JOURNAL OF VIROLOGY, Aug. 1993, p. 4848–4855
0022-538X/93/084848-08$02.00/0

Vol. 67, No. 8

# Location of Intrachain Disulfide Bonds in the VP5* and VP8* Trypsin Cleavage Fragments of the Rhesus Rotavirus Spike Protein VP4

JOHN T. PATTON,* JIAN HUA, AND ERIC A. MANSELL

*Department of Microbiology and Immunology, University of Miami School of Medicine, P.O. Box 016960 (R-138), Miami, Florida 33101*

Received 11 March 1993/Accepted 18 May 1993

**Because the rotavirus spike protein VP4 contains conserved Cys residues at positions 216, 318, 380, and 774 and, for many animal rotaviruses, also at position 203, we sought to determine whether disulfide bonds were structural elements of VP4. Electrophoretic analysis of untreated and trypsin-treated rhesus rotavirus (RRV) and simian rotavirus SA11 in the presence and absence of the reducing agent dithioerythritol revealed that VP4 and its cleavage fragments VP5* and VP8* possessed intrachain disulfide bonds. Given that the VP8* fragments of RRV and SA11 contain only two Cys residues, those at positions 203 and 216, these data indicated that these two residues were covalently linked. Electrophoretic examination of truncated species of VP4 and VP4 containing Cys→Ser mutations synthesized in reticulocyte lysates provided additional evidence that Cys-203 and Cys-216 in VP8* of RRV were linked by a disulfide bridge. VP5* expressed in vitro was able to form a disulfide bond analogous to that in the VP5* fragment of trypsin-treated RRV. Analysis of a Cys-774→Ser mutant of VP5* showed that, while it was able to form a disulfide bond, a Cys-318→Ser mutant of VP5* was not. These results indicated that the VP4 component of all rotaviruses, except B223, contains a disulfide bond that links Cys-318 and Cys-380 in the VP5* region of the protein. This bond is located between the trypsin cleavage site and the putative fusion domain of VP4. Because human rotaviruses lack Cys-203 and, hence, unlike many animal rotaviruses cannot possess a disulfide bond in VP8*, it is apparent that VP4 is structurally variable in nature, with human rotaviruses generally containing one disulfide linkage and animal rotaviruses generally containing two such linkages. Considered with the results of anti-VP4 antibody mapping studies, the data suggest that the disulfide bond in VP5* exists within the 2G4 epitope and may be located at the distal end of the VP4 spike on rotavirus particles.**

Rotaviruses are a cause of acute gastroenteritis in children and adults (1, 6). Like the genomes of all members of the *Reoviridae*, the genome of the rotaviruses consists of segmented double-stranded RNA. The rotavirion is a T=13l icosahedron, composed of three concentric layers of protein (34, 40). The innermost layer is made up of the RNA-binding protein VP2 (102 kDa) (2, 20), which with the putative RNA polymerase VP1 (125 kDa) (39), the guanylyltransferase VP3 (98 kDa) (29), and 11 segments of double-stranded RNA forms the virion core. The middle layer of the protein is composed of the group-specific antigen VP6 (45 kDa), a protein arranged in trimeric units within the virion (16, 35). Examination of rotavirions by cryoelectron microscopy indicates that the outer layer is made up of 780 copies of the serotype-specific glycoprotein VP7 (37 kDa) and 120 copies of the trypsin-sensitive hemagglutinin protein VP4 (86.5 kDa) (26, 33). Dimers of VP4 are proposed to form spikes of approximately 120 Å (1 Å = 0.1 nm) that project from the surface of the rotavirion (33). On the basis of studies of the VP4 component of SA11, trypsin treatment cleaves the rotavirus gene 4 product at position 241 or 247, producing the subfragments VP5* (60 kDa) and VP8* (27 kDa), both of which remain associated with the virus (11, 24). Trypsin cleavage of VP4 increases infectivity presumably by enhancing virus penetration into the cell (11). Antibodies directed at either VP5* or VP8* have been shown to neutralize rotavirus infectivity in vitro and are able to passively protect mice from rotavirus challenge (3, 14, 22, 23, 27, 30).

Although the deduced amino acid sequences of VP4 for numerous rotavirus isolates have been reported, our understanding of the physical structure and location of functional domains within VP4 remains limited. Because of the absence of a corresponding amino acid at residue 136, VP4 proteins of human and animal rotaviruses are usually found to differ in size (775 versus 776 amino acids, respectively) (10, 25). Previous studies examining functional regions in VP4 indicate that the hemagglutination domain lies within VP8* (positions 47 to 247) (23) and that VP8* is able to bind glycophorin, the erythrocyte receptor for rotavirus (14). Within VP5* exists a hydrophobic region (positions 385 to 404) that possesses similarity to the putative fusion sequences of Sindbis and Semliki Forest viruses and thus may play a role in mediating the entry of rotavirus into the cell (27). The locations of functional domains in VP4 that are involved with dimerization and interaction with VP7 and VP6 are not known, nor is information on the site of interaction between VP5* and VP8* available. Analysis of the deduced amino acid sequences for VP4 shows that four Cys residues are conserved in all but one rotavirus isolate examined to date (positions 216, 318, 380, and 774) (Table 1). That exception, VP4 of bovine rotavirus B223, lacks a Cys residue at or near position 318 (17). In contrast to nearly all human rotaviruses, the VP4 component of some animal rotaviruses contains an additional Cys residue at position 203. The highly conserved nature of the Cys residues suggests that disulfide bonds may be structural elements of VP4

* Corresponding author. Electronic mail address: jpatton@miasun.med.miami.edu.

TABLE 1. Locations of conserved cysteine residues in the VP4 proteins of group A rotaviruses

| Virus | Cysteine residue at position[a]: | | | | | Type | Accession no.[b] |
|---|---|---|---|---|---|---|---|
| | 203 | 216 | 318 | 380 | 774 | | |
| C486 | + | + | + | + | + | Bovine | P08713 |
| H-2 | + | + | + | + | + | Equine | L04638 (GenB) |
| NCDV[c] | + | + | + | + | + | Bovine | |
| OSU | + | + | + | + | + | Porcine | P11114 |
| RRV | + | + | + | + | + | Simian | P12473 |
| SA11 | + | + | + | + | + | Simian | P04508 |
| SA11-FEM | + | + | + | + | + | Simian | P17463 |
| SA11-SEM | + | + | + | + | + | Simian | P17464 |
| SA11-4F | + | + | + | + | + | Simian | X57319 (EMBL) |
| YM | + | + | + | + | + | Porcine | P25174 |
| AU-1 | − | + | + | + | + | Human | D10970 (GenB) |
| B641 | − | + | + | + | + | Bovine | P25173 |
| DS1 | − | + | + | + | + | Human | P11196 |
| FRV-1 | − | + | + | + | + | Feline | D10971 (GenB) |
| GOTT | − | + | + | + | + | Porcine | P23045 |
| K8 | − | + | + | + | + | Human | D90260 |
| KU | − | + | + | + | + | Human | P13842 |
| L26 | − | + | + | + | + | Human | P21284 |
| L27 | − | + | + | + | + | Human | C36410 (PIR) |
| M37 | − | + | + | + | + | Human | P11197 |
| MCN13 | − | + | + | + | + | Human | P11199 |
| P | − | + | + | + | + | Human | P11195 |
| RV5 | − | + | + | + | + | Human | M32559 (GenB) |
| ST3 | − | + | + | + | + | Human | P11200 |
| UK | − | + | + | + | + | Bovine | P12474 |
| VA70 | − | + | + | + | + | Human | P11194 |
| Wa | − | + | + | + | + | Human | P11193 |
| 1076 | − | + | + | + | + | Human | P11198 |
| B223 | − | + | − | + | + | Bovine | M92986 (GenB) |
| 69M | + | − | + | + | + | Human | P26451 |

[a] +, present; −, absent.
[b] Unless otherwise indicated, sequences were obtained from the University of Geneva protein sequence data bank release 23. GenB, GenBank nucleotide sequence data bank release 23; PIR, Protein Identification Resource data bank release 33.
[c] See reference 18.

and, as such, may be important for maintaining certain biological activities associated with the protein. Indeed, on the basis of electrophoretic analysis of porcine rotavirus OSU under reducing and nonreducing conditions, Gorziglia et al. (16) suggested that VP4 contained an intrachain disulfide bond. In the study presented here, electrophoretic analysis of untreated and trypsin-treated rhesus rotavirus (RRV) demonstrated that VP4 and its cleavage fragments VP5* and VP8* contained intrachain disulfide bonds. Inspection of the VP8* amino acid sequence of RRV revealed that the bond must link Cys-203 and Cys-216. In vitro expression and analysis of Cys→Ser mutants of VP5* showed that Cys-318 and Cys-380 were also linked by a disulfide bond. Together, these data indicate that, while many animal rotaviruses (e.g., RRV, SA11, SA11-4F, NCDV, and OSU) contain two intrachain disulfide bonds, human rotaviruses contain only a single bond, that which is present in the VP5* fragment. Because the disulfide bond in VP5* is contained in the epitope recognized by the antibody 2G4, a site that has been located in VP4 by three-dimensional cryoelectron microscopy (33), we predict that the bond in VP5* is situated at the distal end of the spike protein in rotavirus particles.

## MATERIALS AND METHODS

**Cells and virus.** Rhesus monkey kidney (MA104) cells were grown in Eagle's minimal essential medium (MEM) containing 10% fetal bovine serum. Stocks of rotavirus were propagated in MA104 cells in the presence of serum-free MEM and 5 μg of trypsin per ml (Difco; 1:250) (19). To prepare $^{35}$S-labeled virions, monolayers of MA104 cells were infected with 5 to 10 PFU of trypsin-activated SA11 virus per cell. At 1 h postinfection, the inoculum was replaced with MEM containing 5 μg of actinomycin D per ml. At 3 to 5 h postinfection, the medium was changed to Met-Cys-free MEM containing 25 μCi of $^{35}$S-labeled amino acids per ml (1,200 Ci/mmol). The cells were harvested at 12 to 22 h postinfection by being scraped into the media, recovered by low-speed centrifugation, and lysed by suspension in TN buffer (20 mM Tris-HCl [pH 7.5], 150 mM NaCl) containing 1% Triton X-100, 50 mM iodoacetamide, and 0.5 μg of leupeptin per ml. Nuclei were then removed by low-speed centrifugation, and the supernatant was layered over preformed 4-ml gradients of 20 to 45% (wt/wt) CsCl in TN buffer containing 0.5 μg of leupeptin per ml. After centrifugation at 33,000 rpm in a Beckman SW50.1 rotor for 12 to 18 h, virus was recovered and again centrifuged to equilibrium in a CsCl gradient. The $^{35}$S-labeled virus was collected and dialyzed against TN buffer. The protease inhibitor leupeptin (0.5 μg/ml) was included in the media, lysis buffer, and CsCl gradients to prevent the cleavage of VP4 into VP5* and VP8*.

**Preparation of transcription vectors used to produce wild-type and modified VP4 and VP5*.** A Bluescript KS construct (pYEKS35) containing the full-length gene of VP4 for RRV inserted into the plasmid at the unique *Eco*RI site and positioned downstream from a T3 transcription promoter was generously supplied to us by E. Mackow (State University of New York at Stony Brook). To place RRV gene 4 under the control of the T7 transcription promoter, pYEKS35 was digested with *Not*I, releasing the gene 4 DNA as a 2.4-kb fragment. The fragment was gel purified and ligated into the plasmid Bluescript SK which had been digested with *Not*I and treated with alkaline phosphatase (36). Following transformation of competent *Escherichia coli* DH5α with the ligation products, the inability to produce β-galactosidase and digestion with restriction enzymes were used to identify colonies that contained a plasmid (pSK-VP4) possessing an RRV gene 4 insert under the control of the T7 transcription promoter (Fig. 1).

To generate a plasmid construct (pSK-VP5) encoding only the VP5* fragment of VP4, pSK-VP4 was digested with *Eco*RV, which cleaved the vector in the polylinker immediately downstream from the T7 transcription promoter and the gene 4 cDNA insert at position 769 (Fig. 1). The DNA fragment containing the coding sequences for VP5* linked to the plasmid was gel purified, recircularized by self-ligation, and transformed into *E. coli* DH5α (36).

The nucleotide sequence of the open reading frame (ORF) for RRV VP4 in the transcription vectors pCR-VP4$_{d318}$ and pCR-VP4$_{d774}$ was altered such that codons for Cys-318 and Cys-774, respectively, instead encoded Ser residues. Specific changes in the nucleotide sequence were introduced into gene 4 by polymerase chain reaction using the amplification conditions previously described (7). The locations and sequences of the primers are given in Fig. 1. The internal gene 4 primers were designed to include single nucleotide differences which, if incorporated into the ORF of gene 4, would change the targeted Cys codon to a Ser codon. To

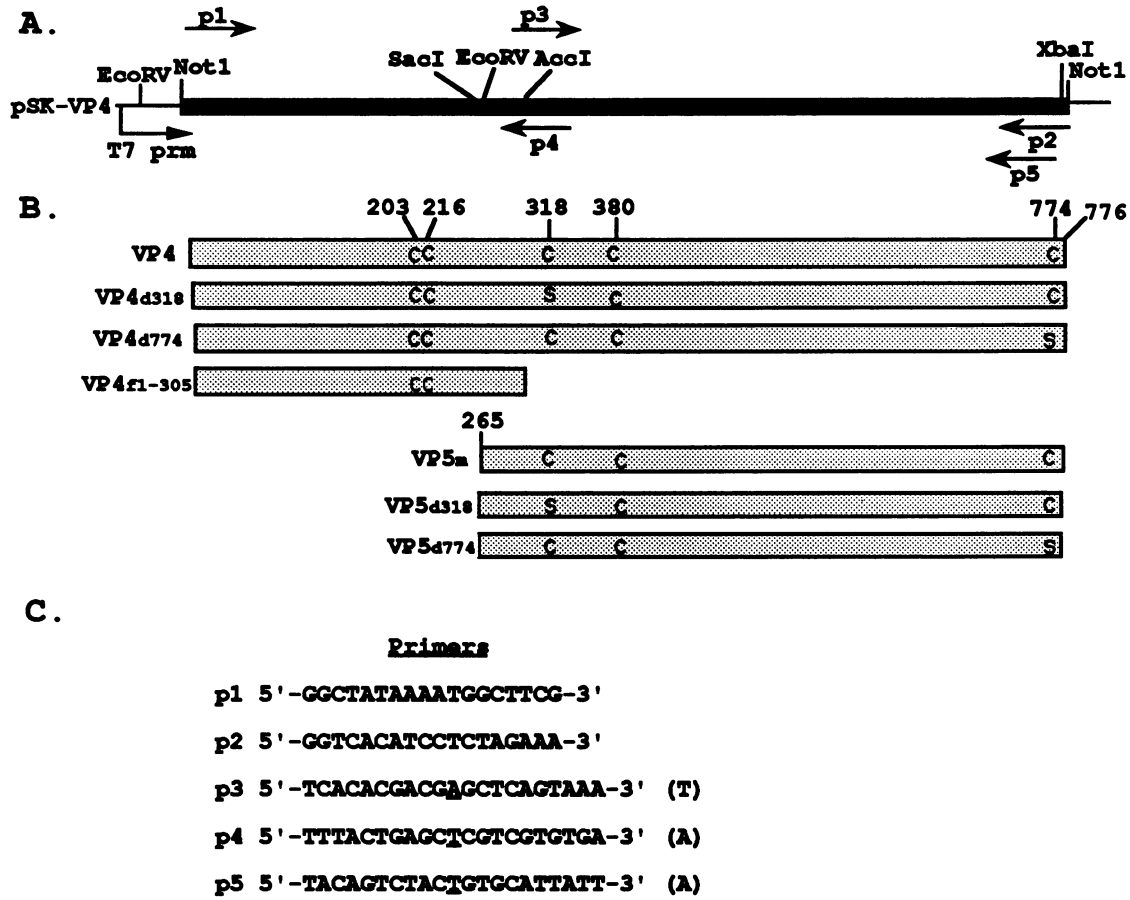


FIG. 1. Site-specific modification of Cys residues in VP4. (A) Diagram of a partial restriction map of the transcription vector, pSK-VP4, that contains a full-length cDNA of gene 4 of RRV positioned downstream from a promoter (prm) for T7 RNA polymerase. Locations along the cDNA for which oligonucleotide primers were prepared and used to mutagenize Cys codons in VP4 are indicated. (B) Locations of Cys residues in authentic and mutagenized VP4. (C) Sequences of primers used in polymerase chain reaction to mutagenize Cys codons in VP4. Bases that are underlined were used in place of those which normally occur (shown in parentheses) so that Ser residues were encoded instead of Cys residues.

produce pCR-VP4$_{d318}$, two amplification reactions were initially performed with both containing pSK-VP4 DNA and *Taq* DNA polymerase but with one reaction containing primers 1 and 4 and the other reaction containing primers 2 and 3. Afterwards, the amplified DNA was gel purified, combined, and added to a second amplification reaction mixture containing primers 1 and 2. Because the DNA products of the two first-step amplification reactions overlapped in sequence, inclusion of these fragments with the terminal primers 1 and 2 resulted in the production of full-length gene 4 DNA mutated such that its ORF specified a Ser residue at position 318. The amplified DNA was gel purified, ligated into the vector pCR1000 according to the directions of the supplier (Invitrogen), and transformed into *E. coli* INVαF' (36). Positive colonies were selected on the basis of kanamycin resistance, size of plasmid, and appropriate orientation of the gene 4 insertion with respect to the T7 transcription promoter as determined by digestion with restriction enzymes. To produce pCR-VP4$_{d774}$, an amplification reaction that contained SKg4 DNA and primers 1 and 5 was carried out (Fig. 1). The DNA product was gel purified, ligated into the vector pCR1000, and transformed into *E. coli* INVαF'. Dideoxynucleotide sequencing was used to confirm site-specific nucleotide changes in purified plasmid DNA (37).

To produce pSP72-VP5$_{d318}$ and pSP72-VP5$_{d774}$, the transcription vectors pCR-VP4$_{d318}$ and pCR-VP4$_{d774}$ were digested, respectively, with the restriction enzymes *Eco*RV and *Xba*I or *Eco*RV and *Sph*I (Fig. 1). *Eco*RV cleaved the gene 4 insert at nucleotide 769, while *Xba*I and *Sph*I cleaved at sequences downstream from the ORF. The DNA fragments containing the 3'-terminal portions of gene 4 were recovered from the digestion products and ligated into pSP72 correspondingly cleaved with either *Eco*RV and *Sph*I or *Eco*RV and *Xba*I (36). The products of ligation were transformed into competent *E. coli* DH5α.

**Synthesis of viral mRNAs and protein in vitro.** To prepare viral mRNAs encoding VP5$_m$, VP5$_{d318}$, and VP5$_{d774}$, the vectors pSK-VP5, pSP72-VP5$_{d318}$, and pSP72-VP5$_{d774}$ were linearized with *Not*I, *Xba*I, and *Sph*I, respectively. To prepare viral mRNAs encoding VP4$_{d318}$, VP4$_{d774}$, and VP4$_{f1-305}$, the vectors pCR-VP4$_{d310}$, pCR-VP4$_{d774}$, and pSK-VP4 were linearized with *Not*I, *Not*I, and *Acc*I, respectively. The cleaved DNAs were purified by phenol-chloroform extraction and then used as templates in runoff transcription reactions containing T7 RNA polymerase as

described previously (29). The mRNA products were recovered by phenol-chloroform extraction and then used to program protein synthesis in reactions containing nuclease-treated rabbit reticulocyte lysates (Promega) and [$^{35}$S]Met (NEN) (7). Translation of the $VP5_m$ mRNA initiated at Met-265 of the gene 4 ORF.

**Electrophoretic analysis of proteins for disulfide bonds.** Radiolabeled viral proteins were assayed for the presence and absence of disulfide bonds by electrophoresis on 12% polyacrylamide gels (30 by 18 cm) containing sodium dodecyl sulfate (SDS) and by fluorography (21). Prior to electrophoresis, samples were adjusted to final concentrations of 1% SDS, 25 mM Tris-HCl (pH 6.8), 10% glycerol, and for samples to be reduced, 50 mM dithioerythritol (DTE) and then boiled for 3 to 5 min.

**Construction of the phylogenetic tree.** The PROTPARS program of Phylip 3.4 was used to estimate the phylogeny of rotavirus VP4 from published amino acid sequences (9, 13). Phylogenetic inference by protein parsimony calculates the minimum number of base changes that would be necessary to construct each possible phylogeny, favoring a tree which requires the fewest substitutions (12). PROTPARS uses a variation of the parsimony method which scores only nucleotide changes that affect amino acid coding and assumes that silent changes are much more common (12). TreeDraw Deck (15) was used to generate a plot of the tree.

## RESULTS

**Detection of intrachain disulfide bonds in VP4.** RRV and SA11-infected MA104 cells were maintained in the presence of $^{35}$S-amino acids and lysed in the presence of iodoacetamide. This compound inhibits the formation of nonauthentic disulfide bonds by blocking free sulfhydryl groups in proteins. Radiolabeled virions were purified by repeated centrifugation on CsCl gradients and then analyzed by electrophoresis on 30-cm SDS-polyacrylamide gels in the presence and absence of the reducing agent DTE (Fig. 2). The results showed that, in contrast to the core protein VP2, the migration rate of the outer-shell proteins VP4 and VP7 and the inner-shell protein VP6 was greater in the absence of DTE than in its presence, indicating that for RRV and SA11 these three proteins possessed intrachain, but not interchain, disulfide bonds. In some cases, an intermediate reduced form of VP4 was seen upon analysis of virions, suggesting the presence of more than one disulfide bond in the protein (Fig. 2A, lane 3).

**Presence of disulfide bonds in VP5* and VP8*.** As an initial step towards locating disulfide bonds within VP4, $^{35}$S-labeled RRV and SA11 were digested with trypsin under conditions that cleave VP4 into the VP5* and VP8* fragments. The protein components of the virions were then analyzed for the presence or absence of disulfide bonds by gel electrophoresis. As shown in Fig. 3, the migration rates of both VP5* and VP8* were greater in the absence of DTE than in its presence, indicating that these fragments possessed intrachain disulfide bonds. As expected, when the nonreduced and reduced forms of VP8* were eluted from the gel and subjected to coelectrophoresis on a second gel under reducing conditions, the eluted proteins were found to comigrate (data not shown). Inspection of the primary amino acid sequences of the VP8* peptides of RRV and SA11 revealed that they contained only two Cys residues, those at positions 203 and 216. Thus, the disulfide bond in the VP8* fragments of these two rotavirus isolates must link these two residues.

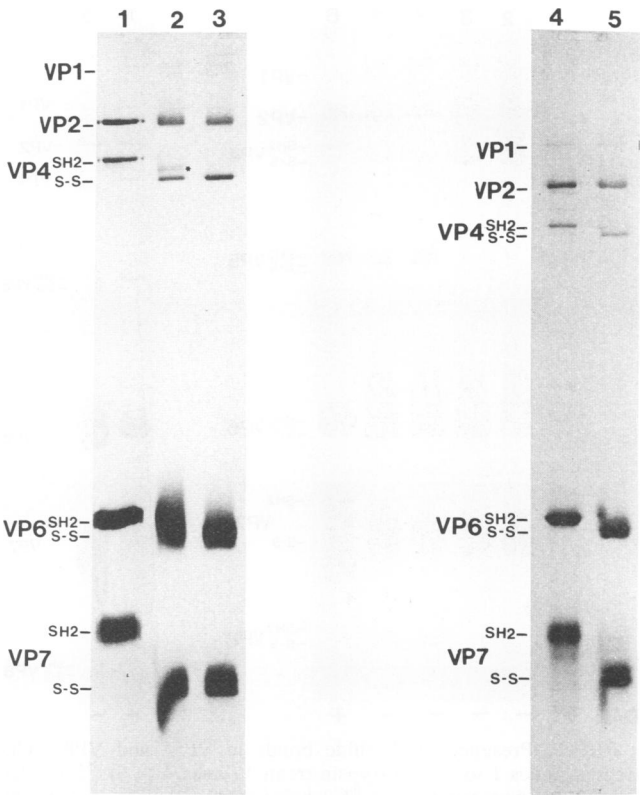


FIG. 2. Detection of disulfide bonds in VP4 of rotavirus. $^{35}$S-labeled RRV (lanes 1 to 3) and SA11 (lanes 4 and 5) were analyzed by electrophoresis on 30-cm 12% polyacrylamide gels in the presence (lanes 1 and 4) and absence (lanes 2, 3, and 5) of 50 mM DTE. Radiolabeled proteins in the gels were detected by fluorography. An intermediate reduced form of the VP4 component of RRV is indicated in lane 2 (*).

**Location of the disulfide bond in VP5*.** The VP5* fragment contains three Cys residues (positions 318, 380, and 774) that are conserved in all rotavirus isolates examined to date with the exception of the bovine rotavirus B223 (Table 1). Like RRV and SA11, most rotaviruses contain no Cys residues in VP5* other than those that are conserved. To determine which residues in VP5* were linked by a disulfide bond, a portion of the RRV gene 4 cDNA that encodes VP5* (positions 265 to 776) was placed into a transcription vector under the control of the promoter for T7 RNA polymerase (pSK-VP5). The polypeptide ($VP5_m$) encoded by this construct begins at Met-265, which is 18 amino acids downstream from the preferred trypsin cleavage site in VP4. Electrophoretic analysis of $VP5_m$ made by in vitro translation showed that most of the protein contained an intrachain disulfide bond (Fig. 4). This result indicated that the primary amino acid sequence of VP5* in itself contains sufficient information necessary to direct the folding of VP4 such that a disulfide bond is formed within it. Two additional constructs were prepared; these constructs contained the same portion of the ORF for gene 4 present in pSK-VP5 except that the codon for Cys-318 (pSP72-$VP5_{d318}$) in one and the codon for Cys-774 (pSP72-$VP5_{d774}$) in the other were altered by site-specific mutagenesis such that they instead encoded Ser residues. In vitro translation of mRNAs derived from the vectors produced the proteins $VP5_{d318}$ and $VP5_{d774}$, which, upon electrophoresis under reducing conditions, migrated in

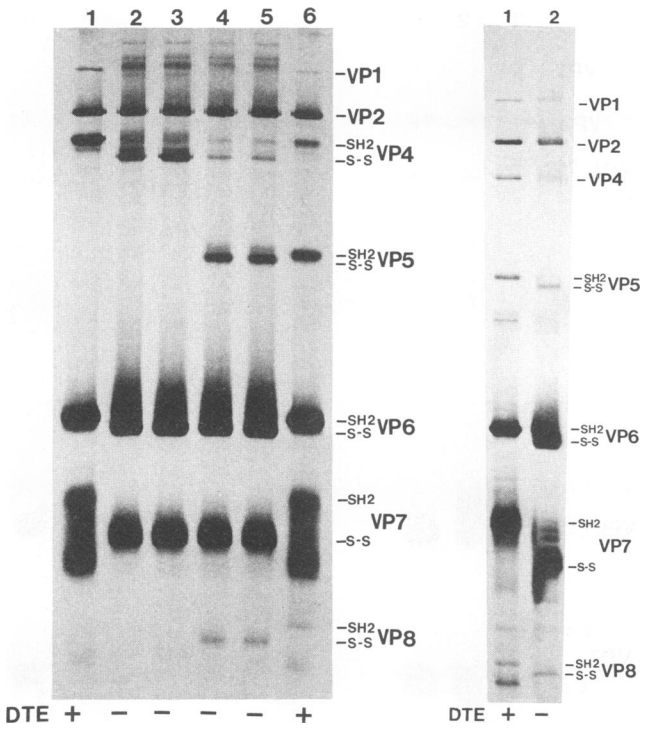


FIG. 3. Presence of disulfide bonds in VP5* and VP8*. Untreated (lanes 1 to 3) and trypsin-treated (lanes 4 to 6) $^{35}$S-labeled RRV (A) and trypsin-treated $^{35}$S-labeled SA11 (B) were analyzed by electrophoresis on 30-cm polyacrylamide gels in the presence and absence of DTE. The positions of the reduced (SH2) and nonreduced (S-S) species of viral proteins are indicated. Partial reduction of VP7 produces a form of the protein that migrates ahead of nonreduced VP7 (panel A, lanes 1 and 6), indicating that the protein possesses more than one disulfide linkage.

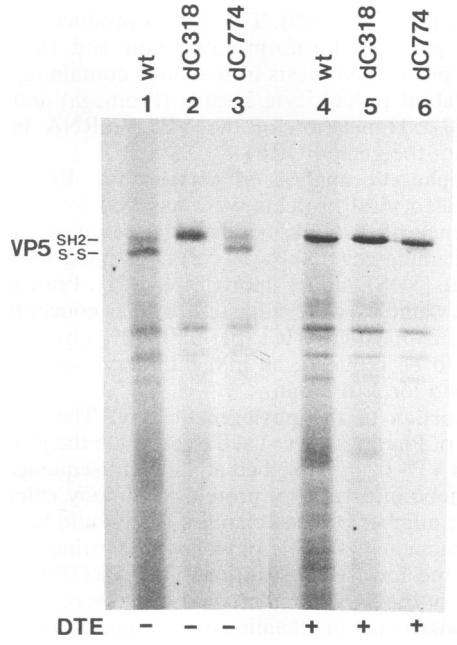


FIG. 4. Location of the disulfide bond in VP5*. The proteins $VP5_m$ (wild type [wt]), $VP5_{dC318}$, and $VP5_{dC774}$ were synthesized in reticulocyte lysates. $VP5_{dC318}$ and $VP5_{dC774}$ contain Ser residues in place of Cys-318 and Cys-774, respectively. The $^{35}$S-labeled products of in vitro translation were resolved by electrophoresis on a 12% polyacrylamide gel in the presence and absence of DTE and were detected by fluorography. The reduced (SH2) and nonreduced (S-S) forms of VP5* are indicated.

parallel with $VP5_m$ (Fig. 4). Electrophoretic analysis of $VP5_{d774}$ under reducing and nonreducing conditions showed that the protein retained the capacity to form an intrachain disulfide bond and that the nonreduced form of the protein comigrated with the nonreduced form of $VP5_m$. Given that $VP5_{d774}$ possesses only two Cys residues, those at positions 318 and 380, these data indicated that a disulfide linkage spans these Cys residues in VP4. Consistent with these results was the finding that the mutant protein $VP5_{d318}$ was unable to generate a disulfide bond, thus supporting the idea that, while Cys-318 serves as one of the anchors of the disulfide bond located in VP5*, Cys-774 plays no role in the formation of the disulfide bond.

To verify that a disulfide bond was present in the VP8* fragment of RRV and linked Cys-203 and Cys-216, the vector pSK-VP4 was linearized with *Acc*I and transcribed with T7 RNA polymerase (Fig. 1). Translation of the mRNAs in vitro produced the 305-amino-acid protein $VP4_{f1-305}$, a truncated species of VP4 that contained the entire VP8* sequence and approximately the first 60 amino acids of VP5*. Electrophoresis of $VP4_{f1-305}$ under reducing and nonreducing conditions showed that the protein contained an intrachain disulfide bond (Fig. 5). Given that $VP4_{f1-305}$ lacks any Cys residues other than those at positions 203 and 216, these data support the conclusion obtained from analysis of trypsin-treated RRV (Fig. 3) that a disulfide bond links the two Cys residues in VP8*.

As an additional test of the accuracy of the map for disulfide bonds in VP4, the mutant proteins $VP4_{d318}$ and $VP4_{d774}$ were synthesized by in vitro translation. $VP4_{d318}$ and $VP4_{d774}$ are identical to VP4 of RRV except that they contain Ser residues in place of Cys residues at positions 318 and 774, respectively. Electrophoretic analysis showed that $VP4_{d318}$ retained the ability to form a disulfide bond despite the absence of a Cys residue shown above to be required for the formation of a bond in VP5 (Fig. 6, lanes 1 to 3). The $VP5_{d318}$ that was made in the same reaction as $VP4_{d318}$ because of initiation of protein synthesis downstream on the $VP4_{d318}$ mRNA at Met-265 indeed showed no evidence of a disulfide linkage. These data taken along with sequence analysis and electrophoretic examination of trypsin-treated RRV lead us to conclude that $VP4_{d318}$ contains a single disulfide bond which is located in the VP8* portion of VP4 which links Cys-203 and Cys-216. In contrast to $VP4_{d318}$, analysis of $VP4_{d774}$ made by in vitro translation showed that two nonreduced forms of the mutant protein were present (Fig. 6, lanes 4 to 6); these forms appeared analogous to those detected upon partial reduction of virion-associated VP4 (Fig. 2, lane 2). One form comigrated with the nonreduced form of $VP4_{d318}$, suggesting that it contained a single disulfide bond and was located in VP8*. As the second nonreduced form migrated upon electrophoresis more rapidly than $VP4_{d318}$, it probably represents $VP4_{d774}$, which contains disulfide bonds in both the VP5* and the VP8* fragments. As expected, $VP5_{d774}$ made during the translation of the $VP4_{d774}$ mRNA because of downstream initiation showed evidence of a disulfide bond linking Cys-318 and Cys-380.

**Relationship between disulfide bonds and the phylogeny of VP4.** A phylogenetic tree was established by protein parsi-

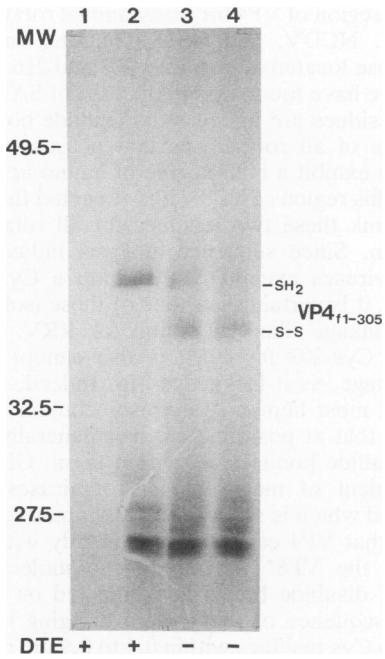


FIG. 5. In vitro-expressed VP8* contains a disulfide bond. The construct pSK-VP4 was linearized by digestion with *Acc*I and transcribed with T7 RNA polymerase. The mRNA was translated in a reticulocyte lysate containing [$^{35}$S]Met, and the radiolabeled product, $VP4_{f1-305}$, was analyzed by electrophoresis in the presence (lane 2) and absence (lanes 3 and 4) of DTE. The products of a translation reaction containing no added mRNA are shown in lane 1. The positions of molecular size markers are indicated (in kilodaltons).

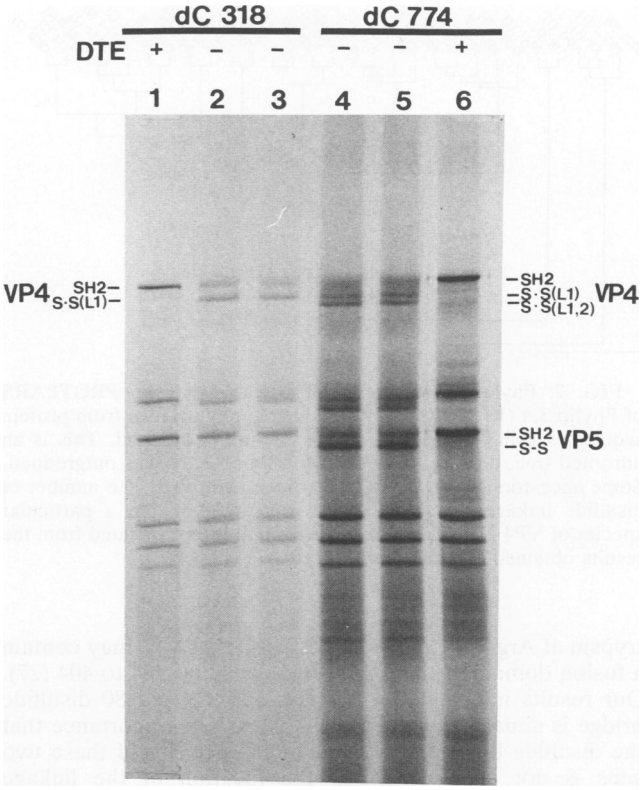


FIG. 6. Analysis of Cys→Ser mutants of VP4 for disulfide linkages. The mutant species $VP4_{dC318}$ (lanes 1 to 3) and $VP4_{dC774}$ (lanes 4 to 6) were synthesized by in vitro translation, and the $^{35}$S-labeled products of that reaction were analyzed by electrophoresis in the presence and absence of DTE. $VP4_{dC318}$ and $VP4_{dC774}$ contain Ser residues in place of Cys-318 and Cys-774, respectively. The reduced (SH2) and nonreduced (S-S) forms of VP4 and VP5 are labeled. The positions of nonreduced VP4 proteins containing a single disulfide linkage, that located in the VP8* region, are indicated (L1). Nonreduced VP4 proteins containing disulfide linkages in VP5* and VP8* are also labeled (L1 and L2). The $^{35}$S-labeled VP5 product was synthesized in the same reaction as VP4 because of downstream initiation of protein synthesis at Met-265 (Fig. 1).

mony for the 31 species of VP4 whose amino acid sequences are available. Interestingly, the tree predicts that all species of VP4 that contain Cys-203 and thus are capable of forming two disulfide bonds are derived from one common ancestor (Fig. 7, Anc II). The progenitor of this common ancestor (Anc II) appears to have been a species of VP4 (Anc III) possessing only a single disulfide bond, that which is located in VP5 between Cys-318 and Cys-380. Indeed, the tree is consistent with the hypothesis that the appearance of Cys-203 in VP4 may have been a primary factor that led to the evolutionary distance between most human species of VP4 (Anc I) and most animal species (Anc II). Clearly, the tree suggests that any species of VP4 that can be shown to have a Cys residue at position 203 can be predicted to be more related to the animal species of VP4, e.g., RRV or SA11, than to the human species, e.g., Wa or DS1. The rotavirus 69M is unique in that its VP4 contains Cys-213 but lacks Cys-216. The phylogenetic tree proposes that VP4 of 69M descended from an ancestor that possessed both these residues and is the only example in the tree of a species of VP4 that appears to have lost a highly conserved Cys residue during evolution of the VP4 gene. The tree also predicts that certain rotavirus strains, e.g., K8, UK, AU-1, and B223, are more related to the ancestor (Anc III) that underwent mutation and gained Cys-203 than to any of the progeny that arose from that ancestor.

## DISCUSSION

Little is known about the structure of the rotavirus capsid proteins, although it is apparent that for most of these viruses the three major components, VP4, VP6, and VP7, all possess intrachain disulfide linkages. Comparison of the amino acid sequences published to date for VP4 shows that all rotaviruses, except B223, contain three Cys residues (positions 318, 380, and 774) in the VP5* portion of the protein (Table 1). Using site-specific mutagenesis, we determined that VP5* of RRV contains a disulfide bond which links Cys-318 and Cys-380. Because the disulfide bond links highly conserved residues, we extrapolate our results with RRV to mean that the VP4 component of all rotaviruses, except for B223, contains a similarly located disulfide linkage. In support of this hypothesis is the finding that the amino acid sequence of VP4 between residues 308 and 385 is highly homologous among all rotaviruses, again with B223 being the lone exception. Hence, the primary amino acid sequence in this region of VP4 can be expected to fold similarly for all rotaviruses except B223 and to stimulate the formation of a disulfide bond linking Cys-318 and Cys-380. The significance of the conserved Cys-774 residue is not known, but its position near the carboxyl end of the protein raises the possibility that the residue is prenylated (38).

Previous studies have shown that VP4 is cleaved by

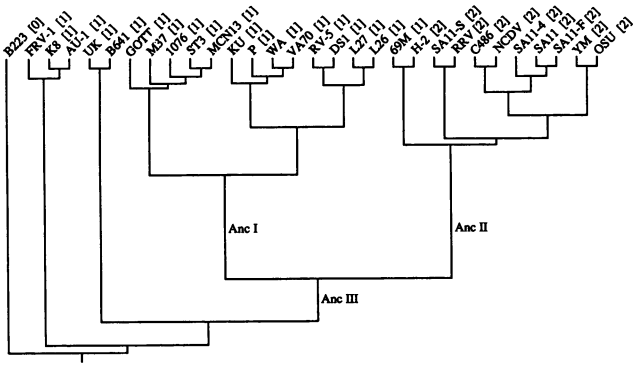


FIG. 7. Phylogenetic tree for VP4 created by using PROTPARS of Phylip 3.4 (13). PROTPARS estimates phylogenies from protein sequences using a variation of the parsimony method. This is an unrooted tree, and the sequence for VP4 of B223 was outgrouped. Some ancestors (Anc I to III) have been identified. The number of disulfide linkages predicted to be contained within a particular species of VP4 is indicated in brackets and is extrapolated from the results obtained for RRV and SA11.

trypsin at Arg-241 or Arg-247 (25) and that VP4 may contain a fusion domain in the region from position 384 to 404 (27). Our results indicate that the Cys-318 to Cys-380 disulfide bridge is situated between these sites. The importance that the disulfide bond might have for the activity of these two sites is not clear, although the location of the linkage indicates that the trypsin cleavage site and the fusion domain are in greater proximity to each other than would be predicted from the primary amino acid sequence of VP4 alone. The VP4 component of B223 lacks Cys-318 and is unusually susceptible to loss from the virion during the standard rotavirus purification techniques, regardless of the background genes of the virus (5). One possible explanation for this phenomenon is that the absence of a disulfide bond in the VP5* region of the protein makes it more prone to destabilization during manipulation of the virus, which leads to the loss of B223 VP4 from the virion.

By immunoprecipitation of truncated molecules of the RRV VP4, Mackow et al. (28) were able to map the epitope for the heterotypic anti-VP4 antibody 2G4 to the region from position 247 to 474. Subsequently, Prasad et al. (33) showed that two molecules of the 2G4 antibody bound to the distal end of the spikes that protrude from rotavirions, thus establishing that each spike consisted of a dimer of VP4. Given that the disulfide linkage in VP5 is centrally located within the 2G4 epitope, it is likely that two such linkages reside near or at the distal end of the VP4 spike and possibly function by stabilizing this region of the structure.

VP4 of the bovine rotavirus B223 lacks Cys-318 (17) and therefore must lack the disulfide linkage that exists between it and Cys-380 for VP4 proteins of RRV and SA11. Yet, in a context-sensitive manner, B223 VP4 can express the 2G4 epitope; that is, while the 2G4 antibody cannot bind to or neutralize B223 virions, the antibody can react with reassortants that express B223 VP4 and a heterologous VP7 protein, e.g., SA11-4F VP7 (4). Because 2G4 can in some instances recognize the B223 VP4, it follows that the Cys-318 to Cys-380 disulfide bridge cannot be an essential feature of the 2G4 epitope. However, it may be possible that in viruses such as RRV and SA11, the expression of the 2G4 epitope in their VP4 proteins may be dependent on a conformation brought about only by the presence of the disulfide bond.

The VP8* region of VP4 for many animal rotaviruses, e.g., SA11, RRV, NCDV, and OSU, contains only two Cys residues, those located at positions 203 and 216 (Table 1). In this study, we have found that in the case of SA11 and RRV, these Cys residues are linked by a disulfide bond. Because VP4 proteins of all rotaviruses that possess Cys-203 and Cys-216 also exhibit a high degree of amino acid homology throughout this region of VP8*, it is expected that a disulfide bridge will link these two residues for all rotaviruses that contain them. Since sequence analysis indicates that no human rotaviruses except 69M contain a Cys residue at position 203, it is certain that none of these isolates contain a disulfide linkage like that found for RRV. Despite the presence of Cys-203 for 69M, it also cannot contain the disulfide linkage, as it lacks Cys-216. Indeed, as the VP8* fragments of most human rotaviruses contain no Cys residues except that at position 216, it is generally impossible that any disulfide bond exists within them. Given that the VP4 component of many animal rotaviruses contains a disulfide bond which is absent from all human rotaviruses, it is apparent that VP4 can vary significantly with respect to structure in the VP8* portion of the molecule. As the formation of disulfide bonds is predicated on the primary amino acid sequence of the protein bringing the α-carbon atoms of two Cys residues within 0.4 to 0.9 nm of each other (8), the tertiary structure of a VP8* that contains the disulfide bond between Cys-203 and Cys-216 probably differs from that of a VP8* which lacks such a bond. As a result, it is possible that both the shape of VP8* and the types and locations of antigenic epitopes expressed in VP8* markedly differ between human rotaviruses and many animal rotaviruses. Our results raise the question of the effectiveness of using the VP4 proteins of the commonly studied animal rotaviruses SA11 and RRV as models for the structure of the human rotaviruses.

In our studies, we obtained no results to suggest that the VP4 monomers making up the rotavirion spike were connected by disulfide bonds, nor did we obtain any results to indicate that the trypsin cleavage fragments VP5* and VP8* were connected to each other by disulfide bonds. The nature of the interchain interactions that stabilize the structure of trypsin-cleaved and uncleaved VP4 spikes on the virion remains to be elucidated.

## ACKNOWLEDGMENTS

We are indebted to Erich Mackow (SUNY at Stony Brook) for providing the RRV strain of virus and for the RRV gene 4 construct pYEKS35. The excellent technical assistance of Larisa Fedoseyeva is acknowledged.

This work was supported by grants from the National Institutes of Health (AI21478) and from the Stanley J. Glaser Foundation. J.T.P. is a recipient of Research Career Development award AI00990.

## REFERENCES

1. **Blacklow, N. R., and H. B. Greenberg.** 1991. Viral gastroenteritis. N. Engl. J. Med. **325**:252–264.
2. **Boyle, J. F., and K. V. Holmes.** 1986. RNA-binding proteins of bovine rotavirus. J. Virol. **58**:561–568.
3. **Burns, J. W., H. B. Greenberg, R. D. Shaw, and M. K. Estes.** 1988. Functional and topographical analyses of epitopes on the hemagglutinin (VP4) of the simian rotavirus SA11. J. Virol. **62**:2164–2172.
4. **Chen, D., M. K. Estes, and R. F. Ramig.** 1992. Specific interactions between rotavirus outer capsid proteins VP4 and VP7 determine expression of a cross-reactive, neutralizing VP4-specific epitope. J. Virol. **66**:432–439.
5. **Chen, D., and R. F. Ramig.** 1992. Determinants of rotavirus

stability and density during CsCl purification. Virology **186:**228–237.
6. **Christensen, M. S.** 1989. Human viral gastroenteritis. Clin. Microbiol. Rev. **2:**51–89.
7. **Clapp, L. L., and J. T. Patton.** 1991. Rotavirus morphogenesis: domains in the major inner capsid protein essential for binding to single-shelled particles and for trimerization. Virology **180:** 697–708.
8. **Creighton, T. E.** 1984. Disulfide bond formation in proteins. Methods Enzymol. **107:**305–307.
9. **Edwards, A. W. F., and L. L. Cavalli-Sforza.** 1963. The reconstruction of evolution. Ann. Hum. Genet. **27:**105.
10. **Estes, M. K., and J. Cohen.** 1989. Rotavirus gene structure and function. Microbiol. Rev. **53:**410–449.
11. **Estes, M. K., D. Y. Graham, and B. B. Mason.** 1981. Proteolytic enhancement of rotavirus infectivity: molecular mechanisms. J. Virol. **39:**879–888.
12. **Felsenstein, J.** 1988. Phylogenies from molecular sequences: inference and reliability. Annu. Rev. Genet. **22:**521–565.
13. **Felsenstein, J. (University of Washington).** 1992. Phylip (Phylogeny Inference Package) version 3.4.
14. **Fiore, L., H. B. Greenberg, and E. R. Mackow.** 1991. The VP8 fragment of VP4 is the rhesus rotavirus hemagglutinin. Virology **181:**553–563.
15. **Gilbert, D. G.** 1990. TreeDraw Deck, a hypercard implementation of Joseph Felsenstein's phylogenetic tree drawing routines. Internet via anonymous ftp. to ftp.bio.indiana.edu.
16. **Gorziglia, M., C. Larrea, F. Liprandi, and J. Esparza.** 1985. Biochemical evidence for the oligomeric possible trimeric structure of the major inner capsid polypeptide (45K) of rotaviruses. J. Gen. Virol. **66:**1889–1900.
17. **Hardy, M. E., M. Gorziglia, and G. N. Woode.** 1992. Amino acid sequence analysis of bovine rotavirus B223 reveals a unique outer capsid protein VP4 and confirms a third bovine VP4 type. Virology **191:**291–300.
18. **Hardy, M. E., G. N. Woode, Z. Xu, and M. Gorziglia.** 1991. Comparative amino acid sequence analysis of VP4 for VP7 serotype 6 bovine rotavirus strains NCDV, B641, and UK. J. Virol. **65:**5535–5538.
19. **Helmberger-Jones, M., and J. T. Patton.** 1986. Characterization of subviral particles in cells infected with simian rotavirus SA11. Virology **155:**665.
20. **Labbe, M., A. Charpilienne, S. E. Crawford, M. K. Estes, and J. Cohen.** 1991. Expression of rotavirus VP2 produces empty corelike particles. J. Virol. **65:**2946–2952.
21. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
22. **Larralde, G., B. Li, A. Z. Kapikian, and M. Gorziglia.** 1991. Serotype-specific epitope(s) present on the VP8 subunit of rotavirus VP4 protein. J. Virol. **65:**3213–3218.
23. **Lizano, M., S. Lopez, and C. Arias.** 1991. The amino-terminal half of rotavirus SA114fM VP4 protein contains a hemagglutination domain and primes for neutralizing antibodies to the virus. J. Virol. **65:**1383–1391.
24. **Lopez, S., C. F. Arias, J. R. Bell, J. H. Strauss, and R. T. Espejo.** 1985. Primary structure of the cleavage site associated with trypsin enhancement of rotavirus SA11 infectivity. Virology **144:**11–19.
25. **Lopez, S., I. Lopez, P. Romero, E. Mendez, X. Soberon, and C. F. Arias.** 1991. Rotavirus YM gene 4: analysis of its deduced amino acid sequence and prediction of the secondary structure of the VP4 protein. J. Virol. **65:**3738–3745.
26. **Mackow, E. R., J. W. Barnett, H. Chan, and H. B. Greenberg.** 1989. The rhesus rotavirus outer capsid protein VP4 functions as a hemagglutinin and is antigenically conserved when expressed by a baculovirus recombinant. J. Virol. **63:**1661–1668.
27. **Mackow, E. R., R. D. Shaw, S. M. Matsui, P. T. Vo, M. N. Dang, and H. B. Greenberg.** 1988. The rhesus rotavirus gene encoding protein VP3: location of amino acids involved in homologous and heterologous rotavirus neutralization and identification of a putative fusion region. Proc. Natl. Acad. Sci. USA **85:**645–649.
28. **Mackow, E. R., M. Y. Yamanaka, M. N. Dang, and H. B. Greenberg.** 1990. DNA amplification-restricted transcription-translation: rapid analysis of rhesus rotavirus neutralization sites. Proc. Natl. Acad. Sci. USA **87:**518–522.
29. **Melton, E. A., P. A. Kreig, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12:**7035–7056.
30. **Offit, P. A., R. D. Shaw, and H. B. Greenberg.** 1986. Passive protection against rotavirus-induced diarrhea by monoclonal antibodies to surface proteins vp3 and vp7. J. Virol. **58:**700–703.
31. **Patton, J. T., L. Salter-Cid, A. Kalbach, E. A. Mansell, and M. Kattoura.** 1993. Nucleotide and amino acid sequence analysis of the rotavirus nonstructural RNA-binding protein NS35. Virology **192:**438–446.
32. **Pizarro, J. L., A. M. Sandino, J. M. Pizarro, J. Fernandez, and E. Spencer.** 1991. Characterization of rotavirus guanylyltransferase activity associated with polypeptide VP3. J. Gen. Virol. **72:**325–332.
33. **Prasad, B. V. V., J. W. Burns, E. Marietta, M. K. Estes, and W. Chiu.** 1990. Localization of VP4 neutralization sites in rotavirus by three-dimensional cryo-electron microscopy. Nature (London) **343:**476–479.
34. **Prasad, B. V. V., G. J. Wang, J. P. M. Clerx, and W. Chiu.** 1988. Three-dimensional structure of rotavirus. J. Mol. Biol. **199:**269–275.
35. **Sabara, M. K., K. F. M. Ready, P. J. Frenchick, and L. A. Babiuk.** 1987. Biochemical evidence for the oligomeric arrangement of bovine rotavirus nucleocapsid protein and its possible significance in the immunogenicity of this protein. J. Gen. Virol. **68:**123–133.
36. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1992. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
37. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
38. **Sinensky, M., and R. J. Lutz.** 1992. The prenylation of proteins. Bioessays **14:**25–31.
39. **Valenzuela, S., J. Pizarro, A. M. Sandino, M. Vazquez, J. Fernandez, O. Hernandez, J. Patton, and E. Spencer.** 1991. Photoaffinity labeling of rotavirus VP1 with 8-azido-ATP: identification of the viral RNA polymerase. J. Virol. **65:**3964–3967.
40. **Yeager, M., K. A. Dryden, N. H. Olson, H. B. Greenberg, and T. S. Baker.** 1990. Three-dimensional structure of rhesus rotavirus by cryoelectron microscopy and image reconstruction. J. Cell Biol. **110:**2133–2144.