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 Culicidae 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 Paleartic-Ethiopian (P/E), which is found in Europe and Africa, the Oriental-Australian (O/A), which is found in Asia and Australia, 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 triaeniocephalus 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 midgut 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 alphavirus and flavivirus 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 pME2S’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...
clones with multiple motifs (e.g., pTE/Mut-1-2-3-4-5-6) were generated by sequentially adding point mutations to pTE/H11032 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/H11032, 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 infections 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). 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).

**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 log_{10} 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

**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.

**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.**

**Vol. 82, 2008 ARBOVIRUS CONSERVED DETERMINANTS OF MOSQUITO INFECTION 2967**
mosquitoes, fixed in 4% paraformaldehyde, and assayed for the presence of SINv E1 antigen by IFA. The MIRs were evaluated by comparing each of the viraemia levels not statistically significantly different from that of TE/5/2J (P > 0.05), with the exception of high-titer (i.e., 9.5 log10 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 (MID50), which is the infectious virus titer at which 50% of the mosquitoes had SINv-infected midguts at 9 dpi. The MID50 can be estimated from Fig. 2 (shaded horizontal bar in MIR), and the data are presented in Table 1. A lower MID50 is an indication of an enhanced efficiency at which a virus establishes an infection in midgut cells. The MID50 was analyzed for viruses propagated in either Vero or C6/36 cells, and the MRE16 virus had a comparably lower MID50 (<5.5 and 7.0 log10, respectively), while TE/5/2J had a higher MID50 (7.5 and >9.5, respectively). All of the SINVs with an E2a portion from MRE16 had relatively low MID50s 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 MID50s 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.

**TABLE 1. Comparison of SINv MID50 in A. aegypti**

<table>
<thead>
<tr>
<th>Virus</th>
<th>Primary residues</th>
<th>A. aegypti MID50</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mut5</td>
<td>Mut6</td>
<td>Vero</td>
</tr>
<tr>
<td>MRE16</td>
<td>SAGD</td>
<td>NH</td>
<td>&lt;5.5*</td>
</tr>
<tr>
<td>TE/5/2J</td>
<td>VSSN</td>
<td>SY</td>
<td>7.5</td>
</tr>
<tr>
<td>TE/ME2</td>
<td>SAGD</td>
<td>NH</td>
<td>5.5</td>
</tr>
<tr>
<td>TE/ME2a</td>
<td>SAGD</td>
<td>NH</td>
<td>5.75</td>
</tr>
<tr>
<td>TE/ME2b</td>
<td>VSSN</td>
<td>SY</td>
<td>8</td>
</tr>
<tr>
<td>TE/Mut-5</td>
<td>SAGD</td>
<td>SY</td>
<td>6.5</td>
</tr>
<tr>
<td>TE/Mut-6</td>
<td>VSSN</td>
<td>NH</td>
<td>6.75</td>
</tr>
<tr>
<td>TE/Mut-5+6</td>
<td>SAGD</td>
<td>SY</td>
<td>6.25</td>
</tr>
</tbody>
</table>

a Data are from shaded horizontal bars in Fig. 2 and 3. Δ = C6/36 MID50 − Vero MID50.

b The MID50 could not be accurately estimated due to virus blood meal titer limits (Fig. 2).
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 peak titer and C6/36 cells, respectively, with one exception. TE/Mut-5 had significantly higher MIR in A. aegypti mosquitoes (Fig. 3A).

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 91 to 119) motif between TE/5’2J and MRE16 did not 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.

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

<table>
<thead>
<tr>
<th>Virus</th>
<th>MRE16 E2 Amino acid position(s)</th>
<th>Schematic of residue exchanges in E2 of TE/5’2J</th>
<th>Mean plaque size (mm)</th>
<th>MIR 9dpi in A. aegypti</th>
</tr>
</thead>
<tbody>
<tr>
<td>TE/5’2J</td>
<td>NA’</td>
<td></td>
<td>2.2</td>
<td>48</td>
</tr>
<tr>
<td>TE/ME2</td>
<td>4-23 From MRE16</td>
<td></td>
<td>8.6d</td>
<td>48</td>
</tr>
<tr>
<td>TE/ME2a</td>
<td>4-213 From MRE16</td>
<td></td>
<td>8.0d</td>
<td>48</td>
</tr>
<tr>
<td>TE/ME2b</td>
<td>194-423 From MRE16</td>
<td></td>
<td>2.4</td>
<td>48</td>
</tr>
<tr>
<td>TE/Mut-1</td>
<td>213 T</td>
<td></td>
<td>2.3</td>
<td>48</td>
</tr>
<tr>
<td>TE/Mut-1-2</td>
<td>197 T</td>
<td>V</td>
<td>1.9</td>
<td>47</td>
</tr>
<tr>
<td>TE/Mut-1-2-3</td>
<td>178 T</td>
<td>T</td>
<td>2.2</td>
<td>48</td>
</tr>
<tr>
<td>TE/Mut-1-2-3-4</td>
<td>153-4c</td>
<td>YI</td>
<td>1.9</td>
<td>48</td>
</tr>
<tr>
<td>TE/Mut-1-2-3-4-5</td>
<td>116-9g SAGD</td>
<td></td>
<td>7.9d</td>
<td>48</td>
</tr>
<tr>
<td>TE/Mut-1-2-3-4-5-6</td>
<td>95-6c NH</td>
<td></td>
<td>8.6d</td>
<td>48</td>
</tr>
<tr>
<td>TE/Mut-6</td>
<td>95-6c NH</td>
<td></td>
<td>2.4</td>
<td>48</td>
</tr>
<tr>
<td>TE/Mut-5</td>
<td>116-9 SAGD</td>
<td></td>
<td>8.1d</td>
<td>48</td>
</tr>
<tr>
<td>TE/Mut-5+6</td>
<td>95-6c, 116-9c NH/SAGD</td>
<td></td>
<td>8.4d</td>
<td>48</td>
</tr>
</tbody>
</table>

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.
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 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 seen previously in SINV strain TR339.
(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

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/5’2/J sequences) identified low and extended sheet structures with an expected average accuracy of >82% or identified no predicted structure ( ). No helices were predicted at this threshold. Bold in the 2° structure indicates predicted solvent-accessible residues at >36%. Vector type: T, tick; M, mosquito.
119) includes a report by Brault et al., who previously identified a genetic determinant in a VEEV 1E subtype at E2 aa 117 (SINV E2 aa 117) that assisted in an enhanced MIR in A. taeniopyrhynus 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 Flavivirus 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 residues) 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