# CHARACTERISATION OF REARRANGED NSP5 GENE OF A HUMAN ROTAVIRUS

E.A. PALOMBO, H.C. BUGG, R.F. BISHOP

Department of Gastroenterology and Clinical Nutrition, Royal Children's Hospital, Flemington Road, Parkville, Victoria 3052, Australia

*Received December 3, 1997; accepted January 3, 1998*

**Summary.** – An atypical human rotavirus strain Z10262, isolated from a chronically infected immunodeficient child, displayed an unusual genomic RNA electrophoretic pattern. Besides, Northern blot analysis indicated that this strain contained an abnormally migrating gene 11 equivalent. Sequencing of this gene showed that it was derived from a genetic rearrangement which involved a partial duplication of the open reading frame (ORF) encoding the non-structural protein NSP5. However, the duplicated region contained a deletion and several point mutations relative to the first copy of the ORF. Phylogenetic analysis of human and animal NSP5 amino acid sequences including Z10262 revealed two groups of human proteins related to different animal proteins. The isolation and analysis of Z10262 strain provides further evidence for the genetic complexity of naturally occurring human rotaviruses.

**Key words:** rotavirus; non-structural protein; genetic rearrangement; epidemiology

## Introduction

Group A rotaviruses are the major cause of severe gastroenteritis in infants and children (Kapikian and Chanock, 1996) and a vaccine against them is a priority. However, effective vaccine strategies require a better understanding of the level of genetic variation in natural rotavirus isolates. The rotavirus particle consists of a two-layered capsid enclosing a core which contains the segmented viral genome (Estes, 1996). The resolution of the eleven rotaviral dsRNA genomic segments by polyacrylamide gel electrophoresis (PAGE) produces characteristic migration patterns known as electropherotypes. Group A rotavirus gene segments can generally be divided into four size classes: I (segments 1-4), II (segments 5 and 6), III (segments 7-9) and IV (segments 10 and 11) (Desselberger, 1996; Estes, 1996). However, several reports have described isolates (either naturally occurring or generated *in vitro* after infection at high multiplicity) with abnormal RNA profiles in which certain segments have disappeared and been replaced by larger, slowly-migrating cognate segments that have arisen from genetic rearrangements (Desselberger, 1996).

The smallest gene segment 11 (approximately 660 bp) has been the most commonly described segment involved in rearrangements (Desselberger, 1996). Nucleotide sequencing revealed two types of the rearranged segments. Type I rearrangements of gene 11 found in both human and animal rotaviruses involve a partial duplication of sequences occurring within the ORF downstream of the stop codon (Gonzalez *et al.*, 1989; Gorziglia *et al.*, 1989; Scott *et al.*, 1989; Matsui *et al.*, 1990; Kojima *et al.*, 1996). The gene 11 of so-called human "short" and "super-short" rotaviruses, however, contains type II rearrangements consisting of an insertion of an A+T rich region of non-rotaviral origin downstream of the stop codon (Matsui *et al.*, 1990; Nutall *et al.*, 1990; Giambiagi *et al.*, 1994). Neither of these types of rearrangement affects the coding region of gene 11, which encodes the non-structural protein NSP5 (Estes, 1996), a phosphorylated and glycosylated protein whose role in the viral replication cycle is unknown (Afrikanova *et al.*, 1996; Poncet *et al.*, 1997). Type I rearrangements have also been described in other gene segments (Desselberger, 1996).

**Abbreviations:** DIG = digoxigenin; EIA = enzyme immunoassay; ORF = open reading frame; PAGE = polyacrylamide gel electrophoresis; RT-PCR = reverse transcription-polymerase chain reaction

In this report, we describe a rotavirus strain Z10262 with an unusual RNA profile, which was isolated from a two-year-old immunocompromised child admitted to the Royal Brisbane Hospital, Australia, with acute gastroenteritis in 1995. Genetic analysis showed that this strain contained an abnormally migrating gene 11 equivalent that had resulted from a type I rearrangement.

## Materials and Methods

*Virus.* Rotavirus strain Z10262 (serotype G1, subgroup II) was detected in the stool of a child admitted to hospital with acute gastroenteritis by routine enzyme immunoassay (EIA). The faeces was collected within 48 hrs of admission and stored at -70°C after transportation to Melbourne.

*RNA extraction, PAGE and Northern blot analysis.* Genomic dsRNA was extracted from a 10%(w/v) faecal extract by phenol-chloroform treatment, electrophoresed in a 10% (w/v) polyacrylamide gel and the separated eleven gene segments were stained with silver nitrate (Dyall-Smith *et al.*, 1984). Northern blot analysis was carried out as previously described (Palombo *et al.*, 1996). Briefly, RNA was electrophoresed, denatured in 0.1 N NaOH and blotted to a positively charged nylon membrane (Boehringer-Mannheim). A pre-hybridisation and hybridisation at 50°C using a digoxigenin (DIG)-labelled gene 11-specific cDNA probe were followed by detection of the bound probe with an anti-DIG antibody conjugated to alkaline phosphatase (Boehringer-Mannheim) and the chemiluminescent substrate CDP-Star (Boehringer-Mannheim).

*RT-PCR and nucleotide sequencing.* The extracted Z10262 dsRNA was purified by adsorption to hydroxyapatite (Gouvea *et al.*, 1991) and used to generate cDNA by reverse transcription-polymerase chain reaction (RT-PCR) by the method of Gentsch *et al.* (1992). Direct cycle sequencing was carried out by use of the fmol DNA Sequencing System (Promega) and gene-specific primers. The sequencing strategy involved the amplification by RT-PCR of three overlapping cDNA fragments of the gene 11 segment because a full-length RT-PCR using gene 11-specific end primers was not successful. Based on published gene 11 nucleotide sequence, outward reading primers were designed to produce a cDNA product only in case gene 11 was present as a head-to-tail direct repeat. Sequencing of this fragment was done using internal primers. These primers in combination with 5'- and 3'- end primers allowed the generation of the 5'- and 3'- ends of the gene 11 segment. These cDNA fragments overlapped the fragment containing the junction of the two NSP5 genes and were sequenced using additional internal primers. The nucleotide sequence data reported in this paper are accessible at the GenBank Database under No. U96698.

*Phylogenetic analysis.* Pairwise comparisons of deduced amino acid sequences were carried out by Clustal W Analysis (Wisconsin Package, version 8.1 UNIX, 1995). Phylogenetic analysis and dendogram construction using distances calculated by the Clustal W Analysis were carried out using the TreeView Programme (Page, 1996).

## Results and Discussion

### *PAGE of Z10262 RNA*

PAGE indicated that, when compared to a normal long RNA profile of a standard serotype G1 subgroup II human rotavirus strain, strain Z10262 displayed an unusual pattern distinguished by the absence of a characteristic eleventh gene segment (Fig. 1). Northern blot analysis using a DIG-labelled gene 11-specific cDNA probe showed that the seventh largest segment exhibited homology to gene 11 (Fig. 1), suggesting that this segment had arisen through a rearrangement of gene 11. The presence of an additional band in the size class I segments was also apparent, possibly resulting from an independent rearrangement of another gene segment.

### *Nucleotide sequence analysis and structure of the rearranged gene 11 of Z10262*

To investigate the nature of the gene 11 rearrangement in strain Z10262, nucleotide sequence analysis of cDNA products derived from RT-PCR of this segment was carried out. The sequence of the rearranged gene 11 (Fig. 2) showed that the segment was 1224 bp in length compared to the usual length of about 660 bp (Estes, 1996). The gene segment was characterised by a reiteration of the sequence starting from nt 44 downstream of the termination codon and extending to the 3'-non-coding region of gene 11. This segment was the product of a type I genetic rearrangement.

In a number of type I rearrangements, direct repeats of sequences have been observed closely upstream of the start of the duplication and of the point of reiteration (Desselberger, 1996; Kojima *et al.*, 1996). Such direct repeats were not a feature of the duplication in Z10262, which is in common with some animal rotavirus type I gene rearrangements (Desselberger, 1996). Although the significance of the direct repeats is not entirely clear, their presence indicated that type I rearrangements occurred at the step of plus strand RNA synthesis (Kojima *et al.*, 1996). Whether this mechanism operates in the absence of direct repeat sequences is not known.

The gene segment included an ORF encoding the non-structural gene NSP5, beginning at the first ATG at positions 23-25 and ending at TAA at positions 613-615. Since the duplicated region commenced at nt 44, a second copy of the NSP5 ORF was not present in the reiterated sequence. Thus, the rearrangement in Z10262 genome generated an abnormally long 3'-non-coding region of 609 bp with respect to the NSP5 ORF. The structure of the rearranged gene 11 and its possible origin from a normal gene 11 is shown in Fig. 3.

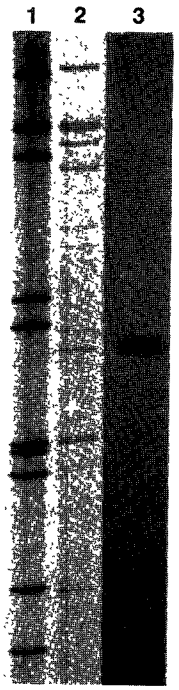


**Fig. 1**
**PAGE and Northern blot analysis of Z10262 RNA**
PAGE of prototype human rotavirus RV4 (lane 1) and strain Z10262 (lane 2) dsRNA, Northern blot analysis of strain Z10262 (lane 3).

The duplicated sequence contained three nucleotide substitutions at the corresponding positions in the first copy of the gene (Fig. 2). This sequence also contained a 12 bp deletion, relative to the first copy, starting downstream of nt 1117. The presence of point mutations and a deletion in the reiterated sequence could indicate that the replication event leading to the formation of the rearranged gene did not occurr recently.

*Amino acid sequence and phylogenetic analysis of Z10262 NSP5*

The deduced amino acid sequence of Z10262 NSP5 was compared to the published amino acid sequences of NSP5 of other human and animal (simian, bovine, porcine and lapine) strains. It displayed greatest identity with those of human strains Wa and Mc323 (normal gene 11) and Mc345 (gene 11 with type I rearrangement), and of porcine strains OSU, C60 and YM (91% – 95% amino acid identity), indicating a possible evolutionary link between NSP5 proteins of porcine and some human strains. In contrast, Z10262 NSP5 exhibited only 82% – 86% amino acid identity with human strains with type II rearrangements in their NSP5 genes (strains RV5, DS1, 69M and B37) and with animal strains of simian, bovine and lapine origin. This indicated

```
                                                    ↓
GGCTTTTAAAGCGCTACAGTGATGTCTCTCAGCATTGACGTGACGAGTCT   50

TCCCTCAATTTCTTCTAGCATTTTCAAAAATGAATCGTCTTCTACAACGT  100

CAACTCTTTCTGGAAAATCTATTGGTAGGAACGAACAGTATGTTTCACCA  150

GATATCGATGCGTTCAATAAATACATGTTGTCGAAGTCTCCAGAGGATAT  200

TGGACCATCTGATTCTGCTTCAAACGATCCACTCACCAGCTTTTCGATTA  250

GATCGAATGCAGTTAAGACAAATGCAGATGCTGGCGTGTCTATGGATTCA  300

TCAACACAATCACGACCTTCAAGCAACGTTGGGTGCGATCAAATGGATTT  350

CTCCTTAACTAAAGGTATTAATGTTAGTGCTAATCTTGATTCATGTGTAT  400

CAATTTCAACTAATCAAAAAAAGGAGAAATCTAAGAAGGATAAAAGTAGG  450

AAACACTACCCAAGAATTGAAGCAGATTCCGACTCTGAGGATTACGTTTT  500

GGATGATTCAGATAGTGATGACGGCAAATGTAAGAATTGTAAATATAAAA  550

AGAAATATTTTGCATTAAGAATGAGGATGAAACAAGTCGCAATGCAATTG  600

                ↓
ATAGAAGATTTGTAACGAGTCTTCCCTCAATTTCTTCTAGCATTTTCAAA  650

                *
AATGAATCGTCTTCTACAACTTCAACTCTTTCTGGAAAATCTATTGGTAG  700

GAACGAACAGTATGTTTCACCAGATATCGATGCGTTCAATAAATACATGT  750

TGTCGAAGTCTCCAGAGGATATTGGACCATCTGATTCTGCTTCAAACGAT  800

CCACTCACCAGCTTTTCGATTAGATCGAATGCAGTTAAGACAAATGCAGA  850

TGCTGGCGTGTCTATGGATTCATCAACACAATCACGACCTTCAAGCAACG  900

TTGGGTGCGATCAAATGGATTTCTCCTTAACTAAAGGTATTAATGTTAGT  950

  *
GCTAGTCTTGATTCATGTGTATCAATTTCAACTAATCAAAAAAAGGAGAA 1000

ATCTAAGAAGGATAAAAGTAGGAAACACTACCCAAGAATTGAAGCAGATT 1050

CCGACTCTGAGGATTACGTTTTGGATGATTCAGATAGTGATGACGGCAAA 1100

                                                  *
TGTAAGAATTGTAAATATTTTGCATTAAGAATGAGGATGAAACAAATCGC 1150

AATGCAATTGATAGAAGATTTGTAATGTCGACCTGAGGACACACTAGGGA 1200

GCTCCCCACTCCCGTTTTGTGACC                            1224
```

**Fig. 2**
**Nucleotide sequence of rearranged gene segment 11 of strain Z10262**
Initiation (nt 22-24) and termination (nt 613-615) codons of the NSP5 ORF are in bold. The gene segment contains a direct repeat starting next to the termination codon and located at nt 44-615 as indicated by arrows. The overlined nt 546-557 indicate the region of the NSP5 ORF that is deleted in the reiterated sequence (downstream of nt 1117). An asterisk above a nucleotide indicates that it differs in the reiterated region from the corresponding one in the NSP5 ORF.

that different human rotaviruses may have derived their NSP5 from diverse animal rotaviruses. This was seen in the phylogenetic analysis (Fig. 4) which showed two distinct clusters of NSP5, one containing human strains with normal or type I rearrangements of gene 11 together with porcine strains while the other included human strains with type

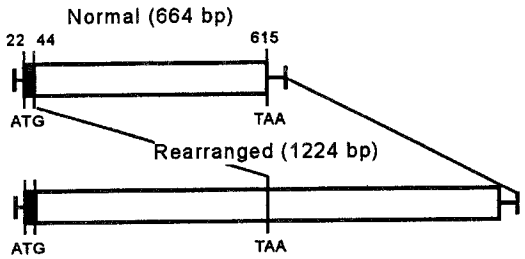


**Fig. 3**
**Proposed derivation of rearranged gene 11 of strain Z10262 from a progenitor normal gene**
The boxed areas (filled and open) indicate the NSP5 ORF in the normal and rearranged gene segments. The open boxes represent the portion of the NSP5 ORF that is duplicated in the rearranged gene, while the filled boxes represent the region of the NSP5 ORF present only once in the normal and rearranged segments.

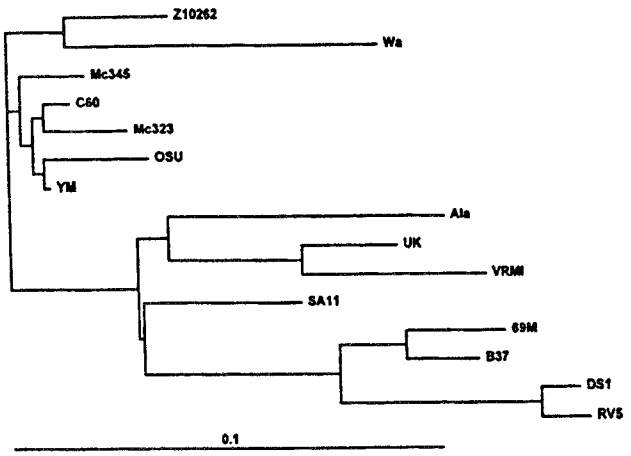


**Fig. 4**
**Phylogenetic tree for NSP5 proteins**
The phylogenetic tree includes NSP5 proteins of human (Z10262, Wa, Mc323, Mc345, 69M, B37, DS1 and RV5), simian (SA11), bovine (VRMI and UK), porcine (C60, OSU and YM) and lapine (Ala) rotaviruses. The sequences are from Lopez and Arias (1993), Mattion *et al.* (1994) and Kojima *et al.* (1996). The length of the abscissa to the connecting node is proportional to the genetic distance between sequences as indicated by the scale bar.

II rearranged gene 11 segments (i.e. strains with a short or super-short electropherotype) and various animal strains.

The description of strain Z10262 is a further example illustrating that persistent natural rotavirus infection in immunocompromised children can result in replication errors that lead to the formation of rearranged gene segments. The description of rotavirus strains with rearranged genes that displayed changes in phenotype or a replication advantage *in vitro* (Tian *et al.*, 1993; Xu *et al.*, 1996) suggests that gene rearrangements can result in viruses exhibiting novel characteristics. With the imminent licensure and expected increase in use of the first human rotavirus vaccine, it is important that surveillance of community strains should continue to determine whether such strains will occur rarely only in unusual clinical presentations or they will emerge steadily as a common component of the rotavirus population. The analysis of such strains may also provide new insights into the replication and morphogenesis of rotaviruses.

**Acknowledgements.** This work was supported by the National Health and Medical Research Council of Australia and the Royal Children's Hospital Research Foundation. We thank Mr M. Witt of the Royal Brisbane Hospital, Brisbane, for sending the faecal specimens and related clinical information, and Mr P. Masendycz of the Royal Children's Hospital, Melbourne, for routine EIA.

## References

Afrikanova I, Miozzo MC, Giambiagi S, Burrone O (1996): Phosphorylation generates different forms of rotavirus NSP5. *J. Gen. Virol.* **77**, 2059–2065.

Desselberger U (1996): Genome rearrangements of rotaviruses. *Adv. Virus Res.* **46**, 69–95.

Dyall-Smith ML, Holmes IH (1984): Sequence homology between human and animal rotavirus serotype-specific glycoproteins. *Nucleic Acids Res.* **12**, 3973–3982.

Estes MK (1996): Rotaviruses and their replication. In Fields BN, Knipe DM, Howley PM (Eds): *Fields Virology*. Lippincott-Raven Publishers, Philadelphia, pp. 1625–1655.

Gentsch JR, Glass RI, Woods P, Gouvea V, Gorziglia M, Flores J, Das BK, Bhan MK (1992): Identification of group A rotavirus gene 4 types by polymerase chain reaction. *J. Clin. Microbiol.* **30**, 1365–1373.

Giambiagi S, Gonzalez Rodriguez I, Gomez J, Burrone O (1994): A rearranged genomic segment 11 is common to different human rotaviruses. *Arch. Virol.* **136**, 415–421.

Gonzalez SA, Mattion NM, Bellinzoni R, Burrone OR (1989): Structure of rearranged genome segment 11 in two different rotavirus strains generated by a similar mechanism. *J. Gen. Virol.* **70**, 1329–1336.

Gorziglia M, Nishikawa K, Fukuhara N (1989): Evidence of amplification and deletion in super short segment 11 of rabbit rotavirus Alabama strain. *Virology* **170**, 587–590.

Gouvea V, Allen JR, Glass RI, Fang Z-Y, Bremont M, Cohen J, McCrae MA, Saif LJ, Sinarachatanant P, Caul EO (1991): Detection of group B and C rotaviruses by polymerase chain reaction. *J. Clin. Microbiol.* **29**, 519–523.

Kapikian AZ, Chanock RM (1996): Rotaviruses. In Fields BN, Knipe DM, Howley PM (Eds): *Fields Virology*. Lippincott-Raven Publishers, Philadelphia, pp. 1657–1708.

Kojima K, Taniguchi K, Urasawa T, Urasawa S (1996): Sequence analysis of normal and rearranged NSP5 genes from human rotavirus strains isolated in nature: implications for the occurrence of the rearrangement at the step of plus and minus strand synthesis. *Virology* **224**, 446–452.

Lopez S, Arias CF (1993): Sequence analysis of rotavirus YM VP6 and NS28. *J. Gen. Virol.* **74**, 1223–1226.

Matsui SM, Mackow ER, Matsuno S, Paul PS, Greenberg HB (1990): Sequence analysis of gene 11 equivalents from "short" and "supershort" strains of rotavirus. *J. Virol.* **64**, 120–124.

Mattion NM, Cohen J, Estes MK (1994): The rotavirus proteins. In Kapikian AZ (Ed): *Viral Infections of the Gastrointestinal Tract.* Marcel Dekker Inc., New York, pp. 169–249.

Nutall SD, Hum CP, Holmes IH, Dyall-Smith ML (1990): Sequences of VP9 genes from short and supershort rotavirus strains. *Virology* **171**, 453–457.

Page, RDM (1996): TREEVIEW: An application to display phylogenetic trees on personal computers. *Comp. Appl. Biosci.* **12**, 357–358.

Palombo EA, Bugg HC, Masendycz PJ, Coulson BS, Barnes GL, Bishop RF (1996): Multiple-gene rotavirus reassortants responsible for an outbreak of gastroenteritis in central and northern Australia. *J. Gen. Virol.* **77**, 1223–1227.

Poncet D, Lindenbaum P, L'Haridon R, Cohen J (1997): In vivo and in vitro phophorylation of rotavirus NSP5 correlates with its localization in viroplasms. *J. Virol.* **71**, 34–41.

Scott GE, Tarlow D, McCrae MA (1989): Detailed structural analysis of a genome rearrangement in bovine rotavirus. *Virus Res.* **14**, 119–128.

Tian Y, Tarlow O, Ballard A, Desselberger U, McCrae MA (1993): Genomic concatemerization/deletion in rotaviruses: a new mechanism for generating rapid genetic change of potential epidemiological importance. *J. Virol.* **67**, 6625–6632.

Xu Z, Tuo W, Clark KI, Woode GN (1996): A major rearrangement of the VP6 gene of a strain of rotavirus provides replication advantage. *Vet. Microbiol.* **52**, 235–247.