

Genetic Determinants of Sindbis Virus Mosquito Infection Are Associated with a Highly Conserved Alphavirus and Flavivirus Envelope Sequence[∇]

Dennis J. Pierro,* Erik L. Powers, and Ken E. Olson

Arthropod-Borne and Infectious Disease Laboratory, Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, Colorado 80523

Received 17 September 2007/Accepted 15 December 2007

Wild-type Sindbis virus (SINV) strain MRE16 efficiently infects *Aedes aegypti* midgut epithelial cells (MEC), but laboratory-derived neurovirulent SINV strain TE/5'2J infects MEC poorly. SINV determinants for MEC infection have been localized to the E2 glycoprotein. The E2 amino acid sequences of MRE16 and TE/5'2J differ at 60 residue sites. To identify the genetic determinants of MEC infection of MRE16, the TE/5'2J virus genome was altered to contain either domain chimeras or more focused nucleotide substitutions of MRE16. The growth patterns of derived viruses in cell culture were determined, as were the midgut infection rates (MIR) in *A. aegypti* mosquitoes. The results showed that substitutions of MRE16 E2 aa 95 to 96 and 116 to 119 into the TE/5'2J virus increased MIR both independently and in combination with each other. In addition, a unique PPF/GDS amino acid motif was located between these two sites that was found to be a highly conserved sequence among alphaviruses and flaviviruses but not other arboviruses.

The majority of medically important mosquito-borne alphaviruses (family *Togaviridae*) and flaviviruses (family *Flaviviridae*) are vectored by two Culicinae mosquito genera, *Aedes* and *Culex*. The mosquito *Aedes aegypti* is a major vector in the epidemic disease cycles of the flaviviruses dengue virus (DENV) and yellow fever virus, as well as being a competent vector for a number of alphaviruses, including the prototypic alphavirus Sindbis virus (SINV). There are three major genotypes of SINV, the Paleoarctic-Ethiopian (P/E), which is found in Europe and Africa, the Oriental-Australian (O/A), which is found in Asia and Australasia, and the Southwest genotype, which is found in the southwestern region of Western Australia (28, 34, 40, 41). SINV strain MRE16 is an O/A genotype isolated between 1966 and 1969 from a pool of *Culex tritaeniorhynchus* mosquitoes in Malaysia (20a). A full-length infectious clone (IC) of MRE16 has been generated, and it is currently the model O/A genotype being investigated in laboratory mosquito infection assays (24, 25, 29, 31, 43). MRE16 was previously found to have a midgut infection rate (MIR) in *A. aegypti* mosquitoes of 100% at 7 days postinfection (dpi) (24). In contrast, recombinant SINV strain TE/5'2J (P/E genotype), a double subgenomic SINV constructed from a chimeric mouse neurovirulent variant of SINV AR339 called TE12, infects less than 15% of *A. aegypti* mosquito midguts when analyzed similarly (31).

The MRE16 E2 gene has been implicated as having genetic determinants of midgut infection (31). Sequence comparison between the SINV TE/5'2J and MRE16 E2 genes identified approximately 316 nucleotide (nt) (24.9%) and 60 amino acid (aa) residue (14.2%) differences (43). In this work, SINV strain

TE/5'2J was mutated to generate viruses containing MRE16 sequences within the E2 gene (as either chimeras or point mutations) and the viruses were evaluated for the ability to productively infect the MEC of *A. aegypti* mosquitoes. The data demonstrated that TE/5'2J viruses having the specific MRE16 E2 residue motif of either NH at positions 95 and 96 or SAGD at positions 116 to 119 conditioned the MIR in *A. aegypti*. When inserted together into TE/5'2J, these residues appeared to have an additive effect on MIR, although this elevated MIR was less than that of the parental MRE16 virus, indicating that there are additional MIR conditioning elements in MRE16. In addition, a unique amino acid sequence, i.e., PPF/GDS, spatially separating these two MRE16 motifs was observed to be highly conserved among the alpha- and flaviviruses in their envelope genes. Moreover, the newly identified PPF/GDS conserved residues, as well as the two MRE16 motifs, are located within a 35-aa region that harbors numerous infection determinants from both the *Flavivirus* and *Alphavirus* genera.

MATERIALS AND METHODS

Cell culture and cDNA plasmids. African green monkey kidney Vero (ATCC CCL-81) and *A. albopictus* C6/36 (ATCC CRL-1660) cells were grown in supplemented Eagle minimal essential medium with 8% fetal bovine serum. Plasmid pTE/5'2J was generously provided by Charles Rice, and its use has been previously described (12, 31). The construction and use of the pMRE16 IC were also described previously (25).

Construction of pTE/ME2a, pTE/ME2b, and point mutant cDNA plasmids. The previously constructed pTE/ME2 IC (also known as pME2/5'2J) used the pTE/5'2J backbone, which was mutated to incorporate the MRE16 E2 sequence (SINV nt 8673 to 9902, E2 aa 4 to 423) (31). For construction of pTE/ME2a, the 5' half of E2 (nt 8673 to 9240, E2 aa 4 to 193) from MRE16 was PCR amplified and ligated into pTE/5'2J between the XmnI site (nt 8673) and an engineered AgeI site (nt 9240); this was followed by site-directed mutagenesis PCR (sPCR) to insert MRE16 E2 residue 213A. For construction of pTE/ME2b, the 3' half of MRE16 E2 (nt 9240 to 9902, E2 aa 194 to 423) was PCR amplified and ligated into pTE/5'2J between the engineered AgeI site (nt 9240) and an engineered AgeI site (nt 9902). For the SINV point mutants, pTE/5'2J was sPCR mutated to generate the desired MRE16 amino acid motifs. Infectious SINV cDNA

* Corresponding author. Present address: Room 6132, Laboratory of Infectious Diseases, NIAID-NIH, 50 South Drive, Bethesda, MD 20892. Phone: (301) 594-2271. Fax: (301) 496-8312. E-mail: pierro@niaid.nih.gov.

[∇] Published ahead of print on 26 December 2007.

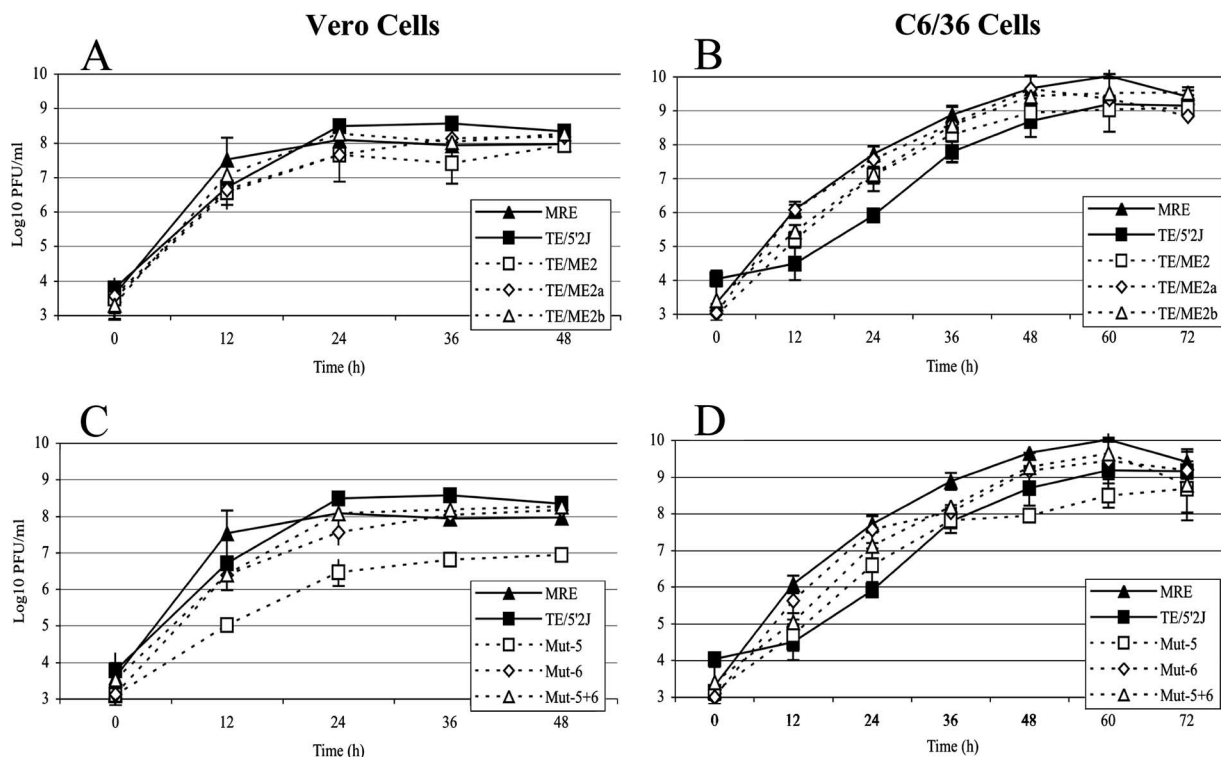


FIG. 1. Growth curves of SINV in vertebrate (A, C) and invertebrate (B, D) cell lines. Triplicate flasks of confluent cell monolayers were initially inoculated at an MOI of 0.01 and incubated at 37°C (vertebrate cells) or 28°C (invertebrate cells) in a 5% CO₂ chamber. Viral titers were determined by plaque titration on Vero cells. Error bars represent the standard error of the mean.

clones with multiple motifs (e.g., pTE/Mut-1-2-3-4-5-6) were generated by sequentially adding point mutations to pTE/5'2J in repeated rounds of sPCR and verified by sequencing of E2.

Generation and characterization of SINV particles. The production of infectious viral particles from the cDNAs of pTE/5'2J, pTE/ME2a, pTE/ME2b, and the sPCR mutant clones in 10^{6.4} Vero cells by reverse genetics was similar to previous descriptions (24, 31). First-in-cell generated virus (i.e., virus obtained by using RNA transcribed from the infectious clone) was termed P-0, and virus passaged once was called P-1. RNA from P-1 virus was isolated, the E2 gene was amplified by reverse transcription-PCR and fully sequenced, and those clones with no adaptive coding mutations were used. Plaque titrations of SINV were performed by infecting confluent monolayers of Vero cells overlaid with agarose as described previously (22). Plaque sizes of P-0 viruses were analyzed on Vero cells at 5 dpi. Growth curve analysis was performed with a confluent monolayer of cells infected in triplicate with P-0 virus at a multiplicity of infection (MOI) of 0.01 in 25-cm² flasks. The indirect immunofluorescence assay (IFA) of SINV-infected cultured cells and mosquito body tissues and midguts was performed with an anti-SINV E1 antibody (30.11a) as the primary antibody as previously described (25).

Mosquitoes and oral infections. Maintenance of *A. aegypti* strain Rexville D (RexD; Rexville D, Puerto Rico), a DENV- and SINV-susceptible colony strain of *A. aegypti*, and the infectious blood meal delivery protocol were similar to those previously described (24). The propagation of virus used in blood meals started with confluent monolayers of either Vero or C6/36 cells that were infected with P-0 virus at an MOI of approximately 0.01 and cultured for 2 days (Vero) or 3 days (C6/36). Freshly harvested (i.e., nonfrozen) P-1 virus was, when applicable, diluted in conditioned cell culture supernatant (0.8-ml final volume) to a predicted viral titer by using the standardized growth curve data as a guide (see Fig. 1). The pretitrated virus was then mixed with 0.8 ml of defibrinated sheep blood, warmed to 37°C, and placed in a water jacket-heated (37°C) glass membrane feeder, where mosquitoes were allowed to probe and feed through a stretched sheet of Parafilm for ≤30 min. Fully engorged mosquitoes were collected and maintained in the insectary with ample food and water until assayed at 9 dpi. The viral blood meal was frozen and later quantified by plaque titration and was considered acceptable at ±0.3 log₁₀ PFU/ml of the predicted viral titer.

Post-blood-meal viral titers were found to be reduced by ≤15% compared with pre-blood-meal titers (data not shown).

Statistical analysis and cluster alignment. The MIR is calculated by the number of IFA-positive mosquito midguts divided by the total number of blood-fed mosquitoes, expressed as a percentage, of two replicates for each virus with the standard error of the mean as error bars. The *n* value for mosquito organs in each analyzed replicate was 21 (total *n* for two replicates = 42). Comparison of the MIR rates for statistical significance was performed by combining the raw data of each replicate and analyzing by chi square test with an alpha cutoff of 0.05. The E2 amino acid cluster alignments (see Fig. 4) used the Pfam database, and the color codes used for clusters are from Jalview and ClustalX (9). The predictions for 2° structure and solvent access (see Fig. 4) were performed by PredictProtein (37–39). Amino acid sequence database searches for the conserved sequence (CS) used the Basic Local Alignment Search Tool (BLAST) with the protein-protein algorithm across the viral Swissprot protein database.

RESULTS

Growth curve characterization of SINV. The replication rates of the parental and chimeric viruses were analyzed by growth curve analysis in both vertebrate (Vero) and invertebrate (C6/36) cell lines (Fig. 1A and B). Maximum titers of the SINV chimeras were achieved at approximately 24 to 36 h postinfection of Vero cells and 48 to 60 h postinfection for propagation in C6/36 cells. The ranges of maximum virus titers were 7.9 to 8.6 and 9.1 to 9.7 log₁₀ PFU/ml in Vero and C6/36 cells, respectively. In addition, a large-plaque phenotype was observed in the MRE16, TE/ME2, and TE/ME2a viruses.

MRE16 E2 division harboring the genetic determinants of MIR. Parental and chimeric SINVs were analyzed in a per os midgut infection assay. At 9 dpi, midguts were dissected from

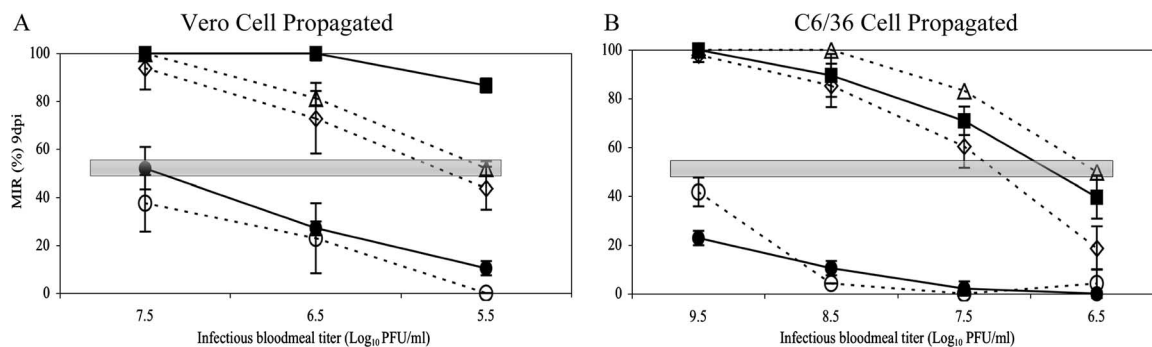


FIG. 2. MIR of chimeric viruses propagated in either Vero (A) or C6/36 (B) cells. The MIR is the average of IFA-positive midguts over n ($n = 21$) from two replicates at 9 dpi. The standard error of the mean is represented by the error bars. The shaded horizontal bar is the estimated MID_{50} . Symbols: ■, MRE16; ●, TE/5'2J; △, TE/ME2; ◇, TE/ME2a; ○, TE/ME2b.

mosquitoes, fixed in 4% paraformaldehyde, and assayed for the presence of SINV E1 antigen by IFA. The MIRs in *A. aegypti* for all of the viruses were dose dependent (Fig. 2A and B). In addition, the MRE16 and TE/5'2J viruses consistently had MIRs significantly different ($P < 0.05$) from each other at the respective blood meal doses. Changing the propagation cell type did not change these observations.

The TE/5'2J viruses containing the N- or C-terminal half of the MRE16 E2 gene (i.e., TE/ME2a and TE/ME2b, respectively) were assayed for oral infectivity. The MIR for TE/ME2a was found to be higher and significantly different from that of TE/5'2J ($P < 0.05$) for both Vero and C6/36 cell line-derived viruses at all of the doses analyzed (Fig. 2). In contrast, the MIR of TE/ME2b was not significantly different from that of TE/5'2J ($P > 0.05$), with the exception of high-titer (i.e., 9.5 \log_{10} PFU/ml) TE/ME2b virus derived from C6/36 cells. The MIRs of both Vero and C6/36 cell-propagated TE/ME2b viruses (low-titer C6/36-derived virus excepted) were statistically significantly lower ($P < 0.05$) than those of MRE16, TE/ME2, and TE/ME2a.

The influence of the infectious dose, as well as the determinants of infection, can be evaluated by comparing each of the virus's specific midgut 50% infective dose (MID_{50}), which is the infectious virus titer at which 50% of the mosquitoes had SINV-infected midguts at 9 dpi. The MID_{50} can be estimated from Fig. 2 (shaded horizontal bar in MIR), and the data are presented in Table 1. A lower MID_{50} is an indication of an

enhanced efficiency at which a virus establishes an infection in midgut cells. The MID_{50} was analyzed for viruses propagated in either Vero or C6/36 cells, and the MRE16 virus had a comparably lower MID_{50} (<5.5 and $7.0 \log_{10}$, respectively), while TE/5'2J had a higher MID_{50} (7.5 and >9.5 , respectively). All of the SINVs with an E2a portion from MRE16 had relatively low MID_{50} s of ≤ 5.75 (Vero cell derived) and ≤ 7.25 (C6/36 cell derived), while those SINVs with an E2a portion from TE/5'2J had MID_{50} s of ≥ 8.0 and >9.5 (Vero and C6/36 cell derived, respectively). These results indicate that determinants for enhanced midgut infection are in the E2a portion (E2 aa 4 to 213) of MRE16.

MRE16 E2 genetic motifs that influence MIR. Of the 209 aa residues within the E2a region, 19 (9%) are nonsynonymous between TE/5'2J and MRE16. To determine which residue conditioned MIR, the MRE16 codons were initially added sequentially by sPCR to the TE/5'2J plasmid, starting at E2 aa 213 (i.e., Mut-1) and moved toward the N terminus until a virus with an enhanced MIR was identified. The goal was to spotlight those residues that positively influence the MIR in *A. aegypti* mosquitoes. From Table 2, it appears that the high-titer TE/Mut-1-2-3-4-5-6 virus, which was propagated in C6/36 cells and fed to mosquitoes at $\geq 9.0 \log_{10}$ PFU/ml, harbors a determinant(s) of midgut infection. The Mut-6 motif was then added by itself to TE/5'2J (i.e., TE/Mut-6) and shown to positively influence the MIR over that of its parental TE/5'2J virus when propagated similarly ($P < 0.05$). The subsequent addition of Mut-5 (i.e., TE/Mut-5+6) further enhanced the MIR to levels not statistically significantly different from that of TE/Mut-1-2-3-4-5-6. The Mut-5 motif was then added to TE/5'2J by itself (i.e., TE/Mut-5), which positively influenced the MIR. Interestingly, the MIR of TE/Mut-5 was not statistically significantly different from those of TE/Mut-1-2-3-4-5-6 and TE/Mut-5+6 ($P > 0.05$), indicating that this mutation may be a major MRE16 determinant in midgut infections. These results indicate that the Mut-5 and Mut-6 residues play a significant role in conditioning the MIR. In addition, the large-plaque phenotype previously observed in MRE16, TE/ME2, and TE/ME2a was associated with TE/5'2J clones harboring the Mut-5 MRE16 amino acid residues (i.e., MRE16 E2 residues SAGD at positions 116 to 119) (Table 2).

Characterization of Mut-5 and Mut-6 motifs. Mutant SINV replication rates characterized by growth curve analysis in ver-

TABLE 1. Comparison of SINV MID_{50} s in *A. aegypti*

Virus	Primary residues		<i>A. aegypti</i> MID_{50} ^a		
	Mut-5	Mut6	Vero	C6/36	Δ
MRE16	SAGD	NH	$<5.5^b$	7	>1.5
TE/5'2J	VSSN	SY	7.5	$>9.5^b$	>2
TE/ME2	SAGD	NH	5.5	6.5	1
TE/ME2a	SAGD	NH	5.75	7.25	1.5
TE/ME2b	VSSN	SY	8	$>9.5^b$	>1.5
TE/Mut-5	SAGD	SY	6.5	8.25	1.75
TE/Mut-6	VSSN	NH	6.75	9	2.25
TE/Mut-5+6	SAGD	SY	6.25	8	1.75

^a Data are from shaded horizontal bars in Fig. 2 and 3. $\Delta = C6/36 \text{ MID}_{50} - \text{Vero MID}_{50}$.

^b The MID_{50} could not be accurately estimated due to virus blood meal titer limits (Fig. 2).

TABLE 2. Analysis of MIR of C6/36-propagated SINV

Virus	MRE16 E2		Schematic of residue exchanges in E2 of TE/5'2J ^a	Mean plaque size (mm)	MIR 9dpi in <i>A. aegypti</i> ^b		
	Amino acid position(s)	Residue(s)			<i>n</i>	+	% MIR (SEM)
TE/5'2J	NA ^c	NA ^c	—————	2.2	48	11	22.9 (2.9)
TE/ME2	4–423	From MRE16	————— —————	8.6 ^d	48	48	100.0 (0) ^d
TE/ME2a	4–213	From MRE16	————— —————	8.0 ^d	48	47	97.9 (2.9) ^d
TE/ME2b	194–423	From MRE16	————— —————	2.4	48	13	27.1 (5.9)
TE/Mut-1	213	T	————— ▼	2.3	48	12	25.0 (11.8)
TE/Mut-1-2	197 ^c	V	————— ▼▼	1.9	47	9	19.1 (8.5)
TE/Mut-1-2-3	178 ^c	T	————— ▼▼▼	2.2	48	13	27.1 (2.9)
TE/Mut-1-2-3-4	153–4 ^c	YI	————— ▼▼▼▼	1.9	48	12	25.0 (5.9)
TE/Mut-1-2-3-4-5	116–9 ^c	SAGD	————— ▼▼▼▼▼	7.9 ^d	48	15	31.3 (8.8)
TE/Mut-1-2-3-4-5-6	95–6 ^c	NH	————— ▼▼▼▼▼▼	8.6 ^d	48	46	95.8 (5.9) ^d
TE/Mut-6	95–6	NH	————— ▼	2.4	48	23	47.9 (14.7) ^d
TE/Mut-5	116–9	SAGD	————— ▼	8.1 ^d	48	41	85.4 (2.9) ^d
TE/Mut-5+6	95–6, 116–9	NH/SAGD	————— ▼▼	8.4 ^d	48	47	97.9 (2.9) ^d

^a The TE/5'2J E2 gene (thin line) was modified to include MRE16 residues as a chimera (thick line) or site-directed point mutations (▼).
^b The blood meal dose was 9.0 log₁₀ PFU/ml. *n* = combination of two replicates of 24 mosquitoes.
^c Position of bold mutation added to the previously generated mutation(s).
^d Statistically significantly different (*P* < 0.05) from the value for TE/5'2J.
^e NA, not applicable.

tebrate (Vero) and invertebrate (C6/36) cell lines showed that peak virus titers were at 24 to 36 and 48 to 60 hpi, respectively (Fig. 1C and D). The ranges of maximum virus titers for these viruses were 7.9 to 8.6 and 8.6 to 10.1 log₁₀ PFU/ml in Vero and C6/36 cells, respectively, with one exception. TE/Mut-5 replicated less efficiently in Vero cells, resulting in lower mean titers at specific time points, as well as a lower mean peak titer than that of the other viruses. These viruses differ only in E2, and altered replication efficiency is speculated to be a result of the maturation of the E2 glycoprotein. Also, TE/5'2J does appear to replicate slower than the other viruses in C6/36 cells, and this may influence the MIR.

The TE/Mut-5, TE/Mut-6, and TE/Mut-5+6 viruses exhibited a dose-dependent MIR in *A. aegypti* mosquitoes (Fig. 3A and B). In addition, the MIRs of both Vero and C6/36 cell line-derived TE/Mut-5, TE/Mut-6, and TE/Mut-5+6 viruses were significantly greater than those of TE/5'2J (*P* < 0.05) at the standardized blood meal dose of 7.5 log₁₀ PFU/ml (Vero cell-derived TE/Mut-5 excepted; Fig. 3C). With this same titer, the Vero cell-derived TE/Mut-5+6 viruses had statistically significantly greater MIRs than the respective Vero cell-derived TE/Mut-6 SINVs (*P* < 0.05). As was the case for the chimeric E2 SINVs, the MID₅₀s of the Mut-5 and Mut-6 viruses were ≥1.0 log₁₀ PFU/ml greater when they were propagated in C6/36 cells compared to Vero cells (Δ values in Table 1). In addition, both Vero and C6/36 cell-generated TE/Mut-5, TE/Mut-6, and TE/Mut 5+6 virus MID₅₀s were lower than those of the parental TE/5'2J virus, indicating an enhanced efficiency of midgut infection. These results indicate that the Mut-5 and

Mut-6 residues independently conditioned the MIR, although at different intensities. These residues also appear to work in conjunction with each other to condition the MIR, although the insensitivity of the MIR assay prohibited a conclusive evaluation.

Cluster alignment of SINV Mut-5+6. Spans of 35 aa residues (E2 aa 91 to 126) from SINV MRE16 and TE/5'2J, which included the Mut-5 and Mut-6 sites, were aligned against each other, as well as against a panel of other alphaviruses, and the amino acids were then clustered on the basis of residue characteristics (Fig. 4). Neither the MRE16 Mut-5 SAGD and Mut-6 NH nor the TE/5'2J Mut-5 VSSN and Mut-6 SY residues clustered as a group with any of the other alphaviruses. Individually, two (33.3%) of the six MRE16 residues in the Mut-5 and Mut-6 motifs did cluster with other alphaviruses. None of the individual TE/5'2J residues in the Mut-5 and Mut-6 motifs clustered with the alphavirus panel.

The predicted secondary structure of MRE16 and TE/5'2J identified only loop structures associated with both Mut-5 and Mut-6 motifs (Fig. 4; characters above and below the aligned SINV sequences). The amino acid residue changes in the Mut-5 (E2 aa 116 to 119) motif between TE/5'2J and MRE16 had no effect on the predicted secondary structure or the predicted solvent accessibility of this area. The amino acid residue changes from TE/5'2J to MRE16 in the Mut-6 (E2 aa 95 to 96) motif resulted in a slight shift in the location of the loop structure, as well as a decrease in the number of loop characters (from two to one) in this motif; however, the predicted solvent accessibility was unchanged. Cumulatively, this

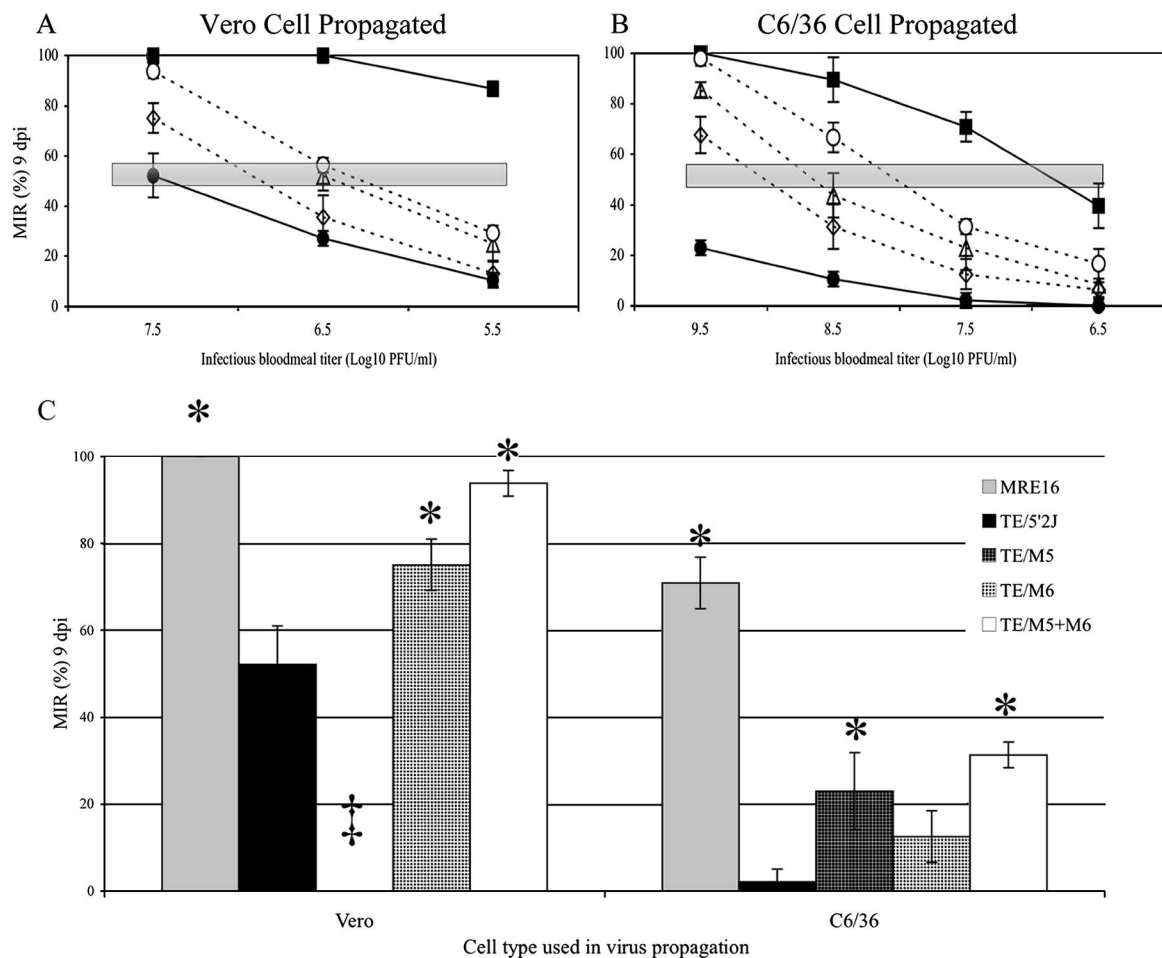


FIG. 3. MIRs of mutant viruses propagated in either Vero (A) or C6/36 (B) cells. The MIR is the average of IFA-positive midguts over n ($n = 21$) from two replicates at 9 dpi. The standard error of the mean is represented by the error bars. The shaded horizontal bar is the estimated MID_{50} . (C) Comparison of MIRs at a standardized infectious blood meal titer of $7.5 \log_{10}$ PFU/ml for viruses derived in either Vero or C6/36 cells. The MIR is the average of IFA-positive midguts over n ($n = 21$) from two replicates at 9 dpi with standard error bars. Symbols: ■, MRE16; ●, TE/5'2J; △, TE/Mut-5; ◇, TE/Mut-6; ○, TE/Mut-5+6. ‡, TE/Mut-5 does not replicate to $7.5 \log_{10}$ PFU/ml. *, statistically significantly different from TE/5'2J.

finding suggests that these changes do not drastically alter the motif's physical characteristics but instead only slightly enhance or reduce each motif's innate characteristics.

As mentioned, Fig. 4 shows that the amino acid residues in the Mut-5 and Mut-6 motifs of SINVs were not well conserved among the alphaviruses. However, a sequence stretch starting at SINV E2 glycine (G) 98 and ending at E2 isoleucine (I) 115, which spatially separates the Mut-5 and Mut-6 motifs, appears to have clustered frequently among the members of the alphavirus panel. Structure prediction analysis of the E2 aa 98-to-115 stretch determined that two beta sheets and one loop structure for the SINV were likely to occur (Fig. 4, characters flanking the SINV sequences). This loop structure was made up of conserved amino acid residues CPPGD at E2 positions 105 to 109. Following protein database analysis, elements of the PPGD motif appear to be part of a highly CS (namely, PPF/GDS) in the envelope glycoprotein ectodomains of the *Alphavirus* and *Flavivirus* genera (CS in Fig. 4). Moreover, the secondary-structure prediction of the larger 35-aa SINV stretch between E2 aa 91 and 126 (Fig. 4, characters above and below the SINV sequences) had a strikingly high degree of

similarity to the known structural conformation of the DENV-2 (Fig. 4; bottom of the flavivirus panel).

DISCUSSION

In this work, we showed that the MRE16 chimeric construct TE/ME2a had a higher MIR than the parental TE/5'2J virus. The amino acid sequence of the MRE16 E2a portion (aa 4 to 213), which contains the putative cell receptor binding domain, was 81% identical to that of TE/5'2J. Mutational exchange of the heterologous MRE16-E2a sites into the TE/5'2J IC resulted in viruses with an MIR significantly higher than that of the parental TE/5'2J virus, as was the case for the TE/Mut-1-2-3-4-5-6, TE/Mut-5+6, TE/Mut-6, and TE/Mut-5 viruses. Interestingly, the TE/Mut-5 virus exhibited an MIR higher than that of the TE/5'2J virus but not statistically significantly different from that of the TE/Mut-5+6 virus, indicating a dominant Mut-5 effect on midgut infections. However, it is likely that the Mut-5 and Mut-6 residues work cumulatively to affect the virus's MIR, as was seen previously in SINV strain TR339

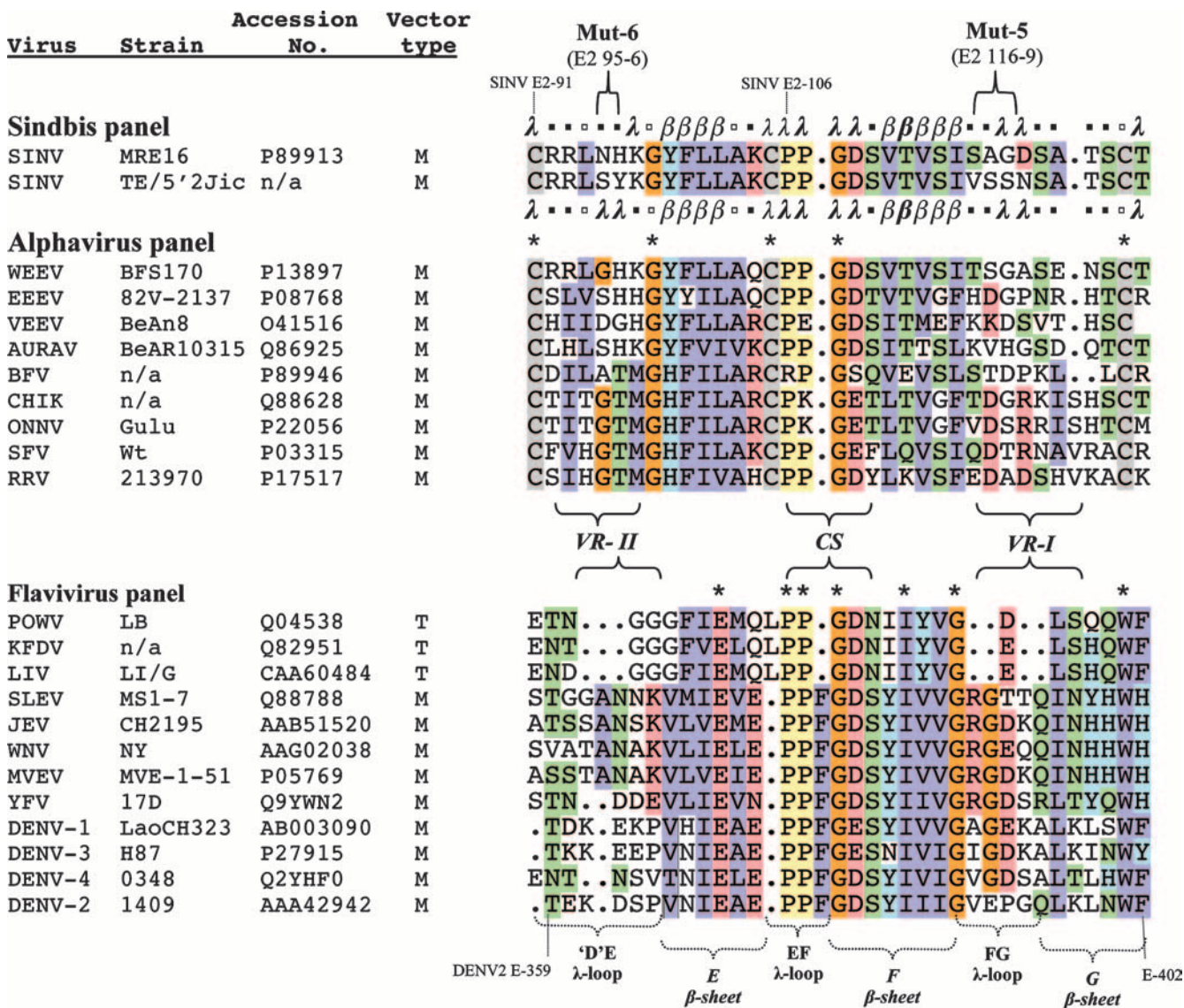


FIG. 4. Cluster alignment of *Alphavirus* and *Flavivirus* envelope glycoproteins. The amino acid alignments used the Pfam database, and the color codes used for clusters are from Jalview. The predictions for 2° structure and solvent access were performed by PredictProtein. The alphavirus and flavivirus amino acid residues were aligned and then color clustered when five or more residues with similar charge characteristics were present at the same location for each respective genus. Alignments are colored using the ClustalX scheme in Jalview (orange, glycine [G]; yellow, proline [P]; blue, small and hydrophobic amino acids [A, V, L, I, M, F, W]; green, hydroxyl and amine amino acids [S, T, N, Q]; red, charged amino acids [D, E, R, K]; cyan, histidine [H] and tyrosine [Y]). Braces indicate respective regions of interest described in the text. n/a, not available; period, no amino acid; asterisk, strictly conserved amino acid residue. The predicted secondary structure (flanking he MRE16 and TE/5'2J sequences) identified low (λ) and extended sheet (β) structures with an expected average accuracy of >82% or identified no predicted structure (□). No helixes were predicted at this threshold. Bold in the 2° structure indicates predicted solvent-accessible residues at >36%. Vector type: T, tick; M, mosquito.

(P/E genotype) but at different locations (E2 aa 55 and 70) (32).

The locations of SINV Mut-5 (E2 aa 116 to 119) and Mut-6 (E2 aa 95 and 96) motifs are unique compared to the majority of previously published alphavirus genetic determinants specifically affecting midgut infections. As mentioned, the genetic determinants of SINV TR339 MIR were found at E2 aa 55 and 70 (32). MRE16 does share the E2 aa 55 Q residue with TR339, and it can be supposed to be another genetic determinant of MRE16's MIR, although its influence was not analyzed

here. E2 aa 117 and 207 in Venezuelan equine encephalitis virus (VEEV) were each shown to assist in the enhancement of the MIR (3, 55). E2 aa 170 to 220 comprise the putative cell receptor binding domain, and studies characterizing this domain predict that it is exposed on the viral surface (6, 45, 46, 48). This information leads to the observation that multiple locations occur in the N-terminal half of E2 of alphaviruses that can influence midgut infections.

Published material concerning influential alphavirus infection residues that are specific to the Mut-5 sites (E2 aa 116 to

119) includes a report by Brault et al., who previously identified a genetic determinant in a VEEV IE subtype at E2 aa 117 (SINV E2 aa 117) that assisted in an enhanced MIR in *A. taeniorhynchus* mosquitoes (3). In addition, a change at VEEV E2 aa 116 affected HS binding (1), a change at SINV E2 aa 114 enhanced viral binding to HS and to BHK cells and altered antibody reactivity and attenuation in mice (5, 20), a change at VEEV vaccine strain TC-83 E2 aa 120 (SINV E2 aa 121) correlated with attenuation in mice (19), and a change at Ross River virus E2 aa 119 (SINV E2 aa 119) was implied to be an attenuation marker (54). In summary, the alphavirus region encompassing all of these mutations (i.e., SINV E2 aa 114 to 121), which includes SINV Mut-5, appears to affect the viral infection potentials of a variety of alphaviruses and implies a structurally important region of the E2. The crystal structure of alphavirus E2 has not been determined, but recent publications have given insight into its basic conformation (23, 57). According to those reports, the E2 glycoprotein stands perpendicular to the viral surface, having a slightly hunched form in 80 homotrimer units. The E2 C terminus is attached to the viral membrane, and the N terminus is thought to be distal from the virus surface (23). The SINV Mut-5 motif appears to be positioned as part of the N terminus hunch of the E2 glycoprotein structure, facing away from the viral surface, with each homotrimer forming a cell surface-exposed pocket of Mut-5 motifs. Synthetic peptide, as well as target-specific antibody (generated from synthetic peptides), antagonists of SINV Mut-5 infection (i.e., anti-E2 aa 112 to 125) were investigated during this work with no effect (data not shown).

No published information is available on the genetic determinants specific to midgut infections in the Mut-6 motif (E2 aa 95 to 96). However, a mutation at SINV E2 aa 96 was shown to affect neutralizing antibody escape, as well as virulence, in neonatal mice (30). Again, with the basic structure proposed by Mukhopadhyay et al., the SINV Mut-6 motif is in close proximity to Mut-5 at the N terminus of E2, also facing away from the viral surface, with each homotrimer forming a cell surface-exposed pocket of these Mut-6 motifs (23).

A significant observation concerning the location of the Mut-5 and Mut-6 amino acid residues along the linear E2 sequence was that they were separated by a PPGDS motif, which appears to be a highly CS in the envelope glycoprotein ectodomains of the *Alphavirus* and *Flavivirus* genera (PPF/.GDS in flaviviruses; Fig. 4). The PPF/.GDS sequence is located at SINV E2 positions 106 to 110 and DENV-2 E positions 371 to 376, respectively. While these two arboviral genera belong to different taxonomic families (i.e., *Togaviridae* and *Flaviviridae*), the motif was not found in the structural genes of other arbovirus families (e.g., *Bunyaviridae*, *Rhabdoviridae*, *Reoviridae*, *Orthomyxoviridae*, and *Asfarviridae*). Interestingly, the CS is absent from the genomes of the insect-only flaviviruses (i.e., Kimiti River and cell fusing agent viruses) and from the genome of the salmon pancreatic disease virus, which is an *Alphavirus* that has no known vector. For the sake of clarity, this newly identified motif (i.e., PPF/.GDS) is termed the CS in Fig. 4.

The function of this newly identified CS is unknown, although some published data suggest that it is an important structural region. Navaratnarajah and Kuhn have recently shown that insertion of a transposon linker (15 to 16 aa resi-

dues) into SINV at E2 aa 105 (just upstream of the CS) resulted in the production of infectious virus while insertions at E2 aa 107 and 109 (in the CS) were lethal to SINV (26). Those authors also observed that various insertions of the transposon linker into the SINV E2 aa 107-to-119 region did not affect E2 expression but did impair proper E2 transport to the plasma membrane. Since the sequence of the CS is dominated by two prolines, a passive structural role rather than an active functional one may be implied since proline is often unreactive and found at very tight turns in protein structures (2). Prolines are also thought to position distant virus reactive sites for optimal accessibility within a protein (7).

According to the flavivirus alignment in Fig. 4, the SINV Mut-5 motif aligns spatially with the DENV-2 FG loop secondary structure, a region exposed on the distal face of domain III of the envelope glycoprotein and thought to be involved in receptor binding (21, 35). The SINV Mut-5 sequence also aligns with the RGD sequences of some of the flaviviruses. The RGD motif is a known participant in the interaction of a number of ligands with cell receptors from the integrin superfamily (8). Residues of the RGD motif, as well as RGD-associated regions, have previously been shown to influence the activity of a number of enveloped and nonenveloped viruses (4, 10, 11, 13, 15, 17, 21, 27, 36, 42, 44, 47, 49, 50, 52, 56). Moreover, as evidence of the importance of this region, a 10-mer DENV-2 synthetic peptide sequence that overlapped the flavivirus Mut-5-like region (i.e., flavivirus amino acid residues at the Mut-5 positions) inhibited the binding of a DENV-2 domain III envelope protein to mosquito but not mammalian cells, indicating that this region has more specificity for insect cells (16). Because the Mut-5-like alphavirus and flavivirus residues align spatially but their sequences are not well conserved (Fig. 4) and even encompass a deletion region that corresponds solely to tick-borne flaviviruses (14), this region may be hypothesized to constitute a variable region (VR) involved in mosquito infections and is termed VR-I in Fig. 4 (SINV E2 aa 116 to 122, DENV-2 E aa 382 to 388).

Also from the flavivirus section of Fig. 4, the Mut-6 region aligned spatially with the flavivirus D'E λ loop (DENV-2 E aa 357 to 365) secondary structure, a region also exposed on the face of domain III of the E glycoprotein. The yellow fever virus E aa 358-to-365 region (DENV-2 E aa 359 to 367) was previously suggested to be involved in a conformational epitope (18, 53). Due to the structural and infection potential conservation, but having a variable sequence, this Mut-6 region constitutes another VR (encompassing SINV E2 aa 92 to 96 and DENV-2 E aa 359 to 365) and is termed VR-II in Fig. 4.

In summary, the specific MRE16 amino acid residue sites involved in midgut infections are SINV E2 aa 95 to 96 and 116 to 119, which are nonconserved within predicted solvent-exposed loop structures of members of the *Alphavirus* genus. In addition, the known alphavirus genetic determinants of midgut infection are variable and appear to be dispersed across the N-terminal half of E2. Lastly, the evidence presented here points to two solvent-exposed viral envelope glycoprotein domains, VR-I and -II, found in members of both the *Alphavirus* and *Flavivirus* genera and involved in mosquito and mammalian cell infections. These two regions are predicted to be loop domains and are separated by the CS, which is also a loop structure. Interestingly, in some group I coronaviruses, the

viral spike glycoprotein binds to its receptor, aminopeptidase N at regions within loop structures that have a highly variable amino acid sequence that has been shown to be a determinant of host range (51). It could be hypothesized that the multiple protruding loop structures on the SINV E2 glycoprotein work in conjunction with each other to enhance the overall affinity of the virus for cellular ligands.

ACKNOWLEDGMENTS

This work was funded by NIH NIAID grant AI046435.

We thank Cindy Meredith for excellent technical help in maintaining our mosquito colonies.

REFERENCES

- Bernard, K. A., W. B. Klimstra, and R. E. Johnston. 2000. Mutations in the E2 glycoprotein of Venezuelan equine encephalitis virus confer heparan sulfate interaction, low morbidity, and rapid clearance from blood of mice. *Virology* **276**:93–103.
- Betts, M. J., and R. B. Russell. 2003. Amino acid properties and consequences of substitutions, p. 289–316. *In* M. R. Barnes and I. C. Gray (ed.), *Bioinformatics for geneticists*. Wiley, New York, NY.
- Brault, A. C., A. M. Powers, D. Ortiz, J. G. Estrada-Franco, R. Navarro-Lopez, and S. C. Weaver. 2004. Venezuelan equine encephalitis emergence: enhanced vector infection from a single amino acid substitution in the envelope glycoprotein. *Proc. Natl. Acad. Sci. USA* **101**:11344–11349.
- Chambers, T. J., and M. Nickells. 2001. Neuroadapted yellow fever virus 17D: genetic and biological characterization of a highly mouse-neurovirulent virus and its infectious molecular clone. *J. Virol.* **75**:10912–10922.
- Davis, N. L., F. J. Fuller, W. G. Dougherty, R. A. Olmsted, and R. E. Johnston. 1986. A single nucleotide change in the E2 glycoprotein gene of Sindbis virus affects penetration rate in cell culture and virulence in neonatal mice. *Proc. Natl. Acad. Sci. USA* **83**:6771–6775.
- Davis, N. L., D. F. Pence, W. J. Meyer, A. L. Schmaljohn, and R. E. Johnston. 1987. Alternative forms of a strain-specific neutralizing antigenic site on the Sindbis virus E2 glycoprotein. *Virology* **161**:101–108.
- Delos, S. E., J. M. Gilbert, and J. M. White. 2000. The central proline of an internal viral fusion peptide serves two important roles. *J. Virol.* **74**:1686–1693.
- D'Souza, S. E., M. H. Ginsberg, and E. F. Plow. 1991. Arginyl-glycyl-aspartic acid (RGD): a cell adhesion motif. *Trends Biochem. Sci.* **16**:246–250.
- Finn, R. D., J. Mistry, B. Schuster-Bockler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S. R. Eddy, E. L. Sonnhammer, and A. Bateman. 2006. Pfam: clans, web tools and services. *Nucleic Acids Res.* **34**:D247–D251.
- Fox, G., N. R. Parry, P. V. Barnett, B. McGinn, D. J. Rowlands, and F. Brown. 1989. The cell attachment site on foot-and-mouth disease virus includes the amino acid sequence RGD (arginine-glycine-aspartic acid). *J. Gen. Virol.* **70**(Pt. 3):625–637.
- Hahn, C. S., J. M. Dalrymple, J. H. Strauss, and C. M. Rice. 1987. Comparison of the virulent Asibi strain of yellow fever virus with the 17D vaccine strain derived from it. *Proc. Natl. Acad. Sci. USA* **84**:2019–2023.
- Hahn, C. S., Y. S. Hahn, T. J. Braciale, and C. M. Rice. 1992. Infectious Sindbis virus transient expression vectors for studying antigen processing and presentation. *Proc. Natl. Acad. Sci. USA* **89**:2679–2683.
- Hiramatsu, K., M. Tadano, R. Men, and C. J. Lai. 1996. Mutational analysis of a neutralization epitope on the dengue type 2 virus (DEN2) envelope protein: monoclonal antibody resistant DEN2/DEN4 chimeras exhibit reduced mouse neurovirulence. *Virology* **224**:437–445.
- Holbrook, M. R., R. E. Shope, and A. D. Barrett. 2004. Use of recombinant E protein domain III-based enzyme-linked immunosorbent assays for differentiation of tick-borne encephalitis serocomplex flaviviruses from mosquito-borne flaviviruses. *J. Clin. Microbiol.* **42**:4101–4110.
- Holzmann, H., F. X. Heinz, C. W. Mandl, F. Guirakhoo, and C. Kunz. 1990. A single amino acid substitution in envelope protein E of tick-borne encephalitis virus leads to attenuation in the mouse model. *J. Virol.* **64**:5156–5159.
- Hung, J. J., M. T. Hsieh, M. J. Young, C. L. Kao, C. C. King, and W. Chang. 2004. An external loop region of domain III of dengue virus type 2 envelope protein is involved in serotype-specific binding to mosquito but not mammalian cells. *J. Virol.* **78**:378–388.
- Hurrelbrink, R. J., and P. C. McMinn. 2001. Attenuation of Murray Valley encephalitis virus by site-directed mutagenesis of the hinge and putative receptor-binding regions of the envelope protein. *J. Virol.* **75**:7692–7702.
- Jackson, R. E., and R. J. Philippotts. 1997. Production of virus-specific antisera using synthetic peptides corresponding to sequences in the yellow fever E protein: brief report. *Viral Immunol.* **10**:11–14.
- Kinney, R. M., G. J. Chang, K. R. Tsuchiya, J. M. Sneider, J. T. Roehrig, T. M. Woodward, and D. W. Trent. 1993. Attenuation of Venezuelan equine encephalitis virus strain TC-83 is encoded by the 5'-noncoding region and the E2 envelope glycoprotein. *J. Virol.* **67**:1269–1277.
- Klimstra, W. B., K. D. Ryman, K. A. Bernard, K. B. Nguyen, C. A. Biron, and R. E. Johnston. 1999. Infection of neonatal mice with Sindbis virus results in a systemic inflammatory response syndrome. *J. Virol.* **73**:10387–10398.
- Kurstak, E. 1979. Arctic and tropical arboviruses. Academic Press, Inc., New York, NY.
- Lee, E., and M. Lobigs. 2000. Substitutions at the putative receptor-binding site of an encephalitic flavivirus alter virulence and host cell tropism and reveal a role for glycosaminoglycans in entry. *J. Virol.* **74**:8867–8875.
- Miller, B. R., and C. J. Mitchell. 1986. Passage of yellow fever virus: its effect on infection and transmission rates in *Aedes aegypti*. *Am. J. Trop. Med. Hyg.* **35**:1302–1309.
- Mukhopadhyay, S., W. Zhang, S. Gabler, P. R. Chipman, E. G. Strauss, J. H. Strauss, T. S. Baker, R. J. Kuhn, and M. G. Rossmann. 2006. Mapping the structure and function of the E1 and E2 glycoproteins in alphaviruses. *Structure* **14**:63–73.
- Myles, K. M., D. J. Pierro, and K. E. Olson. 2004. Comparison of the transmission potential of two genetically distinct Sindbis viruses after oral infection of *Aedes aegypti* (Diptera: Culicidae). *J. Med. Entomol.* **41**:95–106.
- Myles, K. M., D. J. Pierro, and K. E. Olson. 2003. Deletions in the putative cell receptor-binding domain of Sindbis virus strain MRE16 E2 glycoprotein reduce midgut infectivity in *Aedes aegypti*. *J. Virol.* **77**:8872–8881.
- Navaratnarajah, C. K., and R. J. Kuhn. 2007. Functional characterization of the Sindbis virus E2 glycoprotein by transposon linker-insertion mutagenesis. *Virology* **363**:134–147.
- Nickells, M., and T. J. Chambers. 2003. Neuroadapted yellow fever virus 17D: determinants in the envelope protein govern neuroinvasiveness for SCID mice. *J. Virol.* **77**:12232–12242.
- Olson, K., and D. W. Trent. 1985. Genetic and antigenic variations among geographical isolates of Sindbis virus. *J. Gen. Virol.* **66**(Pt. 4):797–810.
- Olson, K. E., K. M. Myles, R. C. Seabaugh, S. Higgs, J. O. Carlson, and B. J. Beaty. 2000. Development of a Sindbis virus expression system that efficiently expresses green fluorescent protein in midguts of *Aedes aegypti* following per os infection. *Insect Mol. Biol.* **9**:57–65.
- Pence, D. F., N. L. Davis, and R. E. Johnston. 1990. Antigenic and genetic characterization of Sindbis virus monoclonal antibody escape mutants which define a pathogenesis domain on glycoprotein E2. *Virology* **175**:41–49.
- Pierro, D. J., K. M. Myles, B. D. Foy, B. J. Beaty, and K. E. Olson. 2003. Development of an orally infectious Sindbis virus transducing system that efficiently disseminates and expresses green fluorescent protein in *Aedes aegypti*. *Insect Mol. Biol.* **12**:107–116.
- Pierro, D. J., E. L. Powers, and K. E. Olson. 2007. Genetic determinants of Sindbis virus strain TR339 affecting midgut infection in the mosquito *Aedes aegypti*. *J. Gen. Virol.* **88**:1545–1554.
- Reference deleted.
- Rentier-Delrue, F., and N. A. Young. 1980. Genomic divergence among Sindbis virus strains. *Virology* **106**:59–70.
- Rey, F. A., F. X. Heinz, C. Mandl, C. Kunz, and S. C. Harrison. 1995. The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. *Nature* **375**:291–298.
- Roehrig, J. T., P. A. Risi, J. R. Brubaker, A. R. Hunt, B. J. Beaty, D. W. Trent, and J. H. Mathews. 1994. T-helper cell epitopes on the E-glycoprotein of dengue 2 Jamaica virus. *Virology* **198**:31–38.
- Rost, B., and C. Sander. 1994. Conservation and prediction of solvent accessibility in protein families. *Proteins* **20**:216–226.
- Rost, B., and C. Sander. 1993. Improved prediction of protein secondary structure by use of sequence profiles and neural networks. *Proc. Natl. Acad. Sci. USA* **90**:7558–7562.
- Rost, B., G. Yachdav, and J. Liu. 2004. The PredictProtein server. *Nucleic Acids Res.* **32**:W321–W326.
- Saleh, S. M., M. Poidinger, J. S. Mackenzie, A. K. Broom, M. D. Lindsay, and R. A. Hall. 2003. Complete genomic sequence of the Australian southwest genotype of Sindbis virus: comparisons with other Sindbis strains and identification of a unique deletion in the 3'-untranslated region. *Virus Genes* **26**:317–327.
- Sammels, L. M., M. D. Lindsay, M. Poidinger, R. J. Coelen, and J. S. Mackenzie. 1999. Geographic distribution and evolution of Sindbis virus in Australia. *J. Gen. Virol.* **80**(Pt. 3):739–748.
- Sánchez, I. J., and B. H. Ruiz. 1996. A single nucleotide change in the E protein gene of dengue virus 2 Mexican strain affects neurovirulence in mice. *J. Gen. Virol.* **77**(Pt. 10):2541–2545.
- Seabaugh, R. C., K. E. Olson, S. Higgs, J. O. Carlson, and B. J. Beaty. 1998. Development of a chimeric Sindbis virus with enhanced per os infection of *Aedes aegypti*. *Virology* **243**:99–112.
- Serafin, I. L., and J. G. Aaskov. 2001. Identification of epitopes on the envelope (E) protein of dengue 2 and dengue 3 viruses using monoclonal antibodies. *Arch. Virol.* **146**:2469–2479.
- Smith, T. J., R. H. Cheng, N. H. Olson, P. Peterson, E. Chase, R. J. Kuhn, and T. S. Baker. 1995. Putative receptor binding sites on alphaviruses as visualized by cryoelectron microscopy. *Proc. Natl. Acad. Sci. USA* **92**:10648–10652.

46. **Stec, D. S., A. Waddell, C. S. Schmaljohn, G. A. Cole, and A. L. Schmaljohn.** 1986. Antibody-selected variation and reversion in Sindbis virus neutralization epitopes. *J. Virol.* **57**:715–720.
47. **Storey, P., J. Theron, F. F. Maree, and H. G. O'Neill.** 2007. A second RGD motif in the 1D capsid protein of a SAT1 type foot-and-mouth disease virus field isolate is not essential for attachment to target cells. *Virus Res.* **124**:184–192.
48. **Strauss, E. G., D. S. Stec, A. L. Schmaljohn, and J. H. Strauss.** 1991. Identification of antigenically important domains in the glycoproteins of Sindbis virus by analysis of antibody escape variants. *J. Virol.* **65**:4654–4664.
49. **Tan, B.-H., E. Nason, N. Staeuber, W. Jiang, K. Monastyrskaya, and P. Roy.** 2001. RGD tripeptide of bluetongue virus VP7 protein is responsible for core attachment to *Culicoides* cells. *J. Virol.* **75**:3937–3947.
50. **Tugizov, S. M., J. W. Berline, and J. M. Palefsky.** 2003. Epstein-Barr virus infection of polarized tongue and nasopharyngeal epithelial cells. *Nat. Med.* **9**:307–314.
51. **Tusell, S. M., S. A. Schittone, and K. V. Holmes.** 2007. Mutational analysis of aminopeptidase N, a receptor for several group 1 coronaviruses, identifies key determinants of viral host range. *J. Virol.* **81**:1261–1273.
52. **van der Most, R. G., J. Corver, and J. H. Strauss.** 1999. Mutagenesis of the RGD motif in the yellow fever virus 17D envelope protein. *Virology* **265**:83–95.
53. **Vlaycheva, L., M. Nickells, D. A. Droll, and T. J. Chambers.** 2005. Neuroblastoma cell-adapted yellow fever virus: mutagenesis of the E protein locus involved in persistent infection and its effects on virus penetration and spread. *J. Gen. Virol.* **86**:413–421.
54. **Vrati, S., C. A. Fernon, L. Dalgarno, and R. C. Weir.** 1988. Location of a major antigenic site involved in Ross River virus neutralization. *Virology* **162**:346–353.
55. **Woodward, T. M., B. R. Miller, B. J. Beaty, D. W. Trent, and J. T. Roehrig.** 1991. A single amino acid change in the E2 glycoprotein of Venezuelan equine encephalitis virus affects replication and dissemination in *Aedes aegypti* mosquitoes. *J. Gen. Virol.* **72**(Pt. 10):2431–2435.
56. **Wu, S. C., W. C. Lian, L. C. Hsu, and M. Y. Liau.** 1997. Japanese encephalitis virus antigenic variants with characteristic differences in neutralization resistance and mouse virulence. *Virus Res.* **51**:173–181.
57. **Zhang, W., S. Mukhopadhyay, S. V. Pletnev, T. S. Baker, R. J. Kuhn, and M. G. Rossmann.** 2002. Placement of the structural proteins in Sindbis virus. *J. Virol.* **76**:11645–11658.