shown in Fig. 2.

- **Cloning of full-length cDNA.** The amplified KB63 gene segments were diluted ten-fold and 1 μl diluted products (500 ng) was ligated to 50 ng linearized pCR II vector (Invitrogen) according to the manufacturer’s instructions. The ligated products were transformed into *E. coli* strain INVαF and colonies were screened by standard procedures (Sambrook et al., 1989).

- **DNA sequencing and analysis.** Dideoxynucleotide chain-termination sequencing with a Sequenase version 2.0 kit (United States Biochemical) or automated sequencing of both strands was carried out with SP6 and T7 primers. Direct sequencing of PCR products was carried out with primers specific for genes 6 and 11 of KB63 as described previously (Shen et al., 1994). Sequence data from gel reading was assembled by using the Staden sequence analysis program (Staden, 1982, 1984). Further analyses were carried out with the GCG suite of programs.

- **Northern blot hybridization.** Northern blot hybridization was performed by using either 32P-labelled probes prepared by reverse transcription of dsRNA of KB63 as described above or DIG-labelled probes prepared by transcription of plasmids pCR42 and pCR39, which contain full-length cDNAs of genes 6 and 11, respectively. DIG-labelled probes were prepared by using the DIG RNA labelling kit (Boehringer Mannheim) according to the manufacturer’s protocol. After PAGE, dsRNA was transferred to nylon membrane (Hybond-N, Amersham) and Northern blotting was carried out as described previously (Lambden et al., 1992).

- **Construction of plasmids.** Plasmids pCR42 and pCR39, which contain the full-length cDNAs of genes 6 and 11, respectively, were obtained by ligating RT–PCR fragments directly into TA vector pCR II as described above.

  For *in vitro* expression, full-length cDNAs of genes 6 and 11 were transferred into plasmid pKT0 (Liu et al., 1994), generating plasmids pKT42 (containing gene 6) and pKT39 (containing gene 11). This was achieved by cloning EcoRV/BamHI-digested cDNA fragments containing genes 6 and 11 into PvuII/BamHI-digested plasmid pKT0.

  Plasmid pKT42A, in which most of the ORF1 region was deleted, was constructed by inserting a PvuII/BamHI-digested cDNA fragment covering the gene 6 sequence from nucleotide 216 to the 3’ terminus into PvuII/BamHI-digested pKT0.

- **Transcription and translation in vitro.** Plasmids were linearized with appropriate restriction enzymes, extracted with phenol–chloroform and precipitated with ethanol. Transcription was performed by using T7 RNA polymerase (Promega) as described by Sambrook et al. (1989). Translation of *in vitro*-synthesized transcripts was carried out in rabbit reticulocyte lysates (RRL, Promega) or wheatgerm extracts (WGE, Promega) according to the manufacturer’s instructions. After incubation at 30 °C for 1 h, aliquots of [35S]methionine-labelled polypeptides were analysed by SDS–PAGE and detected by autoradiography.

- **Dephosphorylation of polypeptides expressed in vitro.** [35S]Methionine-labelled proteins synthesized *in vitro* were dephosphorylated by using calf intestinal alkaline phosphatase (CIP, Promega) in a 20 μl reaction mixture containing 50 mM Tris–HCl (pH 9.3), 10 mM MgCl2, 1 mM ZnCl2 and 10 mM spermidine. After 1 h incubation at 37 °C, the samples were mixed with an equal volume of 2× loading buffer (Laemmli, 1970) and analysed on 15% SDS–polyacrylamide gels.

**Results**

**Preliminary identification of strain KB63 as a group B rotavirus**

A previous report (Wang et al., 1989) showed that isolate KB63 reacted with an antiserum specific for GBR strain ADRV but failed to react in a standard GAR ELISA. The PAGE profiles of KB63 dsRNA (Fig. 1, lane 2) exhibited a 4-2-2-1-1-1 electrophoretotype, one of the typical patterns of GBRs reported so far (Snodgrass et al., 1984; Eiden et al., 1992). PAGE profiles of GAR strain SA11 displayed a 4-2-3-2 pattern common in GARs (Fig. 1, lane 1).

**Cloning of KB63 genes 6 and 11**

cDNA was prepared from KB63 dsRNA and cloned as described in Methods. The initial sequencing of the termini of insert-positive clones revealed that some were flanked by GGXA at the 5’ end and ACCC at the 3’ end of the positive strand. These nucleotides are conserved among the sequenced genes of GBR strains (Eiden et al., 1992), suggesting that such clones may be full-length cDNA copies of KB63 genes. Comparison of nucleotide sequences at the 5’ and 3’ ends of the inserts in plasmids pCR42 and pCR39 with other GBR genes suggested that these two clones contained genes encoding NSP1 and NSP5 of KB63. The identity of these two cDNA clones was further confirmed by Northern blot hybridization (data not shown).

**Sequence analysis of KB63 gene 6**

The nucleotide sequence of gene 6 was deduced initially from the insert contained within plasmid pCR42 and was

![Fig. 1. Silver-stained dsRNA profiles of KB63 and simian GAR SA11. Gene segments were separated on a 10% polyacrylamide gel.](Image)
confirmed in two ways. The sequences of the 5′ and 3′ ends were confirmed by sequencing three other copies of the gene identified during the initial screening of insert-containing plasmids. An internal 998 bp fragment was amplified directly from viral RNA by RT–PCR by using three oligonucleotide primers (SS-1, SS-2 and SS-5) complementary to the sequence initially determined from pCR42 (Fig. 2). RT–PCR with these primers generated fragments of the expected sizes that were subsequently sequenced.

Data generated from these analyses showed that gene 6 of KB63 contained 1275 nucleotides and three long overlapping ORFs (ORFs 1–3) with the potential to encode polypeptides of 101, 175 and 146 amino acids, respectively (Fig. 2). These three ORFs were flanked by untranslated regions of 42 nucleotides at the 5′ end and 60 nucleotides at the 3′ end (Fig. 2). A 95 nucleotide overlap was present between ORFs 1 and 2 and there was a 7 nucleotide overlap between ORFs 2 and 3. The AUG start codon for ORF2 was flanked by a sequence (GAAATGG) associated with efficient initiation (Kozak, 1981). The sequences flanking the start codons for ORF1 and ORF3 contained pyrimidines in the −3 and +4 positions, respectively, and thus would be expected to direct translation less efficiently than that of ORF2.

Sequence comparison (Table 1) revealed nucleotide similarity between KB63 gene 6 and ADRV gene 6 as well as IDIR gene 7 (54.2% and 54.7% identity, respectively), which have been proposed to encode the NSP1 genes of GBRs (Eiden, 1994; E. R. Mackow, unpublished results, GenBank accession no. M91435). As mentioned above, KB63 gene 6 contains three ORFs, while the NSP1 genes of both ADRV and IDIR contain only two. Deduced amino acid sequences of ORFs 1–3 of KB63 were compared and aligned with the deduced amino acid sequences of ORF1 (Table 1 and Fig. 3a) and ORF2 of the ADRV and IDIR NSP1 genes. Amino acid identity ranged from 34.9 to 48.0%, with polypeptide 2 showing the most conservation and polypeptide 3 the least (Table 1). These results confirm the findings of the nucleotide comparisons and indicate that KB63 gene 6 ORF1 is homologous to ORF1 of other GBRs, while ORF3 was probably derived from ORF2 by a frame shift event. A single base insertion (or two-base deletion) upstream of the ORF2 stop codon would lead to the formation of a single ORF encoding a polypeptide just one (or two) amino acid(s) shorter than the equivalent product of IDIR (Fig. 3b). The low level of identity between these sequences made it difficult to define the position of any deletion or insertion. It might have occurred between nucleotides 740 and 750, since an insertion or a deletion in this region would maintain CXXC upstream and RWXXXGXGX downstream, which are conserved among the three strains (Fig. 3b). The sequence of this region was reconfirmed four times from independently amplified RT–PCR products to reduce the uncertainties.

*Table 1. Comparison of the nucleotide and deduced amino acid sequences of KB63 gene 6 and the equivalent genes from ADRV and IDIR*

The deduced amino acid sequences of polypeptides 1–3 encoded by ORFs 1–3 of KB63 gene 6 were compared with those of putative polypeptides encoded by ORF1 and the 5′ and 3′ halves of ORF2, respectively, from the equivalent genes of ADRV and IDIR.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Nucleotide identity (%)</th>
<th>Amino acid identity (similarity) (%)</th>
<th>Polypeptide 1</th>
<th>Polypeptide 2</th>
<th>Polypeptide 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>KB63 vs ADRV</td>
<td>54.2</td>
<td>38.5 (52.5)</td>
<td>46.0 (57.5)</td>
<td>34.9 (39.0)</td>
<td></td>
</tr>
<tr>
<td>KB63 vs IDIR</td>
<td>54.7</td>
<td>40.6 (51.7)</td>
<td>48.0 (59.7)</td>
<td>35.2 (41.4)</td>
<td></td>
</tr>
<tr>
<td>ADRV vs IDIR</td>
<td>66.1</td>
<td>67.3 (73.8)</td>
<td>66.7 (75.3)</td>
<td>59.6 (67.8)</td>
<td></td>
</tr>
</tbody>
</table>
Analysis of group B rotavirus genes 6 and 11

Table 2. Comparison of the nucleotide and deduced amino acid sequences of KB63 gene 11 and the equivalent genes from ADRV and IDIR

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Nucleotide identity (%)</th>
<th>Amino acid identity (%)</th>
<th>Similarity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KB63 vs ADRV</td>
<td>54.2</td>
<td>44.8</td>
<td>54.1</td>
</tr>
<tr>
<td>KB63 vs IDIR</td>
<td>52.2</td>
<td>44.0</td>
<td>54.2</td>
</tr>
<tr>
<td>ADRV vs IDIR</td>
<td>71.7</td>
<td>69.4</td>
<td>79.4</td>
</tr>
</tbody>
</table>

Fig. 3. (a) Alignment of polypeptides 1 (pep1), 2 (pep2) and 3 (pep3) encoded by ORFs 1, 2 and 3, respectively, of KB63 gene 6 with polypeptides encoded by the corresponding NSP1 gene of GBR strains ADRV and IDIR. ADRV.pep2n (IDIR.pep2n) represents the N-terminal half of the polypeptide encoded by ORF2 of strain ADRV (IDIR) and ADRV.pep2c (IDIR.pep2c) represents the C-terminal half of the polypeptide encoded by ORF2 of strain ADRV (IDIR). Completely conserved amino acids are underlined. Gaps introduced to optimize the alignment are indicated by dashes. (b) Alignment of the nucleotide sequence of KB63 gene 6 at the junction area of ORFs 2 and 3 with the corresponding sequences of ADRV and IDIR. Identical nucleotides are indicated by vertical lines, the start codon for ORF3 and the stop codon for ORF2 are indicated in bold, codons for conserved amino acids are underlined and, below the alignment, conserved upstream and downstream amino acids are indicated if a single base was inserted between nucleotide 740 and 750.

350 nucleotides were confirmed by direct sequencing of the products amplified from viral RNA by RT–PCR by using complementary oligonucleotide primers SS-3, SS-4 and SS-6 (Fig. 2). The gene was shown to be 625 nucleotides long and contained an ORF of 501 nucleotides encoding a putative protein of 166 amino acids, which is four and eight amino acids shorter, respectively, than those encoded by gene 11 of ADRV and IDIR (Chen et al., 1990; Petric et al., 1991). The untranslated regions were 58 and 66 nucleotides, respectively, at the 5′ and 3′ ends. Neither of the first two in-frame AUGs was in an optimal Kozak consensus context (Kozak, 1981) and no significant out-of-frame ORF could be detected, in contrast to the cognate gene of GAR SA11 (Mitchell & Both, 1988; Welch et al., 1989). The nucleotide and predicted amino acid sequences of the KB63 gene were compared with those of ADRV and IDIR (Table 2), which is proposed to be the NSP5 gene of GBR. Nucleotide and amino acid identity were significantly lower than those between ADRV and IDIR (54–2 and 44–8% vs 71–7 and 69–4%), which is surprising since, at least among GARs, NSP5 is highly conserved (> 78% identical) (Lopez & Arias, 1993; Blackhall et al., 1997). Nevertheless, alignment of KB63 NSP5 with those of ADRV and IDIR (Fig. 4) revealed three conserved regions (amino acids 1–7, 48–77 and 121 to the C terminus, numbered according to IDIR). Despite the

Sequencing and analysis of KB63 gene 11

KB63 gene 11 was initially sequenced from plasmid pCR39. Once again, the sequences of the 5′ and 3′ ends of the gene were confirmed by sequencing full-length clones derived independently during the cloning procedure and the central likelihood that the proposed frame shift was the result of an artifact introduced during RT–PCR or sequencing.

Fig. 4. Alignment of the deduced amino acid sequence of the product of KB63 gene 11 with those encoded by the NSP5 genes of GBR strains ADRV and IDIR. Completely conserved amino acids are underlined. Gaps introduced to optimize the alignment are indicated by dashes.
sequence diversity, KB63 NSP5 contains 15% acidic and 16–8% basic amino acids and a high proportion of serine and threonine residues (22–9%), typical of NSP5 of GARs, suggesting that, like them, it may be a phosphoprotein. It contains only one potential glycosylation site (Asn–X–Ser/Thr; Spiegelberg et al., 1970), while the deduced sequences of ADRV and IDIR NSP5 contain four and three potential N-glycosylation sites, respectively.

In vitro expression of KB63 gene 6

To investigate the characteristics of RNA derived from KB63 in an in vitro translation assay, the cDNA fragment containing gene 6 was cloned 44 bases downstream of the T7 promoter in pKT0 to generate plasmid pKT42. Plasmid pKT42Δ contains a deletion of nucleotides 1–215 of gene 6, so that the first AUG of this cDNA is the initiation codon for ORF2 (47 bases downstream of the T7 promoter). Linearization of pKT42 with BamHI gave rise to a full-length template (full-length construct). XbaI-digested pKT42 (cut at nucleotide 271) resulted in a template lacking the 70 3′-terminal bases of ORF1 (ORF1 construct). Linearization of pKT42Δ with BamHI gave rise to a template containing ORF2 and ORF3 (ORF2/3 construct). HinclI-digested pKT42Δ (cut at nucleotide 779) resulted in a template containing only ORF2 (ORF2 construct). These templates (see Fig. 2) were used for transcription with T7 polymerase and were all analysed on 1% agarose–0.1% SDS gels to evaluate the amount, purity and size of the transcripts produced.

In the RRL system, no product could be detected when ORF1 RNA was translated (Fig. 5, lane 2). In vitro translation of the full-length RNA (lane 3), ORF2/3 RNA (lane 4) and ORF2 RNA (lane 5) all resulted in the detection of two proteins with apparent molecular masses of 23 and 25 kDa. These results suggest that both the 23 and 25 kDa products may be derived from ORF2, although their apparent molecular masses are higher than the 20 kDa calculated from the amino acid sequence. The 23 kDa protein may represent a premature termination product of ORF2, although post-translational modification of the polypeptide cannot be excluded.

In vitro expression and dephosphorylation of KB63 gene 11 products

In vitro transcription of KB63 gene 11 was performed by using plasmid pKT39 digested with BamHI. The RNA was translated in the RRL system and the translated products were analysed on 15% SDS–polyacrylamide gels. Four major bands, with apparent molecular masses of 19, 21, 23 and 25 kDa, were produced (Fig. 6, lane 2). Since the apparent molecular masses of three of the major bands were higher than the calculated molecular mass of 18.7 kDa, pKT39 was linearized with SspI,
which cuts at nucleotide 483 (133 bases upstream the stop codon), and transcribed and translated in vitro. The result (lane 4) suggested that only the 25 kDa species was the full-length product, while the 19, 21 and 23 kDa bands may be premature termination products. In vitro translation of gene 11 was also carried out in the WGE system and similar results were obtained (data not shown).

Interestingly, at least three faint bands were present above the major 25 kDa band, one above the 19 kDa and one above the 21 kDa proteins. NSP5 of GBRs has been reported to be phosphorylated both in intact cells and in a cell-free system (Afrikanova et al., 1996; Blackhall et al., 1997; Poncet et al., 1997) and the gene product itself has been proposed to be a kinase which autophosphorylates. To determine whether these faint, higher molecular mass bands were the result of phosphorylation, CIP was used to dephosphorylate the in vitro-translated products. Treatment of in vitro-synthesized proteins with CIP resulted in the disappearance of the three faint bands above the 25 kDa species (Fig. 6, lanes 1 and 2). A similar effect was also observed when the truncated products were treated with CIP (lanes 3 and 4). These results indicate that the in vitro-translated products of this gene were indeed phosphorylated.

Discussion

In this communication, we report the complete nucleotide sequences of gene segments 6 and 11 of GBR strain KB63, which encode the putative NSP1 and NSP5, respectively. Comparison of their deduced amino acid sequences with those of the corresponding sequences of other GBR strains provided the first evidence that the substantial diversity of NSP1 found in GBRs also exists among GBR strains and suggested the existence of a high degree of diversity among NSP5 sequences, which has not been observed before in GARs or GBRs.

Alignment of the putative gene 6 sequence with its homologues in other GBRs revealed significant diversity. Amino acid identity with ADRV and IDIR was only 34% and 48%, respectively, while the equivalent value between ADRV and IDIR was more than 63% (NSP1 encoded by ORF2). These values fall well within the range of identities (24%–97%) found when the NSP1 proteins of GARs were compared and suggest that the sequence diversity that characterizes this protein in GARs is duplicated in GBRs. One significant discrepancy between the NSP1 gene of GBRs and those of GARs and GCRs is the absence of the conserved zinc finger motif located in the N-terminal region of the protein in the latter two groups of viruses. Its absence from the sequence of ADRV and IDIR is supported by the more divergent sequence of KB63 reported here and raises questions about the functional relatedness of the NSP1 proteins of these three groups of viruses. Another interesting result yielded by the sequence of KB63 gene 6 is the presence of three overlapping ORFs. Bicistronic genomic segments have been reported in reoviruses, GARs and the two GBR strains sequenced to date. However, while expression of two ORFs has been documented for the reovirus S1 gene and rotavirus (SA11) NSP5 gene, only one ORF (ORF2) has been shown to be expressed from the GBR NSP1 gene (Ernst & Shatkin, 1985; Belli & Samuel, 1991; Matton et al., 1991). We have demonstrated that the first two of the three ORFs of KB63 gene 6 were expressed in the WGE translation system, supporting its designation as a bicistronic genome segment.

On the basis of the sequence analysis, we propose that ORF3 may have resulted from a single base deletion (or two-base insertion) near the 3’ end of ORF2. No expression from this ORF could be detected in either the RRL or WGE system, although adaptation of the virus to growth in cell culture or generation of specific antisera would be required to confirm that it is not expressed in vivo. This apparent C-terminal deletion of the original ORF2 is clearly not a cell culture artifact, since KB63 was isolated from and passaged in young goats prior to analysis. Such C-terminal deletions of gene 6 ORF2 have not been reported previously for GBRs but are corroborated by similar findings in viable GARs. Three non-defective mutants encoding partially or almost totally C-terminally truncated NSP1 have been isolated (Hua & Patton, 1994; Taniguchi et al., 1996). These mutants appear to have emerged as a result of a point mutation, a deletion and a mutation accompanied by a gene rearrangement. The results presented in this study suggest that, as in the case of GARs, a large part of GBR NSP1 is non-essential for virus replication. In addition, the demonstration that the C-terminal portion of NSP1 is non-essential for replication provides some phenotypic support for the classification of gene 6 ORF2 as NSP1, which to date has been based on sequence comparison alone (Eiden, 1994).

Gene 11 of KB63 shows a high degree of sequence diversity at the amino acid level when compared with those of other sequenced GBRs (54%–79% identity). This observation is somewhat surprising, given the high degree of conservation of reported NSP5 sequences from GARs, which exceeds 78% identity (Blackhall et al., 1997). This low level of absolute amino acid identity is all the more interesting since analysis of the amino acids within the protein shows high levels of serine and threonine, typical of NSP5 characterized from GARs and GBRs. Thus, KB63 NSP5 may play a structurally distinct role in rotavirus replication. The diversity of KB63 NSP5 also enables us to identify three conserved and two variable regions that may be useful for future determination of functional domains.

Recent studies have demonstrated that NSP5 from GAR strains SA11 and RF is phosphorylated in vitro and in cell culture and suggested that it functions as a serine/threonine kinase that can autophosphorylate in vitro (Afrikanova et al., 1996; Poncet et al., 1997; Blackhall et al., 1997). In addition, this protein is located in viroplasms and appears to be associated with NSP2, suggesting that it might play a role in the
formation of the virus replication machinery. We report that KB63 NSP5 is phosphorylated in a phenotypically similar manner to its homologues amongst GARs, despite its low level of absolute amino acid identity. Once again, this observation provides phenotypic support for the designation of this gene as NSP5. Investigation of whether this protein is also present within viroplasms will have to await adaptation of the virus to cell culture and generation of other reagents.

We thank Drs I. N. Clarke and P. R. Lambden for their help in the amplification of KB63 genomic RNA using the sequence-independent procedures developed in their laboratory in Department of Microbiology, University of Southampton Medical School, UK.

References


Received 25 January 1999; Accepted 12 April 1999