Primary Structure of the Neutralization Antigen of Simian Rotavirus SA11 as Deduced from cDNA Sequence

CARLOS F. ARIAS,†* SUSANA LÓPEZ,† JOHN R. BELL,‡ AND JAMES H. STRAUSS

Division of Biology, California Institute of Technology, Pasadena, California 91125

DNA sequences complementary to the double-stranded RNA coding for the neutralization antigen (glycoprotein VP7) of simian rotavirus SA11 have been cloned. The VP7 gene consists of 1,062 nucleotides, containing an uninterrupted coding sequence of 978 nucleotides which specifies a glycoprotein of 326 amino acids. The significance of a second possible initiation site 30 nucleotides downstream from the first is discussed. Partial amino acid sequence of this glycoprotein showed unequivocally that the cloned segment (segment 9) codes for glycoprotein VP7 of SA11. The resulting amino acid sequence contained only one carbohydrate acceptor site. Possible sites of membrane interaction and antigenic determinants are discussed based on the analysis of the hydrophobicity and hydrophilicity profiles of VP7.

The rotaviruses are principal etiological agents of acute viral gastroenteritis in the young of a number of mammalian and avian species, including humans, and as such they are of major medical and veterinary importance (10, 13).

These viruses possess a genome that consists of 11 segments of double-stranded RNA (dsRNA) (28). In the simian rotavirus SA11, each dsRNA segment encodes a single viral protein, five of which are incorporated into the mature virion as a double capsid (1, 7). The major outer-capsid polypeptide, VP7, is a glycoprotein (1) coded by RNA segment 9 (see below), which is thought to be responsible for the induction and binding of neutralizing antibodies to the virus (17). To define the antigenic properties of glycoprotein VP7 of simian rotavirus SA11, we cloned a cDNA copy of the glycoprotein gene. The nucleotide sequence of this cDNA allowed us to deduce the amino acid sequence of this glycoprotein and to predict several of its features. Two strategies were used for cloning the SA11 genome. In the first, 10 μg of virion dsRNA was denatured in 90% dimethyl sulfoxide at 70°C for 10 min, precipitated with 2.5 volumes of ethanol, and used as a template for the reverse transcriptase-directed synthesis of cDNA, with DNase I-digested calf thymus DNA used as the random primer (27). Single-stranded cDNA copies of SA11 dsRNAs were treated with 0.1 N NaOH for 60 min at 60°C and separated from degraded RNA in a Bio-Gel A-5M filtration column (27). The single-stranded cDNA was converted to the double-stranded form by utilizing its ability to form 3′-terminal hairpin loops which prime DNA polymerase I (31). Subsequent cleavage of the hairpin loop with nuclease S1 (6) yielded noncovalently joined double-stranded cDNA (dscDNA), which in turn was tailed with polydeoxycytidylic acid by using terminal transferase (29).

In the second strategy, virion dsRNA was denatured and precipitated as above and resuspended in 150 μl of polyadenylation reaction mixture containing 1 U of *Escherichia coli* polyadenylic acid polymerase purified by the method of Sippel (32). An average of 30 adenylate residues were added

to each single-stranded RNA molecule. Polyadenylated RNA was isolated on oligodeoxythymidylic acid-cellulose and precipitated with ethanol (25). After redenaturation as described above, the polyadenylated RNA was transcribed into cDNA by using oligodeoxythymidylic $acid_{12-18}$ as primer (27). Subsequent to alkali digestion of the RNA and separation of the degraded products as above, the single-stranded cDNA was resuspended in 10 mM Tris (pH 7.5), heat denatured at 100°C for 2 min, and self-annealed in 0.25 M NaCl for 6 h at 65°C. The dscDNA formed in this way was recovered by ethanol precipitation and, to increase the probability of obtaining full-length dscDNA, was incubated for 12 h at 14°C in a filling-in reaction mixture containing 12 U of DNA polymerase I (31). This dscDNA was C tailed as above.

The tailed dscDNA obtained by either of the two methods was annealed (24) to pMT21, linearized with *Pst*I, and tailed with dGMP residues. (pMT21, a derivative of pBR322, was obtained from H. Huang; the first 2,520 base pairs were deleted, and a polylinker was inserted which contained the following 10 unique restriction sites: *Sac*II, *Eco*RI, *Ava*I, *Sma*I, *Bam*HI, *Sal*I, *Pst*I, *Bgl*II, *Xba*I, and *Hin*dIII. The *Pst*I site in the ampicillin resistance gene of pBR322 is not present in pMT21.) The hybrid plasmids were transfected into *E. coli* MC1061 (4), and transformants containing plasmids with inserts were selected by screening for resistance to ampicillin. Individual SA11 dsRNA segments isolated by gel electrophoresis (1) were used to identify the gene segments represented in cloned recombinant plasmids. The segments were denatured in dimethyl sulfoxide as described above, spotted onto nitrocellulose strips, baked for 2 h at 80°C, and hybridized to ^{32}P-labeled recombinant plasmid (prepared by nick translation [21]) in 50% formamide–5× SSPE (0.9 M NaCl, 5 mM EDTA, 50 mM $NaPO_4$, pH 7.7)–0.1% sodium dodecyl sulfate–1× Denhardt solution (5)–100 μg of salmon sperm DNA per ml at 42°C for 8 to 16 h. Plasmid pSR9-3 selected in this way was used to screen a larger number of clones by colony hybridization (11).

Sequencing of the inserts in the selected clones was done by the procedure of Maxam and Gilbert (23), starting from the *Eco*RI or *Hin*dIII sites of the plasmid, as described by Ruther et al. (30).

Five clones containing sequences of RNA segment 9 were initially obtained by using the random primer cloning proce-

---

* Corresponding author.

† Present address: Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, México 20, D.F.

‡ Present address: Genentech, Inc., South San Francisco, CA 94080.

FIG. 1. Inserts of cDNA from the simian rotavirus SA11 glycoprotein VP7 gene in the different clones used to obtain the complete nucleotide sequence of this gene. Scale bar. Positive strand numbered from the 5' to 3' end. The sequenced portions and their direction of reading in each clone are indicated by the heavy lines and the arrows, respectively. Clone pSR9-65 was obtained by the oligodeoxythymidylic acid priming procedure, whereas the other clones were products of the random primer cloning strategy.

```
1                                                              M   Y   G   I    4
1    ....TAAAAAGAGAGAATTTCCGTTTGGCTAGCGGTTAGCTCCTTTTA ATG  TAT GGT ATT   60

5    E   Y   T   T   V   L   T   F   L   I   S   T   I   L   L   N    20
61   GAA TAT ACC ACA GTT CTA ACC TTT CTG ATA TCG ACT ATT CTA CTA AAT   108

21   Y   I   L   K   S   L   T   R   I   M   D   F   I   I   Y   R    36
109  TAC ATA CTT AAA TCA TTA ACT AGA ATA ATG GAC TTT ATA ATT TAT AGA   156

37   F   L   F   I   I   V   I   L   S   P   F   L   R   A   Q   N    52
157  TTT CTT TTT ATA ATT GTG ATA TTG TCA CCA TTT CTT AGA GCA CAA AAT   204

53   Y   G   I   N   L   P   I   A   G   S   M   D   T   A   Y   A    68
205  TAT GGT ATT AAT CTT CCA ATC GCA GGC TCC ATG GAC ACT GCA TAC GCT   252

69   N   S   T   Q   E   E   P   F   L   T   S   T   L   C   L   Y    84
253  AAT TCA ACG CAA GAA GAA CCA TTC CTC ACT TCT ACA CTT TGC CTA TAT   300

85   Y   P   T   E   A   A   T   E   I   N   D   N   S   W   K   D    100
301  TAT CCG ACT GAG GCC GCG ACT GAA ATA AAC GAT AAT TCA TGG AAA GAC   348

101  T   L   S   Q   L   F   L   T   K   G   W   P   T   E   S   V    116
349  ACA CTG TCA CAA CTA TTT CTT ACG AAA GGG TGG CCA ACT GAA TCC GTA   396

117  Y   F   K   E   Y   T   N   I   A   S   F   S   V   D   P   Q    132
397  TAT TTT AAA GAA TAT ACT AAC ATT GCA TCG TTT TCT GTT GAT CCG CAG   444

133  L   Y   C   D   Y   N   V   V   L   M   K   Y   D   A   T   L    148
445  TTG TAT TGT GAT TAT AAC GTA GTA CTA ATG AAA TAT GAC GCG ACG TTG   492

149  Q   L   D   M   S   E   L   A   D   L   I   L   N   E   W   L    164
493  CAA TTG GAT ATG TCA GAA CTT GCG GAT CTA ATA TTA AAT GAA TGG TTG   540

165  C   N   P   M   D   I   T   L   Y   Y   Y   Q   Q   T   D   E    180
541  TGT AAT CCA ATG GAT ATT ACT CTG TAT TAT TAT CAG CAA ACT GAC GAA   588

181  A   N   K   W   I   S   M   G   S   S   C   T   I   K   V   C    196
589  GCG AAT AAA TGG ATA TCA ATG GGC TCA TCA TGT ACA ATT AAA GTA TGT   636

197  P   L   N   T   Q   T   L   G   I   G   C   L   T   T   D   A    212
637  CCA CTT AAT ACA CAA ACT CTT GGA ATT GGA TGC TTG ACA ACT GAT GCT   684

213  T   T   F   E   E   V   P   T   A   E   K   L   V   I   T   D    228
685  ACA ACT TTT GAA GAA GTG CCG ACA GCT GAA AAG TTG GTA ATT ACT GAC   732

229  V   V   D   G   V   N   H   K   L   D   V   T   T   A   T   C    244
733  GTG GTT GAT GGC GTT AAT CAT AAG CTG GAT GTC ACA ACA GCA ACG TGT   780

245  T   I   R   N   C   K   K   L   G   P   R   E   N   V   A   V    260
781  ACT ATT AGA AAC TGT AAG AAA TTG GGA CCA AGA GAA AAC GTA GCC GTT   828

261  I   Q   V   G   G   S   D   I   L   D   I   T   A   D   P   T    276
829  ATA CAA GTT GGT GGT TCT GAC ATC CTC GAT ATA ACT GCT GAT CCA ACT   876

277  T   A   P   Q   T   E   R   M   M   R   I   N   W   K   K   W    292
877  ACT GCA CCA CAG ACA GAA CGG ATG ATG CGA ATT AAC TGG AAA AAA TGG   924

293  W   Q   V   F   Y   T   V   V   D   Y   V   D   Q   I   I   Q    308
925  TGG CAA GTT TTT TAT ACT GTA GTA GAC TAT GTA GAT CAG ATA ATA CAA   972

309  V   M   S   K   R   S   R   S   L   N   S   A   A   F   Y   Y    324
973  GTT ATG TCC AAA AGA TCA AGA TCA CTA AAT TCA GCA GCA TTT TAT TAC   1020

325  R   V   X                                                        326
1021 AGA GTG TAG GTATAACTTAGGTTAGAATTGTATGATGTGACC                    1062
```

FIG. 2. DNA sequence of the mRNA sense strand. and the deduced amino acid sequence for simian rotavirus SA11 glycoprotein VP7. Nucleotide residues are numbered from the 5' to 3' end. In clone pSR9-65, the 5'-terminal sequence presented did not extend to the 5' end of the mRNA (four nucleotides were missing; see text). The 3'-terminal sequence of this clone was followed by a polyadenylic acid tract connected with the vector DNA sequence through a polycytidylic acid tract, as expected from the cloning methodology employed, and thus represented the complete sequence of this region. The predicted amino acid sequence is displayed above the nucleotide sequence, with the first of the two in-phase termination codons denoted with an X. Amino acid residues are numbered beginning with the first methionine. Two possible initiator methionines are boxed. The asterisk marks the possible

dure. When these clones were subjected to nucleotide sequence analysis, it was found that the reported sequences for the 5' and 3' ends of this gene (3) were missing (Fig. 1). To obtain the missing sequences, the second cloning strategy was used. After selection, the sizes of 21 clones positive for segment 9 were analyzed in agarose gels. Only one insert, that in clone pSR9-65, was of a length sufficient to represent a full-length copy of the gene. The sequence of this insert was found to be missing only four nucleotides at the 5' end of the plus strand (see below). The portion of the clones analyzed to obtain the nucleotide sequence of the gene is shown in Fig. 1.

Glycoprotein VP7 of simian rotavirus SA11 is thought to be the polypeptide that elicits neutralizing antibody. Genetic studies with the human Wa strain and with the bovine UK strain have led to the assignment of the gene carrying the neutralization activity to dsRNA segment 9 (17). Northern blot analysis has shown that segment 9 in strain SA11 corresponds to segment 9 in strain Wa (5a). Translation experiments in vitro with simian rotavirus SA11 suggested that genome segment 9 encodes the precursor to the viral structural glycoprotein (1, 22, 33). The data we present here confirms that segment 9 indeed encodes glycoprotein VP7. Partial amino acid sequence was obtained directly from the glycoprotein (see below). Glycoprotein VP7 from purified simian rotavirus SA11 was isolated by preparative gel electrophoresis. NH2 terminal sequence of the purified protein revealed that it was blocked, which was somewhat unexpected in view of the fact that the mature glycoprotein is produced by cleavage of the precursor (see below) (7a). To obtain amino acid sequence data from the glycoprotein, we therefore sequenced a fragment obtained after V8 protease digestion. The sequence obtained matched that of amino acids 221 to 237 of the deduced polypeptide (underlined in Fig. 2). This finding showed unequivocally that dsRNA segment 9 of SA11 codes for VP7 and identified the coding

carbohydrate attachment site. The underlined amino acid sequence (from amino acids 221 through 237) was confirmed by direct sequencing of a V8 protease-produced peptide. To obtain this sequence, 10 µg of gel-purified VP7 was digested with 1 µg of V8 protease for 4.5 h at 37°C. The peptides produced were separated in a 15% acrylamide gel and electroeluted separately (1, 15). The protein sequence determination was done as previously described (2, 12). The sequenced peptide began at Val-218, but no assignments could be made for the first three cycles.
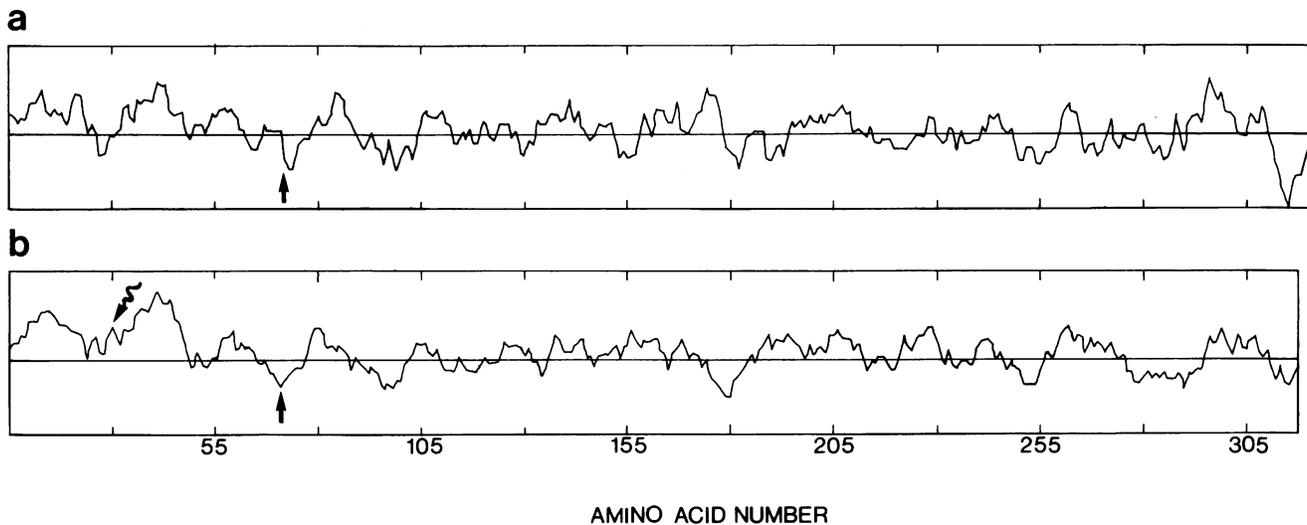
a

b

AMINO ACID NUMBER

FIG. 3. Hydrophobic and hydrophilic regions of the SA11 neutralization antigen. The primary amino acid sequence of glycoprotein VP7 was computer scanned for antigenic determinants with the program of Hopp and Woods (14) (a) or for regions of hydrophobicity with the program of Kyte and Doolittle (19) (b). The degree of hydrophobicity increases above the horizontal line; that of hydrophilicity increases with distance below the horizontal line. The search length was six amino acids for the Hopp and Woods program and nine for the Kyte and Doolittle program. Vertical marks in the horizontal axes correspond to 25 amino acid residues. Straight arrows in (a) and (b), Position of the potential site of glycosylation. Wavy arrow in b, Position of the second possible initiator methionine (see text).

strand of the gene. The complete cDNA sequence (except for the first four bases at the 5' end) is shown in Fig. 2.

The cDNA sequence at the 3' terminus matched exactly the terminal sequence previously reported for gene 9, including the last eight nucleotides, which have been shown to be conserved not only in the different segments of SA11 but also in the genes of different rotavirus strains (3). In comparison with the reported 5'-terminal sequence (3), clone pSR9-65 lacked the first four nucleotides of that end and also lacked the expected deoxyribosyladenine-deoxyribosylthymine homopolymer tract.

From the sequence presented in Fig. 2, the base composition for the plus strand of dsRNA segment 9 was found to be 33% A, 19% G, 31% U, and 17% C. The low GC content of the gene is reflected in the scarcity of the CpG doublet, a feature that has been found in RNA and DNA viruses with vertebrate hosts (9, 26).

The entire nucleotide sequence (including the missing first four bases) was found to be 1,062 base pairs long. The sequence coding for VP7 consisted of an open reading frame of 978 bases which started at nucleotide 49 with the first ATG and ended at nucleotide 1,025 with the first of two in-phase termination codons. In this reading frame, the usage for the possible codons for each amino acid was clearly nonrandom. The resulting amino acid sequence contained only one carbohydrate acceptor site (Asn-Ser-Thr) at amino acid positions 69 through 71. The 5'- and 3'-noncoding sequences were 48 and 36 nucleotides long, respectively. The absence of the sequence AAUAAA, present in eucaryotic polyadenylated mRNAs, from the 3'-noncoding region of this gene is consistent with the lack of polyadenylation in the mRNAs of Reoviridae (16).

An open reading frame of 369 nucleotides was found in the minus strand; however, since the first ATG found in this open region was 235 nucleotides downstream from the start of the region, it is unlikely that it codes for a protein. The other reading frames of both the coding and the complementary strands contained numerous stop codons, such that the longest polypeptide that could be derived from them would be less than 50 amino acids long.

It is interesting to note that the first ATG of the sequence coding for VP7 did not have the consensus sequence which was described by Kozak as being necessary for an efficient initiation of translation (18) but had, instead, pyrimidines in positions +4 and −3. However, one preproinsulin initiation codon engineered in vitro with these characteristics was very efficiently used in translation (20). On the other hand, the second ATG, found 87 nucleotides downstream, had the classical sequence for an initiator ATG. The putative polypeptides resulting from initiation at these two possible sites would be 326 and 297 amino acids long with molecular weights of 37,368 and 33,919, respectively. Because the reported molecular weight for the nonglycosylated precursor of VP7 is 37,000 (7a), translation would appear to start at the first ATG. However, several observations have indicated that VP7 of SA11 is polymorphic, and nonglycosylated precursors of VP7 of smaller molecular weight have been observed (8). SA11 strains with different but homogeneous VP7 have been obtained after plaque purification of the virus (8). The sequence reported here suggests that polymorphism might be the consequence of the existence of virus clones which use either the first or second initiation codon of RNA segment 9.

In an attempt to predict the possible antigenic determinants and to find hydrophobic sequences which might serve as signal sequences or otherwise interact with membranes, we calculated the hydrophilicity curves along the predicted amino acid chain by the methods described by Hopp and Woods (14) and Kyte and Doolittle (19) (Fig. 3). The most hydrophilic peak, which according to Hopp and Woods probably represents an antigenic determinant, was found around amino acid 313, close to the COOH terminus of the glycoprotein (Fig. 3a). Other hydrophilic regions around amino acids 72 (near the recognition site for carbohydrate

attachment) and 180 could also represent antigenic determinants. Knowledge of the antigenic determinants of glycoprotein VP7 is important for the study of the neutralizable regions of the glycoprotein.

The hydrophobicities of the regions from amino acid 4 through 23 and from amino acid 29 through 48 (Fig. 3b), which have an average hydrophobicity of 1.61 and 2.25, respectively, indicate that these regions could be signal peptides or sites for interaction with the membrane or both. The cleavage of the signal peptide, with a size of about 1,500 daltons, from the 37,000-molecular-weight nonglycosylated precursor has been shown by using a plaque-purified virus which contains only one class of glycoprotein, with a molecular weight of 38,000 (7a). If initiation can also start at the second ATG, the second hydrophobic region may also function as a signal peptide. Hence, the polymorphism observed in VP7 may be the result of the existence of two initiation and two cleavage sites in this protein. Sequencing of the 5'-terminal region of the coding strand of segment 9 from virus clones differing in VP7 would provide insight into this problem, which may be involved in virus adaptation.

ADDENDUM

Just before this manuscript was submitted, the sequence of VP7 appeared (G. W. Both, J. S. Mattick, and A. R. Bellamy, Proc. Natl. Acad. Sci. U.S.A. 80:3091-3095, 1983). Our work confirms the coding assignments for VP7 by a different method, the main features of the nucleotide sequence of gene 9, and the inferred protein sequence. Our SA11 was obtained from H. H. Malherbe in 1977, and significant divergence from other preparations might be expected after continuous passage for more than 5 years. However, comparison of the obtained sequences reveals that they have been highly conserved: only 12 nucleotides were found to differ in both strains. Seven of these mutations represented a change in an amino acid, and the other five were silent. Three of the amino acid changes were located in the hydrophobic regions proximal to the NH₂ terminus, but all three were conservative. Only one of the other four amino acid changes is significative, involving nucleotide 389 and changing Gly-114 to Glu-114. None of the nucleotide mutations occurred at the two possible initiation sites.

After submitting this manuscript, a closely related paper describing the nucleotide sequence of the gene encoding the neutralizing antigen of the bovine rotavirus UK appeared (T. C. Elleman, P. A. Hoyme, M. L. Dyall-Smith, I. H. Holmes, and A. A. Azad, Nucleic Acids Res. 11:4689-4701, 1983). The characteristics of the UK gene are strikingly similar to those reported here for the VP7 SA11 gene: it is 1,062 base pairs long, with a long open reading frame of 326 amino acids starting at the first of two in-phase possible initiation codons at nucleotide positions 49 and 136. The hydrophatic profile of the deduced protein showed two NH₂-terminal hydrophobic regions and the presence of a carboxy-terminal highly hydrophilic region that may represent a major antigenic determinant.

In addition to the one possible glycosylation site found in glycoprotein VP7 of simian rotavirus SA11 at amino acid 69, there are two more putative carbohydrate attachment sites in the bovine protein at amino acids 238 and 318. By comparing the UK and SA11 sequences, we found a 77% homology at the nucleotide level and a 86% homology at the amino acid level.

LITERATURE CITED

1. Arias, C. F., S. López, and R. T. Espejo. 1982. Gene protein products of SA11 simian rotavirus genome. J. Virol. 41:42-50.
2. Bell, J. R., M. W. Bond, M. W. Hunkapiller, E. G. Strauss, J. H. Strauss, K. Yamamoto, and B. Simizu. 1983. Structural proteins of western equine encephalitis virus: amino acid compositions and N-terminal sequences. J. Virol. 45:708-714.
3. Both, G. W., A. R. Bellamy, J. E. Street, and L. J. Siegman. 1982. A general strategy for cloning double-stranded RNA: nucleotide sequence of the simian-11 rotavirus gene 8. Nucleic Acids Res. 10:7075-7088.
4. Casadaban, M. J., and S. N. Cohen. 1980. Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. J. Mol. Biol. 138:179-207.
5. Denhardt, D. T. 1966. A membrane filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. 23:641-646.
5a. Dyall-Smith, M. L., A. A. Azad, and I. H. Holmes. 1983. Gene mapping of rotavirus double-stranded RNA segments by Northern blot hybridization: application to segments 7, 8, and 9. J. Virol. 46:317-320.
6. Efstratiadis, A., F. C. Kafatos, A. M. Maxam, and T. Maniatis. 1976. Enzymatic in vitro synthesis of globin genes. Cell 7:279-288.
7. Ericson, B. L., D. Y. Graham, B. B. Mason, and M. K. Estes. 1982. Identification, synthesis, and modification of simian rotavirus SA11 polypeptides in infected cells. J. Virol. 42:825-839.
7a. Ericson, B. L., D. Y. Graham, B. B. Mason, H. H. Hanssen, and M. K. Estes. 1983. Two types of glycoprotein precursors are produced by the simian rotavirus SA11. Virology 127:320-332.
8. Estes, M. K., D. Y. Graham, R. F. Ramig, and B. L. Ericson. 1982. Heterogeneity in the structural glycoprotein (VP7) of simian rotavirus SA11. Virology 122:8-14.
9. Fiers, W., R. Contreras, G. Haegeman, R. Rogiers, A. Van de Voorde, H. Van Heuverswyn, J. Van Herreweghe, G. Volckaert, and M. Ysebaert. 1978. Complete nucleotide sequence of SV40. Nature (London) 273:113-120.
10. Flewett, T. H., and G. N. Woode. 1978. The rotaviruses. Arch. Virol. 57:1-23.
11. Grunstein, M., and D. S. Hogness. 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. U.S.A. 72:3961-3965.
12. Hewick, R. M., M. W. Hunkapiller, L. E. Hood, and W. J. Dreyer. 1981. A gas-liquid solid phase peptide and protein sequenator. J. Biol. Chem. 256:7990-7997.
13. Holmes, I. H. 1979. Viral gastroenteritis. Prog. Med. Virol. 25:1-36.
14. Hopp, T. P., and K. R. Woods. 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828.
15. Hunkapiller, M. W., E. Lujan, F. Ostrander, and L. E. Hood. 1982. Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. Methods Enzymol. 91:227-236.
16. Joklik, W. K. 1981. Structure and function of the reovirus genome. Microbiol. Rev. 45:483-501.
17. Kalica, A. R., H. B. Greenberg, R. G. Wyatt, J. Flores, M. M. Sereno, A. Z. Kapikian, and R. M. Chanock. 1981. Genes of human (strain Wa) and bovine (strain UK) rotaviruses that code for neutralization and subgroup antigens. Virology 112:385-390.
18. Kozak, M. 1981. Possible role of flanking nucleotide in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res. 9:5233-5252.
19. Kyte, J., and R. F. Doolittle. 1982. A simple method for

displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.

20. **Lomedico, P. T., and S. J. McAndrew.** 1982. Eukaryotic ribosomes can recognize preproinsulin initiation codons irrespective of their position relative to the 5' end of mRNA. Nature (London) **299**:221–226.

21. **Maniatis, T., A. Jeffrey, and D. G. Kleid.** 1975. Nucleotide sequence of the rightward operator of phage λ. Proc. Natl. Acad. Sci. U.S.A. **72**:1184–1188.

22. **Mason, B. B., D. Y. Graham, and M. K. Estes.** 1980. In vitro transcription and translation of simian rotavirus SA11 gene products. J. Virol. **33**:1111–1121.

23. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end labeled DNA with base specific chemical cleavage. Methods Enzymol. **65**:499–560.

24. **Okayama, H., and P. Berg.** 1982. High-efficiency cloning of full-length cDNA. Mol. Cell. Biol. **2**:161–170.

25. **Ou, J.-H., E. G. Strauss, and J. H. Strauss.** 1981. Comparative studies of the 3' terminal sequences of several alphavirus RNAs. Virology **109**:281–289.

26. **Porter, A. G., C. Barber, N. H. Carey, R. A. Hallewell, G. Threlfall, and J. S. Emtage.** 1979. Complete nucleotide sequence of an influenza virus haemagglutinin gene from cloned cDNA. Nature (London) **282**:471–477.

27. **Rice, C. M., and J. H. Strauss.** 1981. Synthesis, cleavage and sequence analysis of DNA complementary to the 26S messenger RNA of sindbis virus. J. Mol. Biol. **150**:315–340.

28. **Rodger, S. M., and I. H. Holmes.** 1979. Comparison of the genomes of simian, bovine, and human rotaviruses by gel electrophoresis and detection of genomic variation among bovine isolates. J. Virol. **30**:839–846.

29. **Roychoudhury, R., E. Jay, and R. Wu.** 1976. Terminal labeling and addition of homopolymer tracts to duplex DNA fragments by terminal deoxynucleotidyltransferase. Nucleic Acids Res. **3**:101–116.

30. **Ruther, U., M. Koenen, K. Otto, and B. Müller-Hill.** 1981. PUR222, a vector for cloning and rapid chemical sequencing of DNA. Nucleic Acids Res. **9**:4087–4098.

31. **Seeburg, P. H., J. Shine, J. A. Martial, J. D. Baxter, and H. M. Goodman.** 1977. Nucleotide sequence and amplification in bacteria of structural gene for rat growth hormone. Nature (London) **270**:486–494.

32. **Sippel, A. E.** 1973. Purification and characterization of adenosine triphosphate: ribonucleic acid adenyltransferase from *Escherichia coli*. Eur. J. Biochem. **37**:31–40.

33. **Smith, M. L., I. Lazdins, and I. H. Holmes.** 1980. Coding assignments of double-stranded RNA segments of SA11 rotavirus established by in vitro translation. J. Virol. **33**:976–982.