Molecular Determinants of Alphavirus Neurovirulence: Nucleotide and Deduced Protein Sequence Changes during Attenuation of Venezuelan Equine Encephalitis Virus

By BARBARA J. B. JOHNSON,* RICHARD M. KINNEY, CRYSTLE L. KOST and DENNIS W. TRENT

Division of Vector-Borne Viral Diseases, Center for Infectious Diseases, Centers for Disease Control, Public Health Service, U.S. Department of Health and Human Services, P.O. Box 2087, Fort Collins, Colorado 80522, U.S.A.

(Accepted 21 May 1986)

SUMMARY

The nucleotide and deduced amino acid sequences of the structural proteins of the TC-83 vaccine strain of Venezuelan equine encephalitis (VEE) virus have been determined from a cDNA clone containing the 26S mRNA coding region. A cDNA clone encoding the equivalent region of the virulent parent VEE virus [Trinidad donkey strain (TRD)] has been sequenced previously. Comparison of the sequences of the TC-83 and TRD cDNA clones revealed 13 nucleotide differences. Neither the organization of the structural proteins (5'-capsid-E3-E2-6K-E1-3') nor the length (3762 nucleotides) of the open reading frame coding for the viral polyprotein precursor was altered during attenuation. Of the 13 nucleotide differences between the cDNA clones of TC-83 and TRD, nine occurred in the dominant population of the respective genomic RNAs from plaque-purified viruses. Six of the nine mutations were clustered in the E2 surface glycoprotein gene. All five of the nucleotide changes which produced non-conservative amino acid substitutions in the encoded proteins were located in the E2 gene. Two mutations occurred in the E1 glycoprotein gene; one was silent and the other did not alter the chemical character of the E1 protein. One nucleotide difference was found in the non-coding region immediately preceding the 5'-end of the 26S mRNA. The E2 and non-coding region mutations are candidates for the molecular determinants of VEE virus neurovirulence.

INTRODUCTION

Venezuelan equine encephalitis (VEE) virus has caused periodic catastrophic loss of equines in the Americas since about 1930, with severe economic consequences. Estimated equine deaths in the most recent epizootics spanning 1967 to 1971 exceeded 150000. VEE virus also has had a substantial impact on human health, causing about 60000 cases during the 1960's. Neurological symptoms were present in 4 to 14% of patients and the fatality rate approached 1% (Groot, 1972).

VEE virus is an attractive agent in which to characterize the molecular determinants of neurovirulence. It has a relatively simple RNA genome (about 12 kb; Strauss & Strauss, 1986), and closely related strains exist which differ markedly in their pathogenesis in man and animals. The virulent Trinidad donkey (TRD) strain of VEE virus and its avirulent vaccine derivative, TC-83 virus, were chosen for nucleotide sequencing studies aimed at understanding the basis for the ability of this alphavirus to invade and injure the brain.

The TC-83 vaccine strain of VEE virus was derived from the virulent TRD virus by serial passage of TRD in embryonic guinea-pig heart cells in culture (Berge et al., 1961). TC-83 virus is avirulent for weanling mice by both the intracerebral and intraperitoneal routes of inoculation (Berge et al., 1961) and is also avirulent for congenitally athymic (nude) mice (Le Blanc et al., 1961).
1952  B. J. B. JOHNSON AND OTHERS

The attenuated virus produces viraemias which are 10- to 1000-fold lower than those observed with virulent strains (Jahrling & Scherer, 1973c; Le Blanc et al., 1978) and fails to produce the severe lesions in hamster haematopoietic and neuronal cells which are characteristic of TRD virus infection (Jahrling & Scherer, 1973a). Although developed to protect laboratory and field personnel (McKinney et al., 1963), TC-83 vaccine subsequently proved invaluable in controlling the 1969 to 1971 epizootic of VEE in Central America (Walton et al., 1972; Eddy et al., 1972) and Texas (Calisher & Maness, 1975; Baker et al., 1978). The vaccine induces long-lasting neutralizing antibody to epizootic subtypes IA to IC, but not to variants ID and IE or subtypes II to IV (Burke et al., 1977; Walton & Johnson, 1972; Fillis & Calisher, 1979).

TRD and TC-83 viruses are indistinguishable by classical serology (both are subtype IA) and differ only slightly in their T1 ribonuclease-resistant oligonucleotide fingerprints (Trent et al., 1979). Of the 70 large oligonucleotides of the genomic 42S RNA of the TRD strain, three were absent from the TC-83 42S RNA fingerprint. In addition, TC-83 has three unique oligonucleotides. The oligonucleotide differences observed in the fingerprints have been localized to the subgenomic 26S mRNA which encodes the structural proteins of the virus (Mecham & Trent, 1982a). Mapping of tryptic peptides of the structural proteins (the capsid and the two envelope glycoproteins, E1 and E2) demonstrated that the resolvable differences between TRD and TC-83 viruses are in the E2 membrane glycoprotein (Mecham & Trent, 1982b), which is the haemagglutinin (France et al., 1979) and contains the type-specific neutralization epitope (Roehrig et al., 1982).

We report here the cloning and sequencing of cDNA encoding the structural proteins of TC-83 virus. Comparison of this with that of the virulent TRD strain (Kinney et al., 1986) has resulted in the identification of amino acids in the surface glycoproteins which may determine the neurovirulence of VEE virus.

METHODS

RNA purification, cloning and sequencing methods, and computer analyses. These were done as described previously (Kinney et al., 1986).

Virus and cultured cells. The TC-83 strain of VEE virus used was vaccine lot number 5-LD01 produced by Merrell National Laboratory, Philadelphia, Pa., U.S.A. and was distributed by the Centers for Disease Control, Atlanta, Ga., U.S.A. The virus was plaque-purified in confluent monolayers of Vero cells, and seed stocks prepared as tissue culture fluids obtained by infection of BHK-21 cells at low multiplicity. Viral RNA for use in cDNA synthesis was isolated from purified TC-83 virus grown in BHK-21 cells. Attenuation of the plaque-purified virus was verified in mice and hamsters prior to molecular cloning. RNA was also prepared from the TRD strain of VEE virus, passage history GPB1V6BH21-1 (GPB = guinea-pig brain; V = Vero).

Synthesis and purification of deoxyoligonucleotide primers. Deoxyoligonucleotides for priming cDNA synthesis and for primer extension sequencing of genomic RNA were synthesized with an Applied Biosystems Model 380A DNA synthesizer using phosphoramidite chemistry (Beaucage & Caruthers, 1981). The oligonucleotides were purified by preparative gel electrophoresis through 16% acrylamide/bisacrylamide (19/1) in 7 M-urea and TBE buffer pH 8.3 (89 mM-Tris, 89 mM-boric acid, 2.8 mM-disodium EDTA). The oligomers were located by their u.v. absorbance, the appropriate bands excised, and the DNAs eluted in 2 M-ammonium acetate at 37 °C. Oligonucleotides were then adsorbed to DE-52 cellulose (Whatman), washed with water, eluted with 30% (v/v) triethylammonium bicarbonate pH 8, and desalted in vacuo (C. Naeve, personal communication).

RNA sequencing. Genomic RNA (42S) was sequenced by reverse transcriptase-catalysed extension from synthetic DNA primers using dideoxynucleotide chain terminators (Sanger et al., 1977). The protocol was adapted from Ou et al. (1981). Deoxynucleotidyl primer (about 20-fold molar excess) was hybridized to 42S RNA for 2 min at 95 °C and then placed on ice. The concentrations of dGTP, dCTP and dTTP (P-L Biochemicals) were 3-7 μM and the concentration of [α-32P]dATP (New England Nuclear) was 1 μM (sp. act. 800 Ci/mmol). The ratio of deoxy- to dideoxynucleotides was 2-3. Reaction mixtures contained 44 mM-Tris–HCl pH 8.1, 8.3 mM-MgCl2, 54 mM-KCl, 8.3 mM-dithiothreitol, 9 units RNasin (Promega Biotech) and 2-1 units avian myeloblastosis virus reverse transcriptase (Promega Biotech) in a final volume of 5 μl. The mixtures were incubated for 5 min at 20 °C and then for 30 min at 37 °C. Nascent DNA chains were then chased by addition of 2 μl of a solution containing 0.5 mM of each deoxynucleotide triphosphate and incubated at 37 °C for 15 min. Reactions were stopped with 6 μl formamide/dye, boiled for 5 min, and analysed on 4-4 mm-thick sequencing gels (Sanger & Coulson, 1978) which were dried before autoradiography under Kodak SB-5 film.
RESULTS

Nucleotide sequence of cDNA encoding the 26S RNA of TC-83 virus

DNA complementary to the 26S region of viral 42S RNA was prepared for cloning into pUC18 by priming at the 3'-end of the genome, as previously described (Kinney et al., 1986). Restriction fragments of the cloned cDNA, designated pTC-5, were subcloned into M13 phages mp10 and mp11. The entire sequence was determined on each strand of the pTC-5 cDNA and all the junctions between restriction fragments were confirmed by sequencing other fragments which contained these junctions in an internal position. In the region encoding the capsid, it was necessary to build a synthetic primer to obtain complete data on the second DNA strand and to verify Sau3A and TaqI restriction sites.

The nucleotide sequence of the 26S mRNA region of TC-83 virus and the deduced amino acid sequence of the encoded proteins are shown in Fig. 1. The pTC-5 cDNA clone is 4008 nucleotides long and contains 161 nucleotides to the 5'-side of the beginning of the 26S RNA. The 3'-end of the cDNA clone is marked with a vertical arrow in the region complementary to the oligonucleotide used to prime the cDNA synthesis. The 65 nucleotides at the 3'-end of the genome [excluding the poly(A) tail] are not present in clone pTC-5, but were previously determined by Ou et al. (1982a). The overall organization of the 26S mRNA of TC-83 virus is identical to that of its virulent parent. Starting at the AUG initiation codon, an open reading frame extends for 3762 nucleotides until an opal termination codon (UGA) is reached, exactly as was found for TRD virus (Kinney et al., 1986).

Comparison of the sequences of TC-83 and TRD viruses

Nucleotide differences between the TC-83 and TRD cDNAs are noted in Fig. 1. There are 13 differences between the pTC-5 clone and the pTRD-1 clone of the virulent virus (Kinney et al., 1986). Seven differences are in the E₂ glycoprotein, four in the E₁ glycoprotein, and one each in the 6K protein and in the non-coding region preceding the beginning of the 26S RNA. The nucleotide sequences encoding the capsid and E₃ proteins of the two viruses are identical.

To ascertain whether the differences between the TC-83 and TRD cDNAs reflect differences in their respective genomic RNAs, direct RNA sequencing by primer extension was performed on each of the genomes. Reverse transcriptase errors could have introduced differences in the cDNA clones which did not exist in the viral RNA templates. Furthermore, RNA genomes are highly mutable, and it was possible that we cloned a minor variant from a heterogeneous population of RNA molecules. Viruses used in this study were plaque-purified at terminal dilution before they were grown for RNA isolation to minimize this possibility.

Seven deoxynucleotide primers complementary to the positive strand of the cDNAs were synthesized, purified and used to sequence genomic RNAs across the regions where the 13 differences between the cDNAs were found. The regions sequenced are shown in Fig. 2. Three cDNA differences were unconfirmed at the RNA sequence level; these are marked with an 'X'. (The structure revealed by RNA sequencing is a weighted average of the mixture of mutants present in the sample.) Nucleotide number 1657 from the 5'-start of the 26S RNA was found to be a G in both viruses so that the pTRD-1 clone differs from the majority of the population of RNA molecules at one site in the E₂ glycoprotein gene. Nucleotide number 2370, in the 6K signal peptide, was also found to be a G in both viruses. The mRNA sequences encoding the 6K signal proteins of TC-83 and TRD viruses were therefore identical (see Kinney et al., 1986). Nucleotide number 2822 was a C in both viral RNAs so that clone pTC-5, which contains a U at position 2822, differs from the predominant genomic RNA population at this position.

The sequence of the TC-83 RNA (Fig. 1) is shown with a C at nucleotide position 2822 to reflect the sequence of the TC-83 virus population. This alteration results in Ala at amino acid position E₁-119 instead of the Val encoded by pTC-5. Finally, the question mark in Fig. 2 indicates that the nucleotide difference observed between the pTC-5 and pTRD-1 cDNA clones at position 3366 could not be evaluated by sequencing on RNA templates due to band compression in the sequencing gels. A second, independent TC-83 cDNA clone (pTC-17) was sequenced and found to contain an A at this position, in agreement with clones pTRD-1 and pTRD-7. A G is present.
Fig. 1. Nucleotide sequence of the VEE TC-83 26S mRNA and the deduced amino acid sequence of the encoded proteins. The 5'-end of the 26S mRNA (nucleotide position 1) was assigned by its homology end (...) and the terminal nucleotide in clone pTC-5 is identified (T). 3'-Nucleotides not present in clone nucleotides). Nucleotide differences (13) between cDNA clones pTC-5 and pTRD-1 (Kinney pTC-5, but previously determined (Ou et al., 1982a), are included after the arrow (underlined + ÷). An additional 161 nucleotides which are contained in cDNA clone pTC-5 precede the Y-end of the 26S RNA. The sequence complementary to the oligonucleotide used to prime the cDNA synthesis is identified at the 3'--

sequence. The single letter amino acid code is used. A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.
VEE virus neurovirulence

![Diagram showing the regions of TC-83 and TRD genomic RNA sequenced by extension of primers. Vertical lines indicate nucleotide differences between cDNA clones pTC-5 and pTRD-1. cDNA clone differences which were not observed between genomic RNA populations are denoted 'X'. The nucleotide difference which could not be evaluated by RNA sequencing (denoted '?') was resolved by examining an independently derived cDNA clone (see text).]

**Fig. 2.** Regions of TC-83 and TRD genomic RNA sequenced by extension of primers. Vertical lines indicate nucleotide differences between cDNA clones pTC-5 and pTRD-1. cDNA clone differences which were not observed between genomic RNA populations are denoted 'X'. The nucleotide difference which could not be evaluated by RNA sequencing (denoted '?') was resolved by examining an independently derived cDNA clone (see text).

**Table 1. Nucleotide and deduced amino acid differences between TRD and TC-83 viruses**

<table>
<thead>
<tr>
<th>Position</th>
<th>Nucleotide</th>
<th>Amino acid change</th>
</tr>
</thead>
<tbody>
<tr>
<td>−2, junction region</td>
<td>TRD</td>
<td>TC-83</td>
</tr>
<tr>
<td>1053, E2-7</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>1285, E2-85</td>
<td>G</td>
<td>U</td>
</tr>
<tr>
<td>1391, E2-120</td>
<td>C</td>
<td>U</td>
</tr>
<tr>
<td>1607, E2-192</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>1866, E2-278</td>
<td>U</td>
<td>A</td>
</tr>
<tr>
<td>1919, E2-296</td>
<td>C</td>
<td>U</td>
</tr>
<tr>
<td>2947, E1-161</td>
<td>U</td>
<td>A</td>
</tr>
<tr>
<td>3099, E1-211</td>
<td>A</td>
<td>U</td>
</tr>
<tr>
<td>3874, 3′-non-coding region*</td>
<td>UU</td>
<td>U</td>
</tr>
</tbody>
</table>

* Ou et al. (1982a).

A summary of the nucleotide and deduced amino acid differences between TRD and TC-83 viruses is presented in Table 1. Most of the nucleotide differences (six) are clustered in the E2 glycoprotein gene. Of these six differences, five resulted in non-conservative amino acid changes in the E2 protein. The charge of the E2 protein is altered as a result of three of these changes at nucleotide positions 1053, 1391 and 1607. These three alterations make the E2 of TC-83 vaccine virus slightly less basic than that of its virulent parent. Neither of the nucleotide differences between the E1 genes resulted in an alteration in the chemical character of the E1 proteins. One of the E1 mutations is silent and the other resulted in the substitution of an Ile for Leu.

The 26S junction region, that portion of the genomic RNA containing the 3′-end of the non-structural protein genes and the 5′-start of the structural protein genes, contains a highly conserved nucleotide sequence in other alphaviruses (Ou et al., 1982b). Both TC-83 and TRD viruses deviated from this pattern (Fig. 3). An ochre (UAA) termination codon occurs eight nucleotides before the 26S RNA start site, which slightly shortens the carboxyl terminus of nsP4. This terminator may be a general distinguishing feature of VEE viruses because we have also found it in the junction regions of two enzootic ID strains from Colombia (our unpublished results). Two nucleotides preceding the 26S RNA 5′-end is a confirmed difference between TC-83 and TRD genomic RNAs (Fig. 3). This A to G transition (Fig. 3) may alter the hypothetical viral replicase recognition site.
Fig. 3. Comparison of alphavirus sequence in the region containing the start of the 26S RNA. Data for Sindbis (SIN), Semliki Forest (SF), Ross River (RR) and Middelburg (MID) viruses were obtained from Ou et al. (1982b). The amino acid sequence shown is the carboxyl terminus of non-structural protein 4 (nsP4). Dots indicate nucleotide identity with the VEE TRD sequences. Solid bars mark termination codons.

Fig. 4. Hydrophobicity profiles of the TC-83 virus glycoproteins. The average hydrophobicity of every nine amino acid-long segment of the E₂ and E₁ proteins was computed using the indices of Kyte & Doolittle (1982). The positions of amino acid differences between TC-83 and TRD viruses are marked by 'X'. Potential Asn-linked glycosylation sites are shown (○).
For completeness, the one other known nucleotide difference between TC-83 and TRD viruses is presented in Table 1. One uridine is deleted from TC-83 virus (Ou et al., 1982a) at nucleotide 3874. This deletion makes the 3'-non-coding region 120 nucleotides long, compared with 121 nucleotides in TRD virus. The deleted uridine does not fall within the highly conserved string of 19 nucleotides immediately adjacent to the poly(A) tract, which has been postulated to serve as a replicase recognition signal (Ou et al., 1982 a).

The six deduced amino acid differences between TRD and TC-83 viruses are position on hydrophobicity plots of the TC-83 E2 and E1 proteins (Fig. 4). The amino acid differences, marked 'X', are predominantly in hydrophilic domains of E2. One change (amino acid position E2-296) affects a hydrophobic region of E2. The one amino acid change in E1, a conservative Leu to Ile substitution, occurs in a hydrophobic region. Closed circles in Fig. 4 indicate potential N-glycosylation sites, Asn-X-Ser or Asn-X-Thr sequences (where X is any amino acid), which are a necessary but not sufficient condition for glycosylation. None of the three potential glycosylation sites in E2, nor the single site in E1, was directly altered by the observed nucleotide changes which occurred during attenuation of TRD virus.

DISCUSSION

RNA viral genomes are so highly mutable that the nucleotide sequence of virus propagated from a clone may not be singular. The sequence of an RNA virus population is probabilistic whenever the number of bases in the genome exceeds the reciprocal of the base misincorporation frequency per genome doubling (Holland, 1984). Base-substitution frequencies for vesicular stomatitis at one site in a highly conserved region of the genome average between $10^{-4}$ and $4 \times 10^{-4}$ per base incorporated at that site (Steinhauer & Holland, 1986). A similar base-misincorporation frequency (between $10^{-3}$ and $10^{-4}$) has been calculated for RNA phage Qβ (Domingo et al., 1978). Therefore, progeny of a VEE virus, which has a 12 kb genome, will probably differ both from their parent and from each other by at least one nucleotide.

The nucleotide sequence of cDNA encoding the structural protein region of the benign TC-83 virus was determined in order to compare it with a similar clone derived from the virulent TRD strain (Kinney et al., 1986). Sequencing of cDNA rather than genomic RNA was performed because of the greater confidence which results from sequencing both strands of a cDNA and because we intend to use these cDNAs to express their encoded viral proteins. Thirteen nucleotide differences were found between the two VEE cDNA clones. Four of these differences were shown to result from reverse transcriptase errors or the cloning of a minor variant from the 42S RNA population. The nine remaining nucleotide differences were evaluated as potential determinants of VEE virus neurovirulence.

The nine mutations in TC-83 virus are not randomly distributed over the structural gene region. They are clustered in the gene for the E2 glycoprotein. All five of the non-conservative amino acid changes deduced from the nucleic acid sequences are located in E2. Although two mutations occurred in the E1 glycoprotein gene, one was silent and one resulted in a conservative Leu to Ile change in the amino acid incorporated. No mutations were found in the capsid nor in the E3 or 6K signal sequences, which together comprise 32% of the structural protein genes. The absence of mutations in the signal sequences probably is not remarkable in view of their short lengths.

Three mutations in E2 alter the charge of hydrophilic domains, which are likely to be on the surface of the virus and accessible to the immune system of the host. All five amino acid substitutions in E2 are potentially capable of altering its surface topography and, thereby, its surface-mediated functions. Both the ability to elicit virus type-specific neutralizing antibody and haemagglutination have been associated with the E2 of VEE virus (Pederson & Eddy, 1974; France et al., 1979; Kinney et al., 1983). The antigenic makeup of VEE E2 glycoproteins has been explored in detail with monoclonal antibodies (Roehrig et al., 1982; Roehrig & Mathews, 1985). Four monoclonal antibodies demonstrate differential reactivity with the E2 of TC-83 and TRD viruses; two have neutralization and haemagglutination-inhibition activity with TRD virus but only neutralization activity with TC-83 virus.

One mutation was located in the non-coding junction region, two nucleotides before the 5'-end
of the 26S mRNA. This A-to-G transition may possibly affect the postulated viral replicase recognition site. Replication of TC-83 virus has been compared with its virulent parent \textit{in vitro} (Mecham & Trent, 1983). TC-83 and TRD viruses have similarly shaped growth curves in Vero cells, but 2-5 times less TC-83 virus is produced early in the replication cycle concomitant with a depressed production of 26S mRNA. The relationship of this finding to virulence is unknown. There is precedent, however, for a single nucleotide change in a non-coding region to alter neurovirulence substantially in the Sabin type 3 poliovaccine genome (Evans et al., 1985).

Although the two translationally silent mutations (one each in E\textsubscript{3} and E\textsubscript{1}) do not, of course, perturb the structure of the glycoprotein spike, they may not be irrelevant to the modification of neurovirulence. RNAs may replicate more slowly if they have translationally silent changes, as was suggested by the \textit{Qp2} studies of Domingo et al. (1978). The competitive disadvantage of a \textit{Qp2} variant containing a silent mutation could not be specifically associated with the silent nucleotide substitution, however, since other mutations may have been concurrently present in the genome in unanalysed regions.

TRD and TC-83 viruses do not differ in the number or position of potential \textit{N}-glycosylation sites in their surface proteins. Three Asn–X–Thr sequences occur in E\textsubscript{3} and one in E\textsubscript{1}. One is also present in E\textsubscript{2}. E\textsubscript{3} forms part of the glycoprotein spike of Semliki Forest virus, but it is not detectable in VEE virions (Simons et al., 1980). We believe that the E\textsubscript{2} glycoproteins of both TC-83 and TRD viruses are glycosylated at all three Asn–X–Thr sites. Mecham & Trent (1982b) suggested that TRD and TC-83 viruses differ in the extent of glycosylation of their E\textsubscript{2} proteins because the TRD E\textsubscript{2} had three glycosylated tryptic peptide peaks resolvable by reverse-phase high-pressure liquid chromatography, whereas TC-83 E\textsubscript{2} had only two. The deduced amino acid sequence of the TC-83 E\textsubscript{2} reveals that its second glycosylation site occurs in a tryptic peptide fragment whose composition (NPTYLITR) differs from the cognate tryptic fragment of TRD virus (NPTYLTTR). The substitution of Ile for Thr in this peptide alters the total hydrophobicity index (calculated from the amino acid values suggested by Kyte & Doolittle, 1982) from \(-9.2\) (TRD) to \(-40\) (TC-83). The increased hydrophobicity of the TC-83 glycopeptide probably decreases its mobility in reverse-phase chromatography. In the work of Mecham & Trent (1982b), the first (most hydrophilic) of the three glycosylated tryptic peptide peaks of TRD E\textsubscript{2} is missing in the TC-83 E\textsubscript{2} chromatogram, and the second TC-83 peak is larger in area than the second TRD peak by an amount equal to that of the first TRD peak. This profile indicates that the Thr to Ile change shifted the first glycopeptide peak seen in TRD E\textsubscript{2} into the more hydrophobic second peak, and that the TRD and TC-83 E\textsubscript{2} proteins are equally glycosylated. No differences were seen between the tryptic glycopeptide chromatograms of the E\textsubscript{1} proteins, although more than one glycopeptide was resolved, perhaps because of incomplete trypsin digestion of the E\textsubscript{1}. The difference in neurovirulence of the two viruses, therefore, cannot be associated with differences in the extent of glycosylation of their surface proteins.

The five amino acid substitutions seen in the TC-83 virus E\textsubscript{2} glycoprotein, as well as the single nucleotide substitution in the non-coding junction, are candidates for the molecular determinants of VEE neurovirulence. We are evaluating the relative significance of these putative determinants by sequencing the E\textsubscript{2} and junction regions of phenotypic revertants of TC-83 vaccine virus. TC-83-derived viruses that kill hamsters and mice have been isolated from the throats of several human vaccinees and from hamsters directly after a single passage of the vaccine (Jahrling & Scherer, 1973b; P. B. Jahrling, personal communication). Vaccine production lots that are neurovirulent for mice and hamsters (E. R. Aycardi, personal communication) will also be examined.

This study augments the growing body of data implicating surface protein structure in the modulation of alphavirus virulence. Neuroadaptation of the Sindbis alphavirus was accompanied by alterations in E\textsubscript{1} and E\textsubscript{2} conformation detected by monoclonal antibodies. Most of the changes detected were clustered on E\textsubscript{2} (Stanley et al., 1985). Sequence analysis of the Sindbis virus E\textsubscript{2} glycoprotein genes of several closely related attenuated, virulent and virulent-revertant strains has revealed two sites which affect neurovirulence in newborn mice (N. L. Davis, F. J. Fuller & R. E. Johnston, personal communication). A temperature-sensitive, attenuated mutant of VEE virus exhibited a lower isoelectric point of its E\textsubscript{1} glycoprotein. Reversion of this mutation correlated the E\textsubscript{1} alteration with attenuation (Emini & Wiebe, 1981).
The surface proteins of several viruses have been identified as determinants of pathogenicity. A single amino acid substitution in the rabies glycoprotein prevents rabies viruses from producing lethal infection in adult mice (Dietzschold et al., 1983). Although neurovirulence is a polygenic character in influenza viruses, the cleavage of the haemagglutinin plays a critical role (Bosch et al., 1981; Kawaoka et al., 1984). Reassortant studies of the California group viruses (Bunyaviridae) correlated murine virulence with the M RNA segment. This segment encodes the glycoproteins which mediate virus attachment and perhaps tissue tropism (Shope et al., 1981; Tignor et al., 1983). Neurotropism of mammalian reoviruses is determined by the outer capsid proteins µC and the haemagglutinin µ (Fields & Greene, 1982).

This investigation was confined to the 3'-terminal one-third of the TC-83 and TRD viral genomes. It is in this structural protein region that all of the observable differences in the T1 ribonuclease-resistant oligonucleotides are located (Mecham & Trent, 1982a). Since oligonucleotide fingerprinting resolves fragments comprising about 10% of the genome (Domingo et al., 1978), the actual number of mutations in a virus probably exceeds the number of oligonucleotide spot changes observed. The nine nucleotide differences in TC-83 virus 26S mRNA were reflected in the disappearance of three oligonucleotide spots (Mecham & Trent, 1982a); thus, in this instance, the ratio of actual mutations to fingerprint differences is three. As a consequence, there may be mutations in the genes which encode the non-structural proteins of TC-83 virus which contribute to its avirulent phenotype.

We thank Vickie Brown for technical assistance with portions of this work and Susan Ure for typing the manuscript.

REFERENCES


(Received 7 April 1986)