Volume 11 Number 14 1983 Nucleic Acids Research

Nucleotide sequence of the gene encoding the serotype-specific glycoprotein of UK bovine rotavirus

T.C.Elleman*, P.A.Hoyne*, M.L.Dyall-Smith+, I.H.Holmes+ and A.A.Azad*

*CSIRO Division of Protein Chemistry, and +University of Melbourne, Department of Microbiology, Parkville, Victoria 3052, Australia

Received 6 June 1983; Accepted 29 June 1983

ABSTRACT

The nucleotide sequence of the gene which encodes the major outer-shell glycoprotein of UK bovine rotavirus has been determined. The dsRNA genome segment encoding this protein was converted into ds cDNA and cloned into pBR322 for sequence studies. The gene is 1062 base pairs in length and contains a single, long, open reading-frame capable of coding for a protein of 326 amino-acids. This would leave 5' and 3' non-coding regions of 48 and 36 nucleotides in the mRNA. The predicted amino-acid sequence contains three possible glycosylation sites of the type Asn-X-$\frac{Ser}{Thr}$, and an extremely hydrophobic N-terminal region. This sequence is discussed in the light of the known properties and functions of the protein.

INTRODUCTION

Rotaviruses are a major cause of acute gastroenteritis in infants and in the young of many other animal species (1). They are ubiquitous and serologically form a closely related group belonging to the Reoviridae virus family (2). They possess a genome consisting of eleven segments of double-stranded (ds) RNA enclosed in a non-enveloped, double-shelled isometric capsid (1). At least six proteins form the capsid which is composed of two layers, the inner and outer shell (3). The major component of the outer shell is a glycoprotein (4) which reacts with serotype-specific neutralising antibodies (5,6).

Neutralisation tests have distinguished four serotypes of human rotavirus (7,8). Animal rotaviruses are not as well characterised, but two serotypes of calf rotavirus (G. Woode, personal communication) and three serotypes of avian rotavirus (9) have been identified. Of particular epidemiological interest is the finding that some animal rotaviruses are closely related to certain serotypes of human rotavirus (10, H. Greenberg, personal communication).

Genetic studies have mapped the virus serotype specificity as being coded by one of the group of RNA segments 7,8 and 9 (11,12). These segments have slightly differing relative electrophoretic mobilities in polyacrylamide

gels, depending on the virus strain (13). In the simian rotavirus SA11 and the human strain Wa, serotype-specificity is determined by segment 9, while in UK bovine rotavirus, serotype specificity is determined by segment 8 or 9 (11,12). This correlates with the major outer-shell glycoprotein assignment to gene 9 of SA11 (Kantharidis, P., Dyall-Smith, M.L. and Holmes, I.H. submitted for publication) and to gene 8 of UK bovine rotavirus (14), and the neutralising activity of antibody produced to this glycoprotein (5,6).

As part of a program to compare the genes of different rotaviruses at the structural level we have cloned and sequenced the UK bovine rotavirus gene coding for the major outer-shell glycoprotein (VP7c of ref.14). This gene has been previously referred to as UK bovine gene segment 8 on the basis of electrophoretic mobility in long polyacrylamide gels (14). It is anticipated that comparisons with the homologous gene sequences of rotaviruses of different serotypes will provide information on the nature of rotavirus serotypic variation and lead to the identification of the antigenic determinants.

MATERIALS AND METHODS

Preparation of viral RNA, and cloning of cDNA

The procedure has been previously described by Dyall-Smith *et al.*, (15). Viral dsRNA was prepared from UK bovine rotavirus grown in MA 104 cells (16,17). Following thermal denaturation the dsRNA was poly A-tailed and reverse transcribed using oligo dT_{12}. Complementary strands were annealed and end-repaired using *E. coli* DNA polymerase 1 (Klenow), and ds cDNA was fractionated according to size on agarose gel. The ds cDNA of *ca.* 1100 bp which corresponded to gene segments 7, 8 and 9 was dC-tailed and annealed with Pst I-cleaved, dG-tailed pBR 322, and the preparation was used to transform *E. coli* MC 1061 (18).

Screening of bacterial colonies for the presence of Gene 8

Bacteria harbouring recombinant (Tet^r, Amp^s) plasmids were screened for the presence of gene 8 sequences using procedures described previously for identifying gene 7 clones (15). Briefly, UK bovine rotavirus dsRNA was separated by polyacrylamide gel electrophoresis and the band containing gene segments 8 and 9, which co-migrate, was excised and the RNA eluted. The RNA was 5' end-labelled with $\gamma[^{32}P]$ ATP using polynucleotide kinase and was used to probe lysed bacterial colonies on nitrocellulose filters (19). The insert from one of the clones which hybridised to the segment 8/9 probe was isolated by Pst I digestion and agarose gel electrophoresis, labelled by nick

translation using $\alpha[^{32}P]$ dATP (20), and used to probe a replicate filter of the lysed bacterial colonies described above. This allowed the clones from segment 8 and segment 9 to be separated into two groups, corresponding to the two RNA segments. A representative clone from each group was used for further analyses.

The group of clones containing gene 8, the gene encoding the major outer-shell glycoprotein, was identified by cross hybridisation with RNA from SA11 and Wa rotavirus strains in a 'northern blot' analysis (21). These strains have electrophoretically separable RNA segments 8 and 9 of known protein coding assignments. This identification procedure, described in detail elsewhere (13), involved separation of UK bovine, SA11 and Wa rotavirus RNA segments by polyacrylamide gel electrophoresis and transfer to activated 2-aminophenylthioether paper (22). Purified plasmid DNA was then labelled by nick translation using $\alpha[^{32}P]$ dATP and used to probe the immobilised RNA bands. The clone selected for sequence study hybridised strongly to the segment 8/9 band of UK bovine rotavirus, to segment 9 of SA11 and to segment 9 of Wa rotavirus.

Sequence determination

The recombinant plasmid was prepared from cleared lysates of transformant bacteria (23) by cesium chloride density equilibrium centrifugation in the presence of ethidium bromide (24). Large restriction fragments of the plasmid were isolated from agarose gels by electrophoresis onto DEAE membrane and were subsequently eluted from the membrane at pH 9.5 with 1.0 M sodium chloride. Smaller fragments were isolated from acrylamide gels (0.4 mm) by diffusion into the gel elution buffer of Maxam and Gilbert (25). All sequence determinations were performed by the dideoxy chain-termination method (26). The larger fragments were cloned into the appropriate M13 mp2 vectors (see Fig. 1) for preparation of single-stranded template (30). Sequences were determined using either universal primers (31, 32), or the smaller restriction fragments as primers. In addition $T_{10}GG$ (20 ng) was used as a primer on linearised recombinant plasmid (2µg) and a single unidirectional sequence was obtained since the plasmid contained, due to a copying deficiency, only a single added poly dA tract (this was useful since terminal sequences were difficult to obtain with universal primer, the long C/A tracts disrupting sequence copying of the M13 templates). The entire nucleotide sequence with the exception of the terminal 37 nucleotides at the 5' end of the +ve strand was determined from both complementary directions of the cloned copy. This 5' terminal sequence was determined in a single direction, the first nine

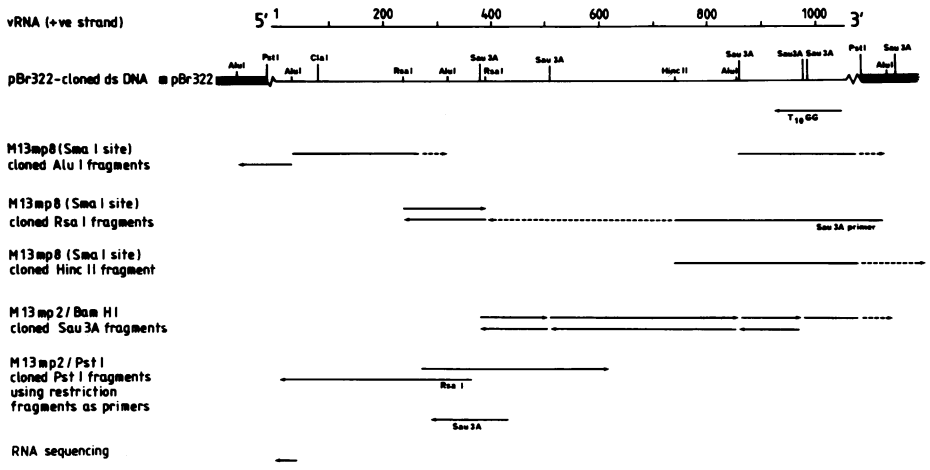

Figure 1. Summary of sequencing strategy. The restriction sites used to determine the sequence of a cloned copy of gene segment 8 are indicated. The sense shown is that of the plus (mRNA) strand. Alu I fragments, Rsa I fragments and Hinc II fragments were subcloned into the Sma I site of M13mp8 (27), Sau 3A fragments into the Bam HI site of M13mp2/Bam HI (28), and the Pst I fragment into M13mp2/Pst I (29). Solid lines indicate the extent to which the sequence was determined in each subclone.

nucleotides using mRNA transcripts (4µg) as templates (33) and an exonuclease-treated (34) ClaI/AluI restriction fragment as primer. The ClaI/AluI fragment (residues 40-91) was prepared from the cloned copy of gene 8 (the yield from 3µg of plasmid being used per reaction). Single-stranded RNA transcripts which display mRNA activity, were produced from UK bovine rotavirus cores by the method of Cohen et al. (35).

RESULTS

Single-stranded cDNA transcripts of rotavirus RNA corresponding in size to all eleven RNA segments were produced. In addition a considerable number of incomplete species were present as demonstrated by denaturing agarose gels (36). To avoid cloning these incomplete species the annealed ds cDNA was end-repaired and fractionated by size on agarose gel prior to cloning. Large numbers of clones containing complete or near complete copies of gene segments 7, 8 and 9, as judged by insert size (ca. 1100bp), were obtained.

Although gene 8 of UK bovine rotavirus has been resolved on long polyacrylamide gels and identified as encoding the major, outer-shell glycoprotein (14), we were unable to separate segments 8 and 9 by electro-

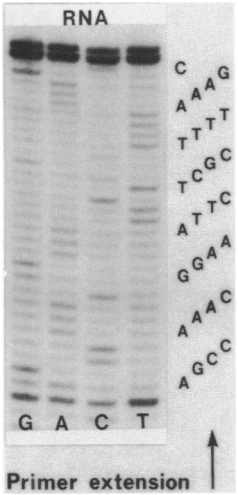

Figure 2. Autoradiograph of a sequencing gel to determine the 5' terminal sequence of gene 8 mRNA. The sequence determined using the chain-termination method is a complementary copy of the 5' terminal sequence of gene segment 8 mRNA. Whole mRNA transcripts were used as a template mixture, and a specific primer (residues 40-91) was prepared from the cloned copy of gene 8 by restriction endonuclease digestion.

phoresis to produce a specific gene 8 probe. These gene segments were separated by cloning. A clone was selected which hybridised to segment 9 of SA11 and Wa rotavirus, as well as hybridising to the segment 8/9 band of UK bovine rotavirus. This clone thus represented the gene encoding the major, outer-shell glycoprotein since it showed homology to the gene segments of SA11 and Wa rotavirus strains which code for this protein (12, Kantharidis, P., Dyall-Smith, M.L. and Holmes, I.H., manuscript submitted for publication).

The cloned copy of gene 8 is 1053 nucleotides long, and has the conserved 3' terminal sequence of rotavirus genes viz. ...$^{U}_{A}$UGUGACC 3' (37) but is 9 nucleotides short of a full-length copy at the 5' end. Missing residues 2-9 were determined by primed synthesis on mRNA (Fig. 2) and the sequence corresponds with the conserved 5' terminal sequence of rotavirus genes, viz. 5' GGCUUU$^{A}_{U}$AAAGC....(37). The capped 5' terminal residue, which is not determined by primed synthesis has previously been determined by direct RNA sequencing (M. McCrae, personal communication). The complete gene sequence of 1062 nucleotides is shown in Figure 3 in the mRNA (+ve) sense of the gene. This sense was assigned by hybridisation studies using single-stranded copies of the gene cloned in M13 in either orientation. Only the M13-cloned sequence complementary to mRNA hybridises to single-stranded rotavirus transcripts. The nucleotide composition of the gene 8 sequence is A 33.6%, T 28.7%, G 20.3% and C 17.3%, very similar to that of the recently completed gene 7 from UK bovine rotavirus (15).

```
                                                                    5'-GGCTTTAAAAGCGAGAATTTCCGTTTGGCTAGCGGTTAGCTCCTTTTA 48

ATG TAT GGT ATT GAA TAT ACC ACA ATT CTA ATC TTC TTG ACA TCG ATT ACA TTG TTG AAT TAT ATC TTA AAA TCA 123
Met Tyr Gly Ile Glu Tyr Thr Thr Ile Leu Ile Phe Leu Thr Ser Ile Thr Leu Leu Asn Tyr Ile Leu Lys Ser 25

ATA ACG AGA ATA ATG GAC TAT ATA ATT TAC AGA TTT CTG CTT ATA GTA GTG GTC TTG GCC ACC ATG ATA AAT GCG 198
Ile Thr Arg Ile Met Asp Tyr Ile Ile Tyr Arg Phe Leu Leu Ile Val Val Val Leu Ala Thr Met Ile Asn Ala 50

CAG AAC TAT GGA GTG AAT TTG CCA ATT ACA GGT TCA ATG GAT ACT GCG TAC GCA AAC TCT ACG CAA AGT GAG CCA 273
Gln Asn Tyr Gly Val Asn Leu Pro Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Ser Glu Pro 75

TTT TTG ACA TCA ACT CTT TGT TTG TAT TAT CCT GTT GAG GCA TCA AAC GAA ATA GCT GAT ACC GAA TGG AAG GAT 348
Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Val Glu Ala Ser Asn Glu Ile Ala Asp Thr Glu Trp Lys Asp 100

ACC TTA TCA CAA TTG TTC TTG ACA AAA GGA TGG CCA ACA GGA TCG GTG TAC TTT AAA GAA TAT ACT GAT ATA GCG 423
Thr Leu Ser Gln Leu Phe Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr Asp Ile Ala 125

GCC TTT TCA GTG GAA CCA CAG TTA TAC TGT GAT TAT AAT TTA GTT TTA ATG AAA TAT GAT TCT ACA CAG GAA CTA 498
Ala Phe Ser Val Glu Pro Gln Leu Tyr Cys Asp Tyr Asn Leu Val Leu Met Lys Tyr Asp Ser Thr Gln Glu Leu 150

GAT ATG TCT GAA TTG GCC GAT CTT ATA CTG AAC GAA TGG CTG TGC AAT CCA ATG GAC ATA ACG CTG TAT TAT TAT 573
Asp Met Ser Glu Leu Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr Leu Tyr Tyr Tyr 175

CAG CAG ACT GAT GAA GCA AAT AAA TGG ATA TCA ATG GGT TCT TCT TGC ACA GTC AAA GTG TGT CCA TTA AAT ACG 648
Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met Gly Ser Ser Cys Thr Val Lys Val Cys Pro Leu Asn Thr 200

CAA ACA CTT GGT ATT GGA TGT CTA ATA ACT AAT CCA GAC ACG TTT GAA ACG GTT GCG ACA ACG GAG AAG TTA GTG 723
Gln Thr Leu Gly Ile Gly Cys Leu Ile Thr Asn Pro Asp Thr Phe Glu Thr Val Ala Thr Thr Glu Lys Leu Val 225

ATT ACA GAT GTT GTA GAT GGT GTC AAC CAT AAA TTA AAC GTC ACA ACA GCA ACG TGC ACC ATA CGC AAC TGT AAA 798
Ile Thr Asp Val Val Asp Gly Val Asn His Lys Leu Asn Val Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys Lys 250

AAA TTA GGA CCA AGG GAA AAC GTA GCA ATC ATA CAG GTA GGC GGC GCA AAT GTT TTA GAC ATC ACA GCT GAT CCA 873
Lys Leu Gly Pro Arg Glu Asn Val Ala Ile Ile Gln Val Gly Gly Ala Asn Val Leu Asp Ile Thr Ala Asp Pro 275

ACA ACT GCA CCA CAG ACA GAG AGA ATG ATG CGA ATA AAC TGG AAA AAA TGG TGG CAA GTG TTT TAC ACA GTA GTC 948
Thr Thr Ala Pro Gln Thr Glu Arg Met Met Arg Ile Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Val Val 300

GAT TAC GTC AAT CAA ATA ATT CAA ACG ATG TCC AAA AGA TCT AGA TCA CTT AAT TCG TCA GCG TTC TAT TAC AGA 1023
Asp Tyr Val Asn Gln Ile Ile Gln Thr Met Ser Lys Arg Ser Arg Ser Leu Asn Ser Ser Ala Phe Tyr Tyr Arg 325

GTA TAGGTGCATGTTAGATTAGAGTTGTATGATGTGACC-3' 1062
Val 326
```

Figure 3. Nucleotide sequence of a cloned copy of the gene encoding the serotype-specific glycoprotein from UK bovine rotavirus. The sequence corresponds to the plus strand (mRNA sense) of the dsRNA in the 5'→3' direction. Nucleotide residues 1 to 9 of the gene were absent from the cloned copy and were deduced from mRNA transcripts. Amino acid residues in the longest potential coding region are numbered from the first ATG triplet.

DISCUSSION

The sequence contains only a single, long, open reading-frame. There are no protein sequence data available to indicate which AUG triplet initiates translation. Early in-phase AUG sequences occur at nucleotide positions 49 and 136. Since the latter AUG codon (AXXAUGG) is a 'strong' initiation sequence in eukaryotes and eukaryotic viruses (see ref. 38) which has not as yet been found in 5' non-coding regions, it is unlikely that initiation of translation occurs at any later point in the sequence. The first AUG triplet at position 49 in the nucleotide sequence is of the PyXXAUGPy type and this type is not generally found among the functional initiators in eukaryotes, but is the most common sequence among non-functional AUG triplets that occur upstream from the initiation site. However codon usage between these two in-phase AUG triplets is consistent with coding sequence (39), suggesting that initiation of translation occurs at the first AUG triplet in the mRNA sequence.

Molecular weight determinations of the non-glycosylated or *in vitro* translation product of gene 8 from SDS polyacrylamide gel electrophoresis cannot be relied upon to yield an accurate molecular weight and so indicate the site which initiates translation. Molecular weights of hydrophobic proteins are underestimated by as much as 40% in this technique (40,41). This effect probably accounts for the differing mobilities of the equivalent primary translation product from several rotavirus strains. The non-glycosylated form of the structural glycoprotein from cells infected by UK bovine rotavirus in the presence of tunicamycin (42) and the primary translation product of *in vitro* translation of gene 8 (14) both migrate more rapidly than the non-structural proteins VP8 (MW 35K) and VP9 (MW 34K), while in other viral strains the precursor of the structural glycoprotein migrates more slowly than or between the non-structural proteins (43).

No AAUAAA polyadenylation signal (44) sequence nor any similar variation was found near the 3' end of the RNA sequence. This sequence is generally located 11-30 nucleotides from the poly A tail of eukaryotic mRNAs. Its absence is consistent with the fact that rotavirus mRNAs are not polyadenylated (45).

Initiation of translation at the first AUG triplet would produce a protein of 326 amino acids. The composition of this protein is shown in Table 1. It contains a preponderance of acidic over basic amino acid residues (32:24) and has three potential glycosylation sites (Asn-X-$\begin{matrix}\text{Thr}\\\text{Ser}\end{matrix}$) at amino acid residue positions 69, 238 and 318. None of these positions are

Table 1. Amino acid composition of the longest potential coding region of gene 8 from UK bovine rotavirus

Amino Acid	Total No.	Amino Acid	Total No.
Ala	17	Leu	32
Arg	9	Lys	14
Asn	20	Met	11
Asp	17	Phe	9
½ Cys	8	Pro	11
Gln	13	Ser	20
Glu	15	Thr	38
Gly	12	Trp	7
His	1	Tyr	21
Ile	27	Val	24

Total amino acid residues: 326
Calculated molecular weight: 38,292

Table 2. Codon usage of the longest potential coding region of UK bovine rotavirus gene 8 (gene 7 usage in brackets)

		Second Position U		Second Position C		Second Position A		Second Position G	First Position Totals
U	Phe	6 UUU(10)	Ser	6 UCU (4)	Tyr	14 UAU(12)	Cys	5 UGU (3)	85
		3 UUC (6)		1 UCC (0)		7 UAC (1)		3 UGC (2)	
	Leu	10 UUA (4)		9 UCA(12)	Term	0 UAA (1)	Term	0 UGA (0)	
		10 UUG (9)		3 UCG (2)		1 UAG (0)	Trp	7 UGG (4)	
C	Leu	5 CUU (2)	Pro	1 CCU (0)	His	1 CAU (6)	Arg	0 CGU (0)	39
		0 CUC (3)		0 CCC (1)		0 CAC (4)		1 CGC (1)	
		3 CUA (6)		10 CCA (5)	Gln	6 CAA (9)		1 CGA (1)	
		4 CUG (4)		0 CCG (2)		7 CAG (3)		0 CGG (0)	
A	Ile	8 AUU(15)	Thr	6 ACU (6)	Asn	11 AAU(19)	Ser	1 AGU (2)	118
		4 AUC (0)		5 ACC (0)		9 AAC (7)		0 AGC (1)	
	Met	15 AUA (7)		18 ACA (4)	Lys	12 AAA(22)	Arg	6 AGA (9)	
		11 AUG (9)		9 ACG (5)		2 AAG (7)		1 AGG (2)	
G	Val	5 GUU (8)	Ala	2 GCU (8)	Asp	13 GAU (9)	Gly	5 GGU (5)	85
		5 GUC (1)		3 GCC (1)		4 GAC (5)		2 GGC (0)	
		6 GAU (6)		7 GCA (4)	Glu	11 GAA(16)		5 GGA (5)	
		8 GUG (7)		5 GCG (6)		4 GAG (4)		0 GGG (1)	
Second Position Totals		103		85		102		37	

Third position totals (U89, C47, A119, G72)

in regions predicted to be deeply buried (see below), while that at position 69 is in a potential surface region. Analysis of the codon usage in the potential coding region (Table 2) shows that at codon position three the choice of nucleotides is similar to the overall nucleotide frequencies in the gene, while at codon position two the nucleotide frequencies show a high C content and a low G content compared to the gene nucleotide frequencies. This distribution of nucleotides is very similar to that of gene 7, which encodes a non-structural protein yet exhibits a similar allocation of nucleotides between the three codon positions, and a somewhat similar codon usage (Table 2). The similarity of codon usage by gene 7 and gene 8 despite the functionally different products agrees with the general finding of Grantham et al. (46) that genes in the same genome exhibit characteristic codon usage among synonymous codons.

Parasites and hosts often have distinguishable codon usage patterns (47). In eukaryotic cellular genes the hierarchy amongst the third base usage of synonymous codons is C>G>U,A, while for animal viral genes the hierarchy is U,A>C,G (48). This preference is quite obviously sustained by the choice of the degenerate third base in rotavirus genes (Table 2).

The translation product of gene 8 is the serotype-specific glycoprotein present as the major component of the outer shell of the virus capsid. Ultrastructural immunocytochemistry studies (49) have indicated that whereas inner capsid proteins are synthesized throughout the cytoplasm and become concentrated in viroplasmic inclusions, the outer capsid glycoprotein is synthesized primarily on ribosomes of the rough endoplasmic reticulum, with glycosylation occurring cotranslationally in the lumen of the endoplasmic reticulum (42, 43). The outer capsid layer appears to be acquired either during or subsequent to virus budding into the cisternae of the endoplasmic reticulum. Although the mature infectious virion is not enveloped, a transient envelope is acquired from the endoplasmic reticulum during budding (50). These observations would imply the presence, within the protein sequence, of a signal peptide function to initiate translocation and a membrane insertion sequence. Both of these features of a protein sequence are generally associated with hydrophobic regions. In this context a hydropathy profile (Fig. 4) which indicates the local hydrophilic or hydrophobic tendencies of a polypeptide chain (51) is of interest. The highest hydropathy value is found within a contiguous length of 27 uncharged amino acids (residues 37-63), the earlier part of which (residues 37-44) is highly enriched in hydrophobic residues. The hydropathy index here exceeds,

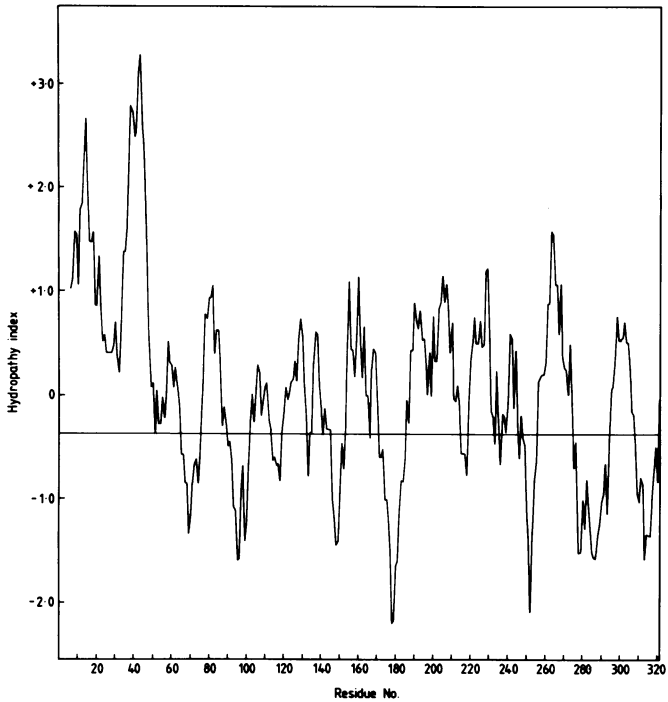

Figure 4. Hydropathy profile of the longest potential protein sequence encoded by gene segment 8 from UK bovine rotavirus. The mean hydropathy index at each sequence position was calculated using the amino acid parameters of Kyte and Doolittle (51) and a span of eleven residues. The mean hydropathy index for soluble proteins is indicated as a mid-line, hydrophobic regions have positive hydropathy. The sequence numbering corresponds with Figure 3.

by three standard deviations, the values generally associated with peptides that traverse the interior of a protein (51). The value resembles more those values associated with hydrophobic membrane-spanning segments as found in integral membrane proteins (51) or those same segments in the now lipid-free protein coats around fd-type bacteriophage (52,53). This hydrophobic region of the rotavirus outer-shell protein therefore has the potential for forming the intermolecular contacts within the capsid of the virus.

A shorter, uncharged and again exceptionally hydrophobic segment is located closer to the potential N-terminus of the protein (residues 6-23). This region could also be involved in intermolecular contacts within the capsid, or it might function as a signal sequence for translocation (54).

Neutralising antibody in rotavirus infection is directed mainly against

antigenic determinants on the major, outer-shell glycoprotein. Antigenic determinants are surface features of proteins frequently found on highly exposed surface regions of a molecule. Using the amino acid parameters of Hopp and Woods for locating antigenic determinants (55), the region 311-316 (Ser-Lys-Arg-Ser-Arg-Ser) is predicted as a likely location of an antigenic determinant. Other regions of high hydrophilicity which might be considered as locating antigenic determinants are 95-100, 179-183 and 251-256. Often highly conserved regions of a protein molecule are found adjacent to antigenic determinants to maintain structural integrity (56). In this regard the presence of two potentially homologous cystine-containing peptides, viz. $\overset{190}{\text{Ser}}$-Cys-Thr-Val-Lys-Val-Cys and $\overset{243}{\text{Thr}}$-Cys-Thr-Ile-Arg-Asn-Cys found adjacent to regions of sequence predicted as highly exposed is interesting. It implies a conserved structure in spite of the high rate of mutation in RNA viruses that is generally sufficient to preclude such internal sequence homologies. It will be interesting to see to what extent these features are present in the major, outer-shell glycoproteins of other rotavirus serotypes.

ACKNOWLEDGEMENTS

M.D.S. is supported by the National Health and Medical Research Council and the Diarrheal Diseases Programme of the W.H.O.

REFERENCES

1. Holmes, I.H. (1979) Prog.Med.Virol. 25, 1-36.
2. Holmes, I.H. (1983) in The Reoviruses, Joklik W.K., Ed., Plenum Publ. Co., New York, in press.
3. Rodger, S.M., Schnagl, R.D. and Holmes, I.H. (1975) J.Virol. 16, 1229-1235.
4. Rodger, S.M., Schnagl, R.D. and Holmes, I.H. (1977) J.Virol. 24, 91-98.
5. Bastardo, J.W., McKimm-Breschkin, J.L., Sonza, S., Mercer, L.D. and Holmes, I.H. (1981). Infection and Immunity 34 641-647.
6. Killen, H.M. and Dimmock, N.J. (1982) J.Gen.Virol. 62, 297-311.
7. Flewett, T.H., Thouless, M.E., Pilford, J.N., Bryden, A.S. and Candeias, J. (1978) Lancet ii, 632.
8. Sato, K., Inaba, Y., Miura, Y., Tokuhisa, S. and Matumoto, M. (1982) Arch. Virol. 73, 45-50.
9. McNulty, M.S., Allan, G.M., Todd, D., McFerran, J.B., McKillop, E.R., Collins, D.S. and McCracken, R.M. (1980) Avian Pathol. 9, 363-375.
10. Wyatt, R.G., Greenberg, H.B., James, W.D., Pittman, A.L., Kalica, A.R., Flores, J., Chanock, R.M. and Kapikian, A.Z. (1982) Infect. Immun. 37, 110-115.
11. Kalica, A.R., Greenberg, H.B., Wyatt, R.G., Flores, J., Sereno, M.M., Kapikian, A.Z. and Chanock, R.M. (1981). Virology 112, 385-390.
12. Greenberg, H.B., Kalica, A.R., Wyatt, R.G., Jones, R.W., Kapikian, A.Z. and Chanock, R.M. (1981) Proc. Nat. Acad. Sci. USA 78, 420-424.
13. Dyall-Smith, M.L., Azad, A.A. and Holmes, I.H. (1983) J. Virology 46, 317-320.

14. McCrae, M.A. and McCorquodale, J.G. (1982) Virology 117, 435-443.
15. Dyall-Smith, M.L., Elleman, T.C., Hoyne, P.A., Holmes, I.H. and Azad, A.A. (1983) Nucleic Acid Res. 11 No.10, in press.
16. Smith, M.L., Lazdins, I. and Holmes, I.H. (1980) J. Virol. 33, 976-982.
17. Dyall-Smith, M.L. and Holmes, I.H. (1981) J. Virol. 38, 1099-1103.
18. Casadaban, M.J. and Cohen, S.N. (1980) J.Mol.Biol. 138, 179-207.
19. Grunstein, M. and Hogness, D.S. (1975) Proc.Natl.Acad.Sci. USA 72, 3961-3965.
20. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J.Mol.Biol. 113, 237-251.
21. Alwine, J.C., Kemp, D.J., Parker, B.A., Reisert, J., Renart, J., Stark, G.R. and Wahl, G.M. (1979) Methods Enzymol. 68, 220-242.
22. Seed, B. (1982) Nucleic Acid Res. 10, 1799-1810.
23. Kupersztoch, Y.M. and Helinski, D.R. (1973) Biochem.Biophy.Res.Comm. 54, 1451-1459.
24. Radloff, R., Bauer, W. and Vinograd, J. (1967) Proc.Natl.Acad.Sci. USA 57 1514-1521.
25. Maxam, A.M. and Gilbert, W. (1977) Proc.Natl.Acad.Sci. USA 74, 560-564.
26. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc.Natl.Acad.Sci. USA 74, 5463-5467.
27. Messing, J. and Vieira, J. (1982) Gene 19, 269-276.
28. Winter, G. and Fields, S. (1980) Nucleic Acids Res. 8, 1965-1974.
29. Bentley, D.L. and Rabbits, T.H. (1980) Nature (Lond.) 288, 730-733.
30. Schreier, P.H. and Cortese, R. (1979) J.Mol.Biol.129, 169-172.
31. Anderson, S., Gait, M.J., Mayol, L. and Young, I.G. (1980) Nucleic Acids Res. 8, 1731-1745.
32. Duckworth, M.L., Gait, M.J., Goelet, P., Hong, G.F., Singh, M., and Titmus, R.C. (1981) Nucleic Acids Res. 9, 1691-1706.
33. Smith, A.J.H. (1980) Methods Enzymol. 65, 560-580.
34. Alton, N.K. and Vapneck, D. (1979) Nature 282, 864-869.
35. Cohen, J., Laporte, J., Charpilienne, A. and Scherrer, R. (1979) Arch.Virol. 60, 177-186.
36. McMaster, G.K. and Carmichael, G.G. (1977) Proc.Natl.Acad.Sci. USA, 74, 4835-4838.
37. Imai, M., Richardson, M.A., Ikegami, N., Shatkin, A.J. and Furuichi, Y. (1983) Proc.Natl.Acad.Sci., 80, 373-377.
38. Kozak, M. (1982) Messenger RNA and Ribosomes in Protein Synthesis, Biochem.Soc.Symp. 47 pp.113-128. Phelps, C.F. and Arnstein, H.R.V. Eds. The Biochemical Soc. London.
39. Staden, R., and McLachlan, A.D. (1982) Nucleic Acid Res. 10, 141-156.
40. Bonitz, S.G., Coruzzi, G., Thalenfeld, B.E., Tzagoloff, A. and Macino, G. (1980) J. Biol. Chem. 255, 11927-11941.
41. Hudspeth, M.E.S., Ainley, W.M., Shumard, D.S., Butow, R.A. and Grossman, L.I. (1982) Cell 30 617-626.
42. McCrae, M.A. and Faulkner-Valle, G.P. (1981) J. Virology 39, 490-496.
43. Ericson, B.L., Graham, D.Y., Mason, B.B. and Estes, M.K. (1982) J. Virology 42, 825-839.
44. Fitzgerald, M. and Shenk, T. (1981) Cell 24, 251-260.
45. Bernstein, J.M. and Hruska, J.F. (1981) J. Virology 37, 1071-1074.
46. Grantham, R., Gautier, C., Gouy, M., Mercier, R. and Pave, A. (1980) Nucleic Acid Res. 8 r49-r62.
47. Grantham, R., Gautier, C. and Gouy, M. (1980) Nucleic Acid Res. 8, 1893-1912.
48. Wain-Hobson, S., Nussinov, R., Brown, R.J. and Sussman, J.L. (1981) Gene 13, 355-364.
49. Petrie, B.L., Graham, D.Y., Hanssen, H. and Estes, M.K. (1982) J. Gen. Virol. 63, 457-467.

50. Petrie, B.L., Graham, D.Y., Estes, M.K. (1981) Intervirology 16, 20-28.
51. Kyte, J., and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105-132.
52. Asbeck, F., Beyreuther, K., Köhler, H., von Wettstein, G. and Braunitzer, G. (1969) Hoppe-Seyler's Z. Physiol. Chem. 350, 1047-1066.
53. Smilowitz, H., Carson, J. and Robbins, P.W. (1972) J. Superamol. Struct. 1, 8-18.
54. Blobel, G., Walter, P., Chang, C.N., Goldman, B.M., Erickson, A.H. and Lingappa, V.R. (1979). Symposium of the Society of Experimental Biology, (Great Britain), (Hopkins, C.R. and Duncan, C.J. Eds.) vol.33, pp.9-36, Cambridge University Press, London.
55. Hopp, T.P. and Woods, K.R. (1981) Proc.Natl.Acad.Sci. USA 78, 3824-3828.
56. Wiley, D.C., Wilson, I.A. and Skehel, J.J. (1981) Nature (London) 289, 373-378.