

AD-A239 227



2

AD _____

CONTRACT NO: DAMD17-90-C-0050

TITLE: MOLECULAR STUDIES OF ALPHAVIRUS IMMUNOGENICITY

PRINCIPAL INVESTIGATOR: James H. Strauss, Ph.D.

CONTRACTING ORGANIZATION: California Institute of Technology
1201 E. California Boulevard
Pasadena, California 91125

REPORT DATE: May 1, 1991

TYPE OF REPORT: Annual Report

PREPARED FOR: U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

DTIC
SELECTED
AUG 6 1991
S B D

91-06858



9 1 8 05 029

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 1 May 1991	3. REPORT TYPE AND DATES COVERED Annual Report (3/30/90 - 3/29/91)	
4. TITLE AND SUBTITLE MOLECULAR STUDIES OF ALPHAVIRUS IMMUNOGENICITY			5. FUNDING NUMBERS Contract No. DAMD17-90-C-0050	
6. AUTHOR(S) James H. Strauss, Ph.D.			61102A 3M161102BS12.AB.117 WUDA346101	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) California Institute of Technology 1201 E. California Boulevard Pasadena, California 91125			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Development Command Fort Detrick Frederick, Maryland 21702-5012			10. SPONSORING MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) <p style="margin-left: 40px;">Ockelbo virus, first isolated in 1982 in Sweden, causes arthritis, fever and rash in man. We have obtained the complete nucleotide sequence of Ockelbo virus and compared this sequence to that of other strains of Sindbis virus. Partial sequence analysis of five other strains of Sindbis virus was also performed. Three principal conclusions arise from our data. (1) Ockelbo is virtually identical to the causative agents of Karelian Fever of Russia and of Pogosta disease of Finland. (2) These agents are closely related to South African strains of Sindbis virus, and Ockelbo was probably introduced into northern Sweden from Africa in the 1960's, followed by spread to Russia and Finland. (3) There exist an European-African group of closely related Sindbis viruses and an Asian-Australian group of Sindbis viruses.</p>				
14. SUBJECT TERMS Alphavirus; Sindbis Virus; Ockelbo Disease; BD; RAI; Antigenic Epitope; Immunogenicity			15. NUMBER OF PAGES	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited	

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

_____ Where copyrighted material is quoted, permission has been obtained to use such material.

_____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

_____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

_____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

_____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

JH In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

JH Strauss
PI - Signature

5/1/91
DATE

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification _____	
By _____	
Distribution _____	
Availability Codes	
Dist	Special
<i>A-1</i>	

Table of Contents

	Page
Front Cover	1
Report Documentation Page	2
Foreword	3
Table of Contents	4
Report	
Introduction	5
Methods Used	7
Complete Nucleotide Sequence of Ockelbo Virus	7
3' Terminal Nucleotide Sequence of Other Strains of Sindbis Virus	15
Conclusions	18
References	19

Introduction

The alphaviruses are a widespread group of human pathogens that are endemic and epidemic in many parts of the world (Chamberlain, 1980; Griffin, 1986; Peters and Dalrymple, 1990). They are mosquito-borne and are particularly prevalent in tropical and subtropical areas of the world, but alphaviruses pathogenic for man are also present in temperate and even Arctic areas. Many alphaviruses are capable of causing fever, rash and arthralgia in man that in some cases can be disabling for extended periods of time. Many of the New World alphaviruses can cause encephalitis in man. We wish to determine the relationships of alphaviruses and strains of alphaviruses to one another and to search for emerging viruses.

The prototype alphavirus, Sindbis virus, is widespread throughout the Old World, occurring in Australia, Southeast Asia, India, Europe and Africa (Niklasson, 1988). Strains of Sindbis virus in Northern Europe, referred to as Ockelbo virus and Karelian fever virus, cause an illness characterized by polyarthritits whose symptoms can persist for months or years. This illness first appeared in Northern Europe in the latter half of the 1960's and a virus was first isolated from female *Culiseta* mosquitos collected in the endemic area in the summer of 1982 (Niklasson et al., 1984). Serological studies with convalescent sera from patients with Ockelbo disease showed that this virus was the causative agent of Ockelbo disease and that it was closely related to Sindbis virus. Similar diseases were also found in Finland and the Karelian region of the Soviet Union in the early 1980's and were called Pogosta disease and Karelian fever, respectively. Karelian fever virus was also isolated in 1982 and shown to be nearly identical to Ockelbo virus (Lvov et al., 1984,1988). The causative agent of Pogosta disease was also found to be closely related to Ockelbo virus and Sindbis virus (Calisher et al., 1985). Thus, we have the phenomenon of a new virus appearing in Northern Europe and apparently spreading, infecting increasing numbers of

people. It is also of interest that this Northern European strain of Sindbis virus causes an illness whose clinical features are most closely related to those caused by Ross River virus, a virus restricted to Australia and the Southern Pacific region that is not closely related to Sindbis virus.

Strains of Sindbis virus in Southern Africa are also known to be virulent for man and have caused epidemics of human illness. The pattern of clinical features in these illnesses is similar to that caused by Old World alphaviruses such as chickengunya and O'Nyong-Nyong, but the severity of the disease is, in general, less.

Sindbis virus is antigenically related to the New World Western equine encephalitis virus which, as the name implies, is capable of causing encephalitis in man. We recently found that Western equine encephalitis virus is recombinant between a Sindbis-like virus, presumably found somewhere in the New World, and the New World Eastern equine encephalitis virus. Its encephalogenic properties probably originated from the Eastern equine encephalitis virus parent (Hahn *et al.*, 1988).

We wish to explore the relationships of these various strains of Sindbis virus to one another and to determine, if possible, why Sindbis virus disease seems to be emerging as a more widespread human threat. For this purpose, we obtained the complete nucleotide sequence of the prototypic Ockelbo virus (Edsbyn 82-5) and obtained partial sequences of several geographical isolates of Sindbis virus. A full description of these results has appeared in *Virology* **182**:753-764 (1991). A preprint of this paper, entitled "Structure of the Ockelbo virus genome and its relationship to other Sindbis viruses", by Y. Shirako, B. Niklasson, J. M. Dalrymple, E. G. Strauss, and J. H. Strauss, was submitted to the U.S. Army Medical Research and Development Command at the time of submission to the journal.

Methods Used

The seven strains of Sindbis virus used in this study and their original sources are described in Table 1. In addition to the Edsbyn 82-5 strain of Ockelbo isolated in 1982 in Edsbyn Village, Sweden (Niklasson, et al., 1984), strains of Sindbis isolated in 1952 in Egypt (Taylor et al., 1955), 1953 in India (Shah et al., 1960), 1963 in South Africa (Malherbe et al., 1963), 1975 in Australia, and strains of Ockelbo/Karelian fever virus isolated in 1983 in Sweden (Niklasson et al., 1984) and the USSR (Lvov et al., 1984), were included. Viruses were grown either in BHK-21 cells or in secondary chicken embryo fibroblasts and RNA was isolated from sucrose density gradient purified virus preparations with SDS-phenol as previously described (Ou et al., 1981).

cDNA clones from the viruses were produced using standard methods (Sambrook et al., 1989). First-strand cDNA was made using oligo(dT) as a primer and second-strand synthesis was by the method of Gubler and Hoffman (Gubler and Hoffman, 1983). In some cases *Hind*III fragments of the cDNA were cloned into vector pGEM3Z. In other cases, *Eco*RI vectors were added to the double-stranded cDNA and the cDNA cloned into the *Eco*RI site of pGEM3Z.

DNA sequencing and RNA sequencing used standard technology that is in common use in our laboratory (Hahn et al., 1989; Rice et al., 1985; Shirako and Strauss, 1990; Strauss et al., 1984).

Complete Nucleotide Sequence of Ockelbo Virus

cDNA representing the complete genome of Ockelbo virus was obtained and sequenced. In addition, the sequence of the 5' end, the 3' end and some internal sequences were obtained by directly sequencing virus genomic RNA using a dideoxy method. The complete nucleotide sequence obtained and the deduced amino acid sequence of the proteins encoded in the viral

Table 1 Sindbis-like Viruses Pertinent to this Study

Name	Strain	Source	Year	Location	Reference
Ockelbo	Edsbyn 82-5	Pooled mosquitos (<i>Culiseta</i> spp.)	1982	Edsbyn village, Sweden	Niklasson et al., (1984)
Ockelbo	Edsbyn 83M107	Mosquito (<i>Culiseta morsitans</i>)	1983	Edsbyn village, Sweden	
Karelian Fever	LEIV 9298	Mosquito (<i>Aedes communis</i>)	1983	Central Karelia, USSR	Lvov et al., (1988)
Sindbis	Girdwood	Human	1963	South Africa	Malherbe et al., (1963)
Sindbis	A-1036	Mite (<i>Bdellonyssus bursa</i>)	1953	India	Shah et al., (1960)
Sindbis	MRM18520	Mosquito (unidentified)	1975	Queensland, Australia	
Sindbis	AR339	Mosquito (<i>Culex univittatus</i>)	1952	Egypt	Taylor et al., (1955)

genome of Ockelbo virus (Edsbyn 82-5) are shown in Figure 1. The viral genome is 11,708 nucleotides in length excluding the 5' terminal cap and the 3' terminal poly(A) tract. The genome organization is virtually identical to that of the Sindbis virus AR339 strain (Strauss et al., 1984) isolated in Sindbis, Egypt in 1952 (Taylor et al., 1955). Compared to AR339, Ockelbo nsP3 contains a deletion of 9 nucleotides and 2 separate insertions of 6 and 9 nucleotides in the C-terminal half, and there are 3 single nucleotide insertions and deletions of the 3' nontranslated region. Otherwise the numbers of nucleotides translated into amino acids in each region were exactly the same for these two strains of Sindbis virus. Overall, as illustrated schematically in Figure 2, there are 672 nucleotide differences between the two viruses (5.7% divergence) that result in 97 amino acid changes (2.6% divergence). Thus, more than 85% of the nucleotide changes are silent. Only proteins E2 and 6K and the C-terminal domain of nsP3 show an amino acid sequence divergence that is significantly higher than the average divergence between the two viruses (2.6% divergence averaged over the entire genome or 1.7% excluding these three domains). The C-terminal half of nsP3 is not conserved in alphaviruses (Strauss et al., 1988) and the 6K protein exhibits relatively low conservation so that the divergence in these regions is probably not significant. However, E2 is a major antigenic determinant, and changes in E2 have been associated with changes in virulence (Lustig et al., 1988; Olmsted et al., 1986; Strauss et al., 1991; Tucker and Griffin, 1991).

These changes in glycoprotein E2 are examined in more detail in Table 2. Differences in E2 among six different strains of Sindbis virus (Ockelbo virus, four strains of Sindbis virus derived from isolate AR339, isolated in Sindbis, Egypt in 1952, and Sindbis isolate SAAR86, isolated in South Africa in 1963) are shown. The residues at positions 172, 209, 212, and 216 are known to be important determinants of the antigenicity of the virus (Strauss et al., 1991), and the differences at these positions have important implications for the reactivity of the different viruses with neutralizing antibodies. The residues at 55 and 172 are known to be important determinants of the neurovirulence of the virus in a mouse model (Lustig et al., 1988), and it is possible that the

FIGURE 1A

nsP1

1 ①-AUUGCGGGUAGUACACACUUAUGAAUCAAAACAGCCGACCAAUUGCACUACCAUCACAUAUGGAGAAAGCCAGUAGUUAACUGUAGACCCUCAGAGUCCGUUUGUGGUGCAACUG 20
21 QAKSFFPQGFEVVAAGQA TPNDHANARAFVSHLSKLI ELEVPTT 60
120 CAAAAGSGFCCGAAUUGAGGUAAGCAGCAGCACUCCCAAUGGACCAUUAUGCCAGAGAUUUUCGCAUCUGGCCAGUAAACUAUUCUGGAGGUGUCCU:CCACA 239

nsP2

1 ALVETVETPRGGVRIIPQGANADRMIGGQVIVVSPS VLVKKNACKLAP 40
1580 GCACGCGUGAAACCCCGCGGCUAGUAAAGGAUUAUACCCUAAGCAAUGACCPLUAGUAGCCAGUUAUUGGUGUCCCLAAACCCUCUGUGUGAAGAAAGCCGCAACCGACCA 1799

nsP3

801 TRGGVGGAA PSYRRTKRENIAADCQEEA VVNAANPLG RPPGEGV 33
4090 ACAAGAGAGGAGUGGAGGCCGACCCGUUACCGCAGUAAAAGGGAGAAAUUGGUGAGUGUCAGUAGGAGGAAAGGAGUUGUCAUUGCAGCCAAUUCGGUGGCCAGACCAGGCGAGGAC 4199

34 CRAIYKRWPN SFTDASA TETGTAK T V C GK K V I H A V G P D F 73
4200 UGCGI GCCAUUUAACGUGCCGACAGUUAUACGCAACATAGACAGCCGCAAAUGCAGUAGGAGGAGUAGGAGGAGGAGUAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4319

74 RYHPEAEGAAEKLLQNAYHAVAADLVNEHN I K SVAIPL L S T G I 113
4320 CGGAAACUCPEAGAAAGCCUGAAUUGCGCAAAAACCGUACCAUGGAGGCAGACUGVUAUAGAACAAUUAUAGGUCUGUCCCAACUUGCUGUACCAUGGAGU 4439

114 YAAAGKADRLLEVSLNCLTALDRDAGAACUGGAG 153
4440 UACGCGCCGAAAGACCCGCUAGAUAUUAUACUUAUACUGGAG 4559

154 LQLKESVTELEKDEDEMEI DDELVM I H F D S C I K G R K G F S T 193
4560 GCGUCCCAACUCAGGAGUCUGUAACAGAGCUGAAGGAGUAGGAGUAGGAGUAGGAGUAGGAGUAGGAGUAGGAGUAGGAGUAGGAGUAGGAGUAGGAGUAGGAGUAGGAG 4679

194 XGKGLS Y F E G T K F H Q A A K D M A E I K V L F P N D G E A N E E A A S A C A G G A A A G C A A G G A A C A A C U G U G G C 233
4580 ACUAAAGAAUGUUAUCUUCGACAGGACCACCAAAUUCACUAAGCAAGAG 4799

234 Y I L G E T M E A I R E K K C P V D H N F S S S P P K I L P C L C M Y A M T P F R 273
4800 UACAUUUGGAGAGACUAUGAAGCAUUCGCGAAAAUUGCCGGUAGCACCACCGGUGUCUAGCCGCAAAAACCGUGCGUGUUGUUGAUGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4919


```

402 L A P N A V I P T S L A L L L C V R S A N A E T F T E T M S Y F W S N S Q P F F 18
9840 CUGGCCCAAAGCCGUAUCCAAUCGUCUGGCAUCUCUGUCGUGUUAGGUCGGCAAACGGCUGAAACGUUACCCGACCAUGAUUUGUGGUAACAGCCAGCCGUUCUUC 9859

19 W V Q L C I P L A A V I V L M R C C S C C L P F L V V A G A Y L A K V D A Y E H 3
9960 UGGGUCCAGUGUGCAUACCCUUGGCCUCUCAUCGUCUAAUUGCGCUGUUGCUCUUGCCUGCCUUUUUAGUGGUUGCCGGCCUACCCUGGGAGGUAAGACGCCUACGAAACAU 10079

4 A T T V P N A V P Q I P Y K A L V E R A G Y A P L N L E I T V M S S E V L P S T N 43
10080 GCGACCACUGUCCAAUUGUGCCACAGAUACCGUAUAAGGCACUUGUUGAAAGGGCAGGGUACGGCCCGCUCAAUCUGGAGAUUACUGUCAUUCUCCGGAGGUUUACCUUCCACCAAC 10199

44 G E Y I T C K F T T V V P S P K V K C C G S L E C Q P A A H A D Y T C K V F G G 83
10200 CAAGAGUACAUCACUUGUAAAUCACUACCGUGGUCCCCCCCCCAAAGUCAAUUGCUGCGCCUUGGAAUUGCAGCCCGCCGUCACGGCAGACUUAACCCUGCAAGGUCUUUGGAGGG 10319

84 V Y P F M W G G A Q C F C D S E N S Q M S E A Y V E L S A D C A T D H A Q A I K 123
10320 GUGUACCCUUGAUGUGGGGAGGAGCCAAUGUUUUUGCGACAGUGAGAACAGCCAGUAGUGAGGCGUACGUCGAAUUGUCAGCAGAUUGCGCGACUGACCACCGCGAGGGCAUUAAG 10439

124 V H T A A A M K V G L R I V Y G N T T S F L D V Y V N G V T P G T S K D L K V I A 163
10440 GUACAUACUGCCGCAUGAAGGAAGGAGGUGUAGGUGUACGGGAACACUACCCAGUUCUUAUGAUGUGUACGUGAACGGAGUACACAGGAACGUCUAAAGACUGAAAGUCUAAGCU 10559

164 G P I S A S F T P F D H K V V I H R G L V Y N Y D F P E Y G A M K P G V F G D I 203
10560 GGACCGAUUUUCAGCAUCGUUUACCAUUCGUAUCAGGUCGUUAUCCUACUGCGCGUGGUGUACAACUUAUGACUUCGCCGAAUACGGAGCGAUGAAACCAAGAGUGUUUGGAGACAUU 10679

204 G A T S L T S K D L I A S T D I R L L K P S A K N V H V P Y T Q A S G F E M W 243
10680 CAAGCUACCCUUGACUAGCAAAGACCUUCAGCCACGACAGACAUUAGGCUACUUAAGCCUUCGCCAAGAACGUGGCAUGUCCGUACACCGCAGGCCCGCAUCAGGAUUGGAGAUUGGG 10799

244 K M N S G R P L Q E T A P F F G C K I A V N P L R A V D C S Y G N I P I S I D I P 283
10800 AAAACAACUCAGCCCGCCACUGCAGGAAACCGCCCUUUUCGGGUGAAAGUUGCGAGUCAAUCCCGUUGGAGCGGUGGACUGCUAUACGGGAACAUUCCAUUUUCUAUUGACAUCCCG 10919

284 N A A F I R T S D A P L V S T V K C D V S E C T Y S A D F G G M A T L G Y V S D 323
10920 AACGCGCCUUUUAUGAGGACAUCAGGACCAACUGGUCACCAACAGUCAAAUUGUAGUAGUGAGUGAGUGCCUUAUUCAGCAGACUUCGCGGGAGUGCCACCCUGCAGUAUVSUDDGAC 11039

324 R E G Q C C P V H S H S S T A T L Q E S T V H V L E K G A V T V H F S T A S P Q A 363
11040 CCGCAAGGCAAAUGCCCGUACAUUCGCAUUCAGCACAGCAACCGUCCAAGAGUGCAGACUUAUGUCCUGGAGAAAGGAGCGGUGACAGUACACUUCAGCACCGCGAGUCCACAGGGC 11159

384 N F I V S L C G K K I T C N A E C K P A D H I V S T P H K N D Q E F G A A I S 403
11160 AACUUCAUUGUAGCUGGCGGGAAGAGACAACAUAGCAAGGAAUUGUAAACCACCACUGACCAUAGUGGAGCACCCCGCACAAAAGACCAAGAAUUCAGCCGCCAUCUCA 11279

404 K T S W S W L F A L F G G A S S L L I I G L I F A C S M M L T S T R R Op 11399
11280 AAAACAUCCAUGGAGUUGGUGUUUGCCUUUUUCGGCGCCUUCGUCGCUAUUUUUUAGGACUUACGAUUUUUGCUUGCAGCAUGAUGCUGACUAGCACAGCAAGAUUCGCGUACCG 11519

11400 CCCCAAUGAGCCGACCAAGCAAAACUCGAGUAUCUUCGGAGGAAACUGAUGGCAUAAUGCAUCAGCCUGGUAUUAUAGAUCCCGCUAACUGCGGGCAUUUAGCAACACCAAAACUCGAGG 11519

11520 UACUCCGAGGAAGCGCAGUGCAUAAUGCUGCGCGAGUGGECACAUUAUCACUAUAUAACCAUUUUAAGCGGACCGCCGAAACUCAAUGUAUUUCUGAGGAAGCAUGGUGCAUAAUG 11639

11840 CCAUGCAGCGUCUGCAAAUUUUUUAAUUUUUUUUAUAAUCAAAAUUUUUUUUUAAAUUUUC 11708

```

1. Complete sequence of the Ockelbo virus. The sequence is shown from 5' to 3' and translated using the single letter amino acid code. Nucleotides different from those in HRSP are underlined, changed amino acids are boxed. Deletions relative to HR are indicated by solid triangles pointing upward and the number of residues deleted. Insertions have both amino acids and nucleotides boxed together, and an open triangle pointing downward. Termination codons are labeled Am (Amber, UAG) or Op (Opal, UGA) as appropriate. Nucleotides are numbered 5' to 3'; no acid numbering begins again at the beginning of each final protein product.

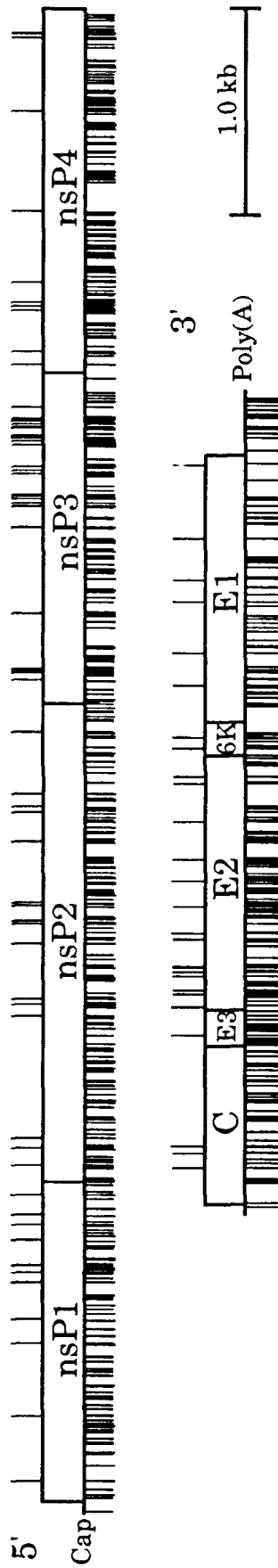


Fig 2. Schematic representation of the differences between HRSP and Ockelbo. The translated portions of the genome are shown as open boxes. Ticks above the box indicate changes in amino acids; ticks below indicate changes in nucleotides.

Table 2. Amino Acid Substitutions in Glycoprotein E2 in Sindbis Strains

RESIDUE	AR339				S.A. AR86 ^e	OCKELBO ^f
	HRSP ^a	DG ^b	AS ^c	RJ ^d		
1	S	S	S	R	S	S
3	I	T	T	T	T	T
23	V	E	E	E	E	E
29	V	V	V	V	I	I
55	Q	Q	Q	Q	Q	K
61	A	A	A	A	S	T
69	L	L	L	L	L	F
70	K	K	E	E	E	E
116	V	V	V	V	A	A
126	L	L	L	L	M	M
172	R	G	G	G	G	G
209	G	R	G	G	G	G
212	S	S	S	S	T	T
216	E	E	K	E	E	E
243	L	L	L	L	S	L
247	D	D	D	D	A	A
277	I	I	I	I	V	I
312	V	V	V	V	I	I
375	T	T	T	T	A	A
386	V	V	V	V	A	A

^a Sequence of HRSP is from Strauss *et al.* (1984).

^b Sequence is from Lustig *et al.* (1988) of the SV1A strain from the laboratory of Diane Griffin.

^c Sequence from Strauss *et al.* (1991 in press) of the strain used by A. Schmaljohn for the isolation of antigenic variants (Stec *et al.* 1986).

^d Sequence from Davis *et al.* (1986) of the laboratory strain of Robert Johnston.

^e Sequence of the SA AR86 strain from Russell *et al.* (1989).

^f Sequence from Shirako *et al.* (1991).

amino acid difference at position 55 between Ockelbo and the other Sindbis strains might be responsible in part for the increased virulence of Ockelbo virus compared to other strains of Sindbis virus.

3' Terminal Nucleotide Sequence of Other Strains of Sindbis Virus

To ascertain the relationships among Sindbis virus strains present in nature, RNA sequence was obtained for a number of isolates of Sindbis virus that differ in their geographic source and in their disease symptomology in man (the strains and their source are shown in Table 1). The 420 nucleotides at the 3'-terminus were sequenced for each isolate. These sequences are shown in Figure 3. The sequence identity throughout this region is greater than 80% for all viruses examined, and the sequence organization is identical in all cases except for a few scattered nucleotide insertions and deletions. In the 3' nontranslated region there are three highly conserved repeats that are 40 nucleotide long (boxed in the figure). The sequences within these repeated elements are more highly conserved than the sequences outside these elements. As an example of this, comparing the Australian and AR339 strains, there are 49 differences in the 3' nontranslated region outside the repeated elements (24.1% divergence), but only 7 changes within the elements (5.8% divergence), for an overall divergence of 18.1%. Similarly, comparing the Ockelbo '82 isolate with AR339, there are 13 changes outside the repeated elements (6.4% divergence) but only 3 within the elements (2.5% divergence), for an overall divergence of 4%.

The relationships among the virus sequences are shown diagrammatically in Figure 4. From this figure the number of nucleotide differences in the 3' terminal 420 nucleotides between any two strains of Sindbis virus can be computed. Three points are immediately obvious from a study of this diagram. One is that the Sindbis strains analyzed can be divided into a European-African group and an Asian-Australian group. The Asian-Australian group differs from the European-African group in 17% of the nucleotides sequenced, whereas within the European-

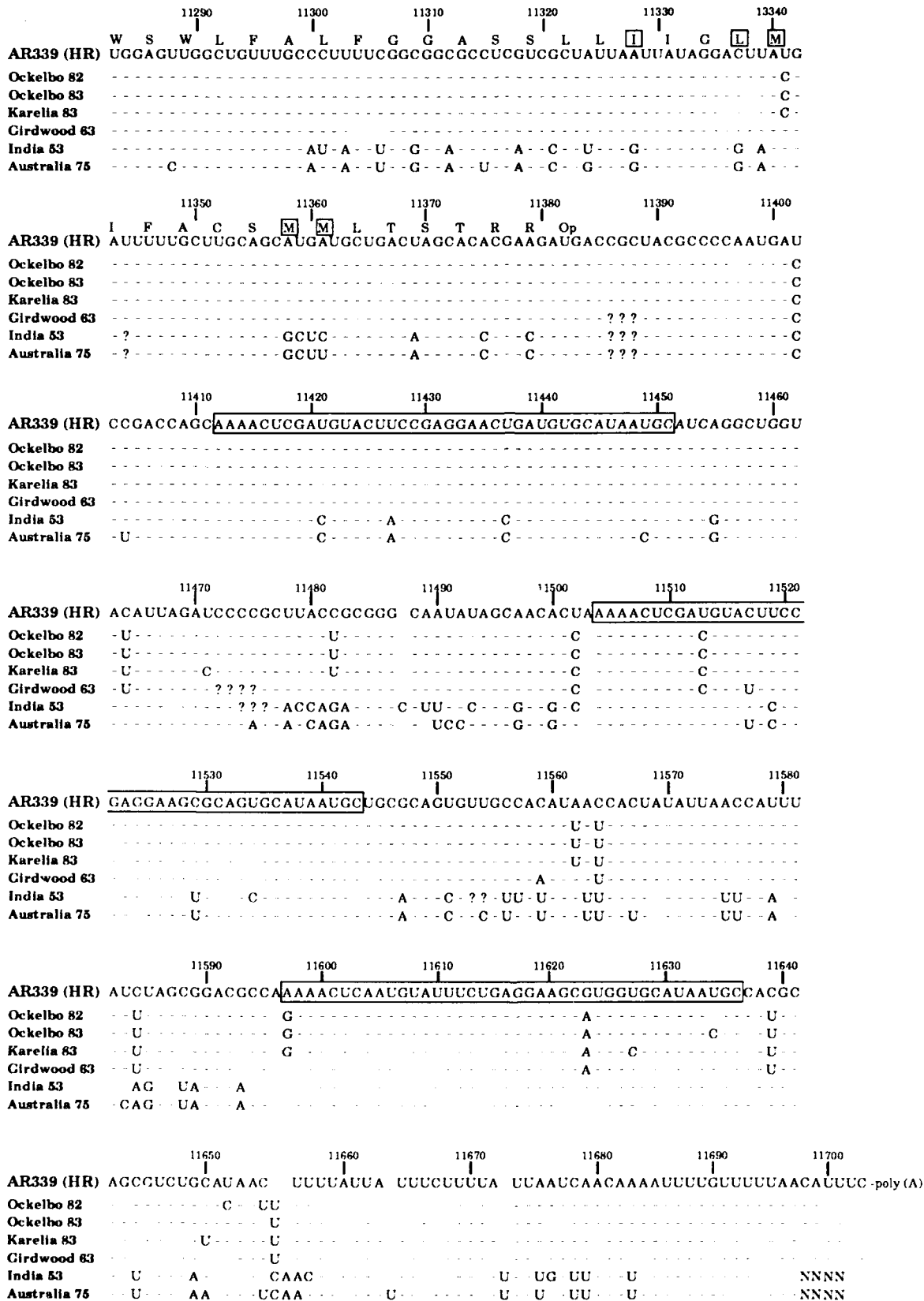


Figure 3. Sequence of the 3' end of several Sindbis viruses. The sequences of Ockelbo 83M107, Karelian fever, and the South African Girdwood were determined from cloned cDNA. Those of Indian A1036 and Australian MRM18520 were determined directly from RNA by dideoxy sequencing using T₁₂GA primer. The sequence of Ockelbo 82 is from Fig. 1 and that for AR339 (HRSP) is from Strauss *et al.* (1984). Gaps have been introduced as necessary to maintain the alignment. Three repeated sequence elements of 40 nucleotides are boxed. The translated amino acid sequence and nucleotide numbers are for AR339 (HR); and amino acid that differs in any of the other viruses is boxed.

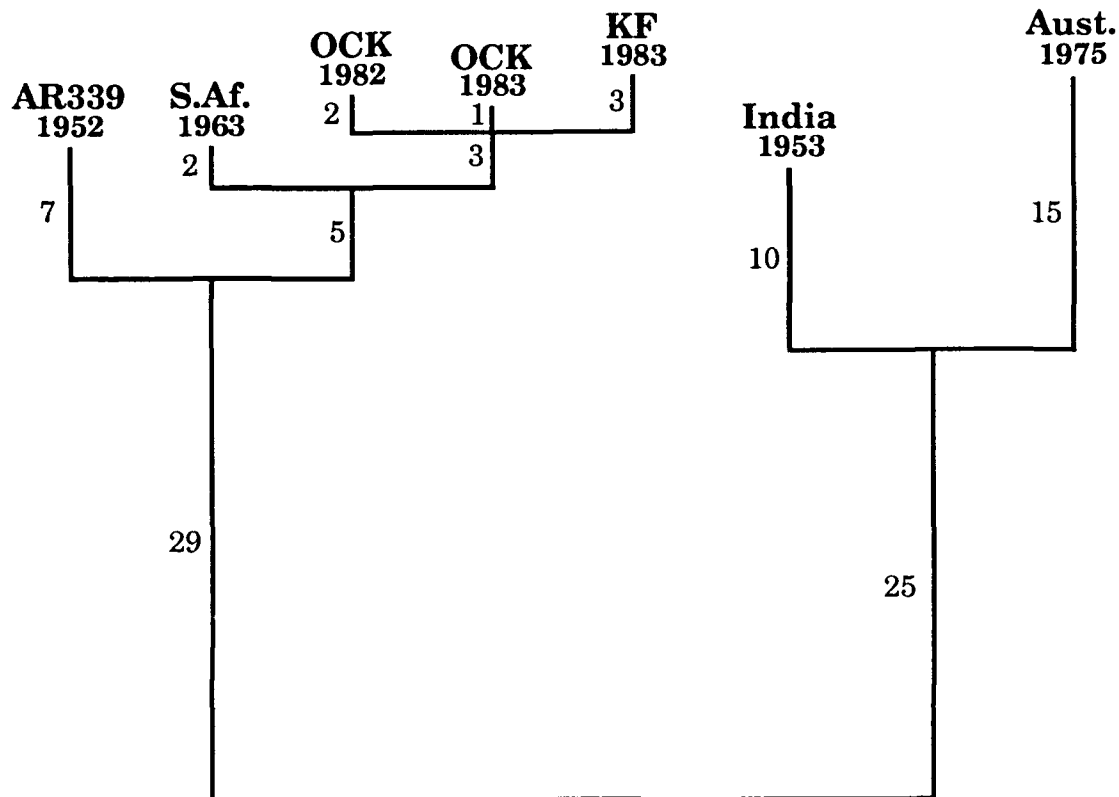


Figure 4. Relationships among strains of Sindbis virus. The vertical distances indicate the number of nucleotide differences between any two strains in the 3' terminal 420 nucleotides. The horizontal distances are arbitrary. Nucleotide differences between any two strains can be determined quite accurately by summing the numbers on the vertical branches between them.

African group the maximum divergence is about 5%. The second obvious point is the relationship between the Karelian fever and Ockelbo viruses. These are different isolates of what is effectively the same virus. The third point is that Ockelbo virus is more closely related to the South African strain isolated in 1963 than it is to the Egyptian strain isolated in 1952, which has important implications for the origin and spread of Ockelbo disease.

Conclusions

Ockelbo virus and Karelian fever virus are the same virus, and we assume that this same strain of Sindbis virus also causes Pogosta disease. The Ockelbo strain of Sindbis virus is very closely related to the South African strains of Sindbis virus, more so than to other European-African subgroup Sindbis viruses such as Egyptian AR339. This close relationship is evident either from comparisons of glycoprotein E2 among various Sindbis viruses, or from comparisons of 3'-terminal sequences. The closeness of this relationship is illustrated by the fact that in the 3' nontranslated region, Ockelbo and the South African strain examined demonstrate a sequence divergence of only 4%. Since these viruses were isolated twenty years apart, the maximum rate of sequence divergence in this region could be no more than 0.2% per year. Since it seems unlikely that the Girdwood strain of South African Sindbis is the direct ancestor of Ockelbo, the actual rate of divergence is probably less. Such a divergence rate is low in comparison to rates established for a number of other RNA viruses (Steinhauer and Holland, 1987; Strauss and Strauss, 1988). Thus, it seems clear that South African and Northern European strains of Sindbis virus have not been separated for long. It is also of note that the South African strains of Sindbis virus have been implicated in human disease, as has Ockelbo virus. It seems likely that Ockelbo originated in South Africa and was introduced into Sweden by migratory birds or by the activities of man, perhaps in the 1960's, and then spread to Finland and the Karelian region of the Soviet Union in the 1980's. Thus, it appears that virulent strains of Sindbis virus have the potential to spread to

other regions of the world and cause epidemics of febrile illness that can be of moderate severity because of associated arthralgia.

It is also important that repeats of a sequence element found in the 3' nontranslated region of Sindbis viruses are much more highly conserved than sequences outside these elements. The function of the 3' nontranslated region, and in particular the repeated sequence elements, is unknown, but the conservation of these elements among Sindbis viruses makes clear that they play an important role in viral replication, and presumably are important for viral RNA replication. We have previously shown that site-specific mutants in the 3' nontranslated region have different effects in mosquito cells and chicken cells, implying that host cell proteins bind to this region to promote replication (Kuhn et al., 1990). The fact that the repeated elements are present in three copies and exhibit a high degree of conservation suggest that they might serve as binding sites for interactions with host cell proteins. These repeated sequence elements differ among different alphaviruses, but equivalent sequences are found in all alphaviruses. Conserved sequence elements could be useful for diagnostic purposes, although other conserved domains of the virus are probably more useful.

References

- Calisher, C. H., Meurman, O., Brummer-Korvenkontio, M., Halonen, P. E., and Muth, D. J. (1985). Sensitive enzyme immunoassay for detecting immunoglobulin M antibodies to Sindbis virus and further evidence that Pogosta disease is caused by a western equine encephalitis complex virus. *J. Clin. Microbiol.* **22**, 566-571.
- Chamberlain, R. W. (1980). Epidemiology of arthropod-borne togaviruses: The role of arthropods as hosts and vectors and of vertebrate hosts in natural transmission cycles. In "The Togaviruses" (R. W. Schlesinger, Ed.), pp. 175-227. Academic Press, New York.
- Davis, N. L., Fuller, F. J., Dougherty, W. G., Olmsted, R. A., and Johnston, R. E. (1986). A single nucleotide change in the E2 glycoprotein gene of Sindbis virus affects penetration

- rate in cell culture and virulence in neonatal mice. *Proc. Natl. Acad. Sci. USA* **83**, 6771-6775.
- Griffin, D. E. (1986). Alphavirus pathogenesis and immunity. In "The Togaviridae and Flaviviridae" (S. Schlesinger and M. J. Schlesinger, Ed.), pp. 209-250. Plenum Publishing Corp., New York.
- Gubler, U., and Hoffman, B. J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263-269.
- Hahn, C. S., Strauss, E. G., and Strauss, J. H. (1989). Dideoxy sequencing of RNA using reverse transcriptase. *Methods in Enzymology* **180**, 121-130.
- Kuhn, R. J., Hong, Z., and Strauss, J. H. (1990). Mutagenesis of the 3' nontranslated region of Sindbis virus RNA. *J. Virol.* **64**, 1465-1476.
- Lustig, S., Jackson, A., Hahn, C. S., Griffin, D. E., Strauss, E. G., and Strauss, J. H. (1988). Molecular basis of Sindbis virus neurovirulence in mice. *J. Virol.* **62**, 2329-2336.
- Lvov, D. K., Berezina, L. K., Yakovlev, B. I., Aristova, V. A., Gushchina, E. L., Lvov, S. D., Myasnikova, I. A., Skvortsova, T. M., Gromashevsky, V. L., Gushchin, B. V., Sidorova, G. A., Klimenko, S. M., Khutoretskaya, N. V., and Khizhnyakova, T. M. (1984). Isolation of Karelian fever agent from *Aedes communis* mosquitoes. *Lancet* **2**, 399-400.
- Lvov, D. K., Vladimirtseva, E. A., Butenko, A. M., Karabatsos, N., Trent, D. W., and Calisher, C. H. (1988). Identity of Karelian fever and Ockelbo viruses determined by serum dilution-plaque reduction neutralization tests and oligonucleotide mapping. *Am J Trop Med Hyg* **39**, 607-610.
- Malherbe, H., Strickland-Cholmley, M., and Jackson, A. L. (1963). Sindbis virus infection in man: Report of a case with recovery of virus from skin lesions. *S. Afr. Med. J.* **37**, 547-552.
- Niklasson, B. (1988). Sindbis and Sindbis-like viruses. In "The Arboviruses, Epidemiology and Ecology" (T. P. Monath, Ed.), pp. 167-176. CRC Press, Inc., Boca Raton, FL.

- Niklasson, B., Aspmark, A., LeDuck, J. W., Gargan, T. P., Ennis, W. A., Tesh, R. B., and A. J. Main, J. (1984). Association of a Sindbis-like virus with Ockelbo disease in Sweden. *Am. J. Trop. Med. Hyg.* **33**, 1212-1217.
- Olmsted, R. A., Meyer, W. J., and Johnston, R. E. (1986). Characterization of Sindbis virus epitopes important for penetration in cell culture and pathogenesis in animals. *Virology* **148**, 245-254.
- Ou, J.-H., Strauss, E. G., and Strauss, J. H. (1981). Comparative studies of the 3'-terminal sequences of several alphavirus RNAs. *Virology* **109**, 281-289.
- Peters, C. J., and Dalrymple, J. M. (1990). Alphaviruses. In "Virology" (B. N. Fields and D. M. Knipe, Ed.), pp. 713-761. Raven Press, New York.
- Rice, C., M., Lenches, E. M., Eddy, S. R., Shin, S. J., Sheets, R. L., and Strauss, J. H. (1985). Nucleotide sequence of yellow fever virus: Implications for flavivirus gene expression and evolution. *Science* **229**, 726-733.
- Russell, D. L., Dalrymple, J. M., and Johnston, R. E. (1989). Sindbis virus mutations which coordinately affect glycoprotein processing, penetration and virulence in mice. *J. Virol.* **63**, 1619-1629.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Second, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Shah, K. V., Johnson, H. N., Rao, T. R., Rajagopalan, P. K., and Lamba, B. S. (1960). Isolation of five strains of Sindbis virus in India. *Ind. J. Med. Res.* **48**, 300-308.
- Shirako, Y., Niklasson, B., Dalrymple, J. M., Strauss, E. G., and Strauss, J. H. (1991). Structure of the Ockelbo virus genome and its relationship to other Sindbis viruses. *Virology* **182**, 753-764.
- Shirako, Y., and Strauss, J. H. (1990). Cleavage between nsP1 and nsP2 initiates the processing pathway of Sindbis virus nonstructural polyprotein P123. *Virology* **177**, 54-64.

- Stec, D. S., Waddell, A., Schmaljohn, C. S., Cole, G. A., and Schmaljohn, A. L. (1986). Antibody-selected variation and reversion in Sindbis Virus neutralization epitopes. *J. Virol.* **57**, 715-720.
- Steinhauer, D. A., and Holland, J. J. (1987). Rapid evolution of RNA viruses. *Ann. Revs. Microbiol.* **41**, 409-433.
- Strauss, E. G., Levinson, R., Rice, C. M., Dalrymple, J. M., and Strauss, J. H. (1988). Nonstructural proteins nsP3 and nsP4 of Ross River and O'Nyong-nyong viruses: Sequence and comparison with those of other alphaviruses. *Virology* **164**, 265-274.
- Strauss, E. G., Rice, C. M., and Strauss, J. H. (1984). Complete nucleotide sequence of the genomic RNA of Sindbis virus. *Virology* **133**, 92-110.
- Strauss, E. G., Stec, D. S., Schmaljohn, A. L., and Strauss, J. H. (1991). Identification of antigenically important domains in the glycoproteins of Sindbis virus by analysis of antibody escape variants. *J. Virol.* **in press**,
- Strauss, J. H., and Strauss, E. G. (1988). Evolution of RNA viruses. *Ann. Rev. Microbiol.* **42**, 657-683.
- Taylor, R. M., Hurlbut, H. S., Work, T. H., Kingston, J. R., and Frothingham, T. E. (1955). Sindbis virus: A newly recognized arthropod-transmitted virus. *Am. J. Trop. Med. Hyg.* **4**, 844-862.
- Tucker, P. C., and Griffin, D. E. (1991). Mechanism of altered Sindbis virus neurovirulence associated with a single-amino-acid change in the E2 glycoprotein. *J. Virol.* **65**, 1551-1557.