

# Nucleotide sequence and expression in *E. coli* of the sequence-specific RNA-binding nonstructural protein NSP3 of a G2 serotype human rotavirus

C. Durga Rao<sup>1,2\*</sup>, Bhargavi S. Rao<sup>1</sup> and Krishne Gowda<sup>1</sup>

<sup>1</sup>Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560 012, India

<sup>2</sup>Centre for Genetic Engineering, Indian Institute of Science, Bangalore 560 012, India

The sequence-specific single-stranded RNA-binding rotavirus nonstructural protein NSP3 plays a central role in the assembly and replication of the 11 double-stranded RNA segments in the nascent viral particles. We have determined the nucleotide sequence of gene 7, encoding the NSP3, from an Indian isolate of G2 serotype human rotavirus strain IS2. This is the first report of the sequence and expression of NSP3 from a human rotavirus. The gene consists of 1066 nucleotides (nt) in length and potentially codes for a polypeptide of 313 amino acids (aa). Comparison with the published, corresponding gene sequences from the simian SA11 and bovine UK strains revealed approximately 75% identity of the IS2 NSP3 at both nucleotide and amino acid levels. An interesting observation is that, in contrast to the nucleotide sequence divergence in the protein coding region, a stretch of about 80 nt in the 3' untranslated region (3'UTR) is highly conserved (>97%). This conserved sequence might play an important role in the regulation of not only the expression of NSP3 but also the viral replication cycle. The 34 kDa NSP3 protein has been expressed in *E. coli* at high levels and purified to homogeneity. Antibodies produced against NSP3 should facilitate studies on the control of expression of NSP3 as well as the replication cycle of the rotavirus.

ROTAVIRUSES are one of the major causes of acute gastroenteritis in infants and young children and represent the leading cause of infant morbidity and mortality in developing countries<sup>1</sup>. Rotaviruses have been classified into 7 groups, A to G<sup>2</sup>, and group A rotaviruses are the subject of intense study as they represent the major pathogens causing diarrhoea in humans. Rotaviruses, members of the family Reoviridae, are composed of a triple-layered capsid enclosing a genome of 11 segments of double-stranded RNA (dsRNA)<sup>3,4</sup>. The genome encodes 6 structural (VP) and 5 nonstructural (NSP) proteins<sup>4</sup>. The outer capsid consists of two proteins, VP4 and VP7, encoded by gene segments 4 and 9,

respectively. The outer capsid proteins exhibit two independent serotype specificities, the G serotype and the P serotype<sup>4</sup>. While the major outer capsid protein, VP7, represents the G serotype, the spike protein, which is also the viral haemagglutinin, specifies the P serotype. To date, at least 14 G serotypes and 18 P types of group A rotaviruses have been identified among humans and animals<sup>5,6</sup>. VP6, the intermediate capsid protein encoded by gene 6, represents the subgroup-specific antigen and two subgroups I and II have been observed among rotaviruses<sup>7</sup>. The inner capsid consists of VP2, encoded by gene 2, which encloses the 11 segments of dsRNA and the core proteins VP1 and VP3 (ref. 3)

The nonstructural proteins encoded by the viral genome play an important role in the regulated expression of the viral genome, viral genome replication and assembly as well as inhibition and/or modification of host cellular functions. Although much of the previous work was directed towards understanding the structure and function of the structural proteins, especially the immunogenic proteins VP4, VP6 and VP7, studies on the nonstructural proteins have begun only in the recent years. The five nonstructural proteins NSP1, NSP2, NSP3, NSP4 and NSP5 are encoded by gene segments 5, 8, 7, 10 and 11, respectively<sup>4</sup>. Although the function of NSP5 is unclear, NSP4 (NS28) was shown to function as the intracellular receptor, in the endoplasmic reticulum, for subviral particles during viral maturation<sup>8</sup>. NSP1, NSP2 and NSP3 were shown to possess RNA-binding activity and to be associated with the replication intermediates<sup>9-13</sup>; but the precise roles of these proteins in viral replication are unclear.

Recently, NSP3 was shown to bind selectively the 3' terminal sequences that are conserved in all the 11 segments of rotaviral RNA<sup>14,15</sup> and to be preferentially localized to the cytoskeletal subcellular compartment in the cell<sup>13</sup>. NSP3 is also the major component of the replication intermediates and has been implicated to play a central role in the assembly and replication of the 11 RNA segments in the nascent replication structures<sup>11</sup>.

\*For correspondence.



Although the genes encoding NSP1, NSP2, NSP4 and NSP5 have been sequenced from several human and animal group A rotaviruses<sup>8, 12, 16, 17</sup>, the sequence of NSP3 from a human rotavirus has not, so far, been reported. The only group A rotavirus strains for which the NSP3 sequence has been reported are the simian SA11 and bovine UK<sup>13, 18, 19</sup>.

In our laboratory, we are studying the genetic variation in human rotaviruses isolated in India to determine the extent of genetic diversity from prototype viruses. Recently, we have reported isolation and characterization of GI0P11 type human asymptomatic neonatal viruses that are multigene reassortants between a bovine serotype 10 rotavirus and a human rotavirus<sup>20, 21</sup>. In this study, we report the complete nucleotide sequence of the NSP3 from an Indian isolate of the human G2 serotype, expression of the protein in *E. coli* and its purification to homogeneity.

## Methods

### *Extraction of viral dsRNA*

The isolation and serotypic characterization of rotaviruses, including the G2 serotype rotavirus isolate IS2, isolated from children suffering from diarrhoea, admitted to various hospitals in Bangalore, has been described earlier<sup>22</sup>. Total nucleic acids from the clarified supernatants of 20% suspension of the stool sample containing IS2 were extracted with phenol-chloroform in the presence of 1% SDS and precipitated with ethanol as described<sup>22</sup>.

### *Purification of the viral dsRNA segments 7, 8 and 9*

The viral double-stranded RNA extracted from the IS2-positive faecal sample was electrophoresed on a 1% agarose gel in the presence of ethidium bromide. The closely migrating dsRNA segments 7, 8 and 9 were eluted together on to a dialysis membrane and the RNA was recovered from the membrane in a buffer containing 0.3 M sodium acetate and 1 mM EDTA. The RNA was precipitated with ethanol after extraction with *n*-butanol, followed by phenol-chloroform.

### *Synthesis of cDNA*

The dsRNA segments were denatured with methyl mercuric hydroxide and the 3' ends were polyadenylated using *E. coli* poly(A)-polymerase<sup>23</sup>. Double-stranded cDNA was synthesized directly on the oligo(dT)-tailed plasmid according to the method described by Okayama and Berg<sup>24</sup>. A cDNA library was constructed in *E. coli* HB101.

### *Identification of NSP3 gene-specific clones*

The cDNA library was screened with [<sup>32</sup>P]-labelled mixed cDNA probes for genes 7, 8 and 9. Clones containing gene-7-specific sequences were identified by comparison of the sequence of the 5' and 3' terminal 100 nt with that of bovine UK and SA11 strains<sup>18, 19</sup>.

### *Oligonucleotide primers*

The different primers used for nucleotide sequence analysis are: 5' GAAGTGTTACTTCTGCTCT 3' (5' primer for Okayama and Berg vector); 5' ATCGTCATCTATCATTA 3' (position 421-437); 5' TAATGATAGATGACGAT 3' (position 437-421); 5' CTACAATTGATCAGGCTT 3' (position 228-245); 5' AAGCCTGATCAATTGTAG 3' (position 245-228); 5' TTGAACAACAATTCAAT 3' (position 810-826) and 5' ATTGAATTGTTGTTCAA 3' (position 826-810). The 5' and 3' primers used for amplification of the gene 7 open reading frame by PCR are 5' ATCCCGG/ATGCTCAAGATGGAGTCT 3' and 5' ATAAGCTT/ATAGGTATAAGTGTA 3', respectively.

### *Nucleotide sequence analysis*

The complete nucleotide sequence was generated by sequencing either the original DNA clones or after subcloning the cDNA insert into the *Bam*HI site of pBluescript KS<sup>+</sup>. Nucleotide sequence was determined by dideoxynucleotide-mediated chain termination method<sup>25</sup>. The original cDNA clones were sequenced at the 5' end using a primer specific for the Okayama and Berg vector<sup>24</sup>. Sequence of both the strands was determined using KS and SK primers on the vector as well as IS2-gene 7-specific primers. The sequence data were processed on a Dec/Vax-6000 main frame using version 6.1 of the GCG application software.

### *Expression of NSP3 in E. coli*

As the complete cDNA of gene 7 contained 5' and 3' untranslated sequences, the gene was tailored for expression in *E. coli* by polymerase chain reaction (PCR)<sup>26</sup> using 5' primer that contained the ATG codon and the sequence downstream of it and the 3' primer containing the sequence upstream of the translational terminator codon. The 5' primer contained a recognition site for *Sma*I and the 3' primer for *Hind*III. The PCR-amplified DNA containing the NSP3 open reading frame was digested with *Sma*I and *Hind*III and was ligated to *Nde*I and *Hind*III-digested pET20b(+) vector<sup>27</sup> (Novagen), in which the *Nde*I site was converted to blunt end by fill-in reaction using Klenow fragment of DNA polymerase I. *E. coli* HB101 was transformed with the chimaeric plasmid DNA and colonies positive



for NSP3 gene sequence were identified by restriction enzyme analysis of the plasmid DNA. The recombinant pETNSP3 plasmid DNA was then used to transform *E. coli* BL21 (DE3) strain for expression of NSP3. Colonies expressing the expected 34 kDa NSP3 were identified by analysis of the bacterial lysates by SDS-polyacrylamide gel electrophoresis (SDS-PAGE)<sup>28</sup>.

### Purification of recombinant NSP3

Recombinant clones expressing NSP3 were grown in Luria broth and were induced in the presence of 0.4 mM isopropyl  $\beta$ -thiogalactoside (IPTG) for 2–3 h when the cultures reached an OD<sub>600</sub> of 0.40. The bacterial pellets were lysed by sonication, the inclusion bodies were purified and dissolved in phosphate buffer containing 8 M urea<sup>29</sup>. NSP3 was then purified by immobilized nickel ion affinity chromatography using Ni<sup>2+</sup>-NTA-agarose resin<sup>29</sup> as the recombinant. NSP3 contains 6 consecutive histidines at the carboxy terminus that form coordination complexes with the immobilized Ni<sup>2+</sup> in the resin. The bound protein was eluted from the resin in phosphate buffer containing 500 mM imidazole<sup>27,29</sup>.

## Results

### Sequence analysis

The NSP3 gene from the G2 serotype human rotavirus strain IS2 was 1066 nt in length and had a single long open reading frame from position 26 to 967 that codes for a polypeptide of 313 amino acids (Figure 1). Two in-frame AUG codons were found at positions 26–29 and 33–35 as found in other NSP3 genes. The second AUG codon at position 33 was in the proper context for translation in eukaryotes<sup>30</sup>. The IS2 gene is shorter by 39 and 29 nucleotides than that of SA114F and bovine UK, respectively. While the 3'UTR of SA114F and SA11 NSP3 gene is 132 and 131, respectively, that of IS2 and bovine UK is 99 and 109 nt, respectively (Table 1). The 3'UTR of SA11 is identical to that of SA114F except for a single nucleotide deletion at position 1011 and both the simian strains contain almost identical NSP3. The difference in the size of 3'UTR of NSP3 gene is due to deletions in two regions downstream of the translational termination codon (Figure 2). The fact that the 3'UTR of gene 7 from human IS2 and bovine UK strains contains deletions in the same region suggests a similar evolutionary origin for the NSP3 of these 2 strains.

The human IS2 gene sequence was 78 and 75.8% homologous at the nucleotide and 78 and 76.9% identical at the amino acid level with that of bovine UK and simian SA114F, respectively. While the NSP3 gene from IS2 and bovine UK potentially codes for a polypeptide

of 313 aa, that of SA11 and SA114F codes for a protein of 315 aa (Table 1). Multiple-sequence alignment of the predicted amino acid sequences using the program MULTALIN<sup>31</sup> revealed a 2-amino-acid deletion in the NSP3 from IS2 and bovine UK viruses following amino acid 308 (Figure 3). Thus, the deletion of amino acids 309 and 310 in the NSP3 of IS2 and bovine UK with reference to that of SA11/SA114F places glutamic acid as the carboxy terminal amino acid in all the NSP3 proteins.

NSP3 is an acidic protein with the isoelectric point ranging from 4.48 to 6.04 as determined by the programme isoelectric (Table 1). The IS2 NSP3 is comparatively more basic than that from other strains, with a pI value of 6.04 (Table 1).

There are four cysteines in the predicted amino acid sequence of NSP3 of IS2. Only two of these cysteines at positions 123 and 139 are conserved in the NSP3 of all the strains whose sequence is known so far (Figure 3). These two cysteines might be critical for intermolecular covalent interaction in the formation of oligomeric structures of NSP (ref. 13). The basic domain implicated in binding to ssRNA is located between amino acids 81 to 150 and the basic amino acids are highly conserved in all the strains. The consensus sequence (I/L) XXM (I/L) (S/T) XXG of ssRNA binding, found in the NS2 protein of orbiviruses and Sigma NS protein of reoviruses, is also conserved in the NSP3 of rotaviruses except that the methionine residue is not conserved in IS2 NSP3. The heptad repeat region, observed in cellular myosin heavy-chain<sup>32</sup> and neurofilament L or M proteins<sup>33</sup>, responsible for oligomerization is also found in the NSP3 from position 181 to 236. The human NSP3 differs from that of simian<sup>11</sup> and bovine UK strains in that the putative leucine zipper domain located near the carboxy terminus<sup>13</sup> is not well-defined due to replacement of most of the leucine residues.

An interesting feature of the 3'UTR of the genes encoding the NSP3 is the high degree of conservation of a stretch of about 80 nt from the 3' end among all the group A rotaviruses (Figure 2). Within this stretch the IS2 gene 7 shared greater than 96% homology with that from other strains. This is in contrast to the observed 76–78% homology observed in the protein coding region of the gene. Interestingly, the 25 nucleotide sequence of the 5'UTR is also conserved in all the strains.

### Expression in *E. coli*

As shown in Figure 4, *E. coli* BL21 (DE3) cells transformed with the recombinant plasmid pETNSP3, which contained the NSP3 coding sequence in-frame with the sequence encoding the six carboxy terminal histidine residues in the vector, expressed the 34 kDa NSP3 at high levels (lanes 4 and 5). The recombinant



```

10          30          50          70          90
1  GCGTTTAAATGCTTTTCAAGTGGTTGATGCTCAAQATGGAGTCTACTCAOCAGATTGGCATCTCTATTATTAACCTCTTCTTTTCAAGCTGC 90
      M L K M E S T Q Q I O I S I I N S S P R A A

110        130        150        170
91 AGTTTGTGCTGCACTTCTACGTTGGAATTAATGGGTATTCAATATGATTATAATGAAGTATATACTAGAGTTAAAAGTAAAGTTTGATT 180
      V C R C T S T L E L M G I Q Y D Y H E V Y T R V K S K F D F

190        210        230        250        270
181 TATAATGGTAGATTCTGGCGTTAAGAATAATCTAATAAGGTAAAGCAOCTACAATTGATCAGGCTTTGATAGGTAAGTTTAGTTTCATCTAT 270
      V M V D S G V K N N L I G K A A T I D Q A L I G K F S S S I

290        310        330        350
271 CAGAAATAGAAATGGATGACTGATTCAAAAACCTGTAGCGAGATTGATGAAAGATGTGAACAACTTAGATTATTATTGTTCATCGAAAAG 360
      R N R N W M T D S K T V A R L D E D V N K L R L L L S S K G

370        390        410        430        450
361 AATTGATCAAAAATGAGAGTTCTTAATGCTTGTTTTAAAGTGTAAAGAGAGTACCTGGAAAATCGTCATCTATCATTAATGTACTAGGTT 450
      I D Q K M R V L M A C F S V K R V P G K S S S I I K C T R L

470        490        510        530
451 AATGAAAAGAGAAAATAGAACGTGGAGAAAGTCGAAGTGGATGATGCAATTCATTGAAGAAAGAAATGGAAATGACACTATAGATTGGAAATC 540
      M K E K I E R G E V E V D D A F I E E R M E I D T I D W K S

550        570        590        610        630
541 CAGATATGATCAACTTGAAGACGATTTGAATCGTTAAAACAGCGAGTTAATGAAAAGTACAATAATTGGGTTATTAAGOCAGGAAAGA 630
      R Y D Q L E R R F E S L K Q R V M E K Y N N W V I K A R K D

650        670        690        710
631 TAACGAGAACATGAACTCTCTTCAGAATGTTATTTTCGCAACAACAAGCTCATATCAATGAATTACAAATATATAATAATAAACTAGAGCG 720
      N E N M N S L Q N V I S Q Q Q A H I N E L Q I Y N M K L E R

730        750        770        790        810
721 TGATTTACAATCAAAAATAGGATCAGTTATTTTCATCCATTGAATGGTACTTACGGTCTATGGAAGTATCAGATGACATTAAATCAGATAT 810
      D L Q S K I G S V I S S I E W Y L R S M E L S D D I K S D I

830        850        870        890
811 TGAACAACAATCAATTCATAGATCATATTAATCCAGTTAATGCTTTTGGATGATTTTGAGTCTATTCTTCGTAATTTAATATCTGATTA 900
      E Q Q F N S I D H I N P V M A F D D F E - S I L R N L I S D Y

910        930        950        970        990
901 TGATAGAATTTTATTATGTTTAAAGGATGTTGTCAGCAAAGTAATTACACTTATACCTATGAGTAAACATAACATATTACCATCTTCAC 990
      D R I F I M F K G L L Q Q S N Y T Y T Y E *

1010       1030       1050
991 GTAACCTCTATGAGCACAAATAGTTAAAAGCTAACACTGTCAAAAACCTAAATGGCTATAGGGGCGTTATGTGACC 1066
    
```

Figure 1. Complete nucleotide and predicted amino acid sequence of gene 7 encoding the NSP3 protein from human rotavirus strain IS2. Nucleotide sequence was generated from several overlapping cDNA clones representing the NSP3 gene (see text for details). cDNA clones were isolated from a cDNA library of genomic RNA segments 7, 8 and 9, purified, directly from a clinical sample containing the symptomatic rotavirus IS2, after electrophoresis in an agarose gel. The nucleotide sequence has been submitted to EMBL database under the accession number X76645.

Table 1. Comparison of the gene and the gene product, NSP3, from different rotaviruses

Rotavirus	Animal	Gene segment			Gene product		
		Length (n)	Length of 3'NCR	Homology (%)	Size	Homology (%)	pl
Group A							
HuIS2 (this work)	Human	1066	99	—	313	—	6.04
BoUK (Ward <i>et al</i> <sup>19</sup> )	Bovine	1076	109	78.1	313	78	5.43
SA11 (Both <i>et al</i> <sup>18</sup> )	Simian	1104	131 (c)	75.8	315	76.7	5.63
SA114F (Mattion <i>et al</i> . <sup>13</sup> )	Simian	1105	132 (c)	75.8	315	77.3	5.63

n, number of nucleotides; c, corrected length

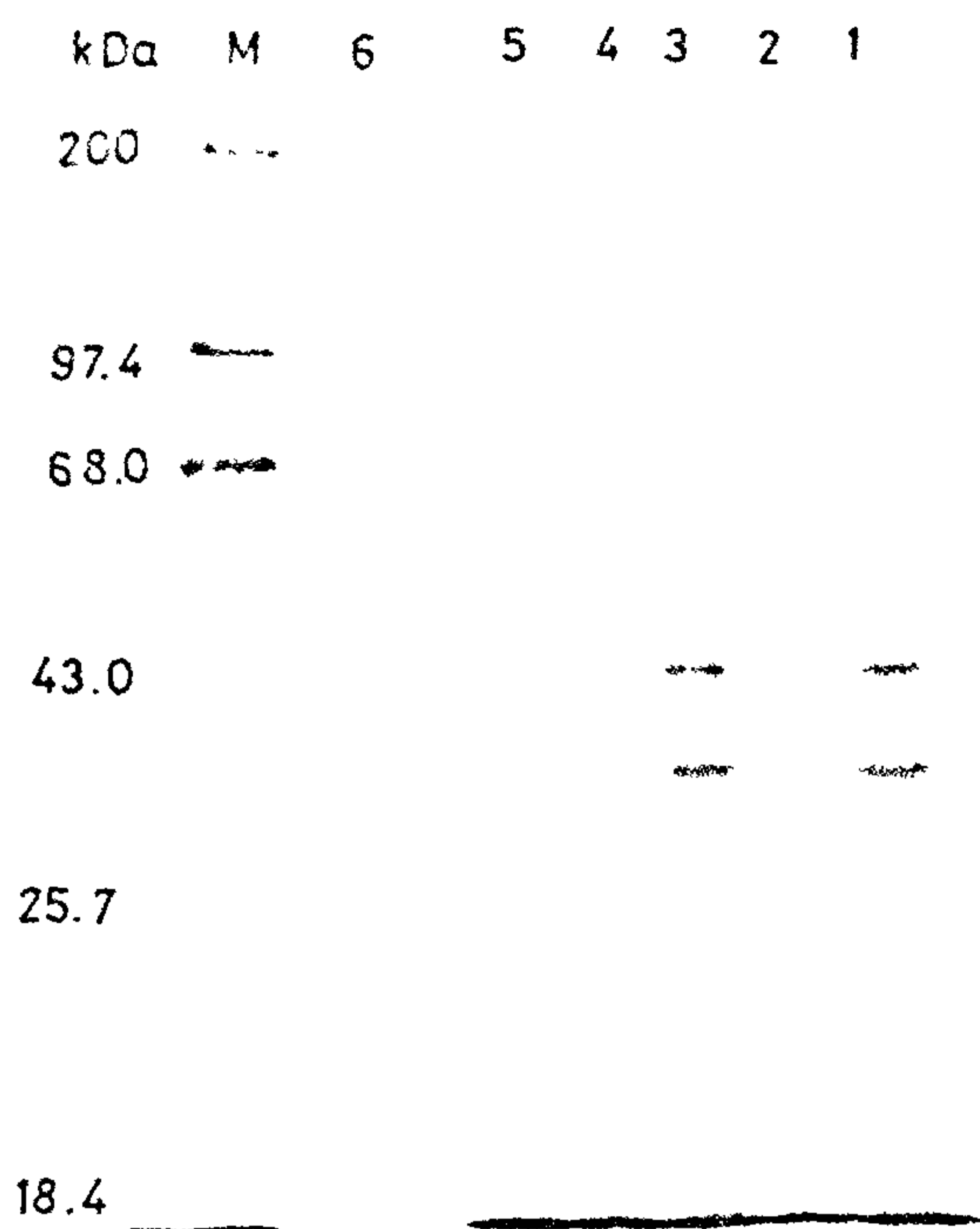
	960	980	1000	1020	
SA114F	TATCAATATTCATTTGGTTGCGAATA <u>ACCATTTT</u> GATACATGTTGAACAATCAAATACAGTGTAGTATGTTGTCATCTA				1026
BoUK	.....GA A CA AT <u>GT</u> .....	A A TTC C T T.....			AC
HuIS2	C.....A T A ACC AT G A .....		C T .....		AC T
	1040	1060	1080	1100	
SA114F	TGCATAACCCTCTATGAGCACAATAGTTAAAGCTAACACTGTCAAAAACCTAAATGGCTATAGGGGCGTTATGTGGCC				1105
BoUK	CA G				A
HuIS2	CA G				A

Figure 2. Nucleotide sequence alignment of the 3' region of the NSP3 genes. The sites of deletions in human IS2 and bovine UK genes compared to SA114F are indicated. The translational termination codon is underlined.

HuIS2	<u>MLKMESTQQIGISII</u> <u>NSSF</u> <u>EA</u> <u>AV</u> <u>CRCT</u> <u>STLELMGIQYDYNEVYTRVKS</u> <u>K</u> <u>FD</u> <u>FVMVDSG</u> <u>VK</u>	60
BoUK	MAS T VAA I Y D	
SA11	MAV VAA A N E QDI S N D	
SA114F	MAV VAA A N E QDI S N D	
HuIS2	<u>NNLIGKAATIDQALIGKFSSIRNRN</u> <u>WMTDSKT</u> <u>VARL</u> <u>DE</u> <u>VNKLRL</u> <u>LLSS</u> <u>KGIDQKMR</u> <u>VL</u>	120
BoUK	L N G VM K R K MM	
SA11	P I NN G A LA TSRP K MM	
SA114F	I NN G A LA TSRP K MM	
HuIS2	<u>NACFSV</u> <u>KRV</u> <u>PGKSS</u> <u>SI</u> <u>IK</u> <u>CTRL</u> <u>MKEK</u> <u>IERGE</u> <u>VEVDDA</u> <u>FIEER</u> <u>MEID</u> <u>TIDW</u> <u>KSR</u> <u>YDQL</u> <u>ERR</u>	180
BoUK	I V D A S V K V V	
SA11	I K RD L S VD K V E Q	
SA114F	I K RD L S VD K V E Q	
HuIS2	<u>FESL</u> <u>KORV</u> <u>NEK</u> <u>YNN</u> <u>WVI</u> <u>KARK</u> <u>DNEN</u> <u>MNSL</u> <u>ON</u> <u>VISO</u> <u>QA</u> <u>HIN</u> <u>EL</u> <u>OI</u> <u>YNN</u> <u>KLER</u> <u>DL</u> <u>QSK</u> <u>IGS</u>	240
BoUK	TT Q K V Y NQ AD N CS A N V	
SA11	S L M H P A V N	
SA114F	S L M H A V N	
HuIS2	<u>VISSIE</u> <u>WYLR</u> <u>SMEL</u> <u>SDDI</u> <u>KSDI</u> <u>EQ</u> <u>FN</u> <u>SIDH</u> <u>IN</u> <u>PV</u> <u>NA</u> <u>FDD</u> <u>FESIL</u> <u>RNL</u> <u>ISD</u> <u>YDR</u> <u>IFIM</u> <u>FK</u>	300
BoUK	LV V K P EV T L T S I I L I L I H T L	
SA11	LT DPE A I A LH L VI KL L	
SA114F	LT DPE A I A LH L VI KL L	
HuIS2	<u>GLL</u> <u>Q</u> <u>S</u> <u>NY</u> <u>..TY</u> <u>TYE</u>	315
BoUK	R C E A	
SA11	I RC QYSFGC	
SA114F	I RC QYSFGC	

Figure 3. Comparison of the deduced amino acid sequences of NSP3 from group A rotaviruses. The regions of predicted structural features of NSP3 are indicated. The basic region is overlined and the heptad repeat regions are underlined. Majority of the leucine residues located in the carboxy terminal leucine zipper domain (overlined by broken line) present in the simian virus SA114F have undergone substitutions in human IS2. The three-amino-acid region where two cysteines are located (position 24-26) near the amino terminal region is overlined. The four-amino-acid stretch conserved between IS2 and bovine UK but different in SA114F in the basic region is marked by a thick bar. The conserved cysteines at positions 123 and 139 are marked by asterisks.





**Figure 4.** SDS polyacrylamide gel electrophoretic analysis of NSP3 expressed in *E. coli*. Lanes 1 and 2 represent total cell lysates of *E. coli* BL21 (DE3), transformed with pET20b(+), uninduced and induced with 0.4 mM. Isopropyl  $\beta$ -D-thiogalactoside, respectively. Lane 3 represents uninduced total lysate from cells containing the recombinant plasmid pETNSP3; Lanes 4 and 5 represent induced total lysates of two individual colonies containing pETNSP3. Lane 6 represents the NSP3 purified by single-step affinity chromatography on Ni-NTA-agarose resin. M, protein molecular weight markers.

protein was found to be predominantly present in the inclusion bodies. NSP3 from the inclusion bodies solubilized in 8 M urea was purified to homogeneity by a single-step affinity chromatography on a Ni-NTA-agarose column (lane 6).

## Discussion

Rotaviruses, besides being the major pathogens of infantile diarrhoea in humans, represent excellent model system for studying the molecular mechanisms of replication and assembly of the multipartite genomes in the nascent viral particles. In general, in all the group A rotaviruses, the first four and the last six nucleotides are highly conserved in all the 11 dsRNA segments<sup>15</sup>. NSP3, by its ability to bind selectively the 3' terminal conserved sequences of the rotaviral RNAs, has been implicated as the central player in the assembly of the 11 positive-sense RNAs into the early replication intermediates, where the ssRNA segments are replicated to form dsRNA segments<sup>13,14</sup>. Moreover, NSP3 has been

shown to form oligomeric structures and was found to be predominantly localized on the cytoskeletal compartment in the cell on which the replication intermediates are assembled<sup>13</sup>. In spite of the importance of NSP3 in rotavirus replication, the gene has not been sequenced, so far, from a human rotavirus.

Amino acid sequence comparisons, in the present study, indicate that while the ssRNA-binding domain (basic region) and the heptad repeat region are highly conserved in all the strains, the carboxy terminal leucine zipper domain is not conserved. This observation corroborates the finding that the leucine zipper domain in the NSP3 of SA11 is not absolutely required for oligomerization<sup>13</sup>. The present studies also suggest, though not experimentally proven, that the cysteines at positions 123 and 139 are critical for the intermolecular covalent interaction in the formation of oligomeric structures of NSP3 as these are the only cysteines that are conserved in all the three strains.

The most interesting observation from our studies is the extreme conservation of a stretch of 80 nt in the 3'UTR of all the NSP3 genes in contrast to the protein coding region (Figure 3). This region might play an important role in the regulated expression of the NSP3. Sequences located in the 3'UTR of certain cellular mRNAs are known to play an important role in the regulation of the stability, translation and intracellular sequestration of the mRNAs<sup>34-38</sup>. NSP3 has been observed to be associated with the cytoskeleton<sup>13</sup> although the molecular basis for such an association is not clear. In this context, it is reasonable to speculate that the highly conserved 3'UTR could mediate selective localization of the mRNA by interaction with specific cellular factors, thereby facilitating the translation of the mRNA directly on the cytoskeleton.

Availability of antibodies against NSP3 should facilitate studies on the structure and function as well as regulation of the expression of NSP3. The role of 3'UTR in the overall regulation of the expression of NSP3 and its possible interaction with cellular factors are being investigated.

1. Kapikian, A. Z. and Chanock, R. M., in *Virology* (eds. Fields, B. N., Knipe, D. N., Chanock, R. M., Melnick, J. L., Roizman, B. and Shope, R. E.), Raven Press, New York, 1990, 2nd edn, pp. 1353-1404.
2. Pedley, S., Bridger, J. C., Chasey, D. and McCrae, M. A., *Arch. Virol.*, 1986, **67**, 131-137.
3. Prasad, B. V. V., Wang, G. L., Clerx, J. P. M. and Chiu, W., *J. Mol. Biol.*, 1988, **199**, 269-275.
4. Estes, M. K. and Cohen, J., *Microbiol. Rev.*, 1989, **53**, 410-449.
5. Browning, G. F., Fitzgerald, T. A., Chalmers, R. M. and Snodgrass, D. R., *J. Clin. Microbiol.*, 1991, **29**, 2043-2046.
6. Dunn, S. J., Burns, J. W., Cross, T. L., Vo, P. T., Ward, R. L., Bremont, M. and Greenberg, H. B., *Virology*, 1994, **203**, 250-259.
7. Greenberg, H. B., McAuliffe, V., Valdesuso, J., Wyatt, R. G., Flores, J., Kalika, R., Hoshino, Y. and Singh, N., *Infect. Immun.*,

- 1983, **39**, 91-99.
8. Au, K. S., Mattion, N. M. and Estes, M. K., *Virology*, 1993, **194**, 665-673.
9. Mitchell, D. B. and Both, G. W., *Virology*, 1990, **174**, 618-621
10. Hua, J., Mansell, E. A. and Patton, J. T., *Virology*, 1993, **196**, 372-378.
11. Kattoura, M. D., Clapp, L. L. and Patton, J. T., *Virology*, 1992, **191**, 698-708
12. Patton, J. T., Cid, L. S., Kalbach, A., Mansell, E. A. and Kattoura, M., *Virology*, 1993, **192**, 438-446.
13. Mattion, N. M., Cohen, J., Aponte, C. and Estes, M., *Virology*, 1992, **190**, 68-83.
14. Poncet, D., Aponte, C. and Cohen, J., *J. Virol.* 1993, **67**, 3159-3165.
15. Imai, M., Akatani, K., Ikegami, N. and Furuichi, Y., *J. Virol.*, 1983, **47**, 125-136.
16. Lopez, S. and Arias, C. F., *J. Gen. Virol.*, 1993, **74**, 1223-1226.
17. Welch, S. W., Crawford, S. E. and Estes, M. K., *J. Virol.*, 1989, **63**, 3974-3982.
18. Both, G. W., Bellamy, A. R. and Siegman, I. J., *Nucleic Acids Res.*, 1984, **12**, 1621-1626.
19. Ward, C. W., Elleman, T. C., Azad, A. A. and Dyll-Smith, M. L., *Virology*, 1984, **134**, 249-253.
20. Dunn, S. J., Greenberg, H. B., Ward, R. L., Nakagomi, O., Burns, J. W., Vo, P. T., Pax, K. A., Das, M., Gowda, K. and Rao, C. D., *J. Clin. Microbiol.*, 1993, **31**, 165-169.
21. Das, M., Dunn, S. J., Woode, G. N., Greenberg, H. B. and Rao, C. D., *Virology*, 1993, **194**, 374-379.
22. Sukumaran, M., Gowda, K., Maiya, P. P., Srinivas, T. P., Kumar, M. S., Aijaz, S., Reddy, R. R., Padilla, L., Greenberg, H. B. and Rao, C. D., *Arch. Virol.*, 1992, **126**, 239-251.
23. Sippel, A., *Eur. J. Biochem.*, 1973, **37**, 31-40.
24. Okayama, H. and Berg, P., *Mol. Cell Biol.*, 1983, **3**, 280-289.
25. Sanger, F., Nicklen, S. and Coulson, A. R., *Proc. Natl. Acad. Sci. USA*, 1977, **74**, 5463-5467.
26. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A., *Science*, 1988, **239**, 487-491.
27. Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W., *Methods Enzymol.*, 1990, **185**, 60-89.
28. Laemmli, U., *Nature*, 1970, **227**, 680-685.
29. Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R. and Stuber, D., *Bio/Technology*, 1988, **6**, 1321-1325.
30. Kozak, M., *Nucleic Acids Res.*, 1984, **12**, 5233-5252.
31. Corpet, F., *Nucleic Acids Res.*, 1988, **16**, 10881-10890.
32. McLachlan, A. D. and Karn, J., *Nature*, 1982, **299**, 226-231
33. Geisler, N., Plessmann, U. and Weber, K., *FEBS Lett.*, 1985, **182**, 475-478
34. Shaw, G. and Kamen, R., *Cell*, 1986, **46**, 659-667.
35. Casey, J. L., Hentze, M. W., Koeller, D. M., Caughman, S. W., Rouault, T. A., Klausner, R. D. and Harford, J. B., *Science*, 1988, **240**, 924-928.
36. Higgins, C. F., *Curr. Opinon in Cell Biol.*, 1991, **3**, 1013-1018.
37. Singer, R. H., *Curr. Opinon in Cell Biol.*, 1992, **4**, 1-19.
38. Hesketh, J., Campbell, G., Piechaczyk, K. M. and Blanchard, J. M., *Biochem J.*, 1994, **398**, 143-148.

ACKNOWLEDGEMENTS. This work was supported in part by a grant from the Department of Biotechnology.

Received 14 October 1994; revised accepted 20 January 1995