# 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

(RNA sequencing/variants/neutralizing monoclonal antibodies/fusion proteins)

ERICH R. MACKOW, ROBERT D. SHAW, SUZANNE M. MATSUI, PHUOC T. VO, MINH-NGOC DANG, AND HARRY B. GREENBERG

Department of Medicine, Stanford University, Stanford, CA 94305; and Palo Alto Veterans Administration Medical Center, Palo Alto, CA 94304

Communicated by Stanley Falkow, September 4, 1987

ABSTRACT The complete gene 4 nucleotide sequence was determined for rhesus rotavirus and each of 11 viral variants selected by neutralizing monoclonal antibodies. Gene 4 is 2362 bases in length and encodes a protein, VP3, of 776 amino acids with a calculated  $M_r$  of 86,500. A conserved trypsin cleavage site, located at amino acid 247, divides VP3 into VP8 and VP5. Neutralizing monoclonal antibodies directed at VP3 were used to select variants that escaped neutralization. Each variant contains a single gene 4 mutation that permits viral growth in the presence of the antibody. Variant mutations were identified in six distinct neutralization regions in VP8 and VP5. Five of the six neutralization regions were found in VP8. The VP8 regions were primarily associated with strain-specific or limited heterotypic rotavirus neutralization. One region was identified in VP5 by three monoclonal antibodies that neutralize a broad range of rotavirus serotypes. The VP5 neutralization region is largely hydrophobic and is similar to putative fusion sequences of Sindbis and Semliki Forest viruses.

Rotavirus is a 70-nm icosahedral virus comprised of two capsid layers. The inner capsid, composed of proteins VP1, VP2, and VP6, contains an endogenous RNA-dependent RNA polymerase and 11 double-stranded RNA segments (1). Gene segments 4 and 9 encode the outer capsid proteins VP3 and VP7, respectively, of rhesus rotavirus (RRV) (2). VP7 (37 kDa) is a glycoprotein, and antibodies directed against VP7 neutralize the virus and specify the viral serotype (3-8). The other viral surface protein, VP3, is an 84- to 88-kDa nonglycosylated protein containing the viral hemagglutinin (9). In the presence of trypsin, VP3 is cleaved to VP5 (60 kDa) and VP8 (28 kDa), which results in enhancement of viral infectivity (10). Preliminary studies indicate that cleavage of VP3 mediates penetration of the virion into the cell (11). VP3 is also associated with restriction of virulence of certain rotavirus strains in mice (12) and humans (13). Antibodies directed at VP3 inhibit viral hemagglutination (14), neutralize the virus in vitro (14), and passively protect mice against heterologous rotavirus challenge in vivo (15). Further studies have demonstrated that VP3 effectively induces protective immunity in animals (16) and is immunogenic in young children (17).

We have studied VP3 neutralization sites of the rhesus strain of simian rotavirus. RRV is serotypically identical to type 3 human rotavirus strains (18). Previously, we described a library of monoclonal antibodies (mAbs) directed against the RRV VP3 protein, which defines its functional and antigenic topography (22). mAbs directed at VP3 were used to delineate distinct neutralization domains and to demonstrate that VP3 contains both heterotypic and ho-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. \$1734 solely to indicate this fact.

motypic neutralizing epitopes. Neutralizing mAbs (N-mAbs) selecting these distinct domains were characterized by competitive binding studies and by their ability to recognize mutants that grew in the presence of N-mAbs.

In this study, we have determined the full-length sequence of the gene 4 RNA segment of RRV\* and have compared the deduced amino acid sequence of RRV to the sequences of simian rotavirus strain SA11 and human strain RV-5. In addition, 11 viral variants selected by individual N-mAbs were sequenced, and the positions of single base, single amino acid changes were defined. Mutations clustered in one VP5 site were selected by mAbs that neutralize a broad range of heterologous rotavirus strains.

# MATERIALS AND METHODS

Cells and Viruses. MA104 cells are a continuous line of rhesus monkey kidney cells, and they were grown in medium 199 as described (23). RRV was grown and purified on CsCl gradients as described (23). We have previously described the production, characterization, selection, and plaque purification of variants made to VP3-specific N-mAb 5C4, 1A9, 7A12, and 2G4 (22). Additional VP3 N-mAbs A1, A15, M11, 5D9, M14, M7, and M2 were chosen to select new variants because they were not serologically related to our previous library of escape mutants (see Table 1), and the selection was carried out as before (22). A sequential series of double, triple, and quadruple variants were also selected. Variant 2G4 was passaged four times in the presence of N-mAb 1A9, and then 2G4/1A9 variants were isolated by plaque purification. Subsequently, the same procedure was employed with N-mAb 7A12 and then N-mAb 5D9. Every variant was resistant to each of the selecting N-mAbs.

**RNA Sequencing.** Plus-stranded RNAs were synthesized from single-shelled cores essentially as described by Flores *et al.* (24). Total single-stranded RNA was initially fractionated on a 1% low-melting agarose gel, and the gene 4 segment was isolated. Approximately 20 adenosine residues were added to the 3' end of gene 4 using poly(A) polymerase as described by Sippel (25). With the knowledge of the conserved 3'-terminal sequence of rotavirus genes (26), the oligonucleotide 5' TTTTTTTTTTTTGG 3' was used to prime the initial Sanger dideoxynucleotide sequencing reactions (27), essentially as described by Zagursky *et al.* (28). Sequence was extended across gene 4 by making negative-strand primers, 18 nucleotides in length, at intervals of 200–250 nucleotides. Primers of positive polarity were an-

Abbreviations: mAb, monoclonal antibody; N-mAb, neutralizing mAb; RRV, rhesus rotavirus.

<sup>\*</sup>This sequence is being deposited in the EMBL/GenBank data base (Bolt, Beranek, and Newman Laboratories, Cambridge, MA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03567).

nealed to double-stranded RNA templates denatured by the method of Bassel-Duby *et al.* (29) and were used to sequence regions of high secondary structure. Sequence in the VP8 portion of gene 4 was also determined by direct plasmid sequencing (28) of two overlapping cDNA clones. The 5'-terminal sequences were verified by direct enzymatic RNA sequencing (30, 31).

### RESULTS

Sequence of RRV Gene 4, VP3. The complete nucleotide sequence of the RRV gene 4 is presented in Fig. 1. Gene 4 contains one long open reading frame that originates 10 nucleotides from the 5' end and terminates with a single stop codon 22 bases from the 3' end. The ATG start codon conforms to Kozak's rules (34) for strong translation initiation (ANN

ATGG), and there is no potential poly(A) addition signal following the termination codon. The encoded protein, VP3, comprises 776 amino acids and has a calculated  $M_r$  of 86,500. The preferred trypsin cleavage site identified by Lopez *et al.* (32) is conserved, yielding a VP8 of 247 amino acids (27 kDa) and VP5 of 529 amino acids (60 kDa). Two other potential trypsin cleavage sites precede the preferred site.

The RRV nucleotide and amino acid sequences were compared with the sequence of simian rotavirus SA11 (32, 33) (Fig. 1) and the human strain RV-5 (35) (data not shown). At the nucleotide level, gene 4 of RRV and SA11 are 74% homologous, whereas RRV and RV-5 are 67% related. At the amino acid level, RRV is more closely related to SA11 than to RV-5. There is 84% total amino acid homology between RRV and SA11, with 78% amino acid identity in VP8 and 88% amino acid identity in VP5. RRV and RV-5

RRV SA11	1 1	GGC	TAI		M	A S	10G 8 A	L	ATT I	Y	AGA R	0	L	L	T	AAT N	TCA S	TAT Y	T	GTT V	GAC D E	L	TCT S	D	GAA	ATA I	CAA O	GAA	ATT I	G G	TCT S	ACT T	AAG K	ACG T	0	AAT N	GTIC V	ACT T	ATT I V	AAT N	CTA L P
RRV SA11	121 38	G G	оос Р	TTC F	603 A	с <b>лл</b> О	ACA T	ggt G	TAT Y	GCT A	P	GTT V	AAC N	796 W	agt G	ост Р	gat G	GAA E	ACT. T	N N	GAT D	TCT 8	ACT T	ACT T	GTA V	GAA E	0003 P	GTA V	L	GAT D	G G	ост Р	TAT Y	0	OCA P	ACT T	100 8 T	TTC F	AAT N	P	P
RRV SA11	241 78	GTA V	GAT D S	TAT Y	TGG W	ATG M	CTA L	L	gca A	P	ACA T	GCA A N	GCT A	GCA G	GTA V	GTA V	GTA V	GAA E	GGA G	ACT.	AAT N	AAT N	ACA T	GAC D N	CGA R	tgg W	CTA L	GCT A	ACA T	ATT I	L	GTT V I	gag E	ост Р	AAC N	GTA V	ACA T Q	TCA S Q	GAA E V	ACC T E	AGA R
RAV SA11	361 118	AGT S T	TAT Y	ACG T	CTA L	TTT F	GGA G	ACG T Q	сал 0	GAG E V	0	ATT I V	ACA T	ATA I V	act A S	TAT Y N	A D	TCC 8	0	ACA T	CAA O K	799 W	AAA K	TTT F	ATT I V	GAT D	GTC V L	GTT V S	AAA K	ACT T Q	T	CAA	AAT N D	gga G	AGC 8 N	TAT Y	TCA S	<b>CAA</b> 0	TAC Y H	gga g	P S
RRV SA11	481 158	TTA L	CAA O L	1CT 8	ACT T	P	AAA K	L	TAT Y	GCC A G	671G V	ATG M	K K	CAT H	AAT N G	gat G	K K	ATT 1	TAT. Y	ACA T	TAT Y	AAT N	GGA G	GAA E	ACT T	003 P	AAT N	GTG V A	ACC T N	ACT T	AAG K G	TAC Y	TAC Y	TCA S	ACT T	ACA T	AAT N	TAT Y F	GAT D	TCA S T	GTA V
RRV SA11	601 198	AAC N	M	ACA T	QCA A	TTT F Y	C	алс   Р	TTT F	TAT Y	ATT	ATA I	ост Р	AGA R L	GAA E A	GAA E Q	GAA E	TCA S A	ACA T K	C	ACC T	GAG E	TAC Y	ATT I	AAT N	AAC N	606 6	TTA L	P	000 P	ATT I	CAG O	AAT N	ACA T	CGA R {	AAC	ATT I	сп V	P	TTG L V	903 A S
RRV SA11	721 238	CTT L I	TCA 8 V	A S	AGA R	AAT N	ATA I	АТА   V	TCA S S	CAT H T	AGA R	GCT	0	903. A P	AAT N	GAA E Q	GAT D	ATC I	GTT V	gtig V	TCA S	AAG K	ACA T	TCA S	L	TGG W	K	GAG E	ATG M	0	TAC Y	AAT N	AGA R	GAC D	ATC I	ACA T V	ATT	R	F	AAA K	F
RRV SA11	841 278	gca A	AG1 S N	TCA S	ATT	U V I	AAA K	5 S	Gat G	6 6	CTA L	GGT G	TAT Y	K K	Tag W	TCA S	GAG E	ATT 1 V	TCA S	TTT F	AAA K	P	GCA A	AAC N F	TAT Y	Q	Y	ACG T	TAT Y	ACA T	CGA R	GAT D	GGA G	GAG E	GAT D E	GΠ V	ACA T	A	CAC H	AOG T	ACG T
RRV SA11	961 318	C	тс <i>и</i> ] 8	GTA V	N	G G	ATG M V	N	GA1 D	F	AAT N	F F Y	AAT N	9996 G	gga G	109 8	TTA L	P	ACG T	GAT D	TTT F	АТА 1 V	I	TCA S	AGA R K	TAT Y	GAA	GTA V	ATT I	AAA K	GAG E	AAT N	TCT S	TAT Y F	GTT V	TAT Y	GTT V I	GAT D	Y	TGG W	GAT D
RRV SA11	1081 358	GA1 D	s tr	0	A GOC	F	AGG R	N	M	ι GTT V	TAT Y	GTA V	AGG R	TCA 8	TTA L	A A	GCT A	AAT N	TTA L	AAC N	TCT S	GTT V	ата 	C		G	GGT G	GAT D		AQC 8	TTT F	GCA A	TTA	¢ P	GTT V	G	CAA O N	TGG W Y	P	GTA V	ATG M
RRV SA11	1201 398		6	G G	A GC/	GTG V	TCA S	L	H	TCA S	A	GGT G	απ V	ACG T	TTA L	TOC S	ACA T	CAG Q	TTC F	ACA T	D	F	GTA V	S S	F	N N	1 TCT 8	L	AGG R	TTC F	AGG R	TTT F	AGA R	L	ACT T S	Υ Υ	GAV E	E	P	S P	F
RRV SA11	1321 438	103 8	i ATC	T L	R R	AC1 T	AGA R	V V	G G S	G	L W	TAT Y	900 G	TTA L	OCT P	gca A	A	TAC Y K	ŝ	AAT N	N	G G S	K Q	GAA E	TAT Y	TAT 1 Y	E	GTG V I	A CT	99C 9	AGA R	L F	TCA S	L	ATA I	5 TCA	L	GT/ V	P	S L	AAT N
RRV SA11	1441 478	GAC D	GA D	T TAC	0 0	T	P	1 1	T	N M	1 TCA 8	V V	ACA T	GTC V	AGA R	CAA	D	L	GAA E	R	¢	L	G G	GAA E	L	R R	E	E	F	AAC N	A N	CTC L	TCA S	CAA Q	GAG E O	I ATA	A (900)	ATG M	8 TOG 8	Q	L
RRV SA11	1561 518	AT I	r ta' Y D	r TTG L	A act	L	L	P	L	D	r Atg M	F	TCG S	M	F	TCT S	G G	ATT I	K	AQC S	ACC T	ATA I	D	A a c	A	K	A TCA S	M	A	ACT T	AGT S N	GTA V	M	K	K R	F	K	K K	S TCA	001 G S	L
RAV SA11	1681 558	gci A	r AAI N	5 TCT 8	GT/ V	5 TC1	T AC	L	T	A GAC D	5 TC/ 5	L	100 8	GAC D	A A	A	TC1 8	TCA S	ATT I	TCA S	R	6 6 5	A GC/	S TCT	1 AT 1 V	r ogi R	r TCJ S	V V	GGA G S	S S	S T	A	S	A 00/	W	T	D E	r GTC V	5 TC/ 8	T N	Q I
RRV SA11	1801 598	ATI I A	CAC T S	T GAT D	r GTT V	r TC1 8 N	r TC/ S V	A TCI S T	r GTR V T	S AG1	r 103 S	I ATC	: TOG 8	T	Q	T ACI	5 TCJ 8	T	1 ATI	AGT S	R	R 030	L	R	L CT/	K	E	M	A	T	0	T	E	( 000 G	M	8 AA' N	r TTC F	D GA	D	I ATA	S S
RRV SA11	1921 638	GC" A	T GC A	A GT/ V	L	K AAG	AC T	T AA K	A AT I	T GA' D	T CGJ R K	s too	T	0	I L	S N	P T	N	T	L	P	D E	T A T/	V V	T	T GA. E	A GC1 A	r TC/ S	E	K	F	I	P	N	R	A (90) A	A TA	C AGU R	A GT/ V	I	N K
RRV SA11	2041 678	AA N D	T GA D	T GA	n Gitt V	F L	r GA E	A GOO	90 G S	T	A GA D	T GG/ G	R R K	Y	F	A	Y D	r ogi R K	ΥGΠ	Ē	T	F	C GA D E	E	A AT	т ос. Р	A TTI	r GA'	V	0	K	F	A	D	L	N GT/ V	A AC T	D	S	F 000	V
RRV SA11	2161 718	AT I	C TC S	A 900 A	TA C	T A T/ 1	A GA D	C TT F	r AA K	G AC	A CTC L	K	N N	L	N	D	N N	T TA' Y	r 001 G	I ATI	r AG1 S	TAGK R	3 CA 0	0	A GC	A TT F L	T AA' N	L		R	8	CAN D	P	R	N GTA V A	L	R	E	n TA	I	N
RRV	2281 758	CA O	A GA	C AA N	T CC P	A AT.	A AT	T CG R	r AA N	C AG. R	A AT	T GAV E	0 8	L	1	M	0			L	i ta/	GC	Ŗ AT'	I IC	AG	a GG	⊼ 1G'	GW													

FIG. 1. The RRV gene 4 RNA sequence is presented in the DNA form along with the encoded VP3 amino acid sequence. Changes in the SA11 VP3 amino acid sequence (32, 33) appear below that of RRV. Cysteines are boxed and long hydrophobic amino acid regions are underlined. Codons are overlined in which single nucleotide substitutions were detected in N-mAb-selected variants. Potential trypsin cleavage sites that divide VP3 into VP8 and VP5 are indicated by arrows.

share 72% identical amino acids-58% in VP8 and 78% in VP5. Only 5 cysteines exist in the RRV VP3, and these are all conserved with those in the SA11 VP3 sequence. RV-5 lacks the cysteine at amino acid 203 but contains the four other cysteine residues in VP3. Unlike SA11 and RV-5, there is no potential chymotrypsin cleavage site located between the RRV VP3 trypsin cleavage sites. The area between the cleavage sites is highly divergent in all three strains. All three viral proteins contain long regions of identical residues at amino acids 4-14, 56-72, 222-235, 257-271, 346-359, 361-378, 407-420, 456-468, 521-540, and 711-736. VP5 also contains one conserved hydrophobic region of 20 amino acids (385-404), which could serve a membrane-spanning role (Fig. 1).

The RRV VP8 amino acid sequence deduced by Gorziglia et al. (36) contains eight discrepancies with the VP8 amino acid sequence presented here, including a deletion and subsequent insertion. We have no explanation for these sequence differences other than the possibility that passage of RRV in separate laboratories has resulted in divergence of the RRV VP3 protein.

Selection of Variants Resistant to N-mAbs. mAbs directed at VP3 were originally selected for their ability to inhibit hemagglutination, neutralize RRV, and immunoprecipitate VP3 from a cell lysate (22). Variants resistant to individual N-mAbs were selected from the parental strain, RRV, and named for the selecting N-mAb (Table 1). Representative variants selected by serologically distinct groups of N-mAbs were chosen for sequence analysis (Table 2).

Each mutant gene 4 was sequenced completely using a series of negative-strand synthetic oligonucleotides generated during sequencing of the parental RRV gene 4. Each variant gene 4 contains a single base change specifying a single amino acid change from that of the parental strain (Table 2). The mutations selected by N-mAbs are listed in Table 2 and have been grouped into six distinct regions.

Variant Mutations Correlate with Their Serologic Grouping. The antigenic characteristics of the variants selected in this study are presented in Table 1. Variants were grouped by their resistance to N-mAbs in hemagglutination inhibition and neutralization assays. The location of base changes in the selected variants was correlated with their serologic grouping. N-mAbs A1, A15, and M11 are closely related by reciprocal hemagglutination inhibition and neutralization analyses, and each selects a variant with a mutation in the same region of VP8 (Table 2). Similar findings were observed for variants M14 and 5C4 (region 4) as well as variants 2G4 and M7 (region 6). Other variants appeared to

have mutations at unique regions on the protein and crossreacted poorly with other serological groups.

Multiple N-mAb Mutant Selection. To obtain variants with multiple mutations, N-mAbs were used sequentially in separate variant selections. The quadruple mutant (GAAD) was expected to be either a composite of the individual mutant changes or a mutant in which one or more of the individual mutations had caused a subsequent N-mAb to select for a new gene 4 mutation. The mutations selected by 2G4 and 1A9 were identified in the same positions as the individual N-mAb variant mutations (Table 2). A mutation at the original 7A12 locus (amino acid 188) was not found in the quadruple mutant; instead a change was identified at amino acid 173. Similarly, a change at amino acid 135 was identified in the quadruple mutant but not at the site of the individual 5D9 mutation (amino acid 114). The intermediate doubleselected (2G4/1A9) and triple-selected (2G4/1A9/7A12) mutants were also sequenced to determine which N-mAb selected the newly observed mutations. The double mutant contained only the 2G4 and 1A9 changes. With the added selection of N-mAb 7A12, the triple mutant contained a mutation at amino acid 173. As a result, the new mutation at amino acid 135 in the quadruple mutant is attributed to the 5D9 selecting N-mAb.

Specificity of the Neutralization Sites. The strain specificity of each N-mAb was determined by neutralization tests with selected viruses of serotypes 1-6 (Table 2). Most of the mAbs neutralize only RRV or combinations of RRV and one other strain of rotavirus. The mutations selected by the mAbs with limited strain specificity occur in VP8. The neutralization region identified in VP5 (region 6) was selected by three N-mAbs that neutralize a wide variety of serotypically distinct rotavirus strains (Table 2).

## DISCUSSION

The primary nucleotide and deduced amino acid sequences of RRV gene 4 are presented in this study. Analysis of the RRV sequence and a comparison with other gene 4 VP3 sequences (32, 33, 35) demonstrates several findings. The RRV gene 4 shares one striking region of nucleic acid conservation with SA11 and RV-5. The first 33 nucleotides of RRV, SA11, and RV-5 gene 4 are identical (except for 2 bases in SA11). In contrast, only 11 bases are conserved at the 3' end of gene 4 of RRV, SA11, and RV-5, and there are no long internal regions of nucleic acid conservation. Analvsis of sequence information from several other rotavirus segments (37-39) also demonstrates long stretches of con-

Region 1 1 1 2 3 4 4 5 6 6 6		Variant														
	N-mAb	M11	A1	A15	1A9	5D9	M14	5C4	7A12	M2	2G4	M7	GAAD			
1	M11	R	R													
1	A1		R	R												
1	A15	R	R	R												
2	1A9			R	R								R			
3	5D9					R										
4	M14						R	R								
4	5C4				R			R					R			
5	7A12								R				R			
6	M2									R						
6	2G4										R	R	R			
6	M7							R			R	R				

Resistance of each variant to hemagglutination inhibition and neutralization was determined as described previously (22). Variant viruses are named for the selecting mAb. GAAD is the quadruple (2G4/1A9/7A12/5D9) variant. Hemagglutination titers were determined for each N-mAb against RRV and the VP3 variant viruses. Variants were considered to be resistant (R) if the hemagglutination inhibition titer and the neutralization titers of the N-mAb against the variant were at least eight times less than the titer of the N-mAb against RRV. Since mAb M2 has little hemagglutination inhibition activity, data for the M2 mAb represents a neutralization test versus RRV and indicated variants. Regions 1-6 refer to the regions where nucleotide changes were observed (see Table 2).

Table 1. Resistance of rotavirus variants to N-mAbs

Table 2. Gene 4 nucleotide and amino acid changes in variants selected by N-mAbs

RRV	(	Change		Selecting mAb					
variant	Nucleotide (bp)	Amino acid (residue)	Region	neutralization specificity					
M11	$A \rightarrow G$ (268)	Thr $\rightarrow$ Ala (87)	1	RRV					
A1	C → A (272)	Thr $\rightarrow$ Lys (88)	1	RRV, NCDV					
A15	$G \rightarrow C$ (274)	Ala $\rightarrow$ Pro (89)	1	RRV					
1A9	$G \rightarrow A (307)$	$Asp \rightarrow Asn (100)$	2	RRV					
5D9	$C \rightarrow T (350)$	Ser $\rightarrow$ Phe (114)	3	RRV					
M14	$A \rightarrow G (452)$	$Gln \rightarrow Arg (148)$	4	RRV, NCDV					
5C4	$G \rightarrow A (458)$	$Gly \rightarrow Glu (150)$	4	RRV					
7A12	$A \rightarrow T$ (572)	Tyr $\rightarrow$ Phe (188)	5	RRV					
M2	$C \rightarrow A (1172)$	Ala $\rightarrow$ Glu (388)	6	RRV, SA11, UK, Wa, Got					
2G4	$A \rightarrow C (1187)$	$Gln \rightarrow Pro(393)$	6	RRV, SA11, OSU, UK					
M7	$A \rightarrow C (1187)$	$Gln \rightarrow Pro(393)$	6	RRV, SA11, UK					
GAAD	$G \rightarrow A (307)$	$Asp \rightarrow Asn (100)$							
	$A \rightarrow G$ (413)	$Gln \rightarrow Arg (135)$							
	$A \rightarrow G (527)$	Lys $\rightarrow$ Arg (173)							
	$A \rightarrow C (1187)$	$Gln \rightarrow Pro(393)$							

N-mAbs that selected variants were tested in plaque-reduction neutralization tests (focus-reduction neutralization for DS1) (22). N-mAbs were titered against RRV, SA11, NCDV, OSU, UK, DS1, WA, and Gottfried (Got) rotavirus strains. Viral neutralization is designated if the titer of the mAb against the indicated viruses is no less than one-eighth of the titer against RRV. GAAD is the quadruple RRV (2G4/1A9/7A12/5D9) variant. bp, Base pair.

servation at the 5' but not at the 3' terminus. Nucleic acid conservation of this type suggests that the 5' end of rotavirus segments may contain recognition signals necessary for segment reassortment or segment packaging.

Further conservation of encoded amino acids is found at positions flanking the VP3 trypsin cleavage sites (Fig. 1). The preferred trypsin cleavage site identified by Lopez et al. (32) occurs after arginine-247. All rotavirus VP3 proteins studied (32, 35, 40) contain this trypsin cleavage site (RA  $\downarrow$ Q). Two other potential cleavage sites ( $R \downarrow NI$ ) precede the preferred trypsin cleavage site. Lopez et al. (32) also identified the use of the trypsin cleavage site at amino acid 241. It is unclear whether or not the site at amino acid 231 is digested, since cleavage at either of the two downstream sites would eliminate this amino terminus from amino acid sequencing reactions. The amino acid sequence between the first and last site is highly divergent, indicating that it is probably removed by trypsin. The regions flanking the trypsin cleavage sites, amino acids 224–236 and 257–271, are 100% conserved in RRV, SA11, and RV-5 and could serve to hold these sites in the proper conformation for cleavage.

RRV, RV-5 (35), and all human strains studied thus far (36, 40) lack a proline following the trypsin cleavage site. Kantharidis *et al.* (35) suggested that the presence of this proline in SA11 could affect the accessibility of the viral trypsin site and account for the efficient growth of SA11 in tissue culture. However, RRV also grows efficiently and readily forms plaques *in vitro*, a growth characteristic more closely related to SA11 than to human rotavirus strains. Additional studies will be required to ascertain the crucial molecular determinants of growth and plaque formation in cell culture.

Five cysteine residues exist in the RRV VP3, and all of these are conserved in SA11. Cysteine-203 is the only cysteine residue that is not conserved in RV-5. In fact, all human rotavirus strains lack a cysteine at amino acid 203 (35, 36). It is conceivable that in simian and other readily cultivable rotavirus strains cysteines at amino acids 203 and 216 in VP8 could form disulfide bonds with cysteines at amino acids 318 and 380 in VP5, which would keep the trypsin cleavage site (amino acid 247) accessible. In human or other fastidious rotavirus strains that lack cysteine-203 but conserve the remaining four cysteines, the trypsin site may be less susceptible to cleavage.

In order to determine the sites involved in viral neutralization, we generated a series of VP3 variants that are resistant to VP3-specific N-mAbs. Sequence analysis of 11 distinct VP3 variants as well as sequential double, triple, and quadruple mutants has identified six regions in VP3 that are involved in rotavirus neutralization (Table 2). Strain-specific neutralization sites were found in VP8 at amino acid residues 100, 114-135, and 173-188, whereas amino acids 87-89 and 148-150 were associated with limited cross-reactive neutralization. Region 6 (amino acids 388-393) in VP5 was linked to highly serotype cross-reactive neutralization. mAbs 7A12 and 5D9 selected new mutations in the quadruple variant when compared to individual mAb-selected mutants (Table 2). Further studies will be required to determine if the 2G4 and 1A9 mutations in the quadruple variant were responsible for new neutralization sites produced by mAbs 7A12 and 5D9 or if such site heterogeneity is an inherent characteristic of the mAb 5D9 and mAb 7A12 binding regions.

All but one of the mAbs used in this study efficiently inhibit viral hemagglutination as well as neutralize infectivity. Since mAb-selected mutations have been identified in several regions of VP3, it seems likely that sites on both VP8 and VP5 contribute to viral hemagglutination and neutralization, especially if the mAbs are subsequently shown to actually bind to the regions of VP3 encoded by the mutations.

Neutralization escape mutations were spread broadly over the coding region of VP8, but no mutants were identified in or near the trypsin cleavage site. In keeping with the strain specificity of the selecting mAbs, the VP8 mutations tended to occur in areas of considerable sequence diversity (Fig. 1). However, VP3-directed mAbs M7, 2G4, and M2 neutralize a variety of serotypically distinct rotavirus strains (Table 2). As might be expected, these mAb-selected mutations are in a conserved area of VP5 (Fig. 1). Taniguchi *et al.* (43) have also identified a variety of heterotypic VP3 mAbs. One of their mAbs, KU6B11, neutralizes RRV as well as most human rotaviruses. This mAb failed to neutralize 2G4 and M7 variants (unpublished results), implying that at least one of the heterotypic neutralizing domains on human VP3 molecules is also probably located in VP5.

Although group A rotaviruses do not normally form syncytia during viral growth, infection with group B rotavirus, avian reovirus, and Nelson Bay virus is associated with membrane fusion (44, 45). Trypsin cleavage of VP3 leads to the rapid enhancement of viral infectivity, apparently in association with the passage of virus through the plasma membrane (11, 21). Trypsin-treated virus can also increase

#### Applied Biology: Mackow et al.

					-					*								
RRV	D	Y	s	F	Α	L	Ρ	۷	G	Q	w	Ρ	V	М	Т	G	G	A
	•	•							•			•		•		•	٠	•
	•	•	•			٠	٠		•		٠	•	٠	•		•	٠	•
Sindbis	D	Y	т	С	к	v	F	G	G	V	Y	Р	F	м	w	G	G	A
	•	•							•		•	٠	l	•		•	٠	•
	•	٠	٠			٠	٠		•		•	٠	•	•		•	٠	•
SA11	D	Υ	s	F	Α	L	Ρ	۷	G	N	Y	Ρ	V	М	ŢΤ	G	G	Α

FIG. 2. Fusion peptide similarity. Positional identity between the putative fusion region of Sindbis virus (41) and the VP3 proteins of RRV and SA11 are presented. The sequence presented corresponds to amino acids 75-92 of the E1 protein of Sindbis virus (19, 41) and to amino acids 384-401 of the RRV and SA11 VP. Identical amino acids are boxed and are separated by two dots. Single dots represent conservative amino acid changes (polar, nonpolar, charged, or aromatic amino acids changes). Mutation sites identified in viral variants 2G4, M7, and M2 are designated by an asterisk.

plasma membrane permeability (11). Because of these observations, we studied the VP3 amino acid sequence to see if we could identify similarities to known fusion proteins. A long hydrophobic region was not found at the VP8 or VP5 amino terminus. However, a region of specific amino acid similarity with internal fusion protein sites from Semliki Forest and Sindbis viruses was detected in VP5 (Fig. 2) (19, 20, 41, 42). The conserved putative fusion sequence of the Sindbis virus E1 protein contains 45% identical plus 27% conserved amino acids with the RRV amino acid 384-401 region of VP5 (Fig. 1). SA11 and RV-5 contain 72% and 61% conserved plus identical amino acids in this region, respectively.

It is surprising that viruses as diverse as rotaviruses and alphaviruses share any amino acid identity in their outer capsid (rotavirus) and envelope (Sindbis) proteins. It is even more striking to find that this peptide conservation occurs in a long hydrophobic region of the alphavirus E1 protein, which has been implicated in cell fusion. It remains to be seen whether this similarity is related to common origins of the genes encoding these surface proteins or common functions of the encoded proteins (or both). This finding does support the hypothesis that the VP3 protein is involved in mediating the entry of rotavirus into the cell. The broadly N-mAbs 2G4, M7, and M2 each select for neutralization escape mutations within this potential fusion site (Fig. 1). Because this site is likely to be critical for viral infection, it may be conserved among many rotavirus strains. It will be interesting to determine if mAbs to region 6 neutralize rotavirus by specifically inhibiting cell penetration.

Genetic and passive transfer studies have clearly shown that antibody to VP3 can efficiently prevent rotavirus illness (15, 16). Recent studies of young children have clearly demonstrated that the epitope defined by mAb 2G4 (region 6) is immunogenic following local immunization with the RRV candidate vaccine (17). The identification of an antigenically conserved region (region 6) in VP3 that participates in viral neutralization may be particularly important when considering a strategy for rotavirus immunization. Considering the essential role that the hydrophobic region in VP5 is likely to play in initiating viral infection, this protein represents an enticing target for future vaccine strategies using either expressed or synthetic viral peptides.

The authors thank Glenn Roldan for preparing the manuscript. This work was supported by Public Health Service Grant R22 AI21362 from the National Institutes of Health and grants from the Veterans Administration and the Thrasher Research Fund.

- Holmes, I. H. (1979) Prog. Med. Virol. 25, 1-36.
- Arias, C. F., Lopez, S. & Espejo, R. T. (1982) J. Virol. 41, 42-50. Bastardo, J. W., McKimm-Breschkin, J. L., Sonza, S., Mercer, 3.
- L. D. & Holmes, I. H. (1981) Infect. Immun. 34, 64-67.
- 4. Kalica, A. R., Greenberg, H. B., Wyatt, R. G., Flores, J., Sereno, M. M., Kapikian, A. Z. & Chanock, R. M. (1981) Virology 112, 385-390.

- 5. Dyall-Smith, M. L., Lazdins, I., Sonza, S., Tregear, G. & Holmes, I. (1985) in Infectious Diarrhoea in the Young: Strategies for Control in Humans and Animals, International Congress Series, ed. Tzipori, S. (Elsevier, Amsterdam), pp. 215-220.
- Greenberg, H. B., Flores, J., Kalica, A., Wyatt, R. & Jones, R. 6. (1983) J. Gen. Virol. 64, 313-324.
- Greenberg, H. B., Valdesuso, J., Van Wyke, K., Midthun, K., Walsh, M., McAuliffe, V., Wyatt, R., Kalica, A., Flores, J. & Hoshino, Y. (1983) J. Virol. 47, 267-275.
- Matsuno, S. & Inouye, S. (1983) Infect. Immun. 39, 155-158.
- Kalica, A. R., Flores, J. & Greenberg, H. B. (1983) Virology 125, 9. 194-205.
- 10. Estes, M. K., Graham, D. Y. & Mason, B. B. (1983) J. Virol. 39, 879-888
- 11. Kaljot, K. T., Shaw, R. D., Rubin, D. H. & Greenberg, H. B. (1986) J. Virol., in press.
- Offit, P. A., Blavat, G., Greenberg, H. B. & Clark, H. F. (1986) J. 12. Virol. 57, 376-378.
- Flores, J., Midthun, K., Hoshino, Y., Green, K., Gorziglia, M., 13. Kapikian, A. Z. & Chanock, R. M. (1986) J. Virol. 60, 972-979.
- 14. Greenberg, H. B., McAuliffe, V., Valdesuso, J., Wyatt, R., Flores, J., Kalica, A., Hoshino, Y. & Singh, N. (1983) Infect. Immun. 39, 91-99.
- 15. Offit, P. A., Shaw, R. D. & Greenberg, H. B. (1986) J. Virol. 58, 700-703.
- Offit, P. A., Clark, F. H., Blavat, G. & Greenberg, H. B. (1986) J. 16. Virol. 60, 491-496.
- Shaw, R. D., Fong, K. J., Losonosky, G. A., Levine, M. M., Maldano, Y., Yolken, R., Flores, J., Kapikian, A. Z., Vo, P. T. & 17. Greenberg, H. B. (1987) Gastroenterology 93, 941-950.
- 18. Hoshino, Y., Wyatt, R. G., Greenberg, H. B., Flores, J. & Kapikian, A. Z. (1984) J. Infect. Dis. 149, 694-701.
- Rice, C. & Strauss, S. (1981) Proc. Natl. Acad. Sci. USA 78, 19. 2062-2066.
- Garoff, H., Frisehauf, A.-M., Simons, K., Lehrach, H. & Delius, 20. H. (1980) Nature (London) 288, 236-241.
- Suzuki, H., Kitaoka, S., Sato, T., Konno, T., Iwasaki, Y., Numa-21. zaki, Y. & Ishida, N. (1986) Arch. Virol. 91, 135-144.
- 22. Shaw, R. D., Vo, P. T., Offit, P. A., Coulson, B. S. & Greenberg, H. B. (1986) Virology 155, 434-451.
- Shaw, R. D., Stoner-Ma, D. L., Estes, M. K. & Greenberg, H. B. 23. (1985) J. Clin. Microbiol. 22, 286-291.
- 24. Flores, J., Greenberg, H. B., Myslinski, J., Kalica, A. R., Wyatt, R. G., Kapikian, A. Z. & Chanock, R. M. (1982) Virology 121, 288-295
- 25. Sippel, A. E. (1973) Eur. J. Biochem. 37, 31-40.
- Both, G. W., Bellamy, A. R., Street, J. E. & Siegman, L. J. (1982) 26. Nucleic Acids Res. 10, 7075-7088.
- 27. Sanger, R., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- 28. Zagursky, R., Baumeister, K., Lomax, N. & Berman, M. (1985) Gene Anal. Tech. 2, 5-11.
- 29. Bassel-Duby, R., Spriggs, D. R., Tyler, K. L. & Fields, B. N. (1986) J. Virol. 60, 64-67.
- 30. Donis-Keller, H. (1977) Nucleic Acids Res. 4, 2527-2535.
- 31.
- Skehel, J. J. & Joklik, W. K. (1969) *Virology* **39**, 822–832. Lopez, S., Arias, C. F., Bell, J. R., Strauss, J. H. & Espejo, R. T. 32. (1985) Virology 144, 11-19.
- Lopez, S. & Arias, C. F. (1987) Nucleic Acids Res. 15, 4691. 33.
- 34. Kozak, M. (1981) Nucleic Acids Res. 9, 5233-5252.
- 35. Kantharidis, P., Dyall-Smith, M. L. & Holmes, I. (1987) Arch. Virol. 93, 111-121.
- Gorziglia, M., Hoshino, Y., Buckler-White, A., Blumenthals, I., 36. Glass, R., Flores, J., Kapikian, A. Z. & Chanock, R. M. (1986) Proc. Natl. Acad. Sci. USA 83, 7039-7043.
- 37. Dyall-Smith, M. L. & Holmes, I. H. (1984) Nucleic Acids Res. 12, 3973-3982
- Ward, C. W., Azad, A. A. & Dyall-Smith, M. L. (1985) Virology 38. 144, 328-336.
- 39. Dyall-Smith, M. L., Elleman, T. C., Hoyne, P. A., Holmes, I. H. & Azad, A. A. (1983) Nucleic Acids Res. 11, 3351-3355.
- 40. Lopez, S., Arias, C. F., Mendez, E. & Espejo, R. T. (1986) Virology 154, 224-227.
- White, J., Kielian, M. & Helenius, A. (1983) Quart. Rev. Biophys. 41. 16, 151-195.
- 42. Kondor-Koch, C., Burke, B. & Garoff, H. (1983) J. Cell Biol. 97, 644-650
- 43. Taniguchi, K., Morita, Y., Urasawa, T. & Urasawa, S. (1987) J. Virol. 61, 1726-1730.
- Theil, K. W. & Saif, L. J. (1985) J. Clin. Microbiol. 21, 844-846. 44. 45.
- Wilcox, G. E. & Compans, R. W. (1982) Virology 123, 312-321.