JOURNAL OF VIROLOGY, Nov. 1991, p. 6358–6361
0022-538X/91/116358-04$02.00/0

Vol. 65, No. 11

Mutational Analysis of a Virulence Locus in the E2 Glycoprotein Gene of Sindbis Virus

JOHN M. POLO† AND ROBERT E. JOHNSTON*

Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7290

Received 7 May 1991/Accepted 31 July 1991

The substitution of arginine for serine at position 114 of glycoprotein E2 in several biological and recombinant Sindbis virus mutants was shown previously to attenuate the virus for neonatal mice and also to accelerate virus penetration into BHK cells. To further examine the genetically linked effects on both virus penetration into cultured cells and pathogenesis in vivo, mutants containing each of 16 different amino acid coding changes at this position were generated by site-directed mutagenesis of a full-length cDNA clone of the Sindbis virus genome. Viable virus was recovered following transfection of RNA transcripts from 14 of the clones. Phenotypic analysis of these virus mutants revealed that specific amino acid residues affected either the pathogenesis or penetration phenotype independently or both phenotypes simultaneously. Thus, both the position of a mutation within the E2 sequence and the particular amino acid encoded at that position are important determinants of the mutant phenotypes.

Sindbis virus (SB) particles are composed of an RNA-containing icosahedral nucleocapsid surrounded by a host-derived lipid envelope. Protruding from the envelope are two viral glycoproteins, E1 and E2, which associate in groups of three heterodimeric spikes (8). The glycoproteins participate in many biological activities, including virus attachment and penetration in cell culture (4, 6, 7, 13, 14) and pathogenesis in animals (5, 12, 15–17, 19, 20). These proteins are also primary targets of recognition by elements of the host immune system. In some cases, neutralizing antigenic sites on E2 are related genetically to the virulence phenotype of the virus (9, 14, 15, 20).

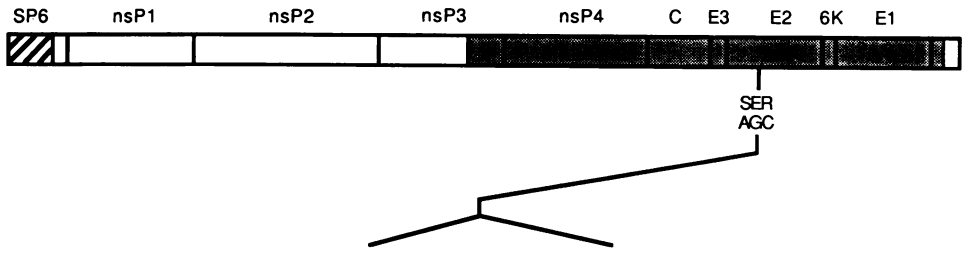
A virulence locus at codon 114 of the E2 glycoprotein gene of SB was identified previously. Mutation at this locus from the wild-type serine codon in SB to an arginine codon in the SB mutant SB-RL (2, 3) altered both the in vivo pathogenesis phenotype of the virus and its ability to penetrate cultured cells. SB-RL was attenuated when inoculated into neonatal mice and displayed an accelerated rate of penetration into BHK cells. Wild-type SB was virulent and more slowly penetrating (5, 13, 14). The effects induced by this mutation in SB-RL were confirmed subsequently by using a full-length cDNA clone of SB from which infectious RNA could be transcribed in vitro (18). Two strains of virus which were isogenic except for E2 codon 114 were constructed. These virus recombinants exhibited the same virulence and penetration phenotypes as the analogous biological strains (16).

E2 residue 114 is located within a hydrophobic sequence of 18 uncharged and nonpolar amino acids (5). One would predict that substitution of a positively charged arginine residue within this region could produce a structural alteration sufficient to account for changes in the pathogenesis and penetration phenotypes. However, this substitution had no discernible effects on other growth properties of the virus, suggesting that specific, localized changes in protein structure and function had occurred. In this study, we have explored functional changes associated with the substitution of other amino acids at E2 residue 114.

An M13-derived mutagenesis vector, mp18SBGP, containing E1 and E2 glycoprotein gene sequences (*Afl*II [nucleotide {nt} 7986] to *Sac*II [nt 11484] from wild-type SB AR339), was used as template for mutagenesis by the technique of Kunkel et al. (10). This construct included an engineered, silent mutation which eliminated a *Pst*I restriction site (nt 9119) within the E2 gene as a convenient marker for sequence replacements into a full-length SB cDNA clone (11, 18). Specific oligonucleotide primers which had been "doped" within the sequences corresponding to E2 codon 114 were synthesized: 5′-GCGTAACGGTTNNCATAGTG AGTA-3′, 5′-GCGTAACGGTTN(A/C/T)CATAGTGAGTA-3′, and 5′-GCGTAACGGTT(G/A)AAATAGTGAGTAG-3′. These were phosphorylated and annealed to single-stranded uracil-containing mp18SBGP DNA that was obtained by propagation in the Dut$^-$ Ung$^-$ *Escherichia coli* strain CJ236 (10). The DNA was converted into double-stranded replicative-form molecules by using Sequenase 2.0 (United States Biochemical Corp.) as described previously (17) and transformed into *E. coli* JM101 (Dut$^+$ Ung$^+$). Individual phage clones were screened for the presence of mutations by sequence analysis using the dideoxy chain-termination method (1) and an E2-specific oligonucleotide, 114SEQ (5′-TAGCTACAAAGGATAC-3′, nt 8912 to 8927), as primer. Mutations which encoded each of 15 different amino acids at E2 codon 114 were identified and included residues of each chemical class: polar uncharged, nonpolar uncharged, negatively charged, and positively charged (Fig. 1).

Replicative-form DNA was isolated from clones containing mutations at codon 114, and restriction fragments encompassing E2 codon 114 (*Stu*I [nt 8571] to *Bss*HII [nt 9804]) were generated. These were used to replace the analogous fragment in the full-length SB cDNA clone pTR4000 prior to transcription in vitro. Plasmid pTR4000 was derived from the full-length SB clone Toto1101 (18) by replacement of all structural gene sequences with the corresponding cDNA sequences from a wild-type SB (AR339) clone (Fig. 1). The

* Corresponding author.
† Present address: Department of Microbiology, University of Southern California, School of Medicine, Los Angeles, CA 90033.



MUTANT	CODON	AMINO ACID	VIABILITY
TR4500	AGA	Arginine	+
TR4114-A	GCC	Alanine	+
TR4114-I	ATC	Isoleucine	+
TR4114-L	CTC	Leucine	+
TR4114-F	TTC	Phenylalanine	+
TR4114-P	CCC	Proline	-
TR4114-V	GTC	Valine	+
TR4114-N	AAC	Asparagine	+
TR4114-G	GGC	Glycine	+
TR4114-T	ACC	Threonine	+
TR4114-Y	TAC	Tyrosine	+
TR4114-D	GAC	Aspartic Acid	+
TR4114-E	GAA	Glutamic Acid	+
TR4114-R	CGC	Arginine	+
TR4114-H	CAC	Histidine	-
TR4114-K	AAA	Lysine	+

FIG. 1. Mutagenesis of E2 codon 114 in the SB recombinant TR4000. The schematic shows the SB genome organization (including genes for nonstructural proteins [nsP1 to nsP4], capsid [C], the two envelope glycoproteins [E2 and E1], and their signal peptides [E3 and 6K]) and the derivation of cDNA sequences in the full-length clone pTR4000. Open areas represent Toto1101 sequences (18), shaded areas represent wild-type SB sequences (16) (*Spe*I site [nt 5262] through *Sac*II site [nt 11484]), and the hatched area represents the bacteriophage SP6 promoter. Mutations generated in the wild-type serine codon and the amino acid substitution encoded are indicated. The viability of the mutants was determined following the transfection of RNA transcripts synthesized in vitro into BHK cells and is represented as follows: +, establishment of a productive infection with cytopathic effect at both 37 and 30°C; −, no productive infection at either temperature.

fidelity of replacement with fragments containing mutated E2 114 codons was confirmed initially by restriction analysis (loss of the *Pst*I site at nt 9119) and subsequently by direct sequence analysis of RNA transcripts synthesized in vitro. Sequence analysis was performed by the dideoxy chain-termination method as adapted for RNA (21) and a virus complementary-sense oligonucleotide primer, E26.5 (5′-TTATCTTGCGGGCCA-3′, nt 9007 to 9021). Full-length pTR4000 clones that contained an altered E2 114 codon were designated the pTR4114 class and identified by the amino acid encoded at position 114 (Fig. 1). Plasmids were linearized with *Xho*I and transcribed in vitro by using SP6 RNA polymerase as described previously (16, 18). RNA transcripts were transfected into BHK cells, using Lipofectin (BRL) as described by Polo and Johnston (17), except that components were scaled down fivefold and transfections were done in 2 wells of a 24-well tissue culture plate. In the cases of isoleucine, threonine, phenylalanine, arginine, lysine, and tryptophan, two independent mutant constructs were examined; one representative of each of the remaining amino acid substitutions was found in the mutagenized phage stocks.

The effect of each mutation on virus viability and growth was determined. BHK monolayers were transfected in duplicate, both at 30°C and at 37°C, to identify potential temperature-sensitive phenotypes. High specific infectivities (comparable to that of TR4000) and cytopathic effect resulted from all but two mutant clones (TR4114-P and TR4114-H) following transfection at either 30 or 37°C (Fig. 1). For the remaining mutants, no significant differences in the onset of cytopathic effect, the resulting virus titers, or the specific infectivities of radiolabeled RNA transcripts were observed (18; also data not shown). Thus, it was unlikely that virus populations containing reversions or compensatory mutations were being selected. The viability of mutant viruses with such a diversity of amino acid substitutions (with the probable exceptions of histidine and proline) suggested a high degree of flexibility in glycoprotein E2 in the vicinity of residue 114.

The virulence of each mutant virus strain was tested by subcutaneous (s.c.) inoculation of 100 PFU into 1-day-old mice (Table 1). The results were compared with those for TR4000, the virulent parental strain (E2 Ser-114) derived from unsubstituted pTR4000, and with those for TR4500, an attenuated derivative of TR4000 which corresponds to the biological mutant SB-RL (E2 Arg-114). TR4000 produced mortality rates of 100% with average survival times (AST) of less than 7 days. In contrast, TR4500, with arginine at E2 114, produced mortality rates of less than 20% with extended survival times. For the experiments described below, mutants characterized by a mortality rate of less than 100% or significantly extended AST or both were considered attenuated. Animals which survived inoculation of each attenuated mutant also survived challenge at 3 to 4 weeks of age with S.A.AR86, an adult neurovirulent strain of SB (19), confirming that they had been previously infected and had

TABLE 1. Virulence of SB E2 114 mutants

Recombinant virus strain	Amino acid encoded at E2 114	Virulence for neonatal mice[a] inoculated with:			
		100 PFU		10 PFU	
		% Mortality (no. dead/no. tested)[b]	AST (days)[c]	% Mortality (no. dead/no. tested)[b]	AST (days)[c]
TR4000	Ser	100 (22/22)	5.4 ± 1.1	100 (10/10)	4.5 ± 0.6
TR4500	Arg	13 (3/23)	9.7 ± 3.2[d]		ND[e]
TR4114-A	Ala	100 (10/10)	4.2 ± 0.8	100 (10/10)	5.3 ± 1.2
TR4114-I	Ile	100 (20/20)	6.3 ± 2.2	100 (9/9)	5.6 ± 1.0
TR4114-N	Asn	100 (11/11)	5.1 ± 1.0	100 (19/19)	4.5 ± 1.4
TR4114-G	Gly	100 (10/10)	6.7 ± 2.4	100 (10/10)	6.3 ± 1.3
TR4114-D	Asp	100 (10/10)	4.7 ± 0.8	100 (9/9)	6.0 ± 1.0
TR4114-E	Glu	100 (9/9)	5.1 ± 1.4	100 (9/9)	5.8 ± 1.0
TR4114-T	Thr	100 (20/20)	6.0 ± 1.3	90 (9/10)	7.4 ± 1.4[d]
TR4114-L	Leu	100 (10/10)	5.9 ± 1.4	60 (6/10)	8.3 ± 1.2[d]
TR4114-V	Val	100 (11/11)	6.7 ± 1.9	57 (4/7)	7.3 ± 0.5[d]
TR4114-Y	Tyr	90 (9/10)	7.7 ± 1.0	13 (2/16)	9.0 ± 1.4[d]
TR4114-F	Phe	85 (28/33)	7.5 ± 1.4	80 (16/20)	7.3 ± 1.3[d]
TR4114-R	Arg	18 (4/22)	9.8 ± 1.0[d]		ND
TR4114-K	Lys	5 (2/40)	11.5 ± 2.1[d]		ND

[a] Virulence following s.c. inoculation of 1-day-old CD-1 mice.
[b] Percent mortality was calculated after a 14-day observation period.
[c] AST ± standard deviation was calculated for animals that died.
[d] AST value which differs significantly from that for TR4000 (Student's t test, $P < 0.001$).
[e] ND, not determined.

mounted a protective immune response. In one case (TR4114-V at 10 PFU), one of three animals surviving the original infection succumbed to the S.A.AR86 challenge, and this individual was not considered in the subsequent analysis.

The virus strains generated by mutagenesis of E2 codon 114 displayed varied effects on virulence and appeared to segregate into several broad groups. The first group, which included the largest proportion of mutants, remained virulent for neonatal mice, with AST similar to those found with TR4000. It has not been established whether the underlying pathogenesis which leads to death is identical for each of these mutants. Another group, which included both the lysine and arginine mutants, produced high degrees of attenuation, similar to that found with TR4500 and roughly equivalent to each other. These mutants were characterized by 50% lethal dose values greater than 10^6 PFU by the s.c. route of inoculation (data not shown). A third group, which included the mutants with phenylalanine and tyrosine residues, appeared to produce an intermediate level of attenuation. The virulence of these mutants was reduced significantly compared with that of the parental virus TR4000; however, the degree of attenuation was substantially less than that produced by the positively charged residues. In addition, three other residues, including threonine, leucine, and valine, exhibited minimal attenuation at small virus inocula. Of the virus mutants which were attenuated by the s.c. route of inoculation, only the lysine and arginine mutants were attenuated following intracerebral inoculation (data not shown). The variations in virulence characteristic of the mutants likely reflect different physical effects exerted on the hydrophobic pocket encompassing residue 114. The presence of a strong positive charge or a large ring structure at this position appears sufficient to induce a structural change leading to attenuation, with positive charge possibly more significant than size and shape constraints.

Because the amino acid substitutions with large ring structures in their side chains constituted one group of attenuating mutations, one would predict that if the resulting virus was viable, a tryptophan substitution might also exhibit a similar attenuated phenotype. In order to explore this possibility, a specific oligonucleotide primer, 5′-GCGTAAC GGTTTGGATAGTGAGTAG-3′, which specifies tryptophan at E2 114, was synthesized and used in mutagenesis reactions as previously described. RNA transcribed in vitro from these templates was shown to produce viable infectious virus as evidenced by times until onset of cytopathic effect, virus titers, and specific infectivities of the RNA transcripts. Following s.c. inoculation of the tryptophan mutant virus into 1-day-old mice, the mortality rates were 97% (38 of 39 mice; AST = 5.9 ± 1.4 days) at a dose of 100 PFU and 33% (3 of 9 mice; AST = 8.3 ± 2.1 days) at a dose of 10 PFU (data not shown). These results were comparable to those of the phenylalanine and tyrosine mutants and consistent with the notion that large ring structures substituted at E2 114 produced an attenuated phenotype.

Previous studies demonstrated that the substitution of arginine for serine at E2 residue 114 affected not only the virulence phenotype but also virus penetration into BHK cells (5, 13, 16). To examine further the relationships between these two phenotypes, we assessed the penetration characteristics of each E2 114 virus mutant. Penetration was defined as the time-dependent acquisition of resistance to neutralization by polyclonal anti-SB antibody and was determined by methods described previously as the percentage of input virus penetrated in a 15-min period (14, 15). The E2 114 mutant viruses, regardless of their virulence in vivo, displayed penetration values ranging from ones similar to that of the slowly penetrating parent TR4000 to ones that were as high as that of the rapidly penetrating mutant TR4500 (Table 2). The tryptophan mutant generated in subsequent experiments had penetration values within the range of those of the phenylalanine and tyrosine mutants, having a penetration rate of 30.5 ± 12.6% (data not shown).

Previous experiments with neutralization escape mutants

TABLE 2. Comparative penetration of SB E2 114 mutants

Recombinant virus strain	Level of virulence[a]	% Penetration[b] ± SD
TR4000	Vir	21.9 ± 11.0
TR4114-N	Vir	25.7 ± 3.8
TR4114-F	Att	32.3 ± 10.5
TR4114-L	Vir	33.8 ± 13.3
TR4114-E	Vir	34.5 ± 11.2
TR4114-D	Vir	35.7 ± 15.9
TR4114-A	Vir	37.7 ± 11.5
TR4114-Y	Att	38.3 ± 6.4[c]
TR4114-V	Vir	38.5 ± 20.2[c]
TR4114-G	Vir	49.0 ± 17.8[c]
TR4500	Att	53.0 ± 7.0[c]
TR4114-T	Vir	53.3 ± 9.1[c]
TR4114-I	Vir	55.7 ± 17.7[c]
TR4114-R	Att	59.3 ± 8.1[c]
TR4114-K	Att	63.0 ± 15.4[c]

[a] Virulence following s.c. inoculation of 100 PFU into 1-day-old CD-1 mice (Table 1). Vir, virulent (i.e., 100% mortality at this dose); Att, attenuated (i.e., less than 100% mortality at this dose).

[b] Values represent mean percent plaque survival after addition of anti-SB antibody and complement 15 min postinfection, as previously described (16). $n = 13$ for TR4000 and TR4500; 4 for Leu, Glu, Val, and Gly mutants; and 3 for all others.

[c] Value which differs significantly from that for TR4000 (Student's t test, $P < 0.01$).

selected with monoclonal antibodies to the E2-c antigenic site showed that attenuation and a rapid penetration phenotype were not invariably linked (14, 15). The results of the experiments reported here clearly demonstrated that the attenuation and penetration phenotypes, linked genetically by the E2 arginine 114 substitution in SB-RL, depended on the particular amino acid encoded at the E2 114 locus. Other amino acid substitutions at E2 residue 114 altered the pathogenesis and penetration phenotypes independently. These data are consistent with the hypothesis that in the vicinity of residue 114, glycoprotein E2 contains both a penetration domain and a domain affecting pathogenesis. The relationship between the penetration phenotype in cell culture and the in vivo phenotype of attenuation is not likely to be one of cause and effect. Rather, the relationship is one in which certain substitutions at a residue common to both domains can affect either phenotype independently or both phenotypes simultaneously, as determined by the structural effects of specific amino acids. In this context, mutants having tyrosine, phenylalanine, or tryptophan at E2 114 constitute a group distinguished from lysine and arginine mutants on the basis of the phenotypic effects of these substitutions. It is likely that the phenotypic differences displayed by each category of substitution are the result of a distinctive alteration in the structure of the glycoprotein spike.

We acknowledge the excellent technical assistance of Katherine Elder.

This work was supported by PHS grants AI22186 and NS26681.

REFERENCES

1. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1987. Current protocols in molecular biology. Wiley, New York.
2. **Baric, R. S., D. B. Moore, and R. E. Johnston.** 1980. In vitro selection of an attenuated variant of Sindbis virus. ICN-UCLA Symp. Mol. Cell. Biol. **18**:685–694.
3. **Baric, R. S., D. W. Trent, and R. E. Johnston.** 1981. A Sindbis virus variant with a cell-determined latent period. Virology **110**:237–242.
4. **Bose, H. R., and B. P. Sagik.** 1970. The virus envelope in cell attachment. J. Gen. Virol. **9**:159–161.
5. **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 of Sindbis virus affects penetration rate in cell culture and virulence in neonatal mice. Proc. Natl. Acad. Sci. USA **83**:6771–6775.
6. **Flynn, D. C., W. J. Meyer, J. M. Mackenzie, Jr., and R. E. Johnston.** 1990. A conformational change in Sindbis virus glycoproteins E1 and E2 is detected at the plasma membrane as a consequence of early virus-cell interaction. J. Virol. **64**:3643–3653.
7. **Flynn, D. C., R. A. Olmsted, J. M. Mackenzie, Jr., and R. E. Johnston.** 1988. Antibody-mediated activation of Sindbis virus. Virology **166**:82–90.
8. **Harrison, S. C.** 1986. Alphavirus structure, p. 21–34. *In* S. Schlesinger and M. J. Schlesinger (ed.), The Togaviridae and Flaviviridae. Plenum Press, New York.
9. **Johnston, R. E., N. L. Davis, D. F. Pence, S. Gidwitz, and F. J. Fuller.** 1987. Nucleic acid sequence analysis of Sindbis pathogenesis and penetration mutants, p. 467–476. *In* M. A. Brinton and R. L. Rueckert (ed.), Positive strand RNA viruses. Alan R. Liss, Inc., New York.
10. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
11. **Lin, S.-C., J. M. Polo, and R. E. Johnston.** Unpublished data.
12. **Lustig, S., A. C. Jackson, C. S. Hahn, D. E. Griffin, E. G. Strauss, and J. H. Strauss.** 1988. Molecular basis of Sindbis virus neurovirulence in mice. J. Virol. **62**:2329–2336.
13. **Olmsted, R. A., R. S. Baric, B. A. Sawyer, and R. E. Johnston.** 1984. Sindbis virus mutants selected for rapid growth in cell culture display attenuated virulence in animals. Science **225**: 424–427.
14. **Olmsted, R. A., W. J. Meyer, and R. E. Johnston.** 1986. Characterization of Sindbis virus epitopes important for penetration in cell culture and pathogenesis in animals. Virology **148**:245–254.
15. **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.
16. **Polo, J. M., N. L. Davis, C. M. Rice, H. V. Huang, and R. E. Johnston.** 1988. Molecular analysis of Sindbis virus pathogenesis in neonatal mice by using virus recombinants constructed in vitro. J. Virol. **62**:2124–2133.
17. **Polo, J. M., and R. E. Johnston.** 1990. Attenuating mutations in glycoproteins E1 and E2 of Sindbis virus produce a highly attenuated strain when combined in vitro. J. Virol. **64**:4438–4444.
18. **Rice, C. M., R. Levis, J. H. Strauss, and H. V. Huang.** 1987. Production of infectious RNA transcripts form Sindbis virus cDNA clones: mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. J. Virol. **61**:3809–3819.
19. **Russell, D. L., J. M. Dalrymple, and R. E. Johnston.** 1989. Sindbis virus mutations which coordinately affect glycoprotein processing, penetration, and virulence in mice. J. Virol. **63**: 1619–1629.
20. **Stanley, J., S. J. Cooper, and D. E. Griffin.** 1985. Alphavirus neurovirulence: monoclonal antibodies discriminating wild-type from neuroadapted Sindbis virus. J. Virol. **56**:110–119.
21. **Zimmern, D., and P. Kaesberg.** 1978. 3′-terminal nucleotide sequence of encephalomyocarditis virus RNA determined by reverse transcriptase and chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **75**:4257–4261.