Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 6771–6775, September 1986
Biochemistry

A single nucleotide change in the E2 glycoprotein gene of Sindbis virus affects penetration rate in cell culture and virulence in neonatal mice

(alphavirus neurovirulence/alphavirus mutant sequence/alphavirus surface proteins)

NANCY L. DAVIS†, FREDERICK J. FULLER‡, WILLIAM G. DOUGHERTY§, ROBERT A. OLMSTED†¶, AND ROBERT E. JOHNSTON†

Departments of †Microbiology and §Plant Pathology, School of Agriculture and Life Sciences, and ‡Department of Microbiology, Pathology and Parasitology, School of Veterinary Medicine, North Carolina State University, Raleigh, NC 27695

Communicated by C. Clark Cockerham, June 5, 1986

ABSTRACT **The nucleotide sequence of the glycoprotein genes of fully virulent Sindbis virus and derived mutants that have reduced neurovirulence for neonatal mice (attenuated mutants) has been determined. A single amino acid difference, arginine instead of serine at position 114 of the mature E2 glycoprotein, distinguished the prototype attenuated mutant from its virulent wild-type parent. Virulent revertants of the attenuated mutant showed same-site reversion to the wild-type sequence. An identical single amino acid substitution, an arginine for the serine at E2 position 114, was found in a second independently selected attenuated mutant. The strains are characterized by genetic linkage between attenuation, accelerated penetration of baby hamster kidney cells, and efficient neutralization by the E2-specific monoclonal antibodies R6 and R13; selection for change in one property simultaneously selected for change in the other two (Olmsted, R. A., Baric, R. S., Sawyer, B. A. & Johnston, R. E. (1984) *Science* 225, 424–427 and Olmsted, R. A., Meyer, W. J. & Johnston, R. E. (1986) *Virology* 148, 1–10). The nucleotide sequence data suggest that a single mutation in the E2 gene is sufficient to cause these coordinate phenotypic changes. These findings identify a single locus in a Sindbis virus surface glycoprotein gene that determines both efficiency of interaction with cultured baby hamster kidney cells and degree of virulence in neonatal mice.**

Sindbis virus, the prototype member of the alphavirus genus (1), has been well studied with respect to pathogenesis in animals and replication in cell culture. A single plaque-forming unit of the virus inoculated into a newborn mouse is sufficient to cause acute encephalitis and death within 5 days (2, 3). A relatively simple structure (4) and high degree of virulence for newborn mice, as well as relatedness to important human and veterinary pathogens (1), make Sindbis virus infection of neonatal mice an appropriate system for the study of viral pathogenesis.

One approach to an understanding of the complex process of pathogenesis is to begin with a specific and quantifiable phenotype related to disease. We have taken this approach in a study of viral virulence, using the measurable properties of percent mortality and mean survival time of neonatal ICR-L$^+$ mice injected subcutaneously with Sindbis virus. Wild-type, virulent Sindbis virus infection of newborn mice by this route leads to 100% mortality and a mean survival time of 4–6 days. Mutant strains of Sindbis virus, developed by serial passage in baby hamster kidney (BHK) cells under stringent pressure for rapid growth, show reduced, or attenuated, virulence; mortality is always less than 100% and the mean survival time is 12–14 days (5). Comparison of attenuated and virulent viruses isolated during the passage series showed that attenuation *in vivo* was consistently correlated with accelerated penetration of BHK cells. Virulent revertants of the attenuated prototype strain SB-RL also reverted with respect to penetration rate. Rapid penetration was, thereby, identified as an *in vitro* marker for attenuation, and a mutant selected directly for accelerated penetration was attenuated (5). A second *in vitro* marker for attenuation was identified in studies with monoclonal antibodies raised to the attenuated SB-RL virus. Two of these antibodies, R6 and R13, which define epitopes on the E2 glycoprotein, preferentially neutralized the attenuated, fast-penetrating strains (5). A mutant of SB-RL selected for resistance to neutralization by one of these antibodies also lost the attenuation and accelerated penetration phenotypes of its SB-RL parent (6). Therefore, the phenotypes of virulence in newborn mice, penetration rate of BHK cells, and neutralization sensitivity to R6 and R13 are genetically linked; selection for any one property simultaneously selected for change in the other two.

Nucleotide sequence analysis of the glycoprotein genes of several virulent and attenuated strains has been performed. These strains include the wild-type, virulent parent virus, attenuated mutants either selected in the original passage series or selected specifically for fast penetration, and virulent revertants of the prototype attenuated strain. The results reported here identify a single mutation in the E2 gene that distinguishes fast-penetrating, attenuated strains from slow-penetrating, virulent strains. This mutation is probably responsible for the coordinate changes in these phenotypes, and thus represents evidence for the existence of a single locus that affects both growth rate *in vitro* and virulence *in vivo*.

MATERIALS AND METHODS

Virus Strains and Cell Culture. The prototype wild-type strain of Sindbis (SB), supplied originally by H. R. Bose (University of Texas), was derived from strain AR339 (7), is virulent in suckling mice, and was the parent virus for the prototype attenuated strain SB-RL (for SB-reduced latent period) and for a fast-penetrating mutant of SB, SB-FP (5). Virulent revertants of SB-RL were plaque-purified from brain tissue of moribund SB-RL-infected mice (5). SIN is the Sindbis strain studied in the laboratories of J. Dalrymple (U. S. Army Medical Research Institute for Infectious Diseases) and A. Schmaljohn (University of Maryland Medical School) and was obtained from A. Schmaljohn. An isolate of the heat-resistant (HR) strain of Sindbis virus was obtained

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.

¶Present address: Laboratory of Infectious Diseases, National Institutes of Health, Bethesda, MD 20892.

from D. T. Brown (University of Texas). Virus stocks were grown from plaque-purified isolates on BHK cell monolayers maintained in Eagle's minimal essential medium (MEM) containing 10% (vol/vol) donor calf serum and 10% (vol/vol) tryptose/phosphate broth. Virus stocks used for RNA extraction represented the third or fourth passage in BHK cells of the original isolates. All passages were tested for virulence phenotype in newborn mice.

Sequence Analysis of Viral Genomic RNA. Virions were purified as described (6). RNA was extracted from NaDodSO$_4$-dissociated virions with a mixture of phenol and chloroform (1:1, vol/vol), and then chloroform. The RNA was concentrated by precipitation with ethanol and quantitated by absorbance at 260 nm.

Oligodeoxynucleotide primers of 14 or 15 nucleotides were synthesized manually (New England Biolabs) or automatically (Applied Biosystems model 380A DNA Synthesizer). The primers used, named for the gene and in order from the 3′ end, were E11 (5′ CTGGTCGGATCATT 3′), E12 (5′ TCGCGGTGCTAAAG 3′), E13 (5′ CCGTATGAACAGTCC 3′), E16 (5′ CATCGCTCCATATTC 3′), E17 (5′ ATCTGCTGACAATTC 3′), E21 (5′ CATGGTCTCGGTGA 3′), E22 (5′ AGATGGTGTACACA 3′), E24 (5′ ATGTATTACATTCG 3′), E26 (5′ ATGATTCTTCCAGGT 3′), E27 (5′ ACCGTTACGCTGTC 3′), and E28 (5′ TGTAGCTAAGCCTT 3′). The five primers used to sequence the E1 glycoprotein gene were a gift from D. T. Brown (University of Texas). Primers were used to sequence virus RNA genomes with the dideoxynucleotide chain-termination method adapted for RNA sequencing (8–10). A "strong stop" artifact in the E1 gene was resolved with the method of Collins *et al.* (11) by using a 5′-end-labeled primer. Sequences obtained for Sindbis virus strains were compared using the SEQALIGN programs (12) on an IBM PC. The computer program used to generate hydropathicity plots was the kind gift of C. Loomis (Duke University).

RESULTS

Nucleotide sequence of the E1, E2, and E3 glycoprotein genes of our prototype virulent parent (SB) and attenuated (SB-RL) Sindbis strains was determined by the dideoxynucleotide chain-termination method using oligonucleotide primers and reverse transcriptase with viral RNA genomes as template. The E1 and E2 genes differed from the published sequence of the HR strain (13). The E1 genes showed two conservative amino acid substitutions (valine-72 and alanine-237), while the E2 genes included five nonconservative coding changes. The complete nucleotide sequence and deduced amino acid sequence of the E2 glycoprotein gene of the virulent SB strain is shown in Fig. 1. Nucleotides are numbered beginning at the 5′ end of the viral genome, and amino acids are numbered from the amino-terminal amino acid of the mature polypeptide, determined by analogy to the published sequence for the HR strain of Sindbis (13). A comparison of the nucleotide sequence of the E1, E2, and E3 genes of the prototype attenuated strain SB-RL with that of its SB parent revealed only a single missense mutation and a single silent change. A change from cytidine to adenosine at nucleotide 8972 resulted in a change from serine in SB to arginine in SB-RL at amino acid 114 of E2 (Figs. 2 and 3). The sequence of the region that includes base 8972 was determined for the SB-RL mutant in three separate experiments with two different preparations of RNA and was determined in two experiments for the SB parent.

At nucleotide 9779 the parent SB population had equal amounts of uridine and cytidine while the SB-RL population uniformly showed a uridine. The SB heterogeneity at this site occurs only at the nucleotide level, since both codons are translated to give threonine (Fig. 1). All of the derived strains discussed below show only a uridine at this position.

The sequence of the 6K genes of SB and SB-RL has been determined with the exception of a single ambiguous nucleotide. Within the known sequence, the SB and SB-RL 6K genes are identical (data not shown).

The mutation at nucleotide 8972 results in the only amino acid sequence difference between the mature glycoproteins of SB and SB-RL. In addition, the substitution of arginine for serine is a nonconservative one. Thus, this difference was a likely candidate for the mutation causing the attenuated, fast-penetrating phenotype of SB-RL. To confirm this, sequences between nucleotides 8800 and 9145 of three virulent, slow-penetrating and five attenuated, fast-penetrating plaque-purified isolates obtained during the selection for SB-RL (5) were determined. The virulent strains carried the SB codon for serine at position 114 of E2, while the attenuated strains showed the arginine substitution characteristic of SB-RL at that position. One of the attenuated strains was isolated from passage four, the first passage in which attenuated mutants were detected. The HR and SIN strains of Sindbis, both of which are virulent in newborn mice, were identical to SB in the region of amino acid 114 (ref. 14 and unpublished data). These results indicated a consistent correlation between the change to arginine at E2 amino acid 114 and the change to an attenuated, fast-penetrating phenotype.

Four virulent revertants, RLvr1-1, RLvr2-1, RLvr6-1, and RLvr6-3, were plaque-purified from brain tissue of moribund SB-RL-infected mice, two from one mouse and one from each of two other mice (5). All four show the same virulence, penetration, and neutralization phenotypes as the SB strain but retain the RNA synthesis phenotype of the SB-RL strain from which they were derived (5). Sequence analysis of the E2 and E1 genes of RLvr1-1 showed that it had regained the SB nucleotide at position 8972 and the serine at position 114 of E2 (Figs. 2 and 3). No other difference between the E2 and E1 genes of RLvr1-1 and SB-RL was found. Sequences of the three other virulent revertants were determined between nucleotides 8800 and 9145. All three also displayed the reversion at position 8972 to give a serine at E2 amino acid 114. Thus, phenotypic reversion is associated with same-site reversion to the original SB nucleotide at position 8972.

Additional evidence that the mutation at amino acid 114 of E2 of the SB-RL strain was responsible for its altered phenotypes was obtained from sequence analysis of an independently isolated fast-penetrating mutant of SB, SB-FP (5). This mutant strain also exhibits the reduced virulence and increased neutralization sensitivity of the SB-RL strain. The E2 gene of SB-FP carries the identical mutation at nucleotide 8972 found in SB-RL (Figs. 2 and 3). No other coding difference was found between the E1 and E2 genes of SB-FP and its parent SB.

The plot of local hydrophobicity and hydrophilicity values (15) for the predicted amino acid sequence of the E2 glycoprotein of SB is shown in Fig. 4*A*. The curve shows a hydrophobic region near the carboxyl end of the protein that includes residues known to span the lipid bilayer of the viral envelope (16). Another smaller hydrophobic region appears between residues 111 and 119, with serine 114 at the most hydrophobic point. Fig. 4*B* illustrates the reduction in local hydrophobicity in this region caused by the substitution of arginine for serine at position 114.

A comparison of the predicted amino acid sequence of the E2 glycoprotein of SB (Fig. 1) with the published sequence of the HR strain of Sindbis (13) showed that the amino-terminal serine of HR E2 has been replaced with an arginine in SB. This change may cause the processing cleavage of the SB PE2 polypeptide to occur between this arginine and the adjacent

```
1
Arg Val Thr Asp Asp Phe Thr Leu Thr Ser Pro Tyr Leu Gly Thr Cys Ser Tyr Cys His His Thr Glu Pro Cys Phe Ser Pro Val Lys
AGA GUC ACU GAC GAC UUU ACC CUG ACC AGC CCC UAC UUG GGC ACA UGC UCG UAC UGC CAC CAU ACU GAA CCG UGC UUC AGC CCU GUU AAG
8631

31
Ile Glu Gln Val Trp Asp Glu Ala Asp Asp Asn Thr Ile Arg Ile Gln Thr Ser Ala Gln Phe Gly Tyr Asp Gln Ser Gly Ala Ala Ser
AUC GAG CAG GUC UGG GAC GAA GCG GAC GAU AAC ACC AUA CGC AUA CAG ACU UCC GCC CAG UUU GGA UAC GAC CAA AGC GGA GCA GCA AGC
8721

61
Ala Asn Lys Tyr Arg Tyr Met Ser Leu Glu Gln Asp His Thr Val Lys Glu Gly Thr Met Asp Asp Ile Lys Ile Ser Thr Ser Gly Pro
GCA AAC AAG UAC CGC UAC AUG UCG CUU GAG CAG GAU CAC ACC GUU AAA GAA GGC ACC AUG GAU GAC AUC AAG AUU AGC ACC UCA GGA CCG
8811

91                                                                                                    *
Cys Arg Arg Leu Ser Tyr Lys Gly Tyr Phe Leu Leu Ala Lys Cys Pro Pro Gly Asp Ser Val Thr Val Ser Ile Val Ser Ser Asn Ser
UGU AGA AGG CUU AGC UAC AAA GGA UAC UUU CUC CUC GCA AAA UGC CCU CCA GGG GAC AGC GUA ACG GUU AGC AUA GUG AGU AGC AAC UCA
8901

121
Ala Thr Ser Cys Thr Leu Ala Arg Lys Ile Lys Pro Lys Phe Val Gly Arg Glu Lys Tyr Asp Leu Pro Pro Val His Gly Lys Lys Ile
GCA ACG UCA UGU ACA CUG GCC CGC AAG AUA AAA CCA AAA UUC GUG GGA CGG GAA AAA UAU GAU CUA CCU CCC GUU CAC GGU AAA AAA AUU
8991

151
Pro Cys Thr Val Tyr Asp Arg Leu Lys Glu Thr Thr Ala Gly Tyr Ile Thr Met His Arg Pro Gly Pro His Ala Tyr Thr Ser Tyr Leu
CCU UGC ACA GUG UAC GAC CGU CUG AAA GAA ACA ACU GCA GGC UAC AUC ACU AUG CAC AGG CCG GGA CCG CAC GCU UAU ACA UCC UAC CUG
9081

181
Glu Glu Ser Ser Gly Lys Val Tyr Ala Lys Pro Pro Ser Gly Lys Asn Ile Thr Tyr Glu Cys Lys Cys Gly Asp Tyr Lys Thr Gly Thr
GAA GAA UCA UCA GGG AAA GUU UAC GCA AAG CCG CCA UCU GGG AAG AAC AUU ACG UAU GAG UGC AAG UGC GGC GAC UAC AAG ACC GGA ACC
9171

211
Val Ser Thr Arg Thr Glu Ile Thr Gly Cys Thr Ala Ile Lys Gln Cys Val Ala Tyr Lys Ser Asp Gln Thr Lys Trp Val Phe Asn Ser
GUU UCG ACC CGC ACC GAA AUC ACU GGU UGC ACC GCC AUC AAG CAG UGC GUC GCC UAU AAG AGC GAC CAA ACG AAG UGG GUC UUC AAC UCA
9261

241
Pro Asp Leu Ile Arg His Asp Asp His Thr Ala Gln Gly Lys Leu His Leu Pro Phe Lys Leu Ile Pro Ser Thr Cys Met Val Pro Val
CCG GAC UUG AUC AGA CAU GAC GAC CAC ACG GCC CAA GGG AAA UUG CAU UUG CCU UUC AAG UUG AUC CCG AGU ACC UGC AUG GUC CCU GUU
9351

271
Ala His Ala Pro Asn Val Ile His Gly Phe Lys His Ile Ser Leu Gln Leu Asp Thr Asp His Leu Thr Leu Leu Thr Thr Arg Arg Leu
GCC CAC GCG CCG AAU GUA AUA CAU GGC UUU AAA CAC AUC AGC CUC CAA UUA GAU ACA GAC CAC UUG ACA UUG CUC ACC ACC AGG AGA CUA
9441

301
Gly Ala Asn Pro Glu Pro Thr Thr Glu Trp Ile Val Gly Lys Thr Val Arg Asn Phe Thr Val Asp Arg Asp Gly Leu Glu Tyr Ile Trp
GGG GCA AAC CCG GAA CCA ACC ACU GAA UGG AUC GUC GGA AAG ACG GUC AGA AAC UUC ACC GUC GAC CGA GAU GGC CUG GAA UAC AUA UGG
9531

331
Gly Asn His Glu Pro Val Arg Val Tyr Ala Gln Glu Ser Ala Pro Gly Asp Pro His Gly Trp Pro His Glu Ile Val Gln His Tyr Tyr
GGA AAU CAU GAG CCA GUG AGG GUC UAU GCC CAA GAG UCA GCA CCA GGA GAC CCU CAC GGA UGG CCA CAC GAA AUA GUA CAG CAU UAC UAC
9621

361
His Arg His Pro Val Tyr Thr Ile Leu Ala Val Ala Ser Ala Thr Val Ala Met Met Ile Gly Val Thr Val Ala Val Leu Cys Ala Cys
CAU CGC CAU CCU GUG UAC ACC AUC UUA GCC GUC GCA UCA GCU ACC GUG GCG AUG AUG AUU GGC GUA ACC GUU GCA GUG UUA UGU GCC UGU
9711

391
Lys Ala Arg Arg Glu Cys Leu Thr Pro Tyr Ala Leu Ala Pro Asn Ala Val Ile Pro Thr Ser Leu Ala Leu Leu Cys Cys Val Arg Ser
AAA GCG CGC CGU GAG UGC CUG ACG CCA UAC GCC CUG GCC CCA AAC GCC GUA AUC CCA ACU UCG CUG GCA CUC UUG UGC UGC GUU AGG UCG
9801

421
Ala Asn Ala
GCC AAU GCU
9891
```

FIG. 1. The nucleotide sequence of the E2 glycoprotein gene of Sindbis virus. The nucleotide sequence was determined using the dideoxynucleotide chain-termination method with reverse transcriptase directly on purified genome RNA. The amino acid sequence was deduced by computer translation. Nucleotides are numbered beginning at the 5′ terminus of the viral genome, and amino acids are numbered from the amino-terminal amino acid of the mature polypeptide, by analogy to the sequence of the HR strain of Sindbis (13). Differences between this sequence and that of the HR strain are underlined. These changes result in five nonconservative amino acid substitutions in E2 (arginine-1. threonine-3, glutamic acid-23, glutamic acid-70, and glycine-172). The amino acid of E2 altered in SB-RL, position 114, is indicated (*).

valine, giving SB a mature E2 protein that is shorter than the E2 of HR by one residue.

DISCUSSION

The development of a group of closely related virulent and attenuated Sindbis virus isolates presents a unique opportunity to study virus adaptation to growth in an unnatural host as it relates to attenuation of virulence. The consistent correlation shown by these isolates between attenuation *in vivo* and accelerated penetration *in vitro* suggested that alteration at a single locus might be responsible for these linked phenotypic changes. We have presented nucleotide sequence data that support this hypothesis. Two different selective protocols for virus growth—one requiring rapid replication in BHK cells, the other requiring rapid penetration of BHK cells—each led to isolation of a spontaneous fast-penetrating, attenuated mutant with an arginine replacing the serine at position 114 of the mature E2 glycoprotein. Except for this one amino acid substitution, the E1, E2, and E3 glycoproteins of the mutants were identical to those of the wild-type virulent parental virus. Slow-penetrating, virulent revertants all had regained the serine at this position. Our findings suggest that a single mutation in the E2 gene is

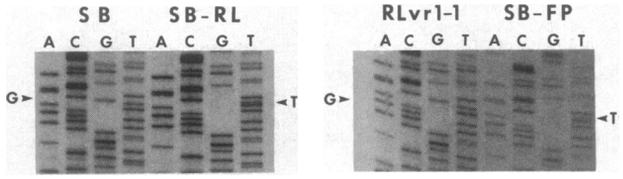

FIG. 2. Identification of base substitutions in the E2 gene of Sindbis virus mutants. ^{32}P-labeled cDNAs synthesized in the presence of a specific dideoxynucleotide (indicated by A, C, G, or T) were resolved by electrophoresis in 8% sequencing gels and visualized by autoradiography. cDNAs span the sequence 5′ ACGTTGCTGAGTTGCTACTCACTATGCTAACCGTTACGCTGTCCCC 3′ (cDNA sense, complementary to nucleotides 8952 through 8997 of the viral genome). Nucleotide substitution for each virus mutant is indicated (◄).

sufficient to alter virulence and penetration properties of Sindbis virus. It is possible that mutation of analogous genes of other viruses was one of the steps in generation of avirulent virus mutants by passage in cells other than the normal target cell (17–21).

The data presented here, in combination with previous studies of the phenotypes of these Sindbis virus mutants (5, 6), form the basis for a working hypothesis to guide our study of Sindbis virus glycoprotein structure and function. The hypothesis includes the following points.

The mutation at amino acid 114 of E2 was the only event required to cause the coordinate changes in virulence, penetration, and neutralization phenotypes shown by SB-RL. Although the capsid protein and nonstructural protein genes of the attenuated and virulent prototypes have not been sequenced yet, we believe it is unlikely that a second mutation was needed to produce the attenuation of virulence shown by the mutant. Given a putative second mutation, phenotypic reversion to virulence would require both the observed same site reversion at E2 position 114 and closely timed reversion or suppression of a second mutation in another gene. In the case of rabies virus, it has been suggested that a significant effect on pathogenicity is caused by the change of a single specific amino acid residue (22). Attenuated antigenic variants of rabies virus consistently showed a substitution for the arginine at position 333 of the surface glycoprotein, while variants that retained virulence also retained the arginine at position 333.

The serine at position 114 of the E2 glycoprotein is an important structural determinant of the virion spike. A plot of local hydropathicity based on the deduced amino acid sequence of E2 places serine 114 at the most hydrophobic point of a short hydrophobic region. Therefore, serine 114 may be an internal residue that interacts with side chains of distal amino acids in the native protein. As would be expected for an important structural determinant, this region is conserved among related viruses. Both Semliki Forest virus and Ross River virus E2 glycoproteins contain a serine at position 114 that occurs at the most hydrophobic point of a short, relatively hydrophobic region (23–25). In Semliki Forest virus the flanking valine-113 and isoleucine-115 also are conserved. In Ross River virus valine-113 is conserved while isoleucine-115 is replaced with phenylalanine-115. The correspondence in this region among the three viruses, whose E2 genes show considerable divergence at the amino acid sequence level (25), suggests an important role for these residues.

A change from serine to arginine at position 114 is sufficient to alter the conformation and/or stability of the virion surface structures. Monoclonal antibodies R6 and R13 bind to a site on the E2 glycoproteins of both SB and SB-RL. However, this binding leads to efficient neutralization only when an arginine is present at position 114, as in SB-RL and SB-FP (6). Also, the half-life of virus infectivity during incubation at 51°C or treatment with 15 mM dithiothreitol is significantly shorter for mutants carrying an arginine at position 114 (S. Gidwitz and R.E.J., unpublished results). The substitution of arginine for serine would reduce the hydrophobicity of this region, would lead to a different pattern of hydrogen or electrostatic bonds, and would necessitate accommodation of the much bulkier arginine R group. Any of these changes might alter the folding and/or stability of the E2 protein in the attenuated mutant virion (26).

Our results indicate that a domain of the E2 protein plays a major role in penetration. We have identified a site in the E2 glycoprotein gene at which mutation causes a significant increase in the penetration rate of BHK cells. The mutation at position 114 of E2 may affect penetration indirectly by destabilizing the conformation of the E1/E2 heterodimer, or alternatively, it may alter a domain that is directly involved in penetration. In studies of mutant influenza virus hemagglutinins, both types of mutation have been identified (27, 28). Single amino acid substitutions that affect fusion-related properties of the viral hemagglutinin occur not only in the conserved hydrophobic region believed to be directly involved in fusion, but also in other regions of the protein. This second type of mutation appears to destabilize the conformation of the native hemagglutinin, thus reducing the energy barrier to a conformational change required for fusion.

Although our studies have pinpointed the residue at position 114 of E2 as one important determinant of virulence and penetration, different selection procedures will reveal other sites at which mutation affects these phenotypes. We have

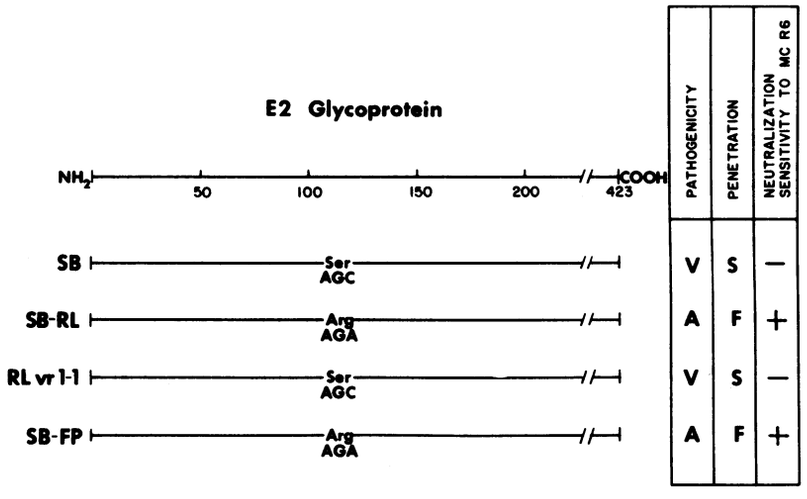

FIG. 3. Map of coding changes in E2 genes of Sindbis virus mutants. Base substitution and the deduced amino acid coding change are listed for each mutant. V, virulent; A, attenuated; S, slow; F, fast; +, efficiently neutralized by R6; −, inefficiently neutralized by R6.

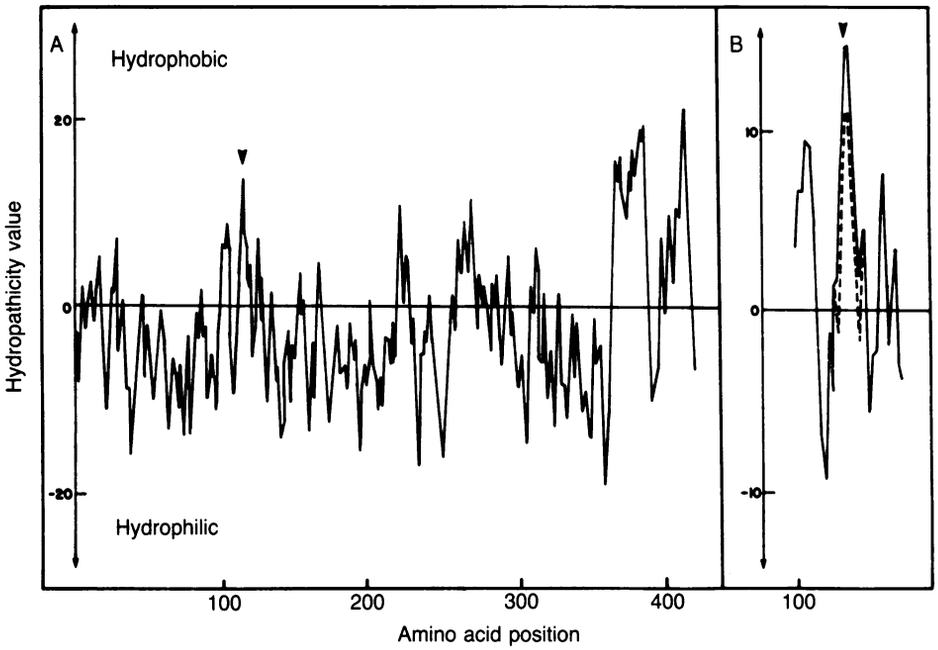

FIG. 4. Hydropathicity plot derived from the amino acid sequence of the Sindbis virus E2 protein. A moving sum of hydropathicity values (15) calculated for groups of seven residues is plotted versus position in the sequence, beginning at the amino terminus. (*A*) Values for the E2 protein of Sindbis virus. Position 114 is indicated (▼). (*B*) Expanded scale plot of region between residues 100 and 125. Solid line, SB; dashed line, SB-RL.

identified such a site in studies of an antigenic variant of SB-RL that is resistant to neutralization by monoclonal antibody R6, and has lost the attenuation and rapid-penetration properties of its SB-RL parent (6). This variant, which will be described in detail elsewhere, retains the parental arginine at position 114 of E2, but carries an additional mutation in E2 that acts as a second-site reversion or suppressor mutation. In this context, it is important to note that other studies of alphavirus pathogenicity have shown that mutations affecting the E1 protein (29) or proteins involved in RNA synthesis (30) can lead to a measurable reduction of virulence.

In summary, a study of closely related attenuated and virulent strains of Sindbis virus has identified a site in the E2 glycoprotein gene at which mutation leads simultaneously to reduced virulence in the neonatal mouse and accelerated penetration of cultured cells. This result defines a domain in the E2 glycoprotein of Sindbis virus that plays a key structural role in maintaining the multiple biological functions of the virion spike.

We thank our colleagues for their help and support, especially David Pence for purifying and testing virus stocks, Dennis Brown for synthesizing the E1 oligonucleotide primers, and Richard Allison for sharing his expertise. This work was supported by Grant A122186 from the National Institutes of Health and Grant 85-G-00703 from the North Carolina Biotechnology Center. This is paper 10357 of the Journal Series of the North Carolina Agricultural Research Service.

1. Andrewes, C. H., Pareira, H. G. & Wildy, P. (1972) *Viruses of Vertebrates* (Bailliere Tindall, London), 4th Ed., p. 76.
2. Reinarz, A. B. G., Broome, M. G. & Sagik, B. P. (1971) *Infect. Immunol.* **3**, 268–273.
3. Johnson, R. T., McFarland, H. F. & Levy, S. E. (1972) *J. Infect. Dis.* **125**, 257–262.
4. Rice, C. M. & Strauss, J. H. (1982) *J. Mol. Biol.* **154**, 325–348.
5. Olmsted, R. A., Baric, R. S., Sawyer, B. A. & Johnston, R. E. (1984) *Science* **225**, 424–427.
6. Olmsted, R. A., Meyer, W. J. & Johnston, R. E. (1986) *Virology* **148**, 1–10.
7. Taylor, R. M., Hurlbut, H. S., Work, T. S., Kingston, J. R. & Frothingham, T. E. (1955) *Am. J. Trop. Med. Hyg.* **4**, 844–862.
8. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
9. Zimmern, D. & Kaesberg, P. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4257–4261.
10. Ahlquist, P., Dasgupta, R. & Kaesberg, P. (1981) *Cell* **23**, 183–189.
11. Collins, P. L., Dickens, L. E., Buckler-White, A., Olmsted, R. A., Spriggs, M. K., Camargo, E. & Coelingh, K. V. W. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4594–4598.
12. Johnston, R. E., MacKenzie, J. M. & Dougherty, W. G. (1986) *Nucleic Acids Res.* **14**, 517–527.
13. Strauss, E. G., Rice, C. M. & Strauss, J. H. (1984) *Virology* **133**, 92–110.
14. Schmaljohn, A. L., Johnson, E. D., Dalrymple, J. M. & Cole, G. A. (1982) *Nature (London)* **297**, 70–72.
15. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132.
16. Rice, C. M., Bell, J. R., Hunkapiller, M. W., Strauss, E. G. & Strauss, J. H. (1982) *J. Mol. Biol.* **154**, 355–378.
17. Waterson, A. P. & Wilkerson, L. (1978) *An Introduction to the History of Virology* (Cambridge Univ. Press, Cambridge, UK), pp. 52–66.
18. Theiler, M. & Smith, H. H. (1937) *J. Exp. Med.* **65**, 767–786.
19. Sabin, A. B., Hennessen, W. A. & Winsser, J. (1954) *J. Exp. Med.* **99**, 551–576.
20. Feinstone, S., Daemer, R. J., Gust, I. D. & Purcell, R. H. (1983) *Dev. Biol. Stand.* **54**, 429–432.
21. Hilleman, M. R., Provost, P. J., Buynak, R. B. & McLean, A. A. (1983) *Dev. Biol. Stand.* **54**, 433–440.
22. Seif, I., Coulon, P., Rollin, P. E. & Flamand, A. (1985) *J. Virol.* **53**, 926–934.
23. Garoff, H., Frischauf, A.-M., Simons, K., Lehrach, H. & Delius, H. (1980) *Nature (London)* **288**, 236–241.
24. Rice, C. M. & Strauss, J. H. (1981) *J. Mol. Biol.* **150**, 315–340.
25. Dalgarno, L., Rice, C. M. & Strauss, J. H. (1983) *Virology* **129**, 170–187.
26. Chothia, C. (1984) *Annu. Rev. Biochem.* **53**, 537–572.
27. Daniels, R. S., Downie, J. C., Hay, A. J., Knossow, M., Skehel, J. J., Wang, M. L. & Wiley, D. C. (1985) *Cell* **40**, 431–439.
28. Doms, R. W., Gething, M.-J., Henneberry, J., White, J. & Helenius, A. (1986) *J. Virol.* **57**, 603–613.
29. Emini, E. A. & Wiebe, M. E. (1981) *Virology* **110**, 185–196.
30. Barrett, P. N. & Atkins, G. J. (1979) *Infect. Immunol.* **26**, 848–852.