Realia et naturalia

DISSERTATIONES
TECHNOLOGIAE
UNIVERSITATIS
TARTUENSIS
22

SIRLE SAUL

Towards understanding the neurovirulence of Semliki Forest virus

DISSERTATIONES TECHNOLOGIAE UNIVERSITATIS TARTUENSIS

22

DISSERTATIONES TECHNOLOGIAE UNIVERSITATIS TARTUENSIS
22

SIRLE SAUL

Towards understanding the neurovirulence of Semliki Forest virus

UNIVERSITY OF TARTU
PRESS
1632

Institute of Technology, Faculty of Science and Technology, University of Tartu, Estonia.

Dissertation was accepted for the commencement of the degree of Doctor of Philosophy in biomedical technology on November 5th, 2015 by the Council of Institute of Technology, Faculty of Science and Technology, University of Tartu, Estonia.

Supervisor:	Andres Merits, PhD, Professor of Applied Virology, Institute of Technology, University of Tartu, Estonia
Co-supervisor:	Eva Žusinaite, MD, Senior Research Fellow in Applied Virology, Institute of Technology, University of Tartu, Estonia
Reviewer:	Kaido Kurrikoff, PhD, Senior Research Fellow in Applied Virology, Institute of Technology, University of Tartu, Estonia
Opponent:	Diane E. Griffin, MD, PhD, Vice President, U.S. National Academy of Sciences University Distinguished Service Professor W. Harry Feinstone Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, USA
Commencement:	Auditorium 121, Nooruse 1, Tartu, Estonia, at 12.15 on December 15th, 2015

Publication of this thesis is granted by the Institute of Technology, University of Tartu and by the Graduate School of Biomedicine and Biotechnology created under auspices of European Social Fund.

ISSN 2228-0855
ISBN 978-9949-32-996-0 (print)
ISBN 978-9949-32-997-7 (pdf)

University of Tartu Press
www.tyk.ee

CONTENTS

LIST OF ORIGINAL PUBLICATIONS 7
LIST OF ABBREVIATIONS 8
1. INTRODUCTION 10
2. REVIEW OF LITERATURE 11
2.1. Alphaviruses 11
2.2. Virion 12
2.3. Genome organization 14
2.4. Alphavirus infection cycle 16
2.4.1. Virus entry 16
2.4.2. RNA replication 17
2.4.3. Sites of replication 19
2.4.4. Synthesis of structural proteins 20
2.4.5 Virion assembly and budding 20
2.5. Functions of individual ns proteins 21
2.6. Virus-host interactions 26
2.6.1. Shutdown of host cell transcription and translation 26
2.6.2. Host factors associated with alphavirus RCs 27
2.7. Pathogenesis of SFV infections 29
2.7.1. Strains of SFV 29
2.7.2. Neuroinvasion and CNS tropism 31
2.7.3. Determinants of SFV virulence 32
2.7.4. Immune response in the mouse CNS 33
3. AIMS OF THE STUDY 35
4. MATERIALS AND METHODS 36
5. RESULTS AND DISCUSSION 38
5.1. The ability of SFV to enter the CNS is determined by the charged amino acid residues on the surface of E2 glycoprotein (I) 38
5.1.1. The phenotypic differences between SFV4 and L10 map to E2 position 162 38
5.1.2. The biological stock of L10 is heterogeneous at position 247 of E2 40
5.1.3. Charged amino acid residues at positions 162 and 247 of E2 modulate the binding of SFV virions to heparan sulfate 41
5.1.4. Amino acid residues 162 and 247 in E2 affect SFV replication in the mouse brain following i.c. inoculation and the effect of virus infection on the integrity of a BBB model 42
5.2. The neurovirulence of SFV is affected by differences in determinants of ns polyprotein processing and by the sequence of nsP3 (II) 43

5.2.1. A single amino acid difference from the A7(74) consensus sequence causes low infectivity of rA774 43
5.2.2. The introduction of synonymous changes reduces recombination between two copies of nsP3-encoding sequences 44
5.2.3. Neurovirulence is the dominant function of nsP3 of SFV6 45
5.2.4. Differences in the P4 position of the 1/2 cleavage site of P1234 and the S4 subsite of nsP2 protease affect the processing of ns polyprotein and virulence 47
5.2.5. The role of nsP3 in SFV neurovirulence and replication 48
5.3. The role of the type I IFN response in SFV neurovirulence (unpublished) 49
5.4. Magnetic fractionation and proteomic dissection provide a useful method to study the composition of SFV replication organelles (III) 54
5.4.1. Vesicles carrying functionally active RCs can be purified via magnetic enrichment 55
5.4.2. The proteome of magnetically isolated fractions can be characterized by using quantitative proteomics approach 55
5.4.3. PCBP1, hnRNP M, hnRNP C and hnRNP K affect alphavirus infection 56
6. CONCLUSIONS 58
REFERENCES 60
SUMMARY IN ESTONIAN 74
ACKNOWLEDGEMENTS 76
PUBLICATIONS 77
CURRICULUM VITAE 133
ELULOOKIRJELDUS 134

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications that are referred to in the text by their Roman numerals:

I. Ferguson MC*, **Saul S***, Fragkoudis R, Weisheit S, Cox J, Patabendige A, Sherwood K, Watson M, Merits A, Fazakerley JK. 2015. The ability of the encephalitic arbovirus Semliki Forest virus to cross the blood brain barrier is determined by the charge of the E2 glycoprotein. J Virol 89:7536–7549.
*Authors contributed equally to this work

II. **Saul S**, Ferguson MC, Cordonin C, Fragkoudis R, Ool M, Tamberg N, Sherwood K, Fazakerley JK, Merits A. 2015. Differences in processing determinants of nonstructural polyprotein and in the sequence of nonstructural protein 3 affect neurovirulence of Semliki Forest virus. J Virol 89:11030–11045.

III. Varjak M, **Saul S**, Arike L, Lulla A, Peil L, Merits A. 2013. Magnetic fractionation and proteomic dissection of cellular organelles occupied by the late replication complexes of Semliki Forest virus. J Virol 87:10295–10312.

The papers listed above are reprinted with the permission of the copyright owners.
Some unpublished data is also presented.

My personal contribution to the papers listed above is as follows:

I. I participated in the experimental design and, together with M. Ferguson, performed most of the experiments; I analyzed the data and helped write the manuscript.

II. I participated in the experimental design and performed most of the experiments; I analyzed the data and wrote the manuscript.

III. I performed the siRNA silencing experiments with SINV and CHIKV. I participated in the growth curve and western blot experiments and analyzed the data.

LIST OF ABBREVIATIONS

BBB	–	blood-brain barrier
cDNA	–	complementary DNA
CHIKV	–	Chikungunya virus
CNS	–	central nervous system
CPV-I	–	type I cytopathic vacuole (alphavirus replication organelle)
CSE	–	conserved sequence element
dsRNA	–	double-stranded RNA
EEEV	–	Eastern equine encephalitis virus
eIF	–	eukaryotic translation initiation factor
ER	–	endoplasmic reticulum
G3BP	–	Ras-GAP SH3-domain-binding protein
GAG	–	glycosaminoglycan
hnRNP	–	heterogeneous ribonucleoprotein
HS	–	heparan sulfate
HVD	–	hypervariable domain
i.c.	–	intracerebral
i.p.	–	intraperitoneal
icDNA	–	infectious cDNA
IFN	–	interferon
m^7GMP	–	7-methyl-guanosine-5'-monophosphate
MOI	–	multiplicity of infection
NLS	–	nuclear localization signal
NRAMP	–	natural resistance-associated macrophage protein
ns	–	nonstructural
nsP	–	nonstructural protein
NTPase	–	nucleoside triphosphatase
ORF	–	open reading frame
PAMP	–	pathogen-associated molecular pattern
PFU	–	plaque forming unit
PI3K-Akt-mTOR	–	phosphatidylinositol-3-kinase-Akt-mammalian target of Rapamycin
PID	–	postinoculation day
PKR	–	protein kinase R, dsRNA-activated protein kinase
RC	–	replicase complex (spherule)
RdRp	–	RNA-dependent RNA polymerase
RIG-I	–	retinoic acid-inducible gene I
RRV	–	Ross River virus
RTPase	–	RNA trisphosphatase
SCID	–	severe combined immunodeficiency
SFV	–	Semliki Forest virus
sg	–	subgenomic

SILAC	–	stable isotope labeling with amino acids in cell culture
SINV	–	Sindbis virus
TATase	–	terminal adenylyltransferase
TF	–	TransFrame protein
UTR	–	untranslated region
VEEV	–	Venezuelan equine encephalitis virus
wt	–	wild type
Y2H	–	yeast two-hybrid system
ZBD	–	zinc binding domain, also known as alphavirus unique domain (AUD)

1. INTRODUCTION

Viruses are obligate intracellular parasites that hijack the host cell's machinery to replicate. They exhibit many different shapes and are present wherever there are cells to infect. In fact, viruses are the most common biological units on Earth, outnumbering all other types combined. One main motivation for the study of viruses is that they cause many important infectious diseases, and some have been shown to contribute to the development of certain forms of cancer.

Alphaviruses are no exception, as this genus includes viruses pathogenic to a wide variety of animals, including humans, causing a spectrum of diseases that ranges from unpleasant cold-like illness and arthritis to fatal encephalitis. Thus, the study of alphaviruses is an important field of investigation. A well-studied virus belonging to this genus is Semliki Forest virus (SFV). Laboratory strains of SFV are relatively harmless to humans and thus have been utilized extensively in genetic engineering. They also provide a well-characterized model system to investigate the pathogenesis of viral encephalitis.

Alphavirus infection consists of an intricate interplay between the virus and the host. The goal of the host is to eradicate the virus and survive, while the goal of the virus is to continue to proliferate regardless of the cost to the host cell. Importantly, viral infection is not equivalent to virus-induced pathogenesis. Alphaviruses can infect a large variety of organisms (from nematodes to humans), but they cause diseases in only a few of their hosts. Investigators have sought to identify the features of viral virulence and host defense that determine the outcome of infection. Different strains of SFV, which are designated as virulent or avirulent according to their effects in adult mice, have provided insight regarding this topic.

The aim of the present study was to assess the determinants and mechanism(s) underlying the contrasting neurovirulence of different strains of SFV. New molecular clones were constructed and used to identify viral factors that contribute to infection. The ability of SFV to enter the central nervous system was found to depend on charged amino acid residues on the surface of viral glycoprotein E2. The importance of nonstructural protein 3 and the rate of nonstructural polyprotein processing in SFV neurovirulence were also investigated. Furthermore, we briefly reviewed the importance of cellular host factors and defense systems that play a role in alphavirus infection.

2. REVIEW OF LITERATURE

The Baltimore classification system places viruses into one of seven groups based on their genome (DNA, RNA) and method of replication. Single-stranded RNA viruses are classified according to the polarity of their genomic RNA as positive-strand and negative-strand viruses (group IV and V, respectively) (1). Positive-strand RNA viruses contain messenger-sense RNA in their virions, they represent more than one-third of known virus genera and include many important human, animal, and plant pathogens. Among the members of this group is the family *Togaviridae*, which consists of two virus genera: the *Alphavirus* and the *Rubivirus*. The sole member of the *Rubivirus* genus is the rubella virus, an airborne virus that only infects humans, causing a disease referred to as "three-day measles". The *Alphavirus* genus currently consists of 31 recognized species, including several important human and animal pathogens (2, 3).

2.1. Alphaviruses

Viruses in the genus *Alphavirus* (hereafter referred to as alphaviruses) are mostly arthropod-borne and are transmitted in nature in a classical arbovirus transmission cycle. Mosquitoes typically of the *Aedes* and *Culex* genera carry them between vertebrate reservoir hosts, including birds and rodents. In insects, alphavirus infection is persistent and lifelong and is assumed to be asymptomatic. In vertebrates, the duration of infection is usually short and ends with the death of the infected host or clearance of the virus by the immune system. These differences are recapitulated in the corresponding cell culture systems (4).

Alphaviruses that infect birds and mammals have historically been divided into Old World and New World alphaviruses. Old World alphaviruses, including Chikungunya (CHIKV), Sindbis (SINV), Ross River (RRV), and Semliki Forest virus (SFV), are found in Europe, Asia, Africa, and Australia. The acute phase of disease associated with these viruses is characterized by fever, chills, headache, myalgia, arthralgia, diarrhea, vomiting, and rash. While the mortality associated with these viruses is low, the diseases associated with some of the Old World alphaviruses can be debilitating, with clinical complications that persist from months to years (5). The recent CHIKV epidemics in Southeast Asia, India, Indian Ocean territories, and Caribbean countries have severely affected millions of people (6). CHIKV transmission has also been reported in Europe (Italy, France and Croatia) and in 45 countries or territories throughout the Americas (7). Consequently, great public attention has been directed towards alphaviruses.

New World alphaviruses, including Eastern equine encephalitis (EEEV), Venezuelan equine encephalitis (VEEV), and Western equine encephalitis virus, are found in North and South America. Symptoms similar to those described above may occur during the acute phase of infection, but, as suggested by their names, these viruses can cause acute encephalitis in humans and

domestic animals. Among the New World alphaviruses, EEEV is particularly virulent in humans, with an associated mortality of 50–75% of symptomatic cases. Due to their potential for emergence/reemergence or use as agents of bioterrorism, EEEV, VEEV, and CHIKV have been declared high priority pathogens by the National Institutes of Health. However, because alphavirus infections were not considered to be medically significant until recently, no effective antiviral drugs or licensed vaccines are available for human use against any alphavirus (4, 8–10).

Historically, the two most studied members of the *Alphavirus* genus have been SFV and SINV. Both of these viruses grow to high titers in cell culture, infect cells from a wide range of invertebrate and vertebrate organisms, and are available as laboratory strains that are not typically associated with serious human illness, making them good model systems. For historical reasons, SFV has been mainly studied in Europe, whereas the primary model of alphavirus research in the USA has been SINV. The availability of complementary DNA (cDNA) clones for both of these viruses (11, 12) has made them a useful tool to study viral replication strategies and virus-host relationships. Studies investigating SFV and SINV have shed light on many aspects of the alphavirus infection cycle and have helped to elucidate the basic characteristics of several cellular processes. SFV was the first virus that was shown to enter cells via endocytosis and membrane fusion (13). In addition to tissue culture, SFV and SINV have been advantageous for the investigation of viral infection at the organism level using mice and rats, making them good models of viral pathogenicity.

2.2. Virion

Alphavirus virions are enveloped spherical particles that are 65–70 nm in diameter. Among enveloped virions, they have an extremely regular structure. The single-stranded positive-sense RNA genome is surrounded by 240 copies of capsid protein that is arranged in a T=4 lattice forming the icosahedral nucleocapsid (Fig. 1A) (14). The N-terminal part of the capsid protein is rich in positively charged amino acid residues and is presumed to bind to the genomic RNA. The nucleocapsid is enveloped by the host-derived lipid bilayer, which is enriched with cholesterol and sphingolipids and embedded with the envelope glycoproteins E1 and E2 (Fig. 1A). The two transmembrane glycoproteins E1 and E2 interact to form 240 rigid heterodimers. For CHIKV, the 3D structures of these proteins and that of their dimer have been resolved; this information, together with cryo-electron microscopy data, has resulted in a high-resolution structure of alphavirus virions (15). Three E1-E2 dimers in turn form 80 spike complexes (16). In addition to E1 and E2, lower numbers of smaller proteins (E3, 6K and TransFrame (TF)) are also present in alphavirus virions; however, the abundance of these proteins differs in different alphaviruses. The E2 protein

mediates binding of alphavirus virions to the host cell and is necessary for virion formation because it interacts with the capsid protein (see 2.4.1 and 2.4.5) (17). E1 mediates fusion of the viral and host cell membranes during entry (18). E3, 6K and TF are important for regulating spike assembly and are necessary for efficient budding of the virus (19, 20).

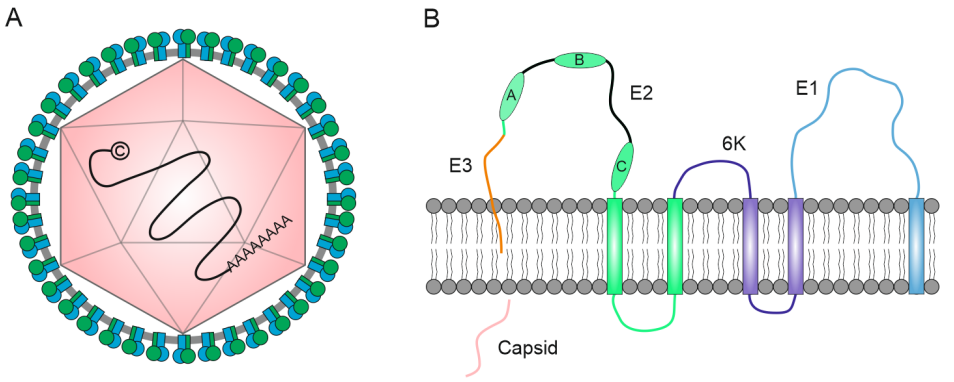

Figure 1. Alphavirus virion and structural proteins. A. The alphavirus virion consists of a nucleocapsid surrounded by a lipid bilayer. The nucleocapsid is composed of a positive-sense single-stranded RNA genome (black) and capsid proteins (pink). The lipid bilayer (grey) is embedded with heterodimers of viral glycoproteins E1 and E2 (blue and green). **B.** Arrangement of structural proteins in cellular membranes. The schematic representation of polypeptide E2 shows domains A, B, and C. The β-ribbon connector between domains A, B, and C is depicted in black.

The E2 protein has a C-terminal transmembrane helix followed by a cytoplasmic domain, which contacts the nucleocapsid (Fig. 1B). This interaction is important for the correct assembly and budding of progeny viruses from the plasma membrane of infected cells (21). The ectodomain of E2 consists of three immunoglobulin-fold domains termed A, B, and C (Fig. 1B). The presence of immunoglobulin folds in E2 is consistent with its function as a cell receptor binding protein. Domain B is at the distal end of the protein to the membrane and domain C is oriented towards the viral membrane. Domain A, the putative receptor-binding domain, is located at the center. In the virion, the E2-E1 heterodimer is oriented such that domains A and B are exposed at the top of the spike and situated at the center and at the periphery, respectively. In the mature virus, domain B covers the fusion loop in the E1 protein. In the linear structure, domain B is connected to domains A and C by long connecting linker peptides (the "β-ribbon connector") (Fig. 1B). An acid-sensitive region in the E2 β-ribbon becomes disordered in the low pH conditions that arise during endosome maturation. This structural transition is required for the virus to become fusogenic (see 2.4.1) (15, 22, 23).

2.3. Genome organization

The genomic RNA (also referred to as 42S RNA for SFV and 49S RNA for SINV) of the typical alphavirus is an approximately 11.5–11.8 kb long single-stranded RNA molecule with a 5' cap0 structure and a 3' poly(A) tail (Fig. 2) (2). The genome has two open reading frames (ORF). The 5' two-thirds of the genome constitutes the first ORF, which is translated directly from genomic RNA and encodes the precursor of the nonstructural (ns) proteins required for RNA synthesis. In the majority of alphaviruses, this region contains a leaky in-frame opal stop codon at the junction of nsP3 and nsP4 (Fig. 2); hence, two polyproteins (designated P123 and P1234) are synthesized (2). The 3' one-third of the genome encodes the structural proteins that function in the assembly of new virus particles and in the attachment and entry of the virus into new cells. The second ORF is expressed through the production of a subgenomic mRNA (sgRNA, also referred to as 26S RNA in SFV) from an internal promoter in the negative-strand RNA replication intermediate (24). The sequence of the 26S RNA overlaps with the last one-third of the 42S RNA. The sgRNAs that are synthesized during replication also have 5' cap0 structures and 3' poly(A) tails (2).

The alphavirus genome contains three untranslated regions (UTRs); one occurs at the 5' end, one is at the 3' end, and one is at the junction region between the ns and the structural ORFs (Fig. 2). All of the UTRs contain *cis*-acting elements called conserved sequence elements (CSEs) that are important for replication and transcription of the virus genome (2, 25).

The 5' end of the genome, or its complement in the 3' end of the negative strand, contains two CSEs (Fig. 2). A conserved stem-loop structure (CSE1) in the 5' UTR is thought to function as a promoter to synthesize genomic RNA from the negative strand and as a co-promoter to synthesize the negative strand from the positive strand template (26). Interestingly, the secondary structure of this region also prevents recognition of the cap0 structure by cellular IFIT1 protein (interferon (IFN)-induced protein with tetratricopeptide repeats) (27). Slightly downstream from the CSE1, in the nsP1-coding region, is the 51-nucleotide CSE2, which forms two stem-loop structures. Both the sequence and the structure of the loops are important for CSE2 to function as a transcriptional enhancer (28). Additionally, CSE2 has been shown to be crucial for alphavirus replication in insect cells, because mutations in the 51-nucleotide element have a greater effect in the mosquito host than in the mammalian host (29). The third CSE is located in the region encoding the C-terminus of nsP4 (Fig. 2) (30). CSE3 forms the conserved part of the sg promoter, which is a *cis*-acting region spanning from position –98 to position +14 with respect to the sgRNA transcription start site. The minimal length of the sg promoter in SINV and most alphaviruses is 24 nucleotides (31); however, it is considerably longer in several SFV stains (32). The final 19 nucleotides preceding the 3' poly(A) tail comprise the CSE4, the sequence of which is highly conserved across the genus and which contains the core negative-strand promoter. The 3' 13 nucleotides of the CSE4 and a poly(A) tail of at least 11 residues were found to be the most

critical for negative strand RNA synthesis, which is initiated at the C-residue immediately upstream of the poly(A) sequence (33).

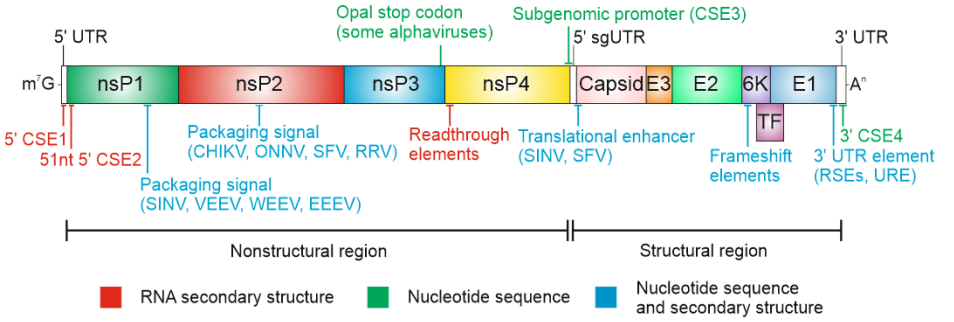

Figure 2. Alphavirus genome organization with the encoded proteins and *cis*-acting sequences. The alphavirus genome is a single-stranded positive sense RNA with a 5' cap0 structure and a 3' poly(A) tail. The 5' two-thirds encodes for ns proteins required for replication and transcription of viral RNA, and the 3' one-third encodes for structural proteins necessary for virion formation. The precursor of the ns proteins is translated directly from the genome, whereas the structural proteins are expressed from sgRNA. The conserved sequence elements (CSEs) are indicated in their relative positions in the genomic RNA. RSEs stands for repeat sequence elements, URE means U-rich element. The *cis*-acting element labels are color coded to indicate whether RNA secondary structure, primary nucleotide sequence, or both, is conserved. The indicated in-frame opal stop codon is present in the genomes of a number of alphaviruses. In most SFV strains and some strains of CHIKV and ONNV, this codon is replaced by an arginine codon. However, a few strains of SFV, including A7(74), contain the opal stop codon.

The genome of alphaviruses also contains *cis*-acting elements that are needed for functions other than RNA replication and transcription. First, the region encoding ns proteins contains a packaging signal for the alphavirus genome. The signal is located in the region encoding the nsP1 protein in SINV and encephalitic viruses (34, 35). In SFV and other members of the SFV clade, the packaging signal is located in the nsP2-encoding region (Fig. 2) (36). Second, there is a stem-loop structure immediately adjacent to the opal stop codon that is present at the end of the sequence encoding the nsP3 protein in SINV, VEEV, and many other alphaviruses (Fig. 2). This structure enhances readthrough of the stop codon by as much as tenfold, leading to more efficient generation of the full-length P1234 polyprotein (37). Third, the 5' end of the capsid gene of SFV and SINV contains a translational enhancer that is needed for the efficient synthesis of structural proteins during later infection (Fig. 2) (38). Fourth, the sequence encoding the 6K protein contains a –1 ribosomal frameshift signal that results in the synthesis of the structural TF protein (Fig. 2) (39). Finally, there are elements in the 3' UTR that interact with host factors during later infection and confer resistance to deadenylation. SINV and VEEV rely on a U-rich element (URE) for binding host factors, whereas other members of the genus depend on the repeat sequence elements (RSEs) (40-42).

2.4. Alphavirus infection cycle

2.4.1. Virus entry

The first step in alphavirus infection involves binding of the virus to a host cell receptor. The viral E2 glycoprotein is an antireceptor primarily responsible for virion binding to the cell surface (43), although the E1 protein may also play a role in receptor engagement. Alphaviruses have a very broad host range and replicate in cells from various tissues within their hosts. This suggests that either a considerably conserved receptor is utilized for attachment or multiple

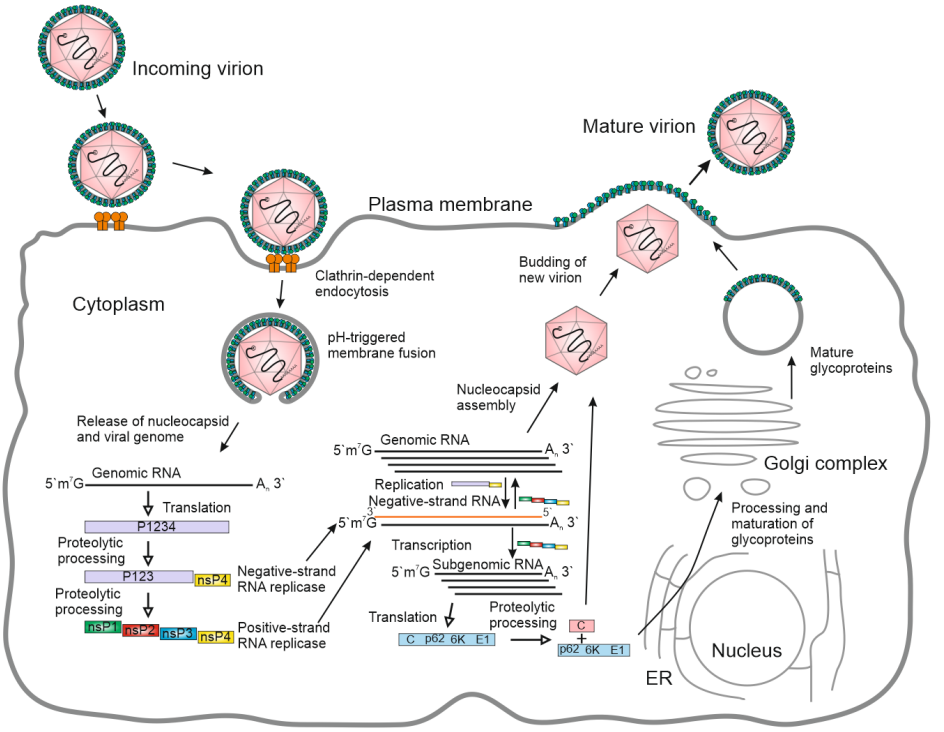

Figure 3. Alphavirus infection cycle. The alphavirus virion enters the host cell by clathrin-dependent endocytosis. The fusion of endosome and viral membranes follows, and the viral genomic RNA is released into cytoplasm. The genomic RNA is immediately translated into ns polyprotein P1234 (grey), which is proteolytically processed and forms first the negative-strand and then positive-strand replicase. The negative-strand replicase produces negative-strand RNA serving as a template for new genomic and sgRNAs. The sgRNA drives the expression of structural polyprotein, which is co- and posttranslationally processed. First, the capsid protein (pink) is released, and it associates with newly synthesized genomic RNA to form the nucleocapsid. Maturation of glycoproteins occurs in the endoplasmic reticulum (ER) and Golgi compartment. Mature glycoproteins (blue and green) are transported to the plasma membrane. The nucleocapsid associates with glycoproteins at the plasma membrane and virion budding occurs.

cellular receptors are employed for virus binding. Specific host receptors are known only for few alphaviruses and even they vary between different alphavirus species. The laminin receptor has been shown to mediate the entry of SINV into mammalian cells and the entry of VEEV into mosquito cells (44, 45). Recently, the natural resistance-associated macrophage protein (NRAMP) (a divalent metal ion transporter) was shown to mediate SINV, but not RRV, entry into both mammalian and insect cells (46). To date, no host cell receptor has been identified for SFV. In addition to proteinaceous receptors, nonprotein attachment factors, including heparan sulfate (HS), might be utilized by alphaviruses to aid initial binding to the host cell surface (17).

Cell-bound virions enter host cells via clathrin-dependent endocytosis (Fig. 3) (47). As the virion-containing endosomal vesicles mature, the pH becomes acidic, triggering conformational rearrangements of the structural proteins. The E1-E2 heterodimers are destabilized, and the previously hidden hydrophobic fusion loop in E1 is exposed (23). The fusion peptide is inserted into the endosomal membrane and E1 heterotrimers are formed, which in turn leads to the fusion of the virion envelope and endosomal membrane and the release of the nucleocapsid into the cytoplasm (22, 48, 49). The released nucleocapsid is disassembled by ribosomes, and the viral RNA genome is liberated into the cytoplasm (50).

2.4.2. RNA replication

Alphavirus replication occurs exclusively in the cytoplasm of the infected cell. The synthesis of viral RNA requires all four ns proteins, both individually, and in the context of ns polyprotein precursors. As for all positive-strand RNA viruses, the genome replication occurs via the synthesis of a negative-strand RNA intermediate. The first step after nucleocapsid disassembly is the translation of viral genomic RNA by host cell ribosomes to produce the ns polyprotein (Fig. 3). The majority of the translation events, as much as ~80–90%, produce the P123 polyprotein; readthrough of the opal stop codon at the junction of nsP3 and nsP4 results in the production of the P1234 polyprotein (51, 52). Some isolates of SFV, CHIKV and ONNV carry an arginine codon instead of the opal stop codon; accordingly, only P1234 is produced by these viruses (53). The synthesized ns polyproteins are autocatalytically cleaved by the protease activity of nsP2 into processing intermediates and thereafter into individual ns proteins in a precisely and temporally regulated manner (54, 55). The P123 polyprotein lacks any intrinsic RNA-synthetic activity due to the absence of the viral RNA-dependent RNA polymerase (RdRp), nsP4 protein. Similarly, the P1234 form of the polyprotein is incapable of RNA synthesis until proteolytic processing releases the nsP4 component of the polyprotein (56).

The full-length polyprotein P1234 cleaves itself *in cis*, yielding P123 and nsP4 (Fig. 4), which form the early replicase (Fig. 3). This P123/nsP4 complex uses genomic RNA as a template to produce negative-strand RNA (Fig. 3),

which, together with the genomic RNA, forms a double-stranded RNA (dsRNA) intermediate. The synthesis of negative-strand RNA occurs early during infection and is rapidly downregulated as the infection proceeds. The early replicase is short-lived; existing data indicate that an individual P123/nsP4 complex may synthesize as few as a single negative-strand RNA molecule before undergoing further processing (56–58). The further cleavage of P123 polyprotein occurs *in cis* and results in the liberation of nsP1 and the formation of the nsP1/P23/nsP4 complex (Fig. 4), which is capable of producing both negative and positive-strand RNAs (57, 59). However, the P23 intermediate is exceptionally short-lived, as it can only be detected following the mutation of the 2/3 cleavage site (60). Therefore, during wild type (wt) virus infection, P23 is quickly processed *in trans* into nsP2 and nsP3 (Fig. 4), which, together with nsP4 and nsP1, form the positive-strand replicase. The nsP1/nsP2/nsP3/nsP4 complex uses the previously synthesized negative-strand as a template to synthesize new positive-strand genomic and sgRNAs (Fig. 3) (54, 56). Regulation of the synthesis of the two positive-sense RNAs depends on nsP4, because distinct sites in nsP4 have been shown to bind the two promoters in the negative-strand RNA (61-63). The effects of some mutations in nsP2 imply that nsP2 may also act as a transcription factor that associates with the sg promoter and recruits the RNA synthesis complex (64, 65). The synthesis of positive-strand RNA continues at a maximal rate until the death of the infected cell. The produced sgRNAs serve as a template for the translation of viral structural proteins, while genomic RNAs interact with the capsid protein and are packed into new virions.

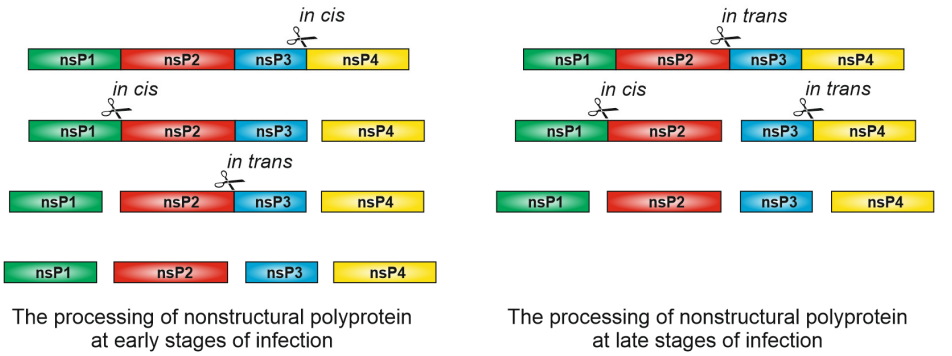

Figure 4. Processing of ns polyprotein by the protease activity of the nsP2 region. Left, the processing of ns polyprotein at the early stages of infection. Initial cleavage occurs *in cis* and results in P123+nsP4, an early replicase that is active in negative strand RNA synthesis. Cleavage of P123 *in cis* yields nsP1+P23. The final cleavage of P23 occurs *in trans*, and the resultant late replicase is active only in the synthesis of genomic and sgRNAs. Right, processing order at the late stages of infection. Because the first cleavage of P1234 occurs between nsP2 and nsP3, there is no formation of early replicase (P123 + nsP4) and, therefore, no synthesis of negative strands. Although P12 and P34 are subsequently processed into mature ns proteins, they do not form the RC.

Cleavage of the 2/3 site serves as an important temporal regulatory step in the replication and infection cycle. It transforms the viral replicase into its late form and prevents the subsequent synthesis of negative-strand RNAs. During the late stage of infection, the amount of free (not included in replicase complexes (RCs)) cytoplasmic nsP2 increases, leading to the rapid *in trans* cleavage of the 2/3 site in newly synthesized ns polyproteins. This generates P12 and P34 polyproteins (Fig. 4), which are incapable of forming new RCs. This cleavage, therefore, prevents the production of P1234 and P123; the formation of new RCs ceases, and the synthesis of negative-strand RNAs is switched off (66). However, the RCs that have already formed are stable and continue to synthesize positive-sense RNAs until the end of the infection cycle.

2.4.3. Sites of replication

Alphaviruses replicate their genomes in association with modified intracellular membranes. They induce the rearrangement of host membranes into cytoplasmic structures known as type I cytopathic vacuoles (CPV-Is) (67, 68) that represent the replication organelles of alphaviruses. CPV-Is are late endosomes and lysosomes with a diameter of 600–2000 nm. These vacuoles contain bulb-shaped invaginations called spherules (~50 nm in diameter), which represent physical forms of alphavirus RCs and are the sites of viral RNA synthesis (69, 70). The spherules first appear on the host plasma membrane; as the infection proceeds, the structures are internalized via endocytosis. Spherule-containing vesicles can fuse with one another and with lysosomes, and they are finally incorporated into CPV-Is (70-72). The significance of this process remains unknown, because blocking spherule internalization with different inhibitors has little or no effect on viral RNA synthesis (71, 72). The formation of these spherules requires ns proteins in the polyprotein stage as well as active RNA synthesis (71, 73). Each spherule appears to contain partly double-stranded replicative intermediate RNAs (74), but the exact copy number, stoichiometry, and locations of different nsPs within the alphavirus RCs remain unknown. In addition to viral proteins, several host proteins are bound to the spherules (71, 72). It is likely that spherules are needed to protect dsRNA replication intermediates from host cell detection and disruption. Additionally, the membrane structures may act as scaffolds and effectively increase the concentration of replication factors at the sites of RNA synthesis. There are also considerable differences among alphaviruses with respect to the size and localization of CPV-Is. For instance, the CPV-Is of SFV are large structures that are located in perinuclear region, whereas the CPV-Is of CHIKV are generally smaller and, in most cases, localized close to the plasma membrane. Recently, the internalization of CPV-Is was shown to be associated with the ability of nsP3 of SFV to activate the phosphatidylinositol-3-kinase (PI3K)-Akt-mammalian target of Rapamycin (mTOR) pathway (75); however, the functional significance of these processes is currently unknown.

2.4.4. Synthesis of structural proteins

Alphavirus structural proteins are translated from sgRNA in the form of C-p62(E3E2)-6K-E1 (Fig. 3) and p62(E3E2)-TF polyprotein precursors. The capsid protein is a serine protease and releases itself from the nascent precursor chain by autoproteolysis (Fig. 3) (76), thereby revealing an N-terminal signal sequence, which is used for p62 chain translocation to the endoplasmic reticulum (ER). In the ER, p62(E3E2)-6K-E1 polyprotein is cotranslationally glycosylated, palmitoylated, and cleaved by cellular proteases into p62, 6K, and E1. Glycosylated p62 and E1 form a stable heterodimer, which is transported through the secretory pathway to the Golgi complex (Fig. 3). In the Golgi, E3 is cleaved from p62-E1 by cellular furin protease, and the resulting E1-E2 heterodimers are transported to the plasma membrane for virion assembly (Fig. 3) (77). For some alphaviruses, including SFV and VEEV, the released E3 protein is also incorporated into the virions (78, 79). However, in the majority of alphaviruses, including SINV, E3 is not present in the virions. It has been shown that binding of E3 to the spike complex protects E1 against low pH-triggered conformational changes during virion biogenesis (80). The 6K protein is also incorporated into virions in smaller numbers (7–30 copies) and likely affects the interactions between E2 and E1 (17). During the translation of SFV structural proteins, a ribosomal frameshift occurs in the sequence encoding for 6K with an efficiency of 10–18% and results in the production of TF protein (39). TF has been shown to be packed into virions, but its exact function(s) remains to be determined (20).

2.4.5 Virion assembly and budding

The formation of new virions begins with the assembly of nucleocapsids in the cytoplasm of the infected cell near the RCs. The encapsidation signal located in the region encoding nsP1 (or, in some viruses, nsP2) of the newly synthesized genomic RNA is recognized by the capsid protein, leading to the multimerization of the protein and the formation of an icosahedral nucleocapsid (Fig. 3) (34, 36). The nucleocapsids diffuse or are transported to patches and extensions on the plasma membrane caused by the accumulation of envelope proteins. During budding, the nucleocapsids are bound by the cytoplasmic tails of E2 glycoproteins, and the plasma membrane curves around the nucleocapsid. These interactions of viral envelope proteins with capsid protein are required to exclude other plasma membrane proteins from the sites of alphavirus budding (81). Finally, the lipoprotein envelope encloses the nucleocapsid, and the virion is released (Fig. 3). This process is facilitated by nsP1, which counteracts the antiviral effects of cellular tetherin (82).

2.5. Functions of individual ns proteins

All four ns proteins are essential for alphavirus replication and have a multi-domain structural organization (Fig. 5). Each protein has a unique and specific role in the viral infection cycle and virus-host interactions.

NsP1 (for SFV: 537 aa, ~60 kDa) has two main functions during alphavirus replication: it serves as a membrane anchor for the RC, and it is responsible for capping positive-strand RNAs. The N-terminal domain of nsP1 exhibits methyltransferase and guanylytransferase activities, which are involved in the capping process. The first reaction in cap synthesis is performed by nsP2, which removes the 5' γ-phosphate of the nascent viral RNA through its RNA triphosphatase (RTPase) activity (83), whereas the subsequent two reactions are performed by nsP1. First, nsP1 acts as a guanine-7N-methyltransferase and transfers a methyl group from S-adenosyl-methionine to GTP. Second, nsP1 guanylyltransferase activity allows it to form a covalent intermediate complex with 7-methyl-guanosine-5'-monophosphate (m^7GMP) (84). The final step in cap synthesis is the transfer of m^7GMP from nsP1 to the viral RNA molecule; this reaction is also most likely performed by nsP1 (85). The alphavirus mode of RNA capping stands in stark contrast to the eukaryotic capping mechanism, in which methylation occurs after the transfer of the guanylate moiety to the substrate RNA.

The N-terminal domain is followed by regions that are necessary for the association of nsP1 (and replicase as whole) with host membranes: an amphipathic helix and site for posttranslational palmitoylation (86-89). The amphipathic helix, which is located between aa 245 and 264 in SFV (Fig. 5), mediates interactions with membrane phospholipids. These interactions are required for the enzymatic activity of nsP1 of SFV (87). Point mutations in the amphipathic helix, which prevent binding to the membrane, are lethal to the virus (89). However, this effect may not be universal for all alphaviruses, because nsP1 of SINV retains enzymatic activity in the absence of phospholipids (90). Membrane binding is further strengthened by posttranslational palmitoylation of nsP1 cysteine residues at position 418–420 in SFV and 420 in SINV (Fig. 5). Nevertheless, palmitoylation is not essential for the enzymatic activities of nsP1, the formation of functional RCs, or the viability of the virus (91, 92). Despite lacking a phenotype in tissue culture models of infection, depalmitoylation mutants exhibit diminished pathogenesis in mice (86). Additionally, a mutation at position 538 (T538I) in nsP1 in the neurovirulent strain of SINV attenuates the virus and plays a key role in regulating viral neurovirulence (93). This attenuating mutation has been shown to modulate type I IFN induction, and an analogous mutation in RRV exerts a similar enhancing effect (94).

NsP1 has been shown to interact with the N-terminus of the nsP4 protein, and this interaction is important for recognition of the negative-strand promoter and elongation of negative-strand RNA (95, 96). NsP1 is also responsible for the induction of filopodia-like structures on the surface of alphavirus-infected

cells (97, 98). The function of these filopodia-like structures, as well as the exact functional significance of palmitoylation, remain unknown.

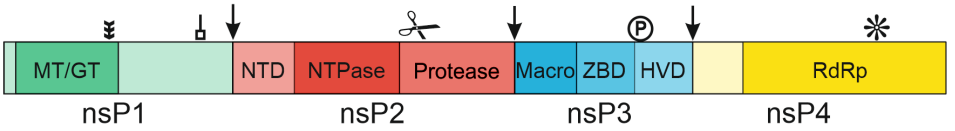

Figure 5. Modular organization of the ns polyprotein of alphaviruses. Arrows indicate sites of proteolytic cleavages within P1234 polyprotein. The N-terminal domain of nsP1 is responsible for methyltransferase and guanylytransferase (MT/GT) activities. An amphipathic helix (triple arrow) and posttranslational palmitoylation (corresponding site is indicated by an empty square) are necessary for the association of nsP1 protein to host membranes. RNA helicase and protease functions of nsP2 are assisted by the N-terminal domain (NTD). The N-terminal part of nsP2 has also NTPase and RTPase activities. Scissors indicate the location of the catalytic residues of the protease. nsP3 has three recognized domains: the macro domain, the zinc binding domain (ZBD), and the hypervariable domain (HVD). The phosphorylation sites are marked with a circled P. The N-terminal region of nsP4 is predicted to be unstructured, while the C-terminal part of nsP4 is homologous to the core of viral RNA-dependent RNA polymerases (RdRp). The catalytic residues of RdRp are indicated by an asterisk.

The multi-domain protein **nsP2** (for SFV: 798 aa, ~90 kDa) exhibits four important enzymatic functions during viral infection, acting as an NTPase, an RNA helicase, an RNA triphosphatase, and a protease. It also possesses a variety of nonenzymatic functions. The N-terminal half of nsP2 contains domains that are required for its nucleoside triphosphatase (NTPase) and RNA-5'-triphosphatase (RTPase) activities (83, 99). The RTPase activity of nsP2 is responsible for removing the γ-phosphate from the 5' end of nascent positive-sense RNA; it is therefore required for the nsP1-mediated capping reaction (83). NsP2 functions as an RNA helicase, likely by unwinding the RNA secondary structures that are formed during viral replication (100). The helicase activity of nsP2 is dependent on the NTPase activity, and unwinding has been shown to occur with a 5'–3' directional bias. All domains are required for the helicase activity of nsP2; helicase activity likely acts in coordination with the polymerase activity of nsP4 (101). Very recently, it was also demonstrated that nsP2 and nsP4 together regulate the replication fidelity of alphaviruses and that mutations in nsP2 allow the virus to overcome intracellular nucleotide depletion (102). The C-terminal region of nsP2 contains a papain-like protease domain that is responsible for processing the ns polyprotein (103). This protease activity is absolutely essential for the replication of the virus, and it has been shown to be functionally discrete from the nsP2 functions described above (104). NsP2 is capable of performing all of the cleavages required for the processing of P1234, this requires a catalytic cysteine residue at position 478 (Fig. 5) (55). Several criteria must be met to ensure that the proteolytic cleavages by nsP2 occur in a

particular order. The amino acid residues surrounding both sides of the scissile bond are important, but the structural placements of the nsP2 domains and other replicase proteins are also critical (66, 104-106). The 3D structures of the VEEV protease (107), SINV protease (108), and CHIKV protease (109) have been determined by X-ray crystallography. In all three proteases, the papain-like domain is followed by an enzymatically nonfunctional methyltransferase-like domain, which plays a role in the regulation of negative-strand RNA synthesis and the development of cellular cytopathic effects (110). It is also essential for the RNA helicase activity of nsP2 (101). Additionally, nsP2 has been shown to function as a transcription factor for sgRNA synthesis by recognizing the sg promoter (65, 111), and it has been proposed to mediate the binding of nsP4 to that promoter (61).

NsP2 is the only ns protein of alphaviruses that has been shown to translocate to the nucleus. In SFV-infected cells, approximately 25% of nsP2s are associated with replication organelles, 25% are localized diffusely throughout the cytoplasm, and 50% are transported to the nucleus. While SINV nsP2 lacks a classical nuclear localization signal (NLS) (112), the nsP2 of SFV carries a pentapeptide PRRRV in its C-terminus (aa 647–651), which is assumed to function as an NLS (113). The mutation of arginine residues in this sequence to aspartate residues renders nsP2 completely cytoplasmic (114), though this effect is observed only at 37 °C and not at 28 °C (115). Regardless of the exact mechanism of entry, nuclear localization of nsP2 is required to turn off cellular transcription, which occurs via degradation of the catalytic subunit of cellular RNA polymerase II (116). Thus, the nuclear fraction of nsP2 is not directly required for viral replication but rather is responsible for the inhibition of cellular transcription. Mutations in nsP2 are also associated with reduced cytotoxicity, a lack of translational shutdown in host cells, and reduced viral pathogenicity in adult mice (114, 115, 117-119).

The functions of **nsP3** (for SFV: 482 aa, ~60 kDa) remain more elusive than those of the other ns proteins. nsP3 has three recognized domains: the macro domain, the zinc-binding domain (ZBD, also known as alphavirus unique domain (AUD)), and the hypervariable domain (HVD) (Fig. 5). The first 160 amino acid residues in the N-terminus of nsP3 form a macro domain that is structurally conserved among alphaviruses, rubiviruses, hepeviruses and corona-viruses (120). The macro domains have detectable homologs in a wide variety of bacteria, archaea, and eukaryotes (121); thus, it is assumed that they play fundamental roles in different organisms. The crystal structures of the macro domain of CHIKV, SINV, and VEEV have been determined. It has been found that the macro domains of CHIKV and VEEV exhibit weak ADP-ribose 1''-phosphate phosphatase activity (122, 123). However, the macro domain of SFV nsP3 lacks this activity, which suggests that this function is not needed for virus replication. The alphavirus macro domain has been shown to bind to RNA, polyADP-ribose, and, in some cases, ADP-ribose. RNA binding might be the true function of the nsP3 macro domain (122, 124). In addition, there is

evidence that the alphaviral macro domain is involved in one or more host protein interactions (125). Finally, the residues located at the C-terminus of the macro domain and/or in the region between the macro domain and the ZBD are essential for 2/3 site processing by nsP2 (105).

The ZBD is located within the central portion of nsP3, a region that shares a strong sequence homology across the *Alphavirus* genus (2). This region was recently crystallized as part of the SINV P23 polyprotein and shown to contain a previously uncharacterized zinc coordination site (108). Genetic manipulation within the ZBD has resulted in defects in negative-strand and sgRNA synthesis, polyprotein processing and neurovirulence (126-128). Recent data indicate a functional interplay between the ZBD and the C-terminal HVD (75).

The C-terminal region of nsP3 is intrinsically disordered. As implied by the name HVD, it exhibits different lengths and sequence compositions among the alphaviruses**.** Despite the lack of sequence conservation, the HVD contains sequence motifs that are present in a few or even in many alphaviruses. Thus, the HVD of the Old World alphaviruses SINV, SFV, and CHIKV exhibits a proline-rich motif that serves as a target site for the Src-homology 3 (SH3) domain of amphiphysin. Mutations in this proline-rich element result in slightly impaired viral RNA replication and significantly decreased virulence in mice (129). The HVD of VEEV nsP3 does not appear to have this motif (130). In the Old World alphaviruses SINV, SFV, and CHIKV, the HVD of nsP3 has also been shown to interact with the host protein Ras-GAP SH3-domain-binding protein (G3BP). This binding leads to the recruitment of G3BP into cytoplasmic foci, which in turn inhibits stress granule formation (131-133). This binding occurs via FGDF motifs that have only been characterized in Old World alphaviruses (134). In mosquito cells, the nsP3 of SFV and CHIKV binds to the mosquito homologue of G3BP proteins, called Rasputin (135, 136), potentially via the same FGDF motifs. The nsP3 of New World alphaviruses lacks FGDF motifs, and hence, VEEV HVD does not bind to G3BP (130, 137). The HVD of VEEV nsP3 has been shown to contain many repeated elements that are indispensable for the replication of the virus in different cell types (other than BHK-21) and the formation of VEEV-specific cytoplasmic protein complexes (137). It should also be noted that interaction with G3BP proteins (and possibly with other cellular proteins) could serve multiple purposes. Thus, while the binding of nsP3 to G3BP proteins counteracts their antiviral activity (stress granule formation), G3BP proteins also possess proviral functions that may facilitate the switch from replicase protein translation to viral RNA replication (138).

Based on our current understanding, nsP3 is the only alphavirus phospho-protein; it undergoes phosphorylation at serine and threonine residues at the junction between the second and third domain (Fig. 5) (139, 140). In SFV, all 16 of the identified phosphorylation sites (among which six of these residues account for the majority of nsP3 phosphorylation) occur in a short 50-amino acid region. SFV, with defective nsP3 phosphorylation exhibits a decreased rate

of RNA synthesis in cultured cells and greatly reduced pathogenicity in mice (140). In SINV, phosphorylation plays a role in negative-strand RNA synthesis (126). As observed for VEEV, the phosphorylation of the nsP3 HVD is important for viral replication in mosquito cells but not in cells of vertebrate origin (137). When expressed alone, nsP3 forms amorphous cytoplasmic non-membranous granules (141). However, when it is expressed as part of the P123 polyprotein, nsP3 is responsible for the relocalization of SFV replicase proteins from the plasma membrane into intracellular vesicles (66). This function is most likely also mediated by the HVD of nsP3, or more precisely by the region involved in the activation of the PI3K-Akt-mTOR pathway. Interestingly, in SFV, this region overlaps with the phosphorylation region. Nevertheless, activation of the PI3K-Akt-mTOR pathway does not depend on the phosphorylation of nsP3, indicating that the same region of HVD has several independent functions (75).

All of the ns proteins are involved in some aspects of alphavirus RNA synthesis; however, **nsP4** (for SFV: 614 aa, ~70 kDa) is solely responsible for the RNA synthesis properties of the viral RC. The 3D structure of nsP4 is not known. However, the C-terminal sequence of the protein includes motifs that are typical of RdRp. Thus, it is very likely that nsP4 has a typical RdRp 3D structure with fingers, a palm containing the GDD motif, and thumb domain. The sequence of the ~100 N-terminal amino acid residues of nsP4 is conserved only among alphaviral nsP4 proteins (142-144).

In comparison to the other ns proteins, nsP4 is scarce in infected cells, for two reasons. First, in most alphaviruses, including several SFV strains, the translation of nsP4 requires readthrough of an opal stop codon that is located at the end of the nsP3-encoding region (2). Second, the N-terminal amino acid in nsP4 is a conserved tyrosine that directs nsP4 to undergo rapid proteasomal degradation via the N-end rule pathway (145). In RCs, nsP4 appears to be stable, but the mechanism(s) responsible for this stabilization is not known. The N-terminal destabilizing tyrosine residue is essential for the polymerase activity of nsP4; accordingly, its replacement with a nonaromatic residue results in poor RNA replication (146).

The expression and purification of functionally active recombinant full-length nsP4 has historically been exceptionally difficult. *In vitro* experiments using N-terminal truncation mutants of nsP4, which are somewhat easier to produce, demonstrated terminal adenylyltransferase (TATase) activity. This function is likely required for the maintenance and repair of the poly(A) tail at the end of genomic and sgRNAs (144). The full-length nsP4 of SINV has been purified with an N-terminal SUMO tag. Full-length recombinant nsP4 exhibits TATase activity and is capable of *de novo* RNA synthetic activity. The latter ability, however, requires the addition of the other viral nsPs from a mammalian cell membrane fraction (143). Thus, although nsP4 is the sole viral protein with RdRp activity, viral replication occurs as a result of the sum of coordinated ns protein activity.

The nonconserved N-terminal region of nsP4 is disordered and is required for interactions with other ns proteins and viral RNA. Genetic evidence suggests that the conserved N-terminal tyrosine interacts with nsP1 for negative-strand synthesis (146). Studies have indicated that the N-terminal region plays an important role in recognizing the promoter at the 3' end of the genomic strand for negative-strand synthesis (143). Crosslinking experiments have shown that nsP4 contains determinants that contact with sg and genomic promoters in the negative strand RNA. NsP4 alone recognizes the genomic promoter (61, 62), but recognition of the sg promoter requires other nsPs, especially nsP2 (147). Genetic evidence suggests that the N-terminal part of nsP4 may also be involved in interactions with host proteins (147).

2.6. Virus-host interactions

At the interface of viral infection and host response lies a complex network of regulated interactions. The host wants to eradicate the virus; the virus seeks to continue its own proliferation. Furthermore, alphaviruses have a limited coding capacity and thus rely extensively on host factors for successful infection and propagation. As a result, these viruses have developed mechanisms to hamper cellular antiviral processes while maintaining their own genome replication and virion production to spread in infected hosts.

2.6.1. Shutdown of host cell transcription and translation

Alphavirus infection of vertebrate cells causes the shutdown of host cell transcription and translation, but at the same time requires the synthesis of viral RNAs and proteins at a high rate. The shutdown of host macromolecule synthesis is essential for limiting the production of antiviral proteins, mainly type I IFNs, and hence diminishing the ability of the innate immune system to attenuate the infection. Old World and New World alphaviruses achieve this effect through the use of different proteins.

For Old World alphaviruses, nsP2 mediates transcriptional shutdown via the degradation of Rpb1, a catalytic subunit of RNA polymerase II. In the presence of nsP2, Rbp1 is ubiquitinated and rapidly degraded, which in turn results in the cessation of host mRNA transcription (116). The New World alphaviruses utilize a different mechanism to shutdown cellular transcription. The VEEV capsid protein forms complexes with importin-α/β and the nuclear export receptor CRM1. These complexes accumulate in the nuclear pores of infected cells and inhibit the nucleocytoplasmic trafficking of proteins, eventually leading to transcriptional shutdown (148). In addition to inducing a general shutdown of host transcription, alphaviruses can also specifically target the expression of genes that are important for the innate immune response. SFV has been shown to specifically target the expression of type I IFNs in a manner that

differs from the virus-induced general shutdown of transcription (149). Viruses can also interfere with cellular antiviral signaling pathways: nsP2 of CHIKV has been shown to inhibit Jak-STAT signaling (150), and a similar phenomenon has also been observed for SINV (151).

The shutdown of host cell translation in both groups is partially mediated by a cellular defense mechanism. The viral replication intermediate dsRNA is recognized by host dsRNA-activated protein kinase R (PKR), which leads to phosphorylation of the translation initiation factor eIF2 and inhibition of the initiation of cellular translation (152). Several alphaviruses, including SINV and SFV, have evolved a unique mechanism to bypass the requirement for eIF2 through the presence of the translational enhancer (stable secondary structure) at the 5' end of the capsid gene, which likely stalls ribosomes and directs them to the appropriate site to initiate translation (38, 153).

PKR-dependent shutdown of translation is not the only mechanism employed by alphaviruses to achieve their goals. Another pathway that leads to eIF2 phosphorylation is the unfolded protein response in the ER (154, 155). Active synthesis of the alphavirus glycoproteins E1 and E2 in the ER likely exceeds the ER folding capacity, resulting in an unfolded proteins response. This process activates one of the major ER stress sensors, eIF2 kinase 3 (PERK), which then phosphorylates eIF2 and thereby inhibits cellular translation (154). To overcome the effects of this inhibition, alphaviruses can suppress the unfolded protein response. In the case of CHIKV, this phenomenon is especially prominent, and viral nsP2 and/or nsP4 participate in this process (156, 157).

A recent study has shown that for the inhibition of cellular protein synthesis to occur, viral RNA replication must take place at controlled levels, leading to the release of nuclear proteins into the cytoplasm (158). The shutdown of transcription and translation of cellular mRNAs are independent events. Although the shutdown of host transcription is the more critical event for viral infection, the shutdown of host translation is also important. In general, mutations in nsP2 disrupt viral shutdown of host cell translation, confirming the central role of this protein in this process (118). However, the mechanism(s) by which nsP2 interferes with host cell translation is poorly understood.

2.6.2. Host factors associated with alphavirus RCs

The identification of relocalized host cell components and host cell components with altered quantities during viral infection is essential for understanding the interplay between the host and positive-strand RNA viruses. Recently, several different approaches have been used to tackle this question, including genome-wide screens, whole-genome searches using siRNA libraries, and yeast two-hybrid (Y2H) screens. These studies, in which several viruses were assessed, revealed hundreds of host-encoded proteins that interact with viral proteins and RNAs or otherwise participate in different stages of viral infection, including

RC assembly, RNA template recruitment, synthesis, and viral RNA stabilization. Host proteins that negatively affect viral infection have also been identified (159, 160). Interestingly, the sets of host proteins that have been identified using different screens for the same virus show a tendency towards limited overlap. Hence, the list of cellular factors that are directly or indirectly involved in positive-strand RNA virus replication is long, and the mechanism of action is known for only a small number of proteins.

Similarly, the list of host components that interact with the alphavirus nsPs and RNAs is slowly but steadily growing. Pull-down experiments using cells infected with SINV carrying GFP-tagged nsP3 (125, 131, 161), GFP-tagged nsP2 (162), and FLAG-tagged nsP4 (163) led to the identification of overlapping sets of co-precipitating cellular proteins, including G3BP1, G3BP2, 14-3-3 proteins, poly(ADP-ribose) polymerase (PARP-1), and different heterogeneous ribonucleoproteins (hnRNPs). The overlap is likely due to precipitation of whole RCs and cellular proteins associated with any of their components, regardless of which ns protein was tagged. Thus, this approach provides limited information regarding the viral protein(s) that specifically interact with the identified host component. For example, G3BPs interact only with nsP3 (134); however, these proteins were also detected in experiments using tagged nsP2 and nsP4 for the precipitation (162, 163).

Very little is known about the functions of identified cellular factors in the viral infection cycle. The interaction of G3BP with viral nsP3 leads to the recruitment of the former into cytoplasmic foci, which in turn inhibits stress granule formation that would otherwise inhibit viral infection (132, 134). Concomitantly, G3BPs also play a proviral role because their depletion hampers CHIKV infection (138). Viral nsP2 has also been found to co-purify with multiple ribosomal proteins, including ribosomal protein S6, and cellular filament components. Interactions between nsP2 proteins and ribosomal components have also been reported for VEEV (164) and likely contribute to the ability of viral factors to hijack the translational machinery.

Another study employed an approach that was based on the comparison of cytoplasmic membrane fractions obtained from mock-infected and infected cells. HnRNP K was identified as an interaction partner of alphavirus RCs. The hnRNP K protein also interacts with SINV sgRNA (165). Another cellular protein, HuR, has been shown to bind to the 3′ UTRs of genomic and sgRNAs, thereby protecting them against degradation. This interaction appears to be specific to the species of virus because it was observed for SINV (41), but not for CHIKV (166). Additionally, nsP3s of CHIKV, SINV and SFV have been shown to interact with host cell amphiphysins (proteins that are prominently involved in cellular membrane dynamics) and to recruit them to RCs (129).

To document the interactions of alphaviruses with host cell components on a genome-wide scale, high-throughput Y2H screening was used to characterize the interactions between nsPs of CHIKV and human proteins (167). Interestingly, only a few of the revealed proteins overlapped with the interaction

partners identified via pull-down experiments using tagged SINV nsP2 mentioned above (162). In addition, RC-bound nsP3 and nsP3 located in cytoplasmic granules has been shown to interact with different host proteins (135).

An obvious approach for mapping virus-host interactions involves siRNA screens. These screens, which target many or all cellular proteins, have been performed for several alphaviruses, but complete data from any of these studies remain to be reported. Nevertheless these screens have been useful; they have led to the identification of NRAMP as the receptor for several alphaviruses (46), as well as to the discovery of the antiviral role of the nonsense-mediated mRNA decay pathway (168). Partners of the EU-funded ICRES (Integrated Chikungunya RESearch) consortium (our laboratory was part of this consortium) performed a whole-genome siRNA screen to identify CHIKV host factors, resulting in the identification of 156 validated proviral and 41 validated antiviral host factors. Among these, six proviral factors could be targeted with approved chemical inhibitors, and the pharmacological targeting of these proviral factors in mouse models of CHIKV also provided antiviral effects (Karlas et al., unpublished).

Thus, multiple studies investigating alphavirus host factors have produced long lists of candidate proteins. However, knowledge concerning their precise roles and functions in the context of alphavirus infections is very limited. Additionally, the poor overlap between the results obtained using different methods suggests that only a fraction of the host proteins that are important for alphavirus infection have been identified. Therefore, many important host components remain to be identified, and almost all of these factors will require further analysis to characterize their precise roles, functions and mechanisms of actions in the context of alphavirus infection.

2.7. Pathogenesis of SFV infections

Virus-host interactions vary between different alphavirus species. Additionally, different strains of the same virus can have contrasting effects on the host. SFV is a prime example of this divergence, comprising strains that are designated as virulent or avirulent according to their pathogenicity in the central nervous system (CNS) of adult mice.

2.7.1. Strains of SFV

The various strains of SFV provide a well-characterized experimental system to study the pathogenesis of viral encephalitis. The most commonly used SFV strains include A7(74), SFV4, and L10. All of these strains infect neonatal and young suckling mice (up to 11 days old), causing lethal encephalitis (117). The ability to cause lethal infection in mice of this age group is not specific to SFV; rather, it is a property shared by many alphaviruses, including those, which do

not efficiently infect adult immunocompetent mice (169). However, the outcome of SFV infection in older mice ranges from asymptomatic for A7(74) to neurovirulent in the case of SFV4 and L10 (117).

L10 originates from virus isolated from a pool of 130 *Aedes africanus* mosquitoes that were captured in the Semliki Forest, Uganda in 1942 (170). It was passaged eight times via intracerebral (i.c.) inoculation in the adult mouse brain, followed by two i.c. passages in the neonatal mouse brain (171). One aliquot that was sent from the Bradish laboratory (Porton Down, United Kingdom) to the Webb laboratory (London, United Kingdom) was further passaged twice in chicken embryo fibroblasts and stored at –80 °C. One of these aliquots was used in the present study (paper **I**) to obtain the consensus sequence of L10 by high-throughput sequencing.

The prototype strain of SFV was isolated from the same pool of mosquitoes as L10; its passage history, however, is different. First, it was passaged four times by i.c. inoculation in the adult mouse brain (171). The subsequent passage history was not recorded, but it most definitely included several (based on some data, as many as sixteen) passages in BHK-21 cells. The prototype strain was the first SFV strain to be sequenced and used for the construction of a full-length infectious cDNA (icDNA) clone. The cDNA clone was designated pSP6-SFV4, and the obtained virus was designated SFV4 (12, 53, 172). Of note, the same isolation (in 1942) gave rise to several other virulent SFV strains, including V13 (available from the ATCC as VR-67; http://www.lgcstandards-atcc.org/ Products/All/VR-67.aspx) and E9. The Osterrieth strain of SFV, the only SFV strain that has caused a fatal case of infection in humans (173), originated from the laboratory of Prof. Paul M. Osterrieth, who himself obtained the SFV stock in 1957 from the Wistar Institute (USA). Given the timing, it is reasonable to assume that this virus must have originated from the 1942 isolation and thus has a common source with SFV4 and L10. However, the passage history of the Osterrieth strain of SFV was different from the other strains'. It is unknown whether the original passage history of this stock is the same as that of the prototype strain; however, before the fatal accident, the virus had been passaged 13 times in primary chicken embryo fibroblasts, seven times i.c. in 4- to 5-week-old mice, and once subcutaneously in 6- to 7-week-old mice. The brain material of these mice was again passaged four times i.c. in 6- to 7-week-old mice and then once in BHK-21 cells. The patient worked with the supernatant collected from this passage (173). There is almost no information on what happened with the virus strain after the fatal accident or even whether or not this strain still exists.

A7(74) was derived from the AR2066 strain by seven passages through neonatal mouse brain and two colony selections on chick embryo fibroblasts. The AR2066 strain was isolated from *Aedes argenteopunctatus* mosquitoes in Namancurra, Mozambique in 1959 (174). The aliquot used in the present study (paper **II**) to obtain the consensus sequence of A7(74) by high-throughput sequencing was received from H. E. Webb, who obtained it from the Bradish

laboratory, and was stored at –80˚C with minimal passaging. Another aliquot of A7(74) from the Webb laboratory was further passaged several times in MBA-13 cells, subsequently plaque-purified three times on the same cells and used for sequencing, molecular analyses, and the construction of the icDNA clone of A7(74), which was designated rA774 (175).

Thus, L10 and SFV4 originate from the same isolation of SFV virus, whereas the ancestor of A7(74) was isolated independently 15 years later from a different location and source. Both groups of viruses have a rather long *in vivo* passage history. In addition, the prototype strain and, accordingly, SFV4 have also undergone a long *in vitro* passage history. Of note, none of these strains has been passaged using mosquitoes/mosquito cells; however, the ability to infect vectors/vector cells has been maintained. Additional isolates of SFV have been collected after 1959 (although not recently), but these isolates have not provided any actively studied laboratory strains. Thus, the vast majority of modern studies of SFV virulence have been conducted using the A7(74), SFV4, and L10 strains of the virus.

2.7.2. Neuroinvasion and CNS tropism

The CNS is separated from circulating blood by the blood-brain barrier (BBB). Inoculation of the virus into the abdominal cavity (intraperitoneal (i.p.) inoculation) is the most commonly used route to study the ability of the virus to cross the BBB, i.e., neuroinvasiveness. Following i.p. inoculation, A7(74) and L10 first replicate in peripheral tissues and produce a high-titer plasma viremia that peaks at 24–48 h postinfection (176). This viremia is controlled by IFN response and then by antibodies (177). High-titer plasma viremia is thought to be a prerequisite of CNS infection; however, in nature, its main significance is to ensure the efficient transmission of the virus to insect vectors during a blood meal. Both A7(74) and L10 are rapidly neuroinvasive and enter the brain by traversing cerebral endothelial cells (178). In contrast, following low-dose i.p. inoculation, SFV4 produces lower plasma viremia and no virus can be detected in the brain (179).

After entering the brain, A7(74) and L10 initially infect and replicate in perivascular oligodendrocytes and neurons but not in astrocytes (117, 177). L10 spreads rapidly throughout the brain, producing fatal panencephalitis in mice of any age (176, 180, 181). Interestingly, following i.c. or intranasal inoculation, SFV4 is also neurovirulent, disseminates rapidly throughout the brain, and efficiently causes lethal encephalitis, as observed with the L10 virus. Thus, SFV4 possesses a defect that reduces its peripheral infection and prevents efficient neuroinvasion. In contrast, the dynamics of A7(74) infection of the CNS varies according to the age of the mice. In neonatal mice, A7(74) disseminates rapidly in the brain and causes fatal infection. In the adult mouse CNS, the spread of A7(74) remains limited and concentrated to the perivascular foci that likely

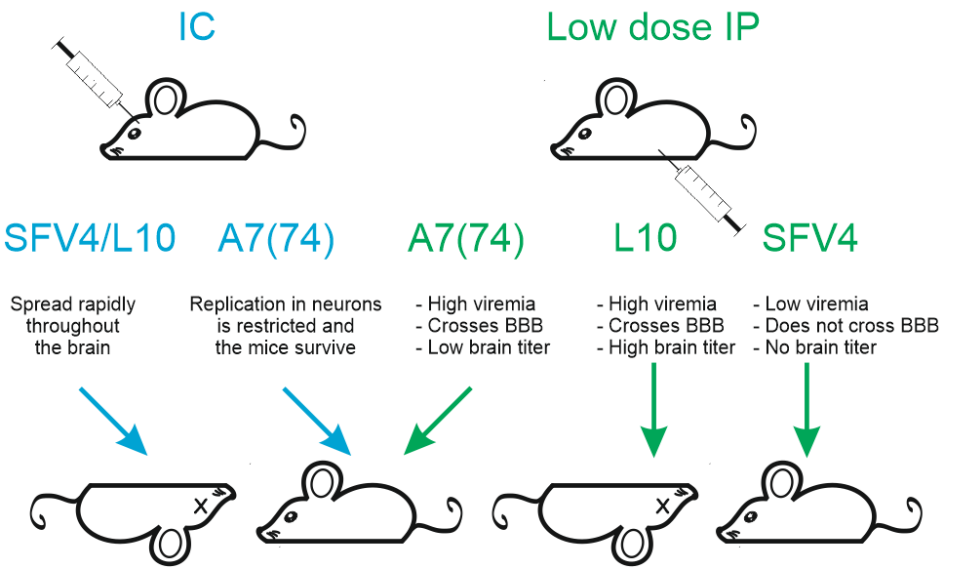

Figure 6. Schematic representation of phenotypes of different SFV strains in an adult mouse model. IC indicates intracerebral inoculation; IP indicates intraperitoneal inoculation.

represent the initial sites of viral entry into the CNS. The replication of A7(74) in mature neurons is severely restricted, and the virus is cleared from the brain by the immune response (176, 181). The avirulent strain does, however, induce a nonlethal demyelinating disease that generally lasts up to 30 days after infection (182). Interestingly, when administered i.c., A7(74) remains avirulent and replication in mature neurons is still restricted; however, widespread replication in oligodendrocytes in major white matter tracts can be observed (183). A schematic overview of SFV phenotypes is depicted in figure 6.

2.7.3. Determinants of SFV virulence

The molecular determinants responsible for the increased susceptibility of younger hosts to severe alphavirus-induced encephalitis are not completely known. This age-related virulence does not appear to depend on the maturity of specific immune responses. This conclusion is based on the results of studies using athymic *nu/nu* mice lacking T-lymphocytes and mice with severe combined immunodeficiency (SCID), which lack both T- and B-lymphocytes, and in which the replication of A7(74) remains restricted to small perivascular foci (176, 184). Rather, the age-related virulence of A7(74) is a function of CNS cell maturity. All three SFV strains replicate efficiently in immature neurons because these cells likely contain suitable membranes and biochemical pathways. During the first postnatal weeks, detailed connections in the neuronal network are finalized, and axogenesis, synaptogenesis, gliogenesis, and

myelination are being completed (181). Due to these changes, the ability of A7(74) to infect mature neurons becomes restricted: viral protein synthesis and RNA replication still occur, but the budding and consequent spread of the virus infection do not follow. Interestingly, when adult mice are pretreated with gold compounds, neurons become permissive to A7(74) replication, and the virus spreads rapidly throughout the brain, resulting in a panencephalitis that is similar to that caused by virulent L10 (185). Gold compounds are known to induce smooth membrane production in neurons, and one can assume that because alphaviruses require suitable membranes for replication and budding, the change in membrane synthesis and/or associated biochemical pathways may be the cause of age-related virulence of A7(74) (117).

Prior to the present study, the viral determinants responsible for the different neurovirulence of SFV strains in adult mice have been shown to reside in the ns region. A chimera containing the structural genes from the molecular clone (rA774) of the avirulent A7(74) and the ns region of the SFV4 produces a highly virulent virus, while the reciprocal recombinant is asymptomatic. Furthermore, replacement of the nsP3 region of rA774 with that of SFV4 reconstitutes the virulent phenotype, demonstrating that nsP3 of the virulent strain is sufficient to restore neurovirulence in an avirulent strain (175). It has been reported that deletions in the nsP3 hypervariable domain attenuate virulence after peripheral inoculation (186). However, the viruses used in this study were also severely attenuated in cell culture. Thus, it is more likely that compromised replication, and not some specific nsP3-related *in vivo* property, was the main contributor to the observed nonvirulent phenotype. Furthermore, it was also clear that the nsP3 gene might not be the only virulence factor and that other determinants must be present in the ns region (128). In contrast to the relatively well-studied differences between the virulence of A7(74) and SFV4, the molecular bases of the different phenotypes of SFV4 and L10 are not known.

2.7.4. Immune response in the mouse CNS

Viral infection triggers the production and secretion of a specific set of pro-inflammatory cytokines, including type I IFNs (IFN-α/β), which play a central role in the host innate immune system and subsequent activation of adaptive immunity. In the case of alphavirus infection, the type I IFN system is important for the initial control of infection, and both the humoral and cellular arms of the adaptive immune response participate in the elimination of the virus (8, 187).

The infection of immunocompetent cells with SFV is followed by a rapid response. Viral dsRNA is recognized by specific pathogen recognition receptors (e.g., toll-like receptors, cytoplasmic RNA helicases), which leads to the induction of type I IFNs (8). Furthermore, SFV replicase has also been shown to convert host cell RNAs into 5′-ppp dsRNA and induce IFN through the retinoic

acid-inducible gene I (RIG-I) pathway (188). Both A7(74) and SFV4 activate type I IFN gene expression in the mature mouse brain, and this expression is proportional to the level of viral RNA (179, 189). A7(74) displays rapid virulence in IFNAR–/– animals lacking functional IFN-α/β receptors. The same is observed for SFV4: IFNAR–/– animals die from infection before the virus can reach the CNS. These data indicate that the type I IFN system strongly and successfully suppresses SFV by preventing widespread dissemination of the virus in extraneural tissues. Consistent with this, the lack of type I IFN response allows a more prominent A7(74) infection of ependymal cells lining the ventricles. In sharp contrast, the inability of A7(74) to infect and spread in adult brain neurons is not affected (179). Thus, the inability of A7(74) to infect neurons is not due to the type I IFN response.

Both avirulent and virulent viral strains induce type I IFNs in *in vitro* cell culture (although to very different extents, see 5.3 and Fig. 7). In addition, they display different sensitivities to these cytokines. In IFN-treated cells, the avirulent strain shows no cytopathic effect and has a 100-fold lower viral yield than the virulent L10. The reduced susceptibility of the L10 strain to the action of IFN allows it to overcome the established IFN-induced antiviral state of the cell, increasing its virulence (190). The virulent strain's tolerance to type I IFN is associated with the nsP3-nsP4 gene region of the virus and is distinct from the genetic loci responsible for SFV neurovirulence (191).

The recovery of adult mice from infection by an avirulent virus requires immune-mediated viral clearance. The clearance is a nonlytic process and, based on results with experiments utilizing immunodeficient animal models, requires both functional T- and B-lymphocytes (117). In athymic *nu/nu* mice lacking T-lymphocytes, the blood titers of A7(74) are reduced to undetectable levels, but the brain titers remain high (192). In μMT mice, which lack functional B-cells and antibodies, the clearance of A7(74) is impaired, and the virus persists in both the serum and brain (177).

In adult animals, A7(74) causes subclinical encephalitis that involves inflammatory demyelination of axons. Despite prolonged brain virus titers, no demyelinating lesions can be observed in infected *nu/nu* or SCID mice (184). Demyelination occurs in the brains of μMT mice, indicating that antibodies are not required for this process (177). The most likely mechanism underlying the generation of these lesions is $CD8^+$ T-cell-mediated elimination of infected oligodendrocytes (176, 193)

3. AIMS OF THE STUDY

Many of the studies investigating SFV neurovirulence have been conducted using reverse genetics, which became possible following the construction of molecular clones of virus strains such as rA774 and SFV4. These icDNA clones were developed long before the availability of next-generation sequencing. Therefore, it was originally not possible to verify whether the sequences of these clones corresponded to the consensus sequences of the respective strains. In addition, although the molecular clone derived SFV4 virus and natural L10 virus are both virulent strains, they are not identical because low-dose i.p. inoculations result in consistent phenotypic differences. Thus, the findings obtained using the molecular clones and natural isolates are often difficult to interpret. Therefore, in the present study, we wanted to identify the consensus sequences of the natural A7(74) and L10 strains, generate new molecular clones corresponding to these sequences to confirm prior findings, and obtain new data regarding the determinants of SFV neurovirulence. The main objectives of this study were as follows:

1. Construct a consensus clone of the L10 strain and assess the molecular determinants and mechanism(s) responsible for the phenotypic differences between SFV4 and L10.
2. Use the consensus clones of A7(74) and L10 to confirm the role of nsP3 in neurovirulence and reveal the molecular bases for the phenotypic difference.
3. Characterize other potential determinants important for SFV neurovirulence and reveal their molecular basis.
4. Use a new approach to analyze the proteins associated with SFV replication organelles to expand knowledge regarding alphavirus-host interactions.

4. MATERIALS AND METHODS

Methods of basic molecular biology, virology and cell biology used in the present study are described in detail in materials and methods sections of publications **I**, **II**, **III** that are included into this dissertation. Original methodology developed for magnetic separation of alphavirus replication organelles is briefly described in section 5.4 and is provided in detail in paper **III**. Therefore, descriptions of these methods are not repeated here.

Methods used for sections of results that are not described in attached publications include:

a. **Analysis of IFN-β production in virus-infected cell cultures**
Mouse fibroblast COP5 cells (194) were maintained in L-glutamine containing Iscove's Modified Dulbecco's *Medium* (IMDM, Gibco) supplemented with 10% fetal bovine serum (FBS). Stocks of SFV6, SFV6-74, SFV6-RE, SFV6-74-RE, A774wt, A774wt-6, A774wt-HV and A774wt-6-HV were obtained and concentrated by ultracentrifugation as described in papers **I** and **II**. COP5 cells were infected at a multiplicity of infection (MOI) of 0.1 with wt or recombinant viruses. At 24 h postinfection the supernatants were collected; the VeriKine-HS^{TM} Mouse IFN Beta Serum ELISA Kit (PBL Assay Science) was used to measure levels of IFN-β according to manufacturer's instructions.

b. **Construction and use of SFV trans-replicase system**
Construction of plasmids for expression of SFV replicase was performed as follows.

First, part of the SFV6 cDNA sequence, consisting of fragments spanning from residue 85 to residue 280 (including *Eco* RV restriction site at position 277) and from residue 6712 to 7422 (including *Bgl* II restriction site at position 6715) flanked with *Bsp* 119I (at 5' end) and *Spe* I (at 3' end) restriction sites was ordered as synthetic DNA (GenScript, USA). The initiation codon for the reading frame of ns proteins was placed in a strong Kozak context. This sequence was cloned into a pUC57Kan vector plasmid. The remaining part of the ns polyprotein encoding region was transferred to this plasmid from plasmids encoding for the full length sequences of SFV6, SFV6-74, SFV6-RE, SFV6-74-RE, A774wt, A774wt-6, A774wt-HV and A774wt-6-HV using *Eco* RV and *Bgl* II restriction sites. In clones encoding for the replicase of A774wt, A774wt-6, A774wt-HV and A774wt-6-HV the Arg2429 residue (numbered according the P1234 of SFV) was replaced with a Lys residue (the natural amino acid residue of A774wt at this position). The inserts corresponding to the regions encoding for ns polyproteins were placed between the immediate-early promoter of human cytomegalovirus (CMV) and the simian virus 40 (SV40) late polyadenylation signal in the pMC-gtGTU2 expression vector plasmid (FIT Biotech Plc, Finland) using *Bsp* 119I and *Spe* I restriction sites. Sequences of all obtained expression constructs were verified using Sanger sequencing.

The expression cassette encoding the truncated SFV mini-genome RNA comprised the following elements: the CMV promoter, the 5' UTR of SFV6 together with first 74 codons of the nsP1 encoding region, the second intron from the human beta globin gene, the sg promoter of SFV (from position –70 to +51 with respect to the sgRNA transcription start site), the coding sequence of *Gaussia* luciferase (Gluc) reporter, the 3' UTR of SFV6 followed by 69 adenine residues, the hepatitis delta virus negative strand ribozyme, and the SV40 late polyadenylation signal. The cassette was obtained as synthetic DNA (GenScript, USA) and cloned into a pBluescript KS vector.

Transfection and IFN-β measurements were performed as described in (188). Briefly, COP5 cells were transfected with constructed plasmids encoding either wt or recombinant SFV replicase using Lipofectamine 2000 (Invitrogen). At 24 h posttransfection the supernatants were gathered and the amount of IFN-β secreted into the cell culture medium was measured using a commercial VeriKine-HS^{TM} Mouse IFN Beta Serum ELISA Kit (PBL Assay Science) according to the manufacturer's instructions. The ability of wt and recombinant replicases to replicate SFV mini-genome was analyzed by measurement of Gluc activity in COP5 cells co-transfected with plasmids encoding for SFV replicase and mini-genome.

5. RESULTS AND DISCUSSION

5.1. The ability of SFV to enter the CNS is determined by charged amino acid residues on the surface of E2 glycoprotein (I)

Alphaviruses, similarly to other RNA viruses, have error-prone polymerases and thus exist as a mixed population of genotypes called quasispecies. Although SFV4 and L10 originate from the same pool of mosquitoes, they have a different passage history and display different phenotypes in mouse model. Following i.c. infection or i.p. inoculation with a very high dose ($\geq 10^6$ plaque forming units, PFUs) of virus, the pathogenesis of SFV4 and L10 is similar: both viruses are virulent, and the infected mice die from encephalitis. However, following low dose i.p. inoculation – a more biologically relevant condition, as mosquitoes cannot deliver a high dose or i.c. infection – L10 still produced high-titer plasma viremia and was rapidly and efficiently neuroinvasive, generating fatal panencephalitis, whereas SFV4 produced low-titer viremia, was rarely neuroinvasive, and displayed low virulence (**I**, Fig. 1 and Fig. 2). All the animal experiments in publications **I** and **II**, unless stated otherwise, were performed using low dose (5000 PFUs) i.p. inoculation.

5.1.1. The phenotypic differences between SFV4 and L10 map to E2 position 162

To determine the genetic differences that are responsible for the different phenotypes, we sequenced an archived stock of L10 strain of known provenance using high-throughput sequencing. The obtained consensus sequence was compared with the published sequence of SFV4 (GenBank accession number KP699763), and twelve nucleotide differences throughout the genome were identified (**I**, Table 1). Six of these changes, one in nsP3, two in the capsid protein, and three in the E2 protein coding sequence, were nonsynonymous. Although nucleotide changes that do not affect the amino acid sequence of virus-encoded proteins could also be responsible for the phenotypic differences, it is more likely that one or more of the six nonsynonymous differences were responsible for the variance in neuroinvasion. Therefore, these six changes were engineered into the SFV4 icDNA clone to generate a new molecular clone; the virus produced from this clone was designated SFV6 (GenBank accession number KT009012). The amino acid sequences of the proteins encoded by SFV6 are identical to the consensus amino acid sequences of the proteins encoded by L10. In BHK-21 cells, the new SFV6 virus as well as other recombinant viruses described below grew to titers similar to those of SFV4, with no significant differences in growth kinetics. However, compared to SFV4, SFV6 produced larger plaques on monolayers of BHK-21 cells. This difference

in the *in vitro* phenotype became an indirect indicator of the neuroinvasive ability of recombinant viruses. Following a low dose i.p. inoculation into BALB/c mice, SFV6 produced high-titer viremia, was efficiently neuroinvasive and all of the infected mice rapidly reached clinically defined endpoints (**I**, Fig. 1). Therefore, it was concluded that we had generated an icDNA molecular clone, which, for the first time, produced SFV virus with the same phenotype as L10 and the prototype biological strains.

To further investigate the effect of these six nonsynonymous differences on the phenotype of the virus, the corresponding substitutions were introduced, individually or in combination, into the SFV4 or SFV6 molecular clones, and recombinant viruses were generated and analyzed. SFV4-nsP3 (SFV4 containing the E1384A substitution in the nsP3 region) remained avirulent. In contrast, SFV4-struct (SFV4, in which the structural region was replaced with that of SFV6) produced high titers in the blood, was detected in the brain, and was virulent (**I**, Fig. 1). Therefore, it was concluded that one or more of the amino acid substitutions in the structural polyprotein had to be responsible for the differences in neuroinvasion.

Replacement of the capsid protein of SFV6 with that of SFV4 (SFV6-C) did not change the neuroinvasive phenotype: the virus remained virulent. However, when the E2 envelope protein of SFV6 was replaced with that of SFV4 (SFV6-E2), the virus became phenotypically similar to SFV4, revealing that the genetic locus responsible for the differences in viremia, neuroinvasion, and virulence mapped to E2 (**I**, Fig. 1). The ability of a single-amino-acid change in E2 to affect neurovirulence in mice has been observed previously for SINV, VEEV, and EEEV (195-198). Indeed, subsequent analyses revealed that all re-combinant viruses carrying a lysine residue at position of 162 of E2 (162K) had small plaques (similar to SFV4, which also has 162K) on BHK-21 cells; in contrast, viruses with glutamic acid residue in the same position (162E) produced large plaques similar to SFV6. Like SFV4, the small plaque virus SFV6-162K generated low-titer viremia and was not detected in the brain. The large plaque SFV4-162E, like SFV6, generated high-titer viremia and was detected in the brain (**I**, Fig. 1). Based on these data, it was concluded that E2 162E is a determinant of SFV6 (and correspondingly, L10) neurovirulence. Interestingly, position 162 of E2 has been described previously as a molecular determinant of SFV virulence (199). However, in that study, infection of pregnant BALB/c mice with high-dose SFV4 was attenuated if E2 162 was changed from lysine to glutamic acid. The brain titers of these viruses were not reported, making it difficult to compare the data to our present findings. It may be hypothesized that the charge at E2 162 affects the ability of the virus to cross the placenta and/or that a virus with enhanced binding to heparin sulfate (see 5.1.3) had some growth advantage in pregnant mice.

5.1.2. The biological stock of L10 is heterogeneous at position 247 of E2

Although the amino acid sequence of SFV6-encoded proteins matched the consensus amino acid sequence of L10 proteins, the inoculation of these two viruses into mice revealed a small but consistent difference in phenotype. Viremia for SFV6 was higher than for L10 on postinoculation day (PID) 1 (**I**, Fig. 1B) and, in general, SFV6-infected mice also reached the clinical endpoint of the experiment earlier than L10-infected mice. The L10 stock used in this study originated from the virus that was isolated in 1942 (170) and underwent various recorded passages, including *in vitro* passages in chicken embryo fibroblasts, after which it was stored frozen for 32 years. The stock was then passaged once on BHK-21 cells to generate new stocks of the virus, which were used to study the acute encephalitis induced by L10 virus (200). To our knowledge, this L10 strain stock had never been plaque-purified and therefore was assumed to contain viruses with different properties. Indeed, the L10 stock produced distinct (small, medium and large) plaques on BHK-21 cells (**I**, Fig. 3A), which indicated that the L10 stock contained at least three different genotypes. In mice, the small-plaque virus produced low viremia, was undetectable in the brain, and was avirulent, whereas the large-plaque virus replicated well in the periphery, entered the brain, and was virulent (**I**, Fig. 3C – 3F). In the small-plaque viruses, one nucleotide difference (compared to the L10 consensus sequence), at position 9160 led to the replacement of a negatively charged glutamic acid (L10 consensus) to a positively charged lysine residue at position 247 of the E2 protein (E247K). This E2 variation was also found by analysis of the high-throughput sequencing data of the L10 stock (**I**, Fig, 3B), demonstrating the relative abundance of the corresponding genomes. As the corresponding recombinant virus (SFV6-247K) was avirulent (**I**, Fig 4), it was concluded that the charge of amino acid residue 247 of E2 is another determinant of SFV viremia and neuroinvasion. The presence of quasispecies with various extents of virulence in the L10 stock highlights the importance of molecular clones in research and, therefore, the usefulness of SFV6. Furthermore, it demonstrates the importance of minimal *in vitro* passaging of wt viruses before engineering a molecular clone.

The L10 large-plaque variant and SFV6 carried a negatively charged glutamic acid at positions 162 and 247 of E2 and were both virulent. Small plaque viruses had a positively charged lysine residue in at least one of these positions (small-plaque variant of the historic L10 stock at position 247 and SFV4 at position 162), produced low-titer viremia and displayed low virulence. Finally, both E162K and E247K mutations in the E2 protein converted SFV6 into a small-plaque, low-level viremia, nonneurovirulent virus (**I**, Fig. 4), confirming that each of these E2 charge reversions was sufficient to cause this phenotypic change. These mutations, which are apparently unfavorable for *in vivo* infection, most likely originate from the passaging of virus stocks in cell

culture and/or i.c. passaging in mouse brain (see 5.1.4). Indeed, serial passaging of four biological replicates of SFV6 on BHK-21 cells confirmed that this process selects for positively charged amino acid residues in E2 (**I**, Table 2). Similar changes, resulting in the appearance of positively charged amino acid residues in the E2 proteins, introduced by site-directed mutagenesis or selected for by passaging, have also been shown to cause a reduction of peripheral virus titers, rapid clearance from the blood, and low virulence in other alphaviruses (195, 201, 202).

5.1.3. Charged amino acid residues at positions 162 and 247 of E2 modulate the binding of SFV virions to heparan sulfate

Based on the analyses of the 3D structure of the alphavirus E1/E2 heterodimer, the amino acid residues 162 and 247 were predicted to lie within the acid-sensitive region of E2. The residues are located on the surface of the heterodimer; similarly, they clearly lie on the surface of the 3D structure of the spike (**I**, Fig. 7). Therefore, positions 162 and 247 of E2 are readily available to bind proteins and other molecules such as cell surface receptors. In SINV the selection for positively charged amino acid residues in E2 promotes binding to the glycosaminoglycan (GAG) heparan sulfate (HS) (201). Furthermore, mutations analogous to those described herein have been shown to affect the binding of virions to HS in other alphaviruses, including RRV, CHIKV, EEEV and VEEV (195, 203-205).

GAGs such as HS are ubiquitously expressed, negatively charged poly-saccharides that are found on the surface or in the extracellular matrix of both vertebrate and invertebrate cells (206-208). In general, proteins bind to negatively charged HS via positively charged lysine and arginine residues. Studies have shown that when HS-binding proteins are injected intravenously, they are rapidly cleared from the circulation through binding to tissue HS (209, 210). This clearance can be interrupted by the injection of heparinase, which digests HS on the tissue surface, or by coinjecting with heparin (209, 211). Thus, virions of alphaviruses that efficiently bind to HS are most likely cleared from the periphery via the same mechanism, resulting in reduced viremia. This phenomenon may be due to absorption of the HS-binding virus in the liver, an organ rich in GAGs (195).

The present study confirmed that different strains of SFV and recombinant viruses vary in their interactions with HS. Positively charged lysine residues either at position 162 or 247 in E2 of SFV augment entry of the virus into CHO-K1 or BHK-21 cells through increased binding to GAGs such as HS (**I**, Fig. 8A, 8B). Furthermore, in a competition assay, preincubation with heparin efficiently reduced the number of plaques of SFV variants harboring E2-162K or E2-247K residues. The SFV variants carrying a glutamic acid residue at either of these positions were less affected (**I**, Fig. 8C). This data is consistent with recent

findings demonstrating that the interactions of CHIKV E2 with GAGs diminish dissemination of the virus to lymphoid tissues and stimulation of inflammatory responses (212). Increased binding to HS has also been associated with a decrease in the neurovirulence of VEEV and the flavivirus Murray Valley encephalitis virus (195, 213).

The data illustrates ability of SFV (and that of RNA viruses in general) to adapt quickly to different conditions (including cell culture conditions) during serial passaging. The SFV strains were isolated a long time ago and went through extensive *in vitro* and *in vivo* passaging before methods allowing construction of stable molecular clones were developed. Changes previously acquired by viruses, such as changes in viral envelope proteins favorable for i.c. (see 5.1.4) and cell culture propagation, were fixed in the sequences of obtained molecular clones. As we saw with SFV and as shown for other alphaviruses, these differences can lead to major changes in the *in vivo* properties of the virus. Therefore, the data obtained with viral strains passaged multiple times in cell culture or with molecular clones obtained from these strains should be interpreted with caution, as properties of such viruses do not necessarily correlate with those of original biological stocks. Now that next generation sequencing is readily available, new virus isolates should first be sequenced, and molecular clones corresponding to the consensus sequence can then be obtained as synthetic DNA and used in subsequent experiments. However, for many model viruses, the original biological samples no longer exist. In the case of stocks with long passage history it is difficult, if not impossible, to be sure that their current phenotypes correspond to those at the time of their isolation.

5.1.4. Amino acid residues 162 and 247 in E2 affect SFV replication in the mouse brain following i.c. inoculation and the effect of virus infection on the integrity of a BBB model

To determine whether amino acid residues 162 and 247 of E2 affect SFV replication in the brain, parental and recombinant viruses were i.c. inoculated into mice, and brain titers were determined. As expected, all of the viruses replicated in the brain; however, SFV4, SFV6-162K, and SFV6-247K had higher titers than SFV6 and SFV4-162E (**I**, Fig. 5). This result is consistent with previous reports showing that wt SFV4 inoculated intranasally (a direct neural route to the CNS) is more virulent than SFV4 mutant harboring residue 162E in E2 (214). Similarly, greater binding to HS is associated with increased neurovirulence in neonatal mice infected with SINV, in adult mice inoculated i.c. with SINV, and in adult mice inoculated i.c. with a natural North American strain of EEEV (197, 198, 215). Thus, lysine residues at positions 162 and 247 of E2 of SFV attenuate peripheral SFV replication but, at the same time, enhance its replication in the CNS. Therefore, it is possible that presence of virus

variants carrying the E2 247K residue in the L10 stock may originate from repeated i.c. passaging (rather than from *in vitro* passaging) of this SFV strain.

To investigate whether neuroinvasion was related solely to the level of viremia or also to the ability of the virus to cross the BBB, the ability of different viruses to traverse an *in vitro* model of the BBB (Corning Transwell co-culture system) was assessed. SFV6 or SFV6-162K was added to the luminal side of the BBB, and the integrity of the BBB was measured. In this experiment, avirulent SFV6-162K had a greater effect on the integrity of the BBB (**I**, Fig. 6A); additionally, a greater amount of virus was detected on the abluminal side of the BBB following infection with SFV6-162K compared to SFV6 infection (**I**, Fig. 6B). However, it is not clear how well these findings reflect the *in vivo* infection because cells in culture have more GAGs on their surface; thus, the increased ability of SFV6-162K to cross the artificial BBB may simply reflect this feature. Similarly, it has been reported that the ability of strains of HIV to cross an *in vitro* BBB correlates with their HS binding efficiency (216, 217). Taken together, these results suggest that the ability of SFV to cross the *in vitro* BBB correlates with its enhanced replication ability in the brain but not with the ability to generate high-titer viremia.

5.2. The neurovirulence of SFV is affected by differences in determinants of ns polyprotein processing and by the sequence of nsP3 (II)

In adult mice, infection with neurovirulent SFV strains leads to lethal encephalitis, whereas mice infected with the avirulent A7(74) strain remain asymptomatic. Previously, a molecular clone of A7(74) designated as rA774 was constructed, and the rescued virus was compared to SFV4. The determinants of virulence were mapped to the ns region of these viruses. Moreover, the importance of nsP3 for SFV virulence was established (175). However, as shown in publication **I,** SFV4 is not truly virulent because its ability to replicate in the periphery is attenuated. In addition, we observed that *in vitro* transcripts of rA774 display an abnormally low quality and infectivity, suggesting that this molecular clone may also carry uncharacterized functional defect(s).

5.2.1. A single amino acid difference from the A7(74) consensus sequence causes low infectivity of rA774

The consensus sequence of a biological stock of A7(74) of known provenance was determined using high-throughput sequencing and compared with rA774 (175). This revealed a single nonsynonymous difference in codon 1347 (corresponding to amino acid residue 11 in nsP3) of the ORF of the ns polyproteins: in the consensus sequence of A7(74) it encodes Ile, in rA774 it encodes Val. The correct molecular clone of A7(74) was obtained by changing

the Val1347 codon in rA774 to the Ile codon. This clone, which was designated A774wt, had efficiency of infectious virus rescue over 100-fold higher than rA774, confirming that the presence of the incorrect amino acid at position 1347 was the cause of the low infectivity of the rA774 clone (**II**, Fig. 1B). Interestingly, this substitution was previously mistakenly thought to represent a natural sequence difference between SFV4 and A7(74), and the correct A7(74) molecular clone (designated rA774-V11I) was actually constructed and to some degree even analyzed (128). Thus, the properties of viruses rescued from icDNA clones of SFV4 and rA774 differ from those of their respective SFV biological isolates, and the results obtained with them are difficult to interpret. Hence, it was important to verify and further characterize the determinants of virulence residing in the ns region of SFV using new consensus clones.

Combined data from three experiments revealed that A774wt killed 5 of 51 mice and was therefore, in contrast to the original avirulent A7(74) (176, 218), slightly virulent. However, this finding is consistent with previously published data for rA774-V11I (128). Our earlier *in vivo* experiments demonstrated that SFV6 is also slightly more virulent than the parental L10 strain, which indicates that slightly enhanced virulence may be a common property of consensus clones of SFV (and possibly other alphaviruses). SFV6 and A774wt correspond to the most common (and possibly the fittest) sequences of the L10 and A7(74) strains, and unlike the native isolates, P_0 stocks (viruses rescued directly from infectious clones) display little sequence variation and may therefore have a higher virulence.

5.2.2. The introduction of synonymous changes reduces recombination between two copies of nsP3-encoding sequences

The analysis of chimeras between SFV4 and rA774 demonstrated that nsP3 is the main virulence determinant of SFV4 (175). Thus, either nsP3 of SFV4 is the factor that causes virulence, or nsP3 of rA774 serves as the factor that restricts SFV virulence. However, it remains unclear which phenotype caused by nsP3 – the virulent or the avirulent phenotype – is dominant. To directly address this question, recombinant viruses carrying two nsP3-encoding regions were constructed and analyzed. In these viruses, the region encoding nsP3 of SFV6, which lacks the opal terminator, was used as upstream copy of nsP3. This order of nsP3 regions ensured that the recombinant viruses would express both nsP3 proteins at equivalent levels.

It should be mentioned that although viruses containing extra sequences represent useful research tools, the duplication of RNA virus genes represents a significant challenge. This is because duplication of nsP3 increases the size of the SFV genome by approximately 12.5%, which slows down its replication. It also creates an unnatural processing site between the two nsP3 regions in the

P12334 polyprotein. A previous study demonstrated that this type of manipulation is tolerated (219) but may still affect the biological properties of the virus.

The first attempt to construct such tools failed: we were unable to obtain homogeneous stocks of viruses containing two identical or highly similar nsP3-encoding regions (the corresponding sequences of A774wt and SFV6 have an identity of 96%). Genetic analysis confirmed that copy-choice recombination between these identical (or nearly identical) nsP3 sequences was highly efficient and generated viruses with a single copy of nsP3. To increase the stability of the recombinant genome, the similarity between two copies of nsP3-encoding regions was reduced to a nonsignificant level by the introduction of numerous synonymous changes into the upstream copy of nsP3 encoding region. The resulting panel of viruses comprised A774wt, SFV6, and double nsP3 viruses designated as A774wt-6/74, A774wt-6/6, SFV6-6/74, and SFV6-6/6. All of the recombinant constructs exhibited an infectious virus rescue efficiency that was similar to the parental clones (**II**, Fig. 1B). However, viruses carrying two copies of nsP3 exhibited a delay in accumulation of infectious progeny and reached titers that were approximately one log lower than the parental virus titers (**II**, Fig. 1C, 1D), likely due to the larger sizes of their genomes.

Genetic analysis of P_0 stocks of recombinant viruses confirmed the presence and stability of viruses harboring two copies of nsP3 (**II**, Fig. 3A). Western blot analysis revealed that these viruses expressed somewhat more nsP3 than A774wt and SFV6 (**II**, Fig. 3B). At the same time, an extra copy of nsP3 did not appear to interfere with the localization of virus replication organelles or individual ns proteins (**II**, Fig. 2A, 2B). Additionally, the presence of an extra copy of nsP3 did not cause major defects in the ns polyprotein processing patterns, as indicated by the results of pulse-chase experiments. The expression of two nsP3 proteins with slightly different mobilities was clearly detected for SFV6-6/74. In SFV6-6/6-infected cells, a band corresponding to the P33 polyprotein was observed (**II**, Fig. 3C). An unexpected finding was the stabilization of nsP4. In sharp contrast to SFV6-infected cells, the amount of free nsP4 in SFV6-6/74- and SFV6-6/6-infected cells increased during the chase period, clearly indicating that nsP4 in these viruses was protected against degradation (**II**, Fig. 3C).

5.2.3. Neurovirulence is the dominant function of nsP3 of SFV6

Following a detailed *in vitro* analysis, the panel of viruses was used in *in vivo* experiments to determine which phenotype associated with nsP3 – virulent (SFV6) or avirulent (A774wt) – was dominant. The animal experiments revealed that all of the mice inoculated with SFV6 reached a clinical endpoint; in contrast, only 3 of 37 mice infected with A774wt developed disease. Both A774wt-6/74, which killed 15 mice, and A774wt-6/6, which killed 19 of 37 mice (**II**, Fig. 4A), were significantly more virulent than the parental A774wt

but less virulent than SFV6. SFV6-6/74 and SFV6-6/6 were virulent and killed five and four out of seven mice, respectively (**II**, Fig. 4B).

Both A774wt and SFV6 produced high-level viremia and were detected in the brain (**II**, Fig. 4E). The brain titers of SFV6 were significantly higher than those of A774wt, confirming that only the virulent strain, SFV6, replicated efficiently and spread rapidly throughout the brain. Due to their slower replication, the viruses harboring two copies of nsP3 produced significantly lower blood titers on PID 1 (approximately 10^4 PFU/ml on average) compared with A774wt and SFV6 (**II**, Fig. 4C). By PID 3, this difference was less prominent (**II**, Fig. 4D). Interestingly, on PID 3, double nsP3 viruses were detected in the brains of 7 of 20 mice, indicating that low viremia did not prevent entry into the brain (**II**, Fig. 4E). By the time animals had reached the terminal endpoint of the experiment and were euthanized (PID 7–10), the brain titers of these viruses were similar to the titers of virulent SFV6. In addition, RT-PCR analysis revealed that the viruses that had replicated in the brain maintained the extra copy of nsP3 (**II**, Fig. 4F). Therefore, double nsP3 viruses were not only able to reach the brain but also replicated and accumulated in brain tissues.

The ability of viruses that were unable to induce high-titer viremia to enter the brain is inconsistent with the findings reported in publication **I**, in which an inability to cause high-titer viremia clearly hampered the virulence of SFV4. SFV4 infection resulted in viremia with a titer $>10^4$ PFU/ml at PID 1 and $\approx 10^2$ PFU/ml at PID 3, which was insufficient for the entry to the brain (**I**, Fig. 1). Infection with A774wt-6/6, which was apparently the most stable among the viruses with two copies of nsP3 (**II**, Fig. 4F), resulted in even lower blood titers ($\approx 10^2$ PFU/ml both at PID 1 and PID 3, **II**, Fig. 4C, 4D). However, in contrast to SFV4, A774wt-6/6 entered and replicated in the brain to high titers, ultimately killing 19 of 37 mice (**II**, Fig. 4A). Thus, a low blood titer *per se* does not prevent entry of SFV to the mouse brain. Several explanations may be provided for this apparent contradiction. In the case of SFV4, the low viremia results from enhanced clearance of the virus from the blood, most likely due to the binding of virions to HS. As a result, the decline in the virus titer (from PID 1 to PID 3) is rapid and irreversible. In contrast, low titers of A774wt-6/6 and other viruses carrying two copies of the nsP3 gene are caused by their slower replication (**II**, Fig. 1C, 1D), and the decline in blood titers from PID 1 to PID 3 is small or even undetectable (**II**, Fig 4C, 4D). Thus, viremia caused by these viruses persists for a longer duration, providing more opportunities to enter the brain. In addition (or alternatively), the different abilities to bind HS result in the attachment of virions to different receptors and cell types, which may affect the entry of SFV into the mouse brain. Finally, short-term but efficient replication of SFV4 and longer but slower replication of A774wt-6/6 (and similar viruses) may cause different antiviral innate immune responses, which may also affect the outcome of infection.

The similar virulence of SFV6-6/74 and SFV6-6/6 indicated that nsP3 of A774wt was unable to efficiently reduce the virulence of SFV6. Consistently,

both A774wt-6/74 and A774wt-6/6 killed approximately 50% of the infected mice (**II**, Fig. 4A). This result clearly demonstrated that the insertion of nsP3 of SFV6 into the A774wt genome increased the virulence of the recombinant viruses regardless of the presence (A774wt-6/74) or absence (A774wt-6/6) of nsP3 from the avirulent strain (**II**, Fig. 4A). Thus, the neurovirulent phenotype caused by the presence of nsP3 of SFV6 was dominant. Alternatively, it is possible that the phenotype associated with the first copy of nsP3 (which was always nsP3 from the virulent SFV6 in the present study) was dominant. To test this hypothesis, viruses carrying the reverse order of the nsP3 copies (nsP3 from A774wt followed by nsP3 from SFV6) should be analyzed. The challenge associated with such an experiment is that the opal stop codon at the end of the nsP3-encoding region of A774wt would drastically reduce the expression level of the second copy of nsP3. This phenomenon would make it virtually impossible to compare the data obtained using such recombinant viruses with that presented above, and thus the experiment was not performed.

Consistent with previously published results (175), we found that A774wt-6 was virulent (**II**, Fig. 5B). The virus produced high-titer viremia and had significantly higher titers in the brain at PID 3 than did A774wt (**II**, Fig. 5). Surprisingly, however, SFV6-74 was as virulent as parental SFV6 and killed all of the infected mice with essentially the same kinetics (**II**, Fig. 5B). These data confirmed the presence of other virulence determinant(s) in the ns region of SFV6 (outside of nsP3) and revealed that nsP3 of A774wt could not eliminate the virulent properties of these determinants.

5.2.4. Differences in the P4 position of the 1/2 cleavage site of P1234 and the S4 subsite of nsP2 protease affect the processing of ns polyprotein and virulence

The sequences of the nsP4 proteins of A7(74) and L10 are nearly identical, and therefore the determinant(s) responsible for the contrasting virulence of A774wt and SFV6-74 most likely reside(s) in the nsP1 and/or nsP2 region(s) (**II**, Table 1). In a parallel study, we observed that SFV4 showed poor tolerance to a His to Arg substitution in the P4 position of the 1/2 cleavage site of P1234. This mutant survived only through the generation of compensatory changes that sometimes occurred at amino acid residue 515 of nsP2 (Lulla et al., unpublished). Another study also predicted the functional linkage between position P4 of the 1/2 site and amino acid residue 515 of nsP2, which was shown to be part of the protease S4 subsite (220). Therefore, we were interested in determining whether these amino acid residues, which are His and Val in L10/SFV6 but Arg and Glu in A7(74)/A774wt, are important for SFV virulence. Two new viruses, designated A774wt-HV and SFV6-74-RE, were constructed and analyzed (**II**, Fig. 6A). A774wt-HV had a virulent phenotype and killed 6 of 7 mice (**II**, Fig. 6B). In contrast, SFV6-74-RE was clearly less pathogenic than SFV6-74 and killed only 2 of 7 mice (**II**, compare Fig. 6B and Fig. 5B). These data demonstrated

that differences in the P4 position of the 1/2 cleavage site and the S4 subsite of the nsP2 protease affect the virulence of SFV.

In the virulent strain of SINV, a mutation in nsP1, which is located close to the 1/2 site of the ns polyprotein, results in the acceleration of P123 processing and significant attenuation of neurovirulence (221). To determine whether substitutions altering the virulence of SFV also affect ns polyprotein processing, pulse-chase experiments with A774wt, A774wt-HV, SFV6-74 and SFV6-74-RE were performed. This experiment revealed that the speed of P123 processing clearly differed between avirulent and virulent viruses: cleavage of the 1/2 site was faster in A774wt and SFV6-74-RE and slower in A774wt-HV and SFV6-74 (**II**, Fig. 7A, 7B). Interestingly, the reverse was observed for processing of the 3/4 site: the P34 polyprotein was efficiently cleaved in viruses with Val515 in nsP2 (A774wt-HV and SFV6-74) but was not so readily cleaved in A774wt and SFV6-74-RE (**II**, Fig. 7A). Most likely the reduction of basal protease activity of nsP2, caused by the Val515 to Glu substitution, compensated for the presence of the Arg residue (the most favorable residue for fast processing) in the P4 position of the 1/2 site.

These data and the results obtained with SINV (221) demonstrate a correlation between the speed of ns polyprotein processing and neurovirulence. However, the mechanism by which the accelerated cleavage of the 1/2 site and/or reduced protease activity of nsP2 affect the *in vivo* phenotype of the virus is less obvious. One possibility is that faster processing of the 1/2 site allows less time for the proper formation of RCs and may reduce the amount of functional RCs in some types of infected cells (possibly in mature neurons). Alternatively (or in addition), it is possible that the reduced protease activity of nsP2 or altered kinetics of P123 processing affect other virus/cell interactions, such as the induction of type I IFN production or interference with IFN signaling. In SINV, accelerated processing of the 1/2 site diminishes the ability of the virus to reduce STAT1/STAT2 phosphorylation (151). Our unpublished results clearly demonstrate that mutations that affect SFV neurovirulence also affect the ability of the virus to induce type I IFN response (see 5.3).

5.2.5. The role of nsP3 in SFV neurovirulence and replication

The mechanism(s) by which nsP3 affects the virulence of SFV remains even more obscure. The most obvious difference between L10 and A7(74) is the opal stop codon that is located close to the 3' end of the nsP3-coding sequence of the latter (**II**, Table 1). However, the effect of the stop codon alone on SFV virulence is limited (175). Hence, the avirulence/ virulence determinants in nsP3 of SFV appear to be more complex. NsP3 has been shown to play a role in interactions with multiple host components (129, 132, 222). It is possible that different strains of SFV interact with different host factors or with the same host factors but with different efficiencies in a cell type/tissue-specific way, which in turn could result in contrasting phenotypes. It is also possible that the speed of

P123 processing (see 5.2.4) and the functions of nsP3 are linked: the two virulence determinants may represent different aspects of the same mechanism (pathway). The initial data concerning type I IFN induction in response to different SFV strains and recombinant viruses is consistent with this hypothesis (see 5.3).

In addition to its rather enigmatic role in SFV virulence, we have described a previously unknown function of nsP3. The analysis of ns polyprotein processing revealed a profound stabilization of nsP4 in viruses carrying a duplicated nsP3 region (**II**, Fig. 3C). This effect was not simply caused by the addition of an extra processing unit into the ns polyprotein of SFV, because the nsP4 of viruses carrying EGFP (or any other marker) between the nsP3 and nsP4 regions (219) is not stabilized. Individual molecules of nsP4 are known to be rapidly degraded (145), whereas, in formed RCs, the nsP4 appears to be stable. However, the mechanisms responsible for this stabilization have not been determined. Our data strongly suggest that nsP3 is actively involved in this phenomenon. The stabilization may occur via the inclusion of excessive amounts of nsP4 into complexes, which could be RCs or some other complexes of nsP3 and nsP4. Analyses of the composition of RCs made by wt SFV and by SFV carrying an extra copy of nsP3 are needed to ascertain whether or not this is the case. To achieve this goal, the methodology described in publication **III** could be applied.

5.3. The role of the type I IFN response in SFV neurovirulence (unpublished)

In the absence of type I IFN response, SFV infection is lethal regardless of the age of the host or strain of the virus (179). Thus, type I IFN is an essential component of the host response to viral infection. Viruses have evolved mechanisms to antagonize the production of and/or cellular response to type I IFN. The contrasting pathogenicity of A7(74) and L10 could depend on the different abilities of these viruses to induce (or to evade) type I IFN production, to interfere with the host type I IFN response, or to resist the antiviral effects of type I IFN.

Previous studies with SINV have demonstrated that a determinant within the 1/2 cleavage region is associated with slower ns polyprotein processing and delayed 26S sgRNA synthesis in the neurovirulent strain AR86. Consistently, accelerated processing and earlier expression from the viral 26S promoter are associated with the nonneurovirulent phenotype of SINV (93, 221). Additional characterization of this determinant demonstrated that the avirulent virus induces higher type I IFN production both *in vivo* and *in vitro* (94), which is probably because, unlike AR86, this mutant cannot suppress STAT1/STAT2 activation in response to type I/type II IFN (151). Interestingly, this phenotype is not specific to the SINV AR86 because a similar mutation introduced into the

1/2 cleavage site of P1234 of RRV also leads to enhanced type I IFN induction (94). Thus, in the case of SINV AR86, pathogenicity correlates with effective suppression of the type I IFN response both *in vivo* and *in vitro*. However, this trend is not a general characteristic of all alphaviruses or even all strains of SINV. There are other determinants of neurovirulence in the E2 glycoproteins of SINV as well as in the 5' UTR (223, 224) that are not associated with altered IFN production. The same is true for phenotypic differences between SFV4 and L10 (**I**). Finally, an opposite effect has been described in two recent papers. First, it was observed that neuroadapted SINV, which kills all C57BL/6 but not BALB/c mice, displayed higher levels of viral replication, higher levels of type I IFN, and slower viral clearance in susceptible mice. This finding strongly suggests that fatal encephalomyelitis in C57BL/6 mice is mediated by the immune response rather than being a direct result of viral infection (225). Second, it has been shown that neurovirulent SFV4 induces greater amounts of IFN-β than the avirulent VA7 strain (variant of A7(74)). In addition, SFV4 is less susceptible to the antiviral effect of type I IFN (191). Thus, enhanced neurovirulence can result either from the ability of the virus to suppress type I IFN production (as in the case of SINV AR86) or from its ability to induce excessive amounts of type I IFN, which can result in an immune-mediated pathology (as in the case of neuroadapted SINV in C57BL/6 mice). Therefore, it became important to analyze the ability of our constructed recombinant viruses (**II**) to induce the production of type I IFN.

First, IFN-β production was analyzed in COP5 cells (murine fibroblasts with an intact IFN response). When the cells were infected with different recombinant viruses at an MOI of 0.1, no IFN-β production was observed at 12 h postinfection (data not shown). This indicated that all of the investigated viruses were capable of counteracting type I IFN production in the initially infected cells. However, at 24 h postinfection, IFN-β was readily detected (Fig. 7). This IFN was produced by cells infected by virus released from the initially infected cells. The same phenomenon has also been described for SINV. It has been proposed that very low, subprotective doses of IFN-β, which are produced by initially infected cells and do not induce the antiviral response in uninfected cells, have a very strong stimulatory effect on the ability of cells to express type I IFN and to activate IFN-stimulated genes during subsequent infection (226). We found that the virulent strain SFV6 induced the highest amount of IFN-β, whereas the avirulent A774wt induced much lower levels of IFN (Fig. 7), which is consistent with data from a previous study (191). However, no direct correlation between the virulence of different recombinant viruses and the amount of IFN-β produced in this experiment was evident. The only conclusion from these data is that all viruses harboring the A774wt backbone induce lower levels of IFN-β than viruses that possess the SFV6 backbone (Fig. 7).

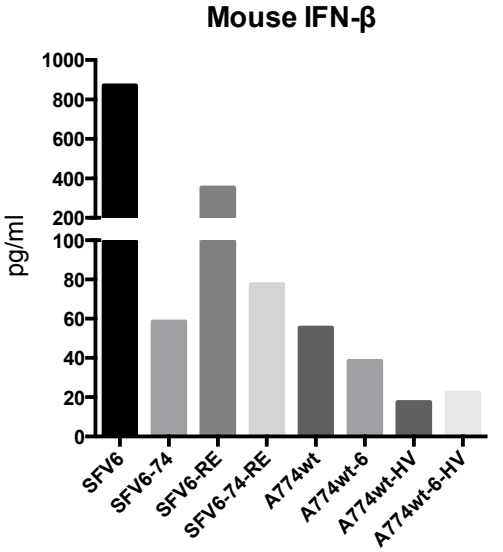

Figure 7. Production of IFN-β by SFV infected cells. COP5 cells were infected at an MOI of 0.1 with wt and the indicated recombinant viruses. At 24 h postinfection the supernatants were gathered and the levels of IFN-β measured.

The most likely explanation for the observed discrepancies is as follows. In this experiment, IFN-β was produced almost exclusively during secondary infection, in cells infected by virions released from the initially infected cells. This makes any standardization of the experimental conditions very difficult, since the magnitude of final IFN-β response depends on the following:

- the MOI (a lower MOI results in a greater number of cells that can be infected by the new generation of virions);
- the ability of different stocks of virus to suppress the production of minor quantities of IFN-β (part of this IFN may, for example, be produced by cells that become infected by defective viruses that are always present in any virus preparation), which prime uninfected cells;
- the precise speed of viral replication/release: IFN-β production during secondary infection is not a synchronized event, it rather represents a self-amplifying cascade, in which most of the IFN-β is produced by efficiently primed (but still susceptible to infection) cells;
- the sensitivity of the virus to the antiviral effects of IFN: all (or almost all) IFN-β is produced by cells that have been previously exposed to IFN (albeit at very low concentrations); thus the spread of A774wt may be diminished due to type I IFN-dependent paracrine signaling (191).

Taken together, this assay is not suitable to analyze the differences in type I IFN production using this panel of constructed recombinant viruses.

Recently, our research team discovered a novel mechanism by which SFV induces type I IFN response (188). Briefly, type I IFN can be induced by SFV

replicase in the absence of replication-competent template RNA (Fig. 8A). This mechanism is highly relevant to natural SFV4 infection, and the large majority of type I IFN-inducing PAMP (pathogen-associated molecular pattern) RNAs that are produced in SFV4-infected cells originate via this pathway (188). Importantly, for the present study, this phenomenon allowed us to use an efficient and far more reliable test system in which type I IFN is produced in transfected cells (Fig. 8A). In addition, in this system the production of PAMP RNAs is not affected by the antiviral effects of type I IFN. Furthermore, the analysis allowed us to verify for the first time whether this novel pathway leading to type I IFN induction was relevant to the important biological properties of SFV.

All of the analyzed replicases were roughly equivalent in their abilities to trigger the replication of the truncated SFV mini-genome (data not shown), which indicated that none of these replicases were defective. This finding is in agreement with data demonstrating that the corresponding viruses grow to similar titers (**II**, Fig. 5A, 6A). Then the experiment was conducted in the absence of replication-competent template, and the amounts of IFN-β released by transfected cells were measured. As shown in figure 8B, a perfect correlation between the *in vivo* phenotypes of the recombinant viruses and the ability of the corresponding replicases to induce IFN-β production in a replication-independent manner was observed. The replicases of avirulent viruses (SFV6-74-RE and A774wt) were poor inducers of IFN-β, while those of virulent ones induced moderate to high levels of IFN-β. Furthermore, perfect gradients of IFN-β production were observed for both replicase backbones:

- SFV6 > SFV6-74 > SFV6-RE > SFV6-74-RE, and
- A774wt < A774wt-6 ≈ A774-HV < A774-6-HV.

This finding strongly suggests that the neurovirulence of SFV has a clear similarity to that of neuroadapted SINV and that the ability of SFV to spread in the CNS of adult mice correlates with excessive induction of type I IFN. The spread of infection may be facilitated by the greater capacity of virulent SFV6 to resist the antiviral effects of IFN. However, this supposition is doubtful because type I IFN tolerance-associated loci (most likely residing in nsP4) are distinct from the loci that are responsible for SFV neurovirulence (191). Thus, it is possible that virulent A774wt-6 and A774-HV still harbor an IFN-sensitive phenotype. Furthermore, as recombinant viruses grow to similar titers (**II** Fig. 5A, 6A) it could be concluded that SFV neurovirulence is not associated with enhanced production of standard PAMP RNAs (dsRNA replication intermediates and negative strand RNAs with 5' ppp groups). Instead, the enhanced spread of the virus and immunopathology appear to be triggered by the production of PAMP RNAs generated by SFV replicase through the use of cellular templates. Finally, this finding clearly demonstrates that two virulence determinants of SFV, which were identified in publication **II**, act via one and the same mechanism. Both slower processing of the 1/2 cleavage site and nsP3 of SFV6 clearly increase the ability of the SFV replicase to trigger IFN-β production;

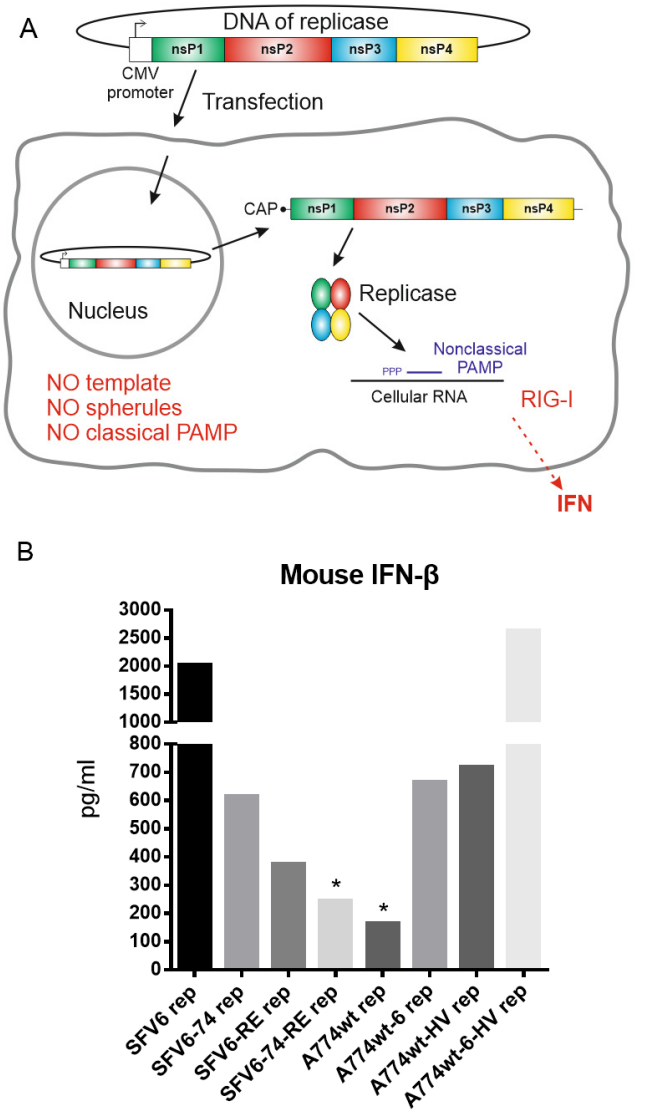

Figure 8. The ability of recombinant SFV replicases to induce IFN-β. A. Schematic representation of the experimental design. Cells are transfected with a plasmid encoding SFV replicase, which is transcribed by cellular machinery; resulting mRNA is translated to replicase proteins. The replicase is able to use cellular templates to produce short RNAs with 5' ppp, which form duplexes with their templates. These nonclassical PAMP RNAs are recognized by RIG-I leading to induction of IFN-β, which can then be detected. **B.** COP5 cells were transfected with plasmid encoding either wt or recombinant SFV replicase. At 24 h posttransfection, the supernatants were gathered and the levels of IFN-β measured. Replicases corresponding to the viruses with an avirulent phenotype in mice are indicated by asterisks.

furthermore, the effects of these determinants are cumulative (Fig. 8B). Validation and assessment of these findings, including the precise roles of the pathogenesis determinants in the induction of IFN-β production, represent topics of further studies.

It should also be noted that there are two main classes of cells in the CNS: nerve cells (neurons) and glial cells (e.g., oligodendrocytes, astrocytes). Avirulent strains such as A7(74) infect oligodendrocytes but display severely restricted replication in mature neurons. Virulent strains (e.g., SFV4, L10) infect both oligodendrocytes and neurons and spread rapidly throughout the brain. These two features of neurovirulence seem to depend on different determinants. In nonneuronal cells, increased virulence correlates with an elevated resistance to the antiviral effects if type I IFN. This is in agreement with previous findings demonstrating that the replication of A7(74) in type I IFN receptor-deficient mice increases dramatically in peripheral tissues and nonneuronal CNS cells but not in neurons (179). The restricted replication of avirulent strains such as A7(74) in neurons, however, appears to depend on the processing speed of the ns polyprotein and the sequence of nsP3, which together affect the ability to produce PAMP RNAs from cellular templates (Fig. 8B). Such PAMP RNAs are mostly recognized by RIG-I (188), and thus it is possible that not only type I IFN but also genes directly induced by activated RIG-I may contribute to the generation of conditions that permit the replication of SFV in mature neurons.

These studies concentrated on characterizing the determinants responsible for the different *in vivo* phenotypes of SFV strains. A large amount of new knowledge was obtained, but further analyses (especially *in vivo* analyses) are needed to elucidate the importance and role of host defense mechanisms in determining the differences in SFV neurovirulence. Similarly, it will be intriguing to determine whether these findings apply to other alphaviruses.

5.4. Magnetic fractionation and proteomic dissection provide a useful method to study the composition of SFV replication organelles (III)

To fully comprehend the complexity of viral infection, it is important not to only examine the virus and host as separate entities but rather see the virus-infected cell as a cooperative functional unit of infection. Alphaviruses encode only ten proteins, and therefore most steps in virus infection involve interactions between relatively few viral components and much more complex pools of host factors. The differences between various strains of SFV are, on the one hand, caused by differences in their genomes, but, on the other hand, determined by the host factors and complex processes that occur in the host. Therefore, it is important to also investigate these aspects of viral infection. Several studies investigating SINV have identified numerous factors that bind ns proteins, and nsP2 of CHIKV has been shown to interact with a number of host

proteins. However, methods such as Y2H screening and immunoprecipitation have limitations because they do not differentiate between host proteins associated with individual nsPs and those associated with RCs and replication organelles. Furthermore, cellular proteins do not necessarily need to interact with virus nsPs directly to affect replication. Hence, multiple alternative methods are needed to understand the complete picture of alphavirus-host interactions.

5.4.1. Vesicles carrying functionally active RCs can be purified via magnetic enrichment

In the present study, we introduced a new approach to identify host proteins that colocalize with mature RCs of SFV. The RCs of alphaviruses form on the plasma membrane and are then internalized via endocytosis (71, 72). For some alphaviruses, such as SFV, they finally localize to large vesicles of endolysosomal origin called CPV-Is (**III**, Fig. 1B). For other alphaviruses, for example CHIKV, most of the vesicles carrying RCs remain close to the plasma membrane and are smaller in size (75). This unique replication organelle biogenesis pathway, coupled with the specific features of SFV replication organelle formation, made it possible to apply a method that was originally used in studies of proteins involved in endocytosis (227) to isolate functional replication organelles of SFV. To the best of our knowledge, this is the first example of the use of such a method in studies of viral infection.

The method involved feeding mock- or SFV4-infected cells with dextran-covered magnetic nanoparticles, which were incorporated into lysosomes, permitting the collection of these specific vesicles via magnetic isolation (**III**, Fig. 2A). The obtained magnetic fractions were enriched for endolysosomes but largely devoid of plasma membrane and ER markers (**III**, Fig. 2C). The vesicles were shown to contain all of the ns proteins of SFV (**III**, Fig. 2C) in the form of functional RCs, evident from their maintained ability to synthesize RNA after isolation (**III**, Fig. 3A). This magnetic isolation possessed several advantages over ultracentrifugation-based separation: it was faster and did not damage the activity of RCs. The limitation of this method is that it can only be applied to study later stages of SFV infection, because CPV-Is are not formed immediately after virus entry. Mature alphavirus RCs only synthesize positive-strand RNAs, and therefore, purified replication organelles may not carry the specific host components that are important for early stages of infection, such as RNA template recruitment, cellular membrane remodeling, and viral negative-strand RNA synthesis.

5.4.2. The proteome of magnetically isolated fractions can be characterized by using quantitative proteomics approach

Prior to this study, the protein composition of SFV replication organelles (beyond the presence of viral proteins and lysosomal markers) was unknown. It is reasonable to assume that some host proteins are included in these structures

accidentally or because they are integral components of lysosomal membranes, while others are included (or excluded) due to specific virus-host interactions. To identify host proteins that are enriched or depleted in the magnetically purified fractions containing SFV replication organelles, a gel-free SILAC-based (stable isotope labeling with amino acids in the cell culture) quantitative proteomics approach (228) was employed. This approach allows the detection of relative changes in the amounts of protein between different biological samples.

In this study, heavy (labeled with heavy arginine and lysine) and light samples from SFV4-infected and uninfected cells (three biological replicates) were analyzed (**III**, Fig. 2B). Comparison of the protein compositions of the magnetically isolated fractions produced a list of proteins that were overrepresented in the lysosomal membranes of infected cells (**III**, Table 1). We did not identify any proteins that were downregulated by a factor of 2.5 or more in samples containing SFV CPV-Is, most likely due to the co-purification of large amounts of normal endosomes (by rough estimation, approximately 50% of the total number of isolated vesicles did not contain viral RCs), making this approach unsuitable for the detection of proteins that were excluded from the replication organelles of the virus. In contrast, a number of proteins enriched in the fraction containing RCs were identified. Importantly, several of these proteins have previously been shown to associate with the alphavirus replicase using alternative methods. These include the RNA-binding proteins G3BP1, G3BP2, hnRNP C, hnRNP M, and hnRNP A1, which are known to interact with ns proteins or genomic and sgRNAs of SINV (131, 161, 162, 165), and hnRNP K, which has been shown to bind CHIKV nsP2 (167). Thus, the SILAC-based proteomics approach is suitable for identifying cellular proteins that colocalize and/or interact with alphavirus RCs.

5.4.3. PCBP1, hnRNP M, hnRNP C, and hnRNP K affect alphavirus infection

The obtained list of host factors included nearly 50 proteins that were not previously known to colocalize with alphavirus RCs and that are likely to be important for SFV replication (**III**, Table 1). These proteins belonged to different functionally connected clusters (**III**, Fig. 4), among which the cluster of RNA-interacting proteins was the most prominent. The impact of four of these RNA-binding proteins, PCBP1, hnRNP M, hnRNP C, and hnRNP K, on alphavirus infection was analyzed.

In uninfected cells, PCBP1, hnRNP M, hnRNP C, and hnRNP K localize predominantly to the nucleus. Confocal microscopy analysis of infected cells showed that all four of these proteins colocalized with dsRNA, nsP3, and therefore, with the SFV RCs (**III**, Fig. 5 and Fig. 6). This confirmed that their relocalization during the late phase of SFV infection was not a random event. To analyze the functional significance of these proteins in alphavirus infection,

siRNA-mediated silencing experiments were conducted. Silencing of hnRNP C and hnRNP M increased SFV gene expression and replication (**III**, Fig. 7B, 8A, 8B). Conversely, PCBP1 silencing reduced the synthesis of reporter proteins expressed both from genomic and sgRNAs of the SFV replicon but failed to affect or only minimally affected viral RNA synthesis and SFV growth in cell culture. Silencing of hnRNP K showed that this protein had little impact on the RNA synthesis and multiplication of SFV (**III**, Fig. 7B, 8A, 8B).

Cellular factors may not affect the infection cycle of different alphaviruses in the same manner. Therefore, the importance of PCBP1, hnRNP M, hnRNP C, and hnRNP K in SINV and CHIKV infection was analyzed. Silencing of PCBP1, hnRNP C and hnRNP M had the same outcome as that observed in SFV-infected cells (**III**, Fig. 9). In contrast to SFV infection, during which it functioned as a repressor, hnRNP K was shown to function as an activator in CHIKV and SINV infection (**III**, Fig. 9), which is consistent with previous publications (165, 167). Thus, not all of the host proteins analyzed in the present study affected different alphaviruses similarly. The observation that the same host factor had contrasting effects on different viruses may reflect differences in the infection cycles of these viruses.

As a next step, the magnetic fractionation and proteomic dissection method introduced in this study could be used to extract and analyze the composition of the replication organelles of cells infected with the previously constructed recombinant SFVs carrying duplicated nsP3 regions. This approach could be used to elucidate the role of nsP3 in the stabilization of nsP4. A comparison of RC-containing fractions from cells infected with wt SFV6 and, for example, cells infected with SFV6-6/6 could shed light on the localization of excessive amounts of nsP4 and also reveal what other changes (if any) are caused by the duplication of the nsP3-encoding regions. It would also be interesting to determine whether different strains of SFV (avirulent and virulent) interact with different sets of host factors or whether the same host factors have different effects on these viruses, in the way hnRNP K has different effects in SINV, CHIKV and SFV infection. Finally, the applicability of the method is not restricted to the viruses used or developed during these studies. For example, SFV mutants that are unable to bind amphiphysins (129) or G3BP proteins (134) have recently been constructed. It would be interesting to determine if and how the loss of one interaction partner will affect the proteome of the alphavirus replication organelle.

6. CONCLUSIONS

A7(74), L10, and SFV4 are the most thoroughly studied strains of SFV. A7(74) is considered to be avirulent because the infection in adult mice is asymptomatic. In contrast, after i.c. or high-dose i.p. inoculation, L10 and SFV4 are both virulent and cause lethal encephalitis. However, following low-dose i.p. inoculation, SFV4 is incapable of reaching the brain, and the infected animals survive. In the present study, biological stocks of A7(74) and L10 strains of known provenance were sequenced, and two new consensus clones, A774wt and SFV6, were constructed. These and the molecular clone of SFV4 were used to study the determinants and mechanism(s) of SFV neurovirulence. The following conclusions were made based on the data, obtained in these studies:

First, a comparison of the consensus sequence of L10 with the published sequence of SFV4 revealed that the latter contains six nonsynonymous changes, which may result from different passage histories and represent adaptations to the cell culture and/or i.c. passaging of the virus stock used to build the SFV4 molecular clone. The different phenotypes of SFV4 and SFV6 are caused by differences in a single amino acid residue located at position 162 of the E2 protein. A negatively charged glutamic acid residue (as in SFV6) increased the viral load in mouse blood and enhanced neuroinvasion and virulence. A positively charged lysine at the same position (as in SFV4) was responsible for the low viral load in the blood. However, this lysine residue facilitated the ability of the virus to cross an *in vitro* BBB model and to replicate in the mouse brain following i.c. infection. These properties correlated with the ability of the corresponding virions to bind GAGs. These findings led to the conclusion that a positive charge at position 162 (or 247) of E2 facilitates the binding of SFV4 virions to HS. This results in rapid clearance of virus from the blood and lower viremia, which in turn prevents the entry of the virus into the CNS.

Second, the determinants responsible for the different neurovirulence of A7(74) and L10 were investigated. A previously available molecular clone of A7(74) was found to have a defect resulting from a change in a single amino acid residue in nsP3; by repairing this defect, a highly infectious consensus clone, A774wt, was obtained. Using viruses with swapped or duplicated nsP3 regions revealed that neurovirulence is the dominant function of nsP3: all of the viruses expressing SFV6 nsP3 were virulent. The *in vivo* phenotype was also dependent on the fine-tuned functions of ns proteins, such as the basic protease activity of nsP2 and the efficiency of P123 processing. The presence of A774wt nsP3 (or rather the absence of SFV6 nsP3) and rapid cleavage of ns polyprotein at the 1/2 site were both required for the avirulent phenotype. Both slower processing of the 1/2 cleavage site or the presence of SFV6 nsP3 enhanced the ability of the virus replicases to produce PAMP RNAs from cellular templates and thereby induce a strong type I IFN response. Thus, the ability of SFV to spread in the CNS of adult mice correlates with the capacity of the virus to cause excessive induction of type I IFN, suggesting a link between SFV

neurovirulence and immunopathology. In addition, a novel function of nsP3, involvement in the stabilization of nsP4, was identified.

The finding that biological differences between SFV strains are determined, at least in part, by virus-host interactions and complex processes that occur in the host (e.g., the IFN response) emphasizes the need to identify and analyze host factors that are involved in alphavirus infection. Thus, a novel approach to determine the host proteins that colocalize with mature RCs of SFV was developed. This approach is based on feeding cells dextran-covered magnetic nanoparticles and subsequently collecting lysosomal vesicles harboring RCs via magnetic isolation. We paired this technique with SILAC-based quantitative proteomics and generated an extensive list of host factors associated with SFV RCs. Further characterization of four of these factors confirmed their significance in SFV infection. Interestingly, three of these factors were found to be antiviral, while only one was confirmed to be proviral.

REFERENCES

1. **Baltimore D.** 1971. Expression of animal virus genomes. Bacteriol Rev **35:**235–241.
2. **Strauss JH, Strauss EG.** 1994. The alphaviruses: gene expression, replication, and evolution. Microbiol Rev **58:**491–562.
3. **King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ.** 2012. Virus Taxonomy: Classification and Nomeclature of Viruses. Ninth Report of the International Committee of Taxonomy of Viruses. Elsevier.
4. **Strauss JH, Strauss EG.** 2008. Plus-strand RNA viruses, p 63–134, Viruses and Human Disease, 2nd ed.
5. **Weaver SC, Lecuit M.** 2015. Chikungunya virus and the global spread of a mosquito-borne disease. N Engl J Med **372:**1231–1239.
6. **Weaver SC.** 2014. Arrival of chikungunya virus in the new world: prospects for spread and impact on public health. PLoS Negl Trop Dis **8:**e2921.
7. **Waggoner JJ, Pinsky BA.** 2015. How great is the threat of chikungunya virus? Expert Rev Anti Infect Ther **13:**291–293.
8. **Ryman KD, Klimstra WB.** 2008. Host responses to alphavirus infection. Immunol Rev **225:**27–45.
9. **Lani R, Hassandarvish P, Chiam CW, Moghaddam E, Chu JJ, Rausalu K, Merits A, Higgs S, Vanlandingham D, Abu Bakar S, Zandi K.** 2015. Antiviral activity of silymarin against chikungunya virus. Sci Rep **5:**11421.
10. **Ahola T, Courderc T, Ng LF, Hallengard D, Powers A, Lecuit M, Esteban M, Merits A, Roques P, Liljestrom P.** 2015. Therapeutics and vaccines against chikungunya virus. Vector Borne Zoonotic Dis **15:**250–257.
11. **Rice CM, Levis R, Strauss JH, Huang HV.** 1987. Production of infectious RNA transcripts from 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.
12. **Liljestrom P, Lusa S, Huylebroeck D, Garoff H.** 1991. In vitro mutagenesis of a full-length cDNA clone of Semliki Forest virus: the small 6,000-molecular-weight membrane protein modulates virus release. J Virol **65:**4107–4113.
13. **Helenius A, Kartenbeck J, Simons K, Fries E.** 1980. On the entry of Semliki forest virus into BHK-21 cells. J Cell Biol **84:**404–420.
14. **Cheng RH, Kuhn RJ, Olson NH, Rossmann MG, Choi HK, Smith TJ, Baker TS.** 1995. Nucleocapsid and glycoprotein organization in an enveloped virus. Cell **80:**621–630.
15. **Voss JE, Vaney MC, Duquerroy S, Vonrhein C, Girard-Blanc C, Crublet E, Thompson A, Bricogne G, Rey FA.** 2010. Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. Nature **468:**709–712.
16. **Paredes AM, Brown DT, Rothnagel R, Chiu W, Schoepp RJ, Johnston RE, Prasad BV.** 1993. Three-dimensional structure of a membrane-containing virus. Proc Natl Acad Sci U S A **90:**9095–9099.
17. **Jose J, Snyder JE, Kuhn RJ.** 2009. A structural and functional perspective of alphavirus replication and assembly. Future Microbiol **4:**837–856.
18. **Omar A, Koblet H.** 1988. Semliki Forest virus particles containing only the E1 envelope glycoprotein are infectious and can induce cell-cell fusion. Virology **166:**17–23.

19. **McInerney GM, Smit JM, Liljestrom P, Wilschut J.** 2004. Semliki Forest virus produced in the absence of the 6K protein has an altered spike structure as revealed by decreased membrane fusion capacity. Virology **325:**200–206.
20. **Snyder JE, Kulcsar KA, Schultz KL, Riley CP, Neary JT, Marr S, Jose J, Griffin DE, Kuhn RJ.** 2013. Functional characterization of the alphavirus TF protein. J Virol **87:**8511–8523.
21. **Jose J, Przybyla L, Edwards TJ, Perera R, Burgner JW, 2nd, Kuhn RJ.** 2012. Interactions of the cytoplasmic domain of Sindbis virus E2 with nucleocapsid cores promote alphavirus budding. J Virol **86:**2585–2599.
22. **Wahlberg JM, Bron R, Wilschut J, Garoff H.** 1992. Membrane fusion of Semliki Forest virus involves homotrimers of the fusion protein. J Virol **66:**7309–7318.
23. **Li L, Jose J, Xiang Y, Kuhn RJ, Rossmann MG.** 2010. Structural changes of envelope proteins during alphavirus fusion. Nature **468:**705–708.
24. **Strauss EG, Rice CM, Strauss JH.** 1984. Complete nucleotide sequence of the genomic RNA of Sindbis virus. Virology **133:**92–110.
25. **Schlesinger S, Schlesinger MJ.** 2001. Togaviridae: The Viruses and their Replication, Fields' Virology, 4th edition. Lippincott Williams & Wilkins.
26. **Frolov I, Hardy R, Rice CM.** 2001. Cis-acting RNA elements at the 5' end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. RNA **7:**1638–1651.
27. **Hyde JL, Gardner CL, Kimura T, White JP, Liu G, Trobaugh DW, Huang C, Tonelli M, Paessler S, Takeda K, Klimstra WB, Amarasinghe GK, Diamond MS.** 2014. A viral RNA structural element alters host recognition of nonself RNA. Science **343:**783–787.
28. **Niesters HG, Strauss JH.** 1990. Mutagenesis of the conserved 51-nucleotide region of Sindbis virus. J Virol **64:**1639–1647.
29. **Fayzulin R, Frolov I.** 2004. Changes of the secondary structure of the 5' end of the Sindbis virus genome inhibit virus growth in mosquito cells and lead to accumulation of adaptive mutations. J Virol **78:**4953–4964.
30. **Ou JH, Rice CM, Dalgarno L, Strauss EG, Strauss JH.** 1982. Sequence studies of several alphavirus genomic RNAs in the region containing the start of the subgenomic RNA. Proc Natl Acad Sci U S A **79:**5235–5239.
31. **Levis R, Schlesinger S, Huang HV.** 1990. Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription. J Virol **64:**1726–1733.
32. **Rausalu K, Iofik A, Ulper L, Karo-Astover L, Lulla V, Merits A.** 2009. Properties and use of novel replication-competent vectors based on Semliki Forest virus. Virol J **6:**33.
33. **Hardy RW.** 2006. The role of the 3' terminus of the Sindbis virus genome in minus-strand initiation site selection. Virology **345:**520–531.
34. **Weiss B, Geigenmuller-Gnirke U, Schlesinger S.** 1994. Interactions between Sindbis virus RNAs and a 68 amino acid derivative of the viral capsid protein further defines the capsid binding site. Nucleic Acids Res **22:**780–786.
35. **Kim DY, Firth AE, Atasheva S, Frolova EI, Frolov I.** 2011. Conservation of a packaging signal and the viral genome RNA packaging mechanism in alphavirus evolution. J Virol **85:**8022–8036.
36. **White CL, Thomson M, Dimmock NJ.** 1998. Deletion analysis of a defective interfering Semliki Forest virus RNA genome defines a region in the nsP2

sequence that is required for efficient packaging of the genome into virus particles. J Virol **72:**4320–4326.
37. **Firth AE, Wills NM, Gesteland RF, Atkins JF.** 2011. Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element. Nucleic Acids Res **39:**6679–6691.
38. **Frolov I, Schlesinger S.** 1996. Translation of Sindbis virus mRNA: analysis of sequences downstream of the initiating AUG codon that enhance translation. J Virol **70:**1182–1190.
39. **Firth AE, Chung BY, Fleeton MN, Atkins JF.** 2008. Discovery of frameshifting in Alphavirus 6K resolves a 20-year enigma. Virol J **5:**108.
40. **Garneau NL, Sokoloski KJ, Opyrchal M, Neff CP, Wilusz CJ, Wilusz J.** 2008. The 3' untranslated region of sindbis virus represses deadenylation of viral transcripts in mosquito and Mammalian cells. J Virol **82:**880–892.
41. **Sokoloski KJ, Dickson AM, Chaskey EL, Garneau NL, Wilusz CJ, Wilusz J.** 2010. Sindbis virus usurps the cellular HuR protein to stabilize its transcripts and promote productive infections in mammalian and mosquito cells. Cell Host Microbe **8:**196–207.
42. **Hyde JL, Chen R, Trobaugh DW, Diamond MS, Weaver SC, Klimstra WB, Wilusz J.** 2015. The 5' and 3' ends of alphavirus RNAs – Non-coding is not non-functional. Virus Res **206:**99–107.
43. **Tucker PC, Griffin DE.** 1991. Mechanism of altered Sindbis virus neurovirulence associated with a single-amino-acid change in the E2 Glycoprotein. J Virol **65:**1551–1557.
44. **Wang KS, Kuhn RJ, Strauss EG, Ou S, Strauss JH.** 1992. High-affinity laminin receptor is a receptor for Sindbis virus in mammalian cells. J Virol **66:**4992–5001.
45. **Ludwig GV, Kondig JP, Smith JF.** 1996. A putative receptor for Venezuelan equine encephalitis virus from mosquito cells. J Virol **70:**5592–5599.
46. **Rose PP, Hanna SL, Spiridigliozzi A, Wannissorn N, Beiting DP, Ross SR, Hardy RW, Bambina SA, Heise MT, Cherry S.** 2011. Natural resistance-associated macrophage protein is a cellular receptor for sindbis virus in both insect and mammalian hosts. Cell Host Microbe **10:**97–104.
47. **DeTulleo L, Kirchhausen T.** 1998. The clathrin endocytic pathway in viral infection. EMBO J **17:**4585–4593.
48. **Bron R, Wahlberg JM, Garoff H, Wilschut J.** 1993. Membrane fusion of Semliki Forest virus in a model system: correlation between fusion kinetics and structural changes in the envelope glycoprotein. EMBO J **12:**693–701.
49. **Justman J, Klimjack MR, Kielian M.** 1993. Role of spike protein conformational changes in fusion of Semliki Forest virus. J Virol **67:**7597–7607.
50. **Singh I, Helenius A.** 1992. Role of ribosomes in Semliki Forest virus nucleocapsid uncoating. J Virol **66:**7049–7058.
51. **Li G, Rice CM.** 1993. The signal for translational readthrough of a UGA codon in Sindbis virus RNA involves a single cytidine residue immediately downstream of the termination codon. J Virol **67:**5062–5067.
52. **Strauss EG, Rice CM, Strauss JH.** 1983. Sequence coding for the alphavirus nonstructural proteins is interrupted by an opal termination codon. Proc Natl Acad Sci U S A **80:**5271–5275.
53. **Takkinen K.** 1986. Complete nucleotide sequence of the nonstructural protein genes of Semliki Forest virus. Nucleic Acids Res **14:**5667–5682.

54. **Vasiljeva L, Merits A, Golubtsov A, Sizemskaja V, Kaariainen L, Ahola T.** 2003. Regulation of the sequential processing of Semliki Forest virus replicase polyprotein. J Biol Chem **278:**41636–41645.
55. **Merits A, Vasiljeva L, Ahola T, Kaariainen L, Auvinen P.** 2001. Proteolytic processing of Semliki Forest virus-specific non-structural polyprotein by nsP2 protease. J Gen Virol **82:**765–773.
56. **Shirako Y, Strauss JH.** 1994. Regulation of Sindbis virus RNA replication: uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. J Virol **68:**1874–1885.
57. **Lemm JA, Bergqvist A, Read CM, Rice CM.** 1998. Template-dependent initiation of Sindbis virus RNA replication in vitro. J Virol **72:**6546–6553.
58. **Lemm JA, Rice CM.** 1993. Assembly of functional Sindbis virus RNA replication complexes: requirement for coexpression of P123 and P34. J Virol **67:**1905–1915.
59. **Lemm JA, Rumenapf T, Strauss EG, Strauss JH, Rice CM.** 1994. Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus- and plus-strand RNA synthesis. Embo j **13:**2925–2934.
60. **Hardy WR, Hahn YS, de Groot RJ, Strauss EG, Strauss JH.** 1990. Synthesis and processing of the nonstructural polyproteins of several temperature-sensitive mutants of Sindbis virus. Virology **177:**199–208.
61. **Li ML, Stollar V.** 2004. Identification of the amino acid sequence in Sindbis virus nsP4 that binds to the promoter for the synthesis of the subgenomic RNA. Proc Natl Acad Sci U S A **101:**9429–9434.
62. **Li ML, Stollar V.** 2007. Distinct sites on the Sindbis virus RNA-dependent RNA polymerase for binding to the promoters for the synthesis of genomic and subgenomic RNA. J Virol **81:**4371–4373.
63. **Li ML, Wang H, Stollar V.** 2010. In vitro synthesis of Sindbis virus genomic and subgenomic RNAs: influence of nsP4 mutations and nucleoside triphosphate concentrations. J Virol **84:**2732–2739.
64. **Sawicki DL, Kaariainen L, Lambek C, Gomatos PJ.** 1978. Mechanism for control of synthesis of Semliki Forest virus 26S and 42s RNA. J Virol **25:**19–27.
65. **Suopanki J, Sawicki DL, Sawicki SG, Kaariainen L.** 1998. Regulation of alphavirus 26S mRNA transcription by replicase component nsP2. J Gen Virol **79 (Pt 2):**309–319.
66. **Salonen A, Vasiljeva L, Merits A, Magden J, Jokitalo E, Kaariainen L.** 2003. Properly folded nonstructural polyprotein directs the semliki forest virus replication complex to the endosomal compartment. J Virol **77:**1691–1702.
67. **Grimley PM, Berezesky IK, Friedman RM.** 1968. Cytoplasmic structures associated with an arbovirus infection: loci of viral ribonucleic acid synthesis. J Virol **2:**1326–1338.
68. **Salonen A, Ahola T, Kaariainen L.** 2005. Viral RNA replication in association with cellular membranes. Curr Top Microbiol Immunol **285:**139–173.
69. **Froshauer S, Kartenbeck J, Helenius A.** 1988. Alphavirus RNA replicase is located on the cytoplasmic surface of endosomes and lysosomes. J Cell Biol **107:**2075–2086.

70. **Kujala P, Ikaheimonen A, Ehsani N, Vihinen H, Auvinen P, Kaariainen L.** 2001. Biogenesis of the Semliki Forest virus RNA replication complex. J Virol **75:**3873–3884.
71. **Frolova EI, Gorchakov R, Pereboeva L, Atasheva S, Frolov I.** 2010. Functional Sindbis virus replicative complexes are formed at the plasma membrane. J Virol **84:**11679–11695.
72. **Spuul P, Balistreri G, Kaariainen L, Ahola T.** 2010. Phosphatidylinositol 3-kinase-, actin-, and microtubule-dependent transport of Semliki Forest Virus replication complexes from the plasma membrane to modified lysosomes. J Virol **84:**7543–7557.
73. **Spuul P, Balistreri G, Hellstrom K, Golubtsov AV, Jokitalo E, Ahola T.** 2011. Assembly of alphavirus replication complexes from RNA and protein components in a novel trans-replication system in mammalian cells. J Virol **85:**4739–4751.
74. **Kaariainen L, Ahola T.** 2002. Functions of alphavirus nonstructural proteins in RNA replication. Prog Nucleic Acid Res Mol Biol **71:**187–222.
75. **Thaa B, Biasiotto R, Eng K, Neuvonen M, Gotte B, Rheinemann L, Mutso M, Utt A, Varghese F, Balistreri G, Merits A, Ahola T, McInerney GM.** 2015. Differential Phosphatidylinositol-3-Kinase-Akt-mTOR Activation by Semliki Forest and Chikungunya Viruses Is Dependent on nsP3 and Connected to Replication Complex Internalization. J Virol **89:**11420–11437.
76. **Choi HK, Tong L, Minor W, Dumas P, Boege U, Rossmann MG, Wengler G.** 1991. Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion. Nature **354:**37–43.
77. **Zhang X, Fugere M, Day R, Kielian M.** 2003. Furin processing and proteolytic activation of Semliki Forest virus. J Virol **77:**2981–2989.
78. **Wu SR, Haag L, Sjoberg M, Garoff H, Hammar L.** 2008. The dynamic envelope of a fusion class II virus. E3 domain of glycoprotein E2 precursor in Semliki Forest virus provides a unique contact with the fusion protein E1. J Biol Chem **283:**26452–26460.
79. **Zhang R, Hryc CF, Cong Y, Liu X, Jakana J, Gorchakov R, Baker ML, Weaver SC, Chiu W.** 2011. 4.4 A cryo-EM structure of an enveloped alphavirus Venezuelan equine encephalitis virus. EMBO J **30:**3854–3863.
80. **Uchime O, Fields W, Kielian M.** 2013. The role of E3 in pH protection during alphavirus assembly and exit. J Virol **87:**10255–10262.
81. **Martinez MG, Snapp EL, Perumal GS, Macaluso FP, Kielian M.** 2014. Imaging the alphavirus exit pathway. J Virol **88:**6922–6933.
82. **Jones PH, Maric M, Madison MN, Maury W, Roller RJ, Okeoma CM.** 2013. BST-2/tetherin-mediated restriction of chikungunya (CHIKV) VLP budding is counteracted by CHIKV non-structural protein 1 (nsP1). Virology **438:**37–49.
83. **Vasiljeva L, Merits A, Auvinen P, Kaariainen L.** 2000. Identification of a novel function of the alphavirus capping apparatus. RNA 5'-triphosphatase activity of Nsp2. J Biol Chem **275:**17281–17287.
84. **Ahola T, Kaariainen L.** 1995. Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. Proc Natl Acad Sci U S A **92:**507–511.
85. **Li C, Guillen J, Rabah N, Blanjoie A, Debart F, Vasseur JJ, Canard B, Decroly E, Coutard B.** 2015. mRNA Capping by Venezuelan Equine Encephalitis Virus nsP1: Functional Characterization and Implications for Antiviral Research. J Virol **89:**8292–8303.

86. **Ahola T, Kujala P, Tuittila M, Blom T, Laakkonen P, Hinkkanen A, Auvinen P.** 2000. Effects of palmitoylation of replicase protein nsP1 on alphavirus infection. J Virol **74:**6725–6733.
87. **Ahola T, Lampio A, Auvinen P, Kaariainen L.** 1999. Semliki Forest virus mRNA capping enzyme requires association with anionic membrane phospholipids for activity. Embo j **18:**3164–3172.
88. **Peranen J, Laakkonen P, Hyvonen M, Kaariainen L.** 1995. The alphavirus replicase protein nsP1 is membrane-associated and has affinity to endocytic organelles. Virology **208:**610–620.
89. **Spuul P, Salonen A, Merits A, Jokitalo E, Kaariainen L, Ahola T.** 2007. Role of the amphipathic peptide of Semliki forest virus replicase protein nsP1 in membrane association and virus replication. J Virol **81:**872–883.
90. **Tomar S, Narwal M, Harms E, Smith JL, Kuhn RJ.** 2011. Heterologous production, purification and characterization of enzymatically active Sindbis virus nonstructural protein nsP1. Protein Expr Purif **79:**277–284.
91. **Laakkonen P, Hyvonen M, Peranen J, Kaariainen L.** 1994. Expression of Semliki Forest virus nsP1-specific methyltransferase in insect cells and in Escherichia coli. J Virol **68:**7418–7425.
92. **Mi S, Stollar V.** 1991. Expression of Sindbis virus nsP1 and methyltransferase activity in Escherichia coli. Virology **184:**423–427.
93. **Heise MT, Simpson DA, Johnston RE.** 2000. A single amino acid change in nsP1 attenuates neurovirulence of the Sindbis-group alphavirus S.A.AR86. J Virol **74:**4207–4213.
94. **Cruz CC, Suthar MS, Montgomery SA, Shabman R, Simmons J, Johnston RE, Morrison TE, Heise MT.** 2010. Modulation of type I IFN induction by a virulence determinant within the alphavirus nsP1 protein. Virology **399:**1–10.
95. **Shirako Y, Strauss EG, Strauss JH.** 2000. Suppressor mutations that allow sindbis virus RNA polymerase to function with nonaromatic amino acids at the N-terminus: evidence for interaction between nsP1 and nsP4 in minus-strand RNA synthesis. Virology **276:**148–160.
96. **Lulla V, Sawicki DL, Sawicki SG, Lulla A, Merits A, Ahola T.** 2008. Molecular defects caused by temperature-sensitive mutations in Semliki Forest virus nsP1. J Virol **82:**9236–9244.
97. **Laakkonen P, Auvinen P, Kujala P, Kaariainen L.** 1998. Alphavirus replicase protein NSP1 induces filopodia and rearrangement of actin filaments. J Virol **72:**10265–10269.
98. **Zusinaite E, Tints K, Kiiver K, Spuul P, Karo-Astover L, Merits A, Sarand I.** 2007. Mutations at the palmitoylation site of non-structural protein nsP1 of Semliki Forest virus attenuate virus replication and cause accumulation of compensatory mutations. J Gen Virol **88:**1977–1985.
99. **Rikkonen M, Peranen J, Kaariainen L.** 1994. ATPase and GTPase activities associated with Semliki Forest virus nonstructural protein nsP2. J Virol **68:**5804–5810.
100. **Gomez de Cedron M, Ehsani N, Mikkola ML, Garcia JA, Kaariainen L.** 1999. RNA helicase activity of Semliki Forest virus replicase protein NSP2. FEBS Lett **448:**19–22.
101. **Das PK, Merits A, Lulla A.** 2014. Functional cross-talk between distant domains of chikungunya virus non-structural protein 2 is decisive for its RNA-modulating activity. J Biol Chem **289:**5635–5653.

102. **Stapleford KA, Rozen-Gagnon K, Das PK, Saul S, Poirier EZ, Blanc H, Vidalain PO, Merits A, Vignuzzi M.** 2015. Viral Polymerase-Helicase Complexes Regulate Replication Fidelity To Overcome Intracellular Nucleotide Depletion. J Virol **89:**11233–11244.
103. **Hardy WR, Strauss JH.** 1989. Processing the nonstructural polyproteins of sindbis virus: nonstructural proteinase is in the C-terminal half of nsP2 and functions both in cis and in trans. J Virol **63:**4653–4664.
104. **Lulla A, Lulla V, Tints K, Ahola T, Merits A.** 2006. Molecular determinants of substrate specificity for Semliki Forest virus nonstructural protease. J Virol **80:**5413–5422.
105. **Lulla A, Lulla V, Merits A.** 2012. Macromolecular assembly-driven processing of the 2/3 cleavage site in the alphavirus replicase polyprotein. J Virol **86:**553–565.
106. **Lulla V, Karo-Astover L, Rausalu K, Merits A, Lulla A.** 2013. Presentation overrides specificity: probing the plasticity of alphaviral proteolytic activity through mutational analysis. J Virol **87:**10207–10220.
107. **Russo AT, White MA, Watowich SJ.** 2006. The crystal structure of the Venezuelan equine encephalitis alphavirus nsP2 protease. Structure **14:**1449–1458.
108. **Shin G, Yost SA, Miller MT, Elrod EJ, Grakoui A, Marcotrigiano J.** 2012. Structural and functional insights into alphavirus polyprotein processing and pathogenesis. Proc Natl Acad Sci U S A **109:**16534–16539.
109. **Cheung J, Franklin M, Mancia F, Rudolph M, Cassidy M, Gary E, Burshteyn F, Love J.** 2011. Structure of Chikungunya virus nsP2 protease.
110. **Mayuri, Geders TW, Smith JL, Kuhn RJ.** 2008. Role for conserved residues of sindbis virus nonstructural protein 2 methyltransferase-like domain in regulation of minus-strand synthesis and development of cytopathic infection. J Virol **82:** 7284–7297.
111. **Hahn YS, Grakoui A, Rice CM, Strauss EG, Strauss JH.** 1989. Mapping of RNA- temperature-sensitive mutants of Sindbis virus: complementation group F mutants have lesions in nsP4. J Virol **63:**1194–1202.
112. **Frolov I, Garmashova N, Atasheva S, Frolova EI.** 2009. Random insertion mutagenesis of sindbis virus nonstructural protein 2 and selection of variants incapable of downregulating cellular transcription. J Virol **83:**9031–9044.
113. **Rikkonen M, Peranen J, Kaariainen L.** 1992. Nuclear and nucleolar targeting signals of Semliki Forest virus nonstructural protein nsP2. Virology **189:**462–473.
114. **Rikkonen M.** 1996. Functional significance of the nuclear-targeting and NTP-binding motifs of Semliki Forest virus nonstructural protein nsP2. Virology **218:**352–361.
115. **Tamm K, Merits A, Sarand I.** 2008. Mutations in the nuclear localization signal of nsP2 influencing RNA synthesis, protein expression and cytotoxicity of Semliki Forest virus. J Gen Virol **89:**676–686.
116. **Akhrymuk I, Kulemzin SV, Frolova EI.** 2012. Evasion of the innate immune response: the Old World alphavirus nsP2 protein induces rapid degradation of Rpb1, a catalytic subunit of RNA polymerase II. J Virol **86:**7180–7191.
117. **Fazakerley JK.** 2002. Pathogenesis of Semliki Forest virus encephalitis. J Neurovirol **8 Suppl 2:**66–74.
118. **Gorchakov R, Frolova E, Frolov I.** 2005. Inhibition of transcription and translation in Sindbis virus-infected cells. J Virol **79:**9397–9409.

119. **Utt A, Das PK, Varjak M, Lulla V, Lulla A, Merits A.** 2015. Mutations conferring a noncytotoxic phenotype on chikungunya virus replicons compromise enzymatic properties of nonstructural protein 2. J Virol **89:**3145–3162.
120. **Koonin EV, Dolja VV.** 1993. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. Crit Rev Biochem Mol Biol **28:**375–430.
121. **Pehrson JR, Fuji RN.** 1998. Evolutionary conservation of histone macroH2A subtypes and domains. Nucleic Acids Res **26:**2837–2842.
122. **Malet H, Coutard B, Jamal S, Dutartre H, Papageorgiou N, Neuvonen M, Ahola T, Forrester N, Gould EA, Lafitte D, Ferron F, Lescar J, Gorbalenya AE, de Lamballerie X, Canard B.** 2009. The crystal structures of Chikungunya and Venezuelan equine encephalitis virus nsP3 macro domains define a conserved adenosine binding pocket. J Virol **83:**6534–6545.
123. **Egloff MP, Malet H, Putics A, Heinonen M, Dutartre H, Frangeul A, Gruez A, Campanacci V, Cambillau C, Ziebuhr J, Ahola T, Canard B.** 2006. Structural and functional basis for ADP-ribose and poly(ADP-ribose) binding by viral macro domains. J Virol **80:**8493–8502.
124. **Neuvonen M, Ahola T.** 2009. Differential activities of cellular and viral macro domain proteins in binding of ADP-ribose metabolites. J Mol Biol **385:**212–225.
125. **Park E, Griffin DE.** 2009. Interaction of Sindbis virus non-structural protein 3 with poly(ADP-ribose) polymerase 1 in neuronal cells. J Gen Virol **90:**2073–2080.
126. **De I, Fata-Hartley C, Sawicki SG, Sawicki DL.** 2003. Functional analysis of nsP3 phosphoprotein mutants of Sindbis virus. J Virol **77:**13106–13116.
127. **LaStarza MW, Lemm JA, Rice CM.** 1994. Genetic analysis of the nsP3 region of Sindbis virus: evidence for roles in minus-strand and subgenomic RNA synthesis. J Virol **68:**5781–5791.
128. **Tuittila M, Hinkkanen AE.** 2003. Amino acid mutations in the replicase protein nsP3 of Semliki Forest virus cumulatively affect neurovirulence. J Gen Virol **84:**1525–1533.
129. **Neuvonen M, Kazlauskas A, Martikainen M, Hinkkanen A, Ahola T, Saksela K.** 2011. SH3 domain-mediated recruitment of host cell amphiphysins by alphavirus nsP3 promotes viral RNA replication. PLoS Pathog **7:**e1002383.
130. **Foy NJ, Akhrymuk M, Akhrymuk I, Atasheva S, Bopda-Waffo A, Frolov I, Frolova EI.** 2013. Hypervariable domains of nsP3 proteins of New World and Old World alphaviruses mediate formation of distinct, virus-specific protein complexes. J Virol **87:**1997–2010.
131. **Cristea IM, Carroll JW, Rout MP, Rice CM, Chait BT, MacDonald MR.** 2006. Tracking and elucidating alphavirus-host protein interactions. J Biol Chem **281:**30269–30278.
132. **Panas MD, Varjak M, Lulla A, Eng KE, Merits A, Karlsson Hedestam GB, McInerney GM.** 2012. Sequestration of G3BP coupled with efficient translation inhibits stress granules in Semliki Forest virus infection. Mol Biol Cell **23:**4701–4712.
133. **Fros JJ, Domeradzka NE, Baggen J, Geertsema C, Flipse J, Vlak JM, Pijlman GP.** 2012. Chikungunya virus nsP3 blocks stress granule assembly by recruitment of G3BP into cytoplasmic foci. J Virol **86:**10873–10879.
134. **Panas MD, Schulte T, Thaa B, Sandalova T, Kedersha N, Achour A, McInerney GM.** 2015. Viral and cellular proteins containing FGDF motifs bind G3BP to block stress granule formation. PLoS Pathog **11:**e1004659.

135. **Gorchakov R, Garmashova N, Frolova E, Frolov I.** 2008. Different types of nsP3-containing protein complexes in Sindbis virus-infected cells. J Virol **82:**10088–10101.
136. **Fros JJ, Geertsema C, Zouache K, Baggen J, Domeradzka N, van Leeuwen DM, Flipse J, Vlak JM, Failloux AB, Pijlman GP.** 2015. Mosquito Rasputin interacts with chikungunya virus nsP3 and determines the infection rate in Aedes albopictus. Parasit Vectors **8:**464.
137. **Foy NJ, Akhrymuk M, Shustov AV, Frolova EI, Frolov I.** 2013. Hypervariable domain of nonstructural protein nsP3 of Venezuelan equine encephalitis virus determines cell-specific mode of virus replication. J Virol **87:**7569–7584.
138. **Scholte FE, Tas A, Albulescu IC, Zusinaite E, Merits A, Snijder EJ, van Hemert MJ.** 2015. Stress granule components G3BP1 and G3BP2 play a proviral role early in Chikungunya virus replication. J Virol **89:**4457–4469.
139. **Vihinen H, Saarinen J.** 2000. Phosphorylation site analysis of Semliki forest virus nonstructural protein 3. J Biol Chem **275:**27775–27783.
140. **Vihinen H, Ahola T, Tuittila M, Merits A, Kaariainen L.** 2001. Elimination of phosphorylation sites of Semliki Forest virus replicase protein nsP3. J Biol Chem **276:**5745–5752.
141. **Varjak M, Zusinaite E, Merits A.** 2010. Novel functions of the alphavirus nonstructural protein nsP3 C-terminal region. J Virol **84:**2352–2364.
142. **O'Reilly EK, Kao CC.** 1998. Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. Virology **252:**287–303.
143. **Rubach JK, Wasik BR, Rupp JC, Kuhn RJ, Hardy RW, Smith JL.** 2009. Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity in vitro. Virology **384:**201–208.
144. **Tomar S, Hardy RW, Smith JL, Kuhn RJ.** 2006. Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity. J Virol **80:**9962–9969.
145. **de Groot RJ, Rumenapf T, Kuhn RJ, Strauss EG, Strauss JH.** 1991. Sindbis virus RNA polymerase is degraded by the N-end rule pathway. Proc Natl Acad Sci U S A **88:**8967–8971.
146. **Shirako Y, Strauss JH.** 1998. Requirement for an aromatic amino acid or histidine at the N terminus of Sindbis virus RNA polymerase. J Virol **72:**2310–2315.
147. **Rupp JC, Jundt N, Hardy RW.** 2011. Requirement for the amino-terminal domain of sindbis virus nsP4 during virus infection. J Virol **85:**3449–3460.
148. **Atasheva S, Fish A, Fornerod M, Frolova EI.** 2010. Venezuelan equine Encephalitis virus capsid protein forms a tetrameric complex with CRM1 and importin alpha/beta that obstructs nuclear pore complex function. J Virol **84:**4158–4171.
149. **Breakwell L, Dosenovic P, Karlsson Hedestam GB, D'Amato M, Liljestrom P, Fazakerley J, McInerney GM.** 2007. Semliki Forest virus nonstructural protein 2 is involved in suppression of the type I interferon response. J Virol **81:**8677–8684.
150. **Fros JJ, Liu WJ, Prow NA, Geertsema C, Ligtenberg M, Vanlandingham DL, Schnettler E, Vlak JM, Suhrbier A, Khromykh AA, Pijlman GP.** 2010. Chikungunya virus nonstructural protein 2 inhibits type I/II interferon-stimulated JAK-STAT signaling. J Virol **84:**10877–10887.

151. **Simmons JD, Wollish AC, Heise MT.** 2010. A determinant of Sindbis virus neurovirulence enables efficient disruption of Jak/STAT signaling. J Virol **84:**11429–11439.
152. **Gorchakov R, Frolova E, Williams BR, Rice CM, Frolov I.** 2004. PKR-dependent and -independent mechanisms are involved in translational shutoff during Sindbis virus infection. J Virol **78:**8455–8467.
153. **Ventoso I, Sanz MA, Molina S, Berlanga JJ, Carrasco L, Esteban M.** 2006. Translational resistance of late alphavirus mRNA to eIF2alpha phosphorylation: a strategy to overcome the antiviral effect of protein kinase PKR. Genes Dev **20:**87–100.
154. **Nivitchanyong T, Tsai YC, Betenbaugh MJ, Oyler GA.** 2009. An improved in vitro and in vivo Sindbis virus expression system through host and virus engineering. Virus Res **141:**1–12.
155. **Barry G, Fragkoudis R, Ferguson MC, Lulla A, Merits A, Kohl A, Fazakerley JK.** 2010. Semliki forest virus-induced endoplasmic reticulum stress accelerates apoptotic death of mammalian cells. J Virol **84:**7369–7377.
156. **Rathore AP, Ng ML, Vasudevan SG.** 2013. Differential unfolded protein response during Chikungunya and Sindbis virus infection: CHIKV nsP4 suppresses eIF2alpha phosphorylation. Virol J **10:**36.
157. **Fros JJ, Major LD, Scholte FE, Gardner J, van Hemert MJ, Suhrbier A, Pijlman GP.** 2015. Chikungunya virus non-structural protein 2-mediated host shut-off disables the unfolded protein response. J Gen Virol **96:**580–589.
158. **Sanz MA, Garcia-Moreno M, Carrasco L.** 2015. Inhibition of host protein synthesis by Sindbis virus: correlation with viral RNA replication and release of nuclear proteins to the cytoplasm. Cell Microbiol **17:**520–541.
159. **Li Z, Nagy PD.** 2011. Diverse roles of host RNA binding proteins in RNA virus replication. RNA Biol **8:**305–315.
160. **Nagy PD, Pogany J.** 2012. The dependence of viral RNA replication on co-opted host factors. Nat Rev Microbiol **10:**137–149.
161. **Frolova E, Gorchakov R, Garmashova N, Atasheva S, Vergara LA, Frolov I.** 2006. Formation of nsP3-specific protein complexes during Sindbis virus replication. J Virol **80:**4122–4134.
162. **Atasheva S, Gorchakov R, English R, Frolov I, Frolova E.** 2007. Development of Sindbis viruses encoding nsP2/GFP chimeric proteins and their application for studying nsP2 functioning. J Virol **81:**5046–5057.
163. **Cristea IM, Rozjabek H, Molloy KR, Karki S, White LL, Rice CM, Rout MP, Chait BT, MacDonald MR.** 2010. Host factors associated with the Sindbis virus RNA-dependent RNA polymerase: role for G3BP1 and G3BP2 in virus replication. J Virol **84:**6720–6732.
164. **Montgomery SA, Berglund P, Beard CW, Johnston RE.** 2006. Ribosomal protein S6 associates with alphavirus nonstructural protein 2 and mediates expression from alphavirus messages. J Virol **80:**7729–7739.
165. **Burnham AJ, Gong L, Hardy RW.** 2007. Heterogeneous nuclear ribonuclear protein K interacts with Sindbis virus nonstructural proteins and viral subgenomic mRNA. Virology **367:**212–221.
166. **Dickson AM, Anderson JR, Barnhart MD, Sokoloski KJ, Oko L, Opyrchal M, Galanis E, Wilusz CJ, Morrison TE, Wilusz J.** 2012. Dephosphorylation of HuR protein during alphavirus infection is associated with HuR relocalization to the cytoplasm. J Biol Chem **287:**36229–36238.

167. **Bourai M, Lucas-Hourani M, Gad HH, Drosten C, Jacob Y, Tafforeau L, Cassonnet P, Jones LM, Judith D, Couderc T, Lecuit M, Andre P, Kummerer BM, Lotteau V, Despres P, Tangy F, Vidalain PO.** 2012. Mapping of Chikungunya virus interactions with host proteins identified nsP2 as a highly connected viral component. J Virol **86:**3121–3134.
168. **Balistreri G, Horvath P, Schweingruber C, Zund D, McInerney G, Merits A, Muhlemann O, Azzalin C, Helenius A.** 2014. The host nonsense-mediated mRNA decay pathway restricts Mammalian RNA virus replication. Cell Host Microbe **16:**403–411.
169. **Griffin DE.** 2005. Neuronal cell death in alphavirus encephalomyelitis. Curr Top Microbiol Immunol **289:**57–77.
170. **Smithburn KC, Haddow WJ.** 1944. Semliki Forest virus. I. Isolation and pathogenic properties. J Immunol **49:**141–145.
171. **Bradish CJ, Allner K, Maber HB.** 1971. The virulence of original and derived strains of Semliki forest virus for mice, guinea-pigs and rabbits. J Gen Virol **12:**141–160.
172. **Garoff H, Frischauf AM, Simons K, Lehrach H, Delius H.** 1980. Nucleotide sequence of cdna coding for Semliki Forest virus membrane glycoproteins. Nature **288:**236–241.
173. **Willems WR, Kaluza G, Boschek CB, Bauer H, Hager H, Schutz HJ, Feistner H.** 1979. Semliki forest virus: cause of a fatal case of human encephalitis. Science **203:**1127–1129.
174. **McIntosh BM, Worth CB, Kokernot RH.** 1961. Isolation of Semliki Forest virus from Aedes (Aedimorphus) argenteopunctatus (Theobald) collected in Portuguese East Africa. Trans R Soc Trop Med Hyg **55:**192–198.
175. **Tuittila MT, Santagati MG, Roytta M, Maatta JA, Hinkkanen AE.** 2000. Replicase complex genes of Semliki Forest virus confer lethal neurovirulence. J Virol **74:**4579–4589.
176. **Fazakerley JK, Pathak S, Scallan M, Amor S, Dyson H.** 1993. Replication of the A7(74) strain of Semliki Forest virus is restricted in neurons. Virology **195:**627–637.
177. **Fragkoudis R, Ballany CM, Boyd A, Fazakerley JK.** 2008. In Semliki Forest virus encephalitis, antibody rapidly clears infectious virus and is required to eliminate viral material from the brain, but is not required to generate lesions of demyelination. J Gen Virol **89:**2565–2568.
178. **Pathak S, Webb HE.** 1974. Possible mechanisms for the transport of Semliki forest virus into and within mouse brain. An electron-microscopic study. J Neurol Sci **23:**175–184.
179. **Fragkoudis R, Breakwell L, McKimmie C, Boyd A, Barry G, Kohl A, Merits A, Fazakerley JK.** 2007. The type I interferon system protects mice from Semliki Forest virus by preventing widespread virus dissemination in extraneural tissues, but does not mediate the restricted replication of avirulent virus in central nervous system neurons. J Gen Virol **88:**3373–3384.
180. **Gates MC, Sheahan BJ, O'Sullivan MA, Atkins GJ.** 1985. The pathogenicity of the A7, M9 and L10 strains of Semliki Forest virus for weanling mice and primary mouse brain cell cultures. J Gen Virol **66 (Pt 11):**2365–2373.
181. **Oliver KR, Scallan MF, Dyson H, Fazakerley JK.** 1997. Susceptibility to a neurotropic virus and its changing distribution in the developing brain is a function of CNS maturity. J Neurovirol **3:**38–48.

182. **Atkins GJ, Sheahan BJ, Liljestrom P.** 1999. The molecular pathogenesis of Semliki Forest virus: a model virus made useful? J Gen Virol **80 (Pt 9):**2287–2297.
183. **Fazakerley JK, Cotterill CL, Lee G, Graham A.** 2006. Virus tropism, distribution, persistence and pathology in the corpus callosum of the Semliki Forest virus-infected mouse brain: a novel system to study virus-oligodendrocyte interactions. Neuropathol Appl Neurobiol **32:**397–409.
184. **Amor S, Scallan MF, Morris MM, Dyson H, Fazakerley JK.** 1996. Role of immune responses in protection and pathogenesis during Semliki Forest virus encephalitis. J Gen Virol **77 (Pt 2):**281–291.
185. **Scallan MF, Fazakerley JK.** 1999. Aurothiolates enhance the replication of Semliki Forest virus in the CNS and the exocrine pancreas. J Neurovirol **5:**392–400.
186. **Galbraith SE, Sheahan BJ, Atkins GJ.** 2006. Deletions in the hypervariable domain of the nsP3 gene attenuate Semliki Forest virus virulence. J Gen Virol **87:**937–947.
187. **Griffin DE, Metcalf T.** 2011. Clearance of virus infection from the CNS. Curr Opin Virol **1:**216–221.
188. **Nikonov A, Molder T, Sikut R, Kiiver K, Mannik A, Toots U, Lulla A, Lulla V, Utt A, Merits A, Ustav M.** 2013. RIG-I and MDA-5 detection of viral RNA-dependent RNA polymerase activity restricts positive-strand RNA virus replication. PLoS Pathog **9:**e1003610.
189. **Bradish CJ, Allner K, Maber HB.** 1972. Infection, interaction and the expression of virulence by defined strains of Semliki forest virus. J Gen Virol **16:** 359–372.
190. **Deuber SA, Pavlovic J.** 2007. Virulence of a mouse-adapted Semliki Forest virus strain is associated with reduced susceptibility to interferon. J Gen Virol **88:**1952–1959.
191. **Martikainen M, Niittykoski M, von Und Zu Fraunberg M, Immonen A, Koponen S, van Geenen M, Vaha-Koskela M, Ylosmaki E, Jaaskelainen JE, Saksela K, Hinkkanen A.** 2015. MicroRNA-Attenuated Clone of Virulent Semliki Forest Virus Overcomes Antiviral Type I Interferon in Resistant Mouse CT-2A Glioma. J Virol **89:**10637–10647.
192. **Fazakerley JK, Webb HE.** 1987. Semliki Forest virus-induced, immune-mediated demyelination: adoptive transfer studies and viral persistence in nude mice. J Gen Virol **68 (Pt 2):**377–385.
193. **Subak-Sharpe I, Dyson H, Fazakerley J.** 1993. In vivo depletion of CD8+ T cells prevents lesions of demyelination in Semliki Forest virus infection. J Virol **67:**7629–7633.
194. **Tyndall C, La Mantia G, Thacker CM, Favaloro J, Kamen R.** 1981. A region of the polyoma virus genome between the replication origin and late protein coding sequences is required in cis for both early gene expression and viral DNA replication. Nucleic Acids Res **9:**6231–6250.
195. **Bernard KA, Klimstra WB, Johnston RE.** 2000. Mutations in the E2 glycoprotein of Venezuelan equine encephalitis virus confer heparan sulfate interaction, low morbidity, and rapid clearance from blood of mice. Virology **276:**93–103.
196. **Davis NL, Fuller FJ, Dougherty WG, Olmsted RA, Johnston RE.** 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 U S A **83:**6771–6775.
197. **Gardner CL, Ebel GD, Ryman KD, Klimstra WB.** 2011. Heparan sulfate binding by natural eastern equine encephalitis viruses promotes neurovirulence. Proc Natl Acad Sci U S A **108:**16026–16031.
198. **Ryman KD, Gardner CL, Burke CW, Meier KC, Thompson JM, Klimstra WB.** 2007. Heparan sulfate binding can contribute to the neurovirulence of neuroadapted and nonneuroadapted Sindbis viruses. J Virol **81:**3563–3573.
199. **Glasgow GM, Killen HM, Liljestrom P, Sheahan BJ, Atkins GJ.** 1994. A single amino acid change in the E2 spike protein of a virulent strain of Semliki Forest virus attenuates pathogenicity. J Gen Virol **75 (Pt 3):**663–668.
200. **Fazakerley JK.** 2004. Semliki forest virus infection of laboratory mice: a model to study the pathogenesis of viral encephalitis. Arch Virol Suppl**:**179–190.
201. **Klimstra WB, Ryman KD, Johnston RE.** 1998. Adaptation of Sindbis virus to BHK cells selects for use of heparan sulfate as an attachment receptor. J Virol **72:**7357–7366.
202. **Byrnes AP, Griffin DE.** 2000. Large-plaque mutants of Sindbis virus show reduced binding to heparan sulfate, heightened viremia, and slower clearance from the circulation. J Virol **74:**644–651.
203. **Heil ML, Albee A, Strauss JH, Kuhn RJ.** 2001. An amino acid substitution in the coding region of the E2 glycoprotein adapts Ross River virus to utilize heparan sulfate as an attachment moiety. J Virol **75:**6303–6309.
204. **Gardner CL, Choi-Nurvitadhi J, Sun C, Bayer A, Hritz J, Ryman KD, Klimstra WB.** 2013. Natural variation in the heparan sulfate binding domain of the eastern equine encephalitis virus E2 glycoprotein alters interactions with cell surfaces and virulence in mice. J Virol **87:**8582–8590.
205. **Gardner CL, Hritz J, Sun C, Vanlandingham DL, Song TY, Ghedin E, Higgs S, Klimstra WB, Ryman KD.** 2014. Deliberate attenuation of chikungunya virus by adaptation to heparan sulfate-dependent infectivity: a model for rational arboviral vaccine design. PLoS Negl Trop Dis **8:**e2719.
206. **Cassaro CM, Dietrich CP.** 1977. Distribution of sulfated mucopolysaccharides in invertebrates. J Biol Chem **252:**2254–2261.
207. **Kjellen L, Lindahl U.** 1991. Proteoglycans: structures and interactions. Annu Rev Biochem **60:**443–475.
208. **Stringer SE, Gallagher JT.** 1997. Heparan sulphate. Int J Biochem Cell Biol **29:**709–714.
209. **Karlsson K, Marklund SL.** 1988. Plasma clearance of human extracellular-superoxide dismutase C in rabbits. J Clin Invest **82:**762–766.
210. **Karlsson K, Sandstrom J, Edlund A, Edlund T, Marklund SL.** 1993. Pharmacokinetics of extracellular-superoxide dismutase in the vascular system. Free Radic Biol Med **14:**185–190.
211. **Ji ZS, Sanan DA, Mahley RW.** 1995. Intravenous heparinase inhibits remnant lipoprotein clearance from the plasma and uptake by the liver: in vivo role of heparan sulfate proteoglycans. J Lipid Res **36:**583–592.
212. **Ashbrook AW, Burrack KS, Silva LA, Montgomery SA, Heise MT, Morrison TE, Dermody TS.** 2014. Residue 82 of the Chikungunya virus E2 attachment protein modulates viral dissemination and arthritis in mice. J Virol **88:**12180–12192.

213. **Lee E, Lobigs M.** 2000. Substitutions at the putative receptor-binding site of an encephalitic flavivirus alter virulence and host cell tropism and reveal a role for glycosaminoglycans in entry. J Virol **74:**8867–8875.
214. **Glasgow GM, Sheahan BJ, Atkins GJ, Wahlberg JM, Salminen A, Liljestrom P.** 1991. Two mutations in the envelope glycoprotein E2 of Semliki Forest virus affecting the maturation and entry patterns of the virus alter pathogenicity for mice. Virology **185:**741–748.
215. **Bear JS, Byrnes AP, Griffin DE.** 2006. Heparin-binding and patterns of virulence for two recombinant strains of Sindbis virus. Virology **347:**183–190.
216. **Banks WA, Robinson SM, Wolf KM, Bess JW, Jr., Arthur LO.** 2004. Binding, internalization, and membrane incorporation of human immunodeficiency virus-1 at the blood-brain barrier is differentially regulated. Neuroscience **128:**143–153.
217. **Bobardt MD, Armand-Ugon M, Clotet I, Zhang Z, David G, Este JA, Gallay PA.** 2004. Effect of polyanion-resistance on HIV-1 infection. Virology **325:**389–398.
218. **Atkins GJ.** 1983. The avirulent A7 Strain of Semliki Forest virus has reduced cytopathogenicity for neuroblastoma cells compared to the virulent L10 strain. J Gen Virol **64 (Pt 6):**1401–1404.
219. **Tamberg N, Lulla V, Fragkoudis R, Lulla A, Fazakerley JK, Merits A.** 2007. Insertion of EGFP into the replicase gene of Semliki Forest virus results in a novel, genetically stable marker virus. J Gen Virol **88:**1225–1230.
220. **Russo AT, Malmstrom RD, White MA, Watowich SJ.** 2010. Structural basis for substrate specificity of alphavirus nsP2 proteases. J Mol Graph Model **29:**46–53.
221. **Heise MT, White LJ, Simpson DA, Leonard C, Bernard KA, Meeker RB, Johnston RE.** 2003. An attenuating mutation in nsP1 of the Sindbis-group virus S.A.AR86 accelerates nonstructural protein processing and up-regulates viral 26S RNA synthesis. J Virol **77:**1149–1156.
222. **Panas MD, Ahola T, McInerney GM.** 2014. The C-terminal repeat domains of nsP3 from the Old World alphaviruses bind directly to G3BP. J Virol **88:**5888–5893.
223. **Tucker PC, Lee SH, Bui N, Martinie D, Griffin DE.** 1997. Amino acid changes in the Sindbis virus E2 glycoprotein that increase neurovirulence improve entry into neuroblastoma cells. J Virol **71:**6106–6112.
224. **Kobiler D, Rice CM, Brodie C, Shahar A, Dubuisson J, Halevy M, Lustig S.** 1999. A single nucleotide change in the 5' noncoding region of Sindbis virus confers neurovirulence in rats. J Virol **73:**10440–10446.
225. **Kulcsar KA, Baxter VK, Abraham R, Nelson A, Griffin DE.** 2015. Distinct Immune Responses in Resistant and Susceptible Strains of Mice during Neurovirulent Alphavirus Encephalomyelitis. J Virol **89:**8280–8291.
226. **Frolov I, Akhrymuk M, Akhrymuk I, Atasheva S, Frolova EI.** 2012. Early events in alphavirus replication determine the outcome of infection. J Virol **86:**5055–5066.
227. **Glebov OO, Bright NA, Nichols BJ.** 2006. Flotillin-1 defines a clathrin-independent endocytic pathway in mammalian cells. Nat Cell Biol **8:**46–54.
228. **Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M.** 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics **1:**376–386.

SUMMARY IN ESTONIAN

Samm lähemale mõistmaks Semliki Forest viiruse neurovirulentsust

Semliki Forest viirus (SFV) on positiivse polaarsusega RNA genoomiga viirus, mis kuulub alfaviiruste perekonda sugukonnas *Togaviridae*. Alfaviiruste seas leidub mitmeid inimese ja loomade patogeene, mille poolt põhjustatavad haigused varieeruvad külmetuse sarnaste sümptomitega tõvest aastaid kestva artriidi või fataalse entsefaliidini. Seega kujutab alfaviiruste uurimine endast olulist valdkonda. SFV on inimesele suhteliselt ohutute laboratoorsete tüvede olemasolu tõttu antud perekonnas üheks enim uuritud esindajaks; seda viirust on laialdaselt kasutatud nii geenitehnoloogia süsteemina kui ka mudelobjektina viirusliku entsefaliidiga kaasneva patogeneesi uurimisel.

Kõige põhjalikumalt uuritud SFV tüvedeks on A7(74), SFV4 ja L10. A7(74) on avirulentne tüvi, sest see põhjustab täiskasvanud hiires asümptomaatilist infektsiooni. SFV4 ja L10 on aga virulentsed tüved, sest otse ajju või kõrge doosiga kõhuõõnde süstituna põhjustavad nad hiirtel surmaga lõppevat entsefaliiti. Kui aga kõhuõõnde süstimiseks kasutada bioloogiliselt mõtestatud (samas suurusjärgus, kui seda sisestavad viirust edasikandvad sääsed) kogust viirust, siis SFV4 ei jõua ajju ning nakatatud hiired jäävad ellu. Käesoleva uurimistöö üldeesmärgiks oli uurida tegureid ja mehhanisme, mis vastutavad nende tüvede erineva neurovirulentsuse eest.

Esimeseks eesmärgiks oli luua L10 tüvele vastav konsensuskloon ning analüüsida SFV4 ja L10 erinevaid fenotüüpe põhjustavaid molekulaarseid determinante ja mehhanisme. Tehtud uurimused näitasid, et SFV4 ja L10 vahelised erinevused hiires on tingitud viiruse ümbrisevalgu aminohapete laengutest. Negatiivse laenguga glutamiinhappejäägid E2 valgu teatud positsioonides võimaldavad viirusel põhjustada kõrget vireemiat. Kõrge vireemia on omakorda eelduseks viiruse jõudmisel ajju ning seetõttu on E2 valgus negatiivse laenguga aminohappejääke sisaldavad viirused (nagu L10) neuroinvasiivsed ning virulentsed. Positiivse laenguga lüsiinijäägid vastavates E2 valgu positioonides vastutavad madala vireemia eest, mistõttu ei suuda sellised viirused (nagu SFV4) ajju siseneda ega entsefaliiti põhjustada. E2 valgul on oluline roll viiruse seondumisel peremehe rakkudele. Me näitasime, et positiivse laenguga lüsiinijäägid E2 valgus soodustavad küll viiruse seondumist koekultuuri rakkudele, kuid põhjustavad *in vivo* tingimustes SFV4 virionide efektiivse seostumise heparaansulfaadiga, mis vähendab viiruse taset veres ja selle kaudu võimekust ajju siseneda.

Järgnevaks eesmärgiks oli kasutada meie loodud A7(74) ja L10 tüvede konsensusjärjestustele vastavaid rekombinantseid viirusi, et heita valgust nende tüvede vahelisi fenotüübilisi erinevusi põhjustavatele molekulaarsetele teguritele. Saadud andmed näitasid, et erinevused A7(74) ja L10 neurovirulentsuses on tingitud erinevustest viiruse mittestruktuurse liitvalgu proteolüütilise lõikamise

kiiruses ning mittestruktuurse valgu 3 (nsP3) järjestuses. Saadud tulemused näitavad, et neurovirulentsuse põhjustamisel on virulentse tüve nsP3 valgul dominantne funktsioon. A7(74) nsP3 valk (või pigem L10 nsP3 valgu puudumine) ja mittestruktuurse liitvalgu 1/2 lõikamisjärjestuse kiire protsessimine vastutavad avirulentse *in vivo* fenotüübi eest. Liitvalgu 1/2 lõikamisjärjestuse aeglane protsessimine või L10 nsP3 valgu olemasolu suurendavad SFV replikaasi võimet sünteesida ebaharilikke PAMP RNA-sid, kasutades matriitsina raku RNA-sid, ja see omakorda põhjustab tugevat interferoonvastust. Seega, SFV tüvede erinev suutlikkus täiskasvanud hiire ajus levida korreleerub siin viiruste erineva võimega indutseerida interferooni tootmist; see viitab seosele SFV neurovirulentsuse ja immunopatoloogia vahel.

Üheks uurimustöö alameesmärgiks oli analüüsida SFV replikatsioonikompleksidega seonduvaid valke, saamaks paremat ülevaadet alfaviiruse ja peremehe vahel esinevatest keerulistest interaktsioonidest. Selleks arendati käesolevas uurimistöös välja uus lähenemine, mis põhineb rakkude söötmisel dekstraaniga kaetud magneetiliste nanopartiklitega ja järgneval viiruse replikatsioonikomplekse sisaldavate lüsosoomide magneetilisel eraldamisel. Mitteradioaktiivsel isotoopmärgistusel põhineva kvantitatiivse proteoomika kasutamine võimaldas meil koostada põhjaliku nimekirja SFV replikatsioonikompleksidega seostuvatest rakulistest valkudest. Nelja sellesse nimekirja kuuluva valgu edasine põhjalikum analüüs kinnitas nende olulisust SFV infektsioonis ja seega ka uudse metoodika usaldusväärsust. Järgneva sammuna võiks antud meetodit kasutada erineva virulentsusega SFV tüvede uurimiseks. Oleks huvitav teada, kas erinevad SFV tüved seostuvad erinevate peremehe poolsete valkudega või mõjutavad samad peremehe valgud erinevaid SFV tüvesid erinevalt. See omakorda avardaks veelgi meie arusaamist SFV patogeensuse põhjustest ja mehhanismidest.

Käesoleva uurimistöö käigus saadud tulemused aitavad ühelt poolt paremini mõista SFV neurovirulentsust, lisaks sellele on mudelviiruse uurimisel kasutatud meetode ja avastatud seaduspärasusi võimalik rakendada ka inimesele ohtlike alfaviiruste mõistmiseks.

ACKNOWLEDGEMENTS

Nelson Mandela has said: “It always seems impossible until it’s done.” My sincere gratitude goes to all who have encouraged and helped me during my doctorate studies and made it possible for me to now say that it’s done!

First, I would like to thank my supervisor professor Andres Merits for his guidance, support, and dedication. I admire his passion for science and his vast knowledge about virology. I am thankful to my co-supervisor Eva Žusinaite for her assistance and advice. I would also like to express my appreciation to the coauthors of the papers for their hard work and dedication.

I want to thank all of the former and current members of the RNA virology group, whom I have had the pleasure and honor to work with over the years. I am grateful to Margus Varjak for his teachings and would like to thank Nele Tamberg for our discussions. In addition, I appreciate all of the wonderful people in Ustav’s, Loog’s, Tenson’s, Ilves’s and Skotheim’s group. Thank you all for your support and the amazing times we spent together.

I am so grateful to all of my dearest friends, who kept me grounded and made my time outside the lab so memorable. I would especially like to thank Age Utt for being there for me since the beginning and through thick and thin.

I would like to express my deepest gratitude to my family for their unconditional love, encouragement, and endless trust in me. I am grateful for Flora for keeping me sane during the years. And last, my beloved Mardo, thank you for being my rock!

PUBLICATIONS

CURRICULUM VITAE

Name: Sirle Saul
Date of birth: 24th June 1986
Citizenship: Estonian
e-mail: sirlesau@ut.ee

Education and professional employment:

2002–2005	Tallinn English College
2005–2008	University of Tartu, BSc in biology
2008–2010	University of Tartu, MSc in gene technology, *cum laude*
2010 June	University of La Reunion, transfer of technologies between common projects with research group of Philippe Gasque
2010–	University of Tartu, doctorate studies in biomedical engineering
2012 December/ 2013 November	University of Edinburgh, collaboration with research group of John Fazakerley
2014–	junior researcher, Institute of Technology, University of Tartu

List of publication:

1. Varjak M, **Saul S**, Arike L, Lulla A, Peil L, Merits A. 2013. Magnetic fractionation and proteomic dissection of cellular organelles occupied by the late replication complexes of Semliki Forest virus. J Virol 87:10295–10312.
2. Ferguson MC*, **Saul S***, Fragkoudis R, Weisheit S, Cox J, Patabendige A, Sherwood K, Watson M, Merits A, Fazakerley JK. 2015. The ability of the encephalitic arbovirus Semliki Forest virus to cross the blood brain barrier is determined by the charge of the E2 glycoprotein. J Virol 89:7536–7549.
3. **Saul S**, Ferguson MC, Cordonin C, Fragkoudis R, Ool M, Tamberg N, Sherwood K, Fazakerley JK, Merits A. 2015. Differences in processing determinants of nonstructural polyprotein and in the sequence of nonstructural protein 3 affect neurovirulence of Semliki Forest virus. J Virol 89:11030–11045.
4. Stapleford KA, Rozen-Gagnon K, Das PK, **Saul S,** Poirier EZ, Blanc H, Vidalain PO, Merits A, Vignuzzi M. 2015. Viral Polymerase-Helicase Complexes Regulate Replication Fidelity To Overcome Intracellular Nucleotide Depletion. J Virol 89:11233–11244.

ELULOOKIRJELDUS

Nimi: Sirle Saul
Sünniaeg: 24.06.1986
Kodakondsus: Eesti
e-mail: sirlesau@ut.ee

Haridus- ja teenistuskäik:

2002–2005	Tallinna Inglise Kolledž
2005–2008	Tartu Ülikool, loodusteaduse bakalaureus (BSc) bioloogia erialal
2008–2010	Tartu Ülikool, loodusteaduse magister (MSc) geenitehnoloogia erialal, *cum laude*
2010 juuni	La Reunion'i Ülikool, ühiste projektide vaheline teadussiire Philippe Gasque'i uurimisrühmaga
2010–	Tartu Ülikool, doktorantuur biomeditsiini tehnoloogia erialal
2012 detsember/ 2013 november	Edinburgh'i Ülikool, koostöö ühiste projektide raames John Fazakerley uurimisrühmaga
2014–	nooremteadur Tartu Ülikooli Tehnoloogiainstituudis

Publikatsioonide nimekiri:

1. Varjak M, **Saul S**, Arike L, Lulla A, Peil L, Merits A. 2013. Magnetic fractionation and proteomic dissection of cellular organelles occupied by the late replication complexes of Semliki Forest virus. J Virol 87:10295–10312.
2. Ferguson MC*, **Saul S***, Fragkoudis R, Weisheit S, Cox J, Patabendige A, Sherwood K, Watson M, Merits A, Fazakerley JK. 2015. The ability of the encephalitic arbovirus Semliki Forest virus to cross the blood brain barrier is determined by the charge of the E2 glycoprotein. J Virol 89:7536–7549.
3. **Saul S**, Ferguson MC, Cordonin C, Fragkoudis R, Ool M, Tamberg N, Sherwood K, Fazakerley JK, Merits A. 2015. Differences in processing determinants of nonstructural polyprotein and in the sequence of nonstructural protein 3 affect neurovirulence of Semliki Forest virus. J Virol 89:11030–11045.
4. Stapleford KA, Rozen-Gagnon K, Das PK, **Saul S,** Poirier EZ, Blanc H, Vidalain PO, Merits A, Vignuzzi M. 2015. Viral Polymerase-Helicase Complexes Regulate Replication Fidelity To Overcome Intracellular Nucleotide Depletion. J Virol 89:11233–11244.

DISSERTATIONES TECHNOLOGIAE UNIVERSITATIS TARTUENSIS

1. **Imre Mäger.** Characterization of cell-penetrating peptides: Assessment of cellular internalization kinetics, mechanisms and bioactivity. Tartu 2011, 132 p.
2. **Taavi Lehto.** Delivery of nucleic acids by cell-penetrating peptides: application in modulation of gene expression. Tartu 2011, 155 p.
3. **Hannes Luidalepp**. Studies on the antibiotic susceptibility of *Escherichia coli*. Tartu 2012, 111 p.
4. **Vahur Zadin**. Modelling the 3D-microbattery. Tartu 2012, 149 p.
5. **Janno Torop**. Carbide-derived carbon-based electromechanical actuators. Tartu 2012, 113 p.
6. **Julia Suhorutšenko.** Cell-penetrating peptides: cytotoxicity, immunogenicity and application for tumor targeting. Tartu 2012, 139 p.
7. **Viktoryia Shyp.** G nucleotide regulation of translational GTPases and the stringent response factor RelA. Tartu 2012, 105 p.
8. **Mardo Kõivomägi.** Studies on the substrate specificity and multisite phosphorylation mechanisms of cyclin-dependent kinase Cdk1 in *Saccharomyces cerevisiae*. Tartu, 2013, 157 p.
9. **Liis Karo-Astover.** Studies on the Semliki Forest virus replicase protein nsP1. Tartu, 2013, 113 p.
10. **Piret Arukuusk**. NickFects–novel cell-penetrating peptides. Design and uptake mechanism. Tartu, 2013, 124 p.
11. **Piret Villo.** Synthesis of acetogenin analogues. Asymmetric transfer hydrogenation coupled with dynamic kinetic resolution of α-amido-β-keto esters. Tartu, 2013, 151 p.
12. **Villu Kasari.** Bacterial toxin-antitoxin systems: transcriptional cross-activation and characterization of a novel *mqsRA* system. Tartu, 2013, 108 p.
13. **Margus Varjak.** Functional analysis of viral and host components of alphavirus replicase complexes. Tartu, 2013, 151 p.
14. **Liane Viru.** Development and analysis of novel alphavirus-based multifunctional gene therapy and expression systems. Tartu, 2013, 113 p.
15. **Kent Langel.** Cell-penetrating peptide mechanism studies: from peptides to cargo delivery. Tartu, 2014, 115 p.
16. **Rauno Temmer**. Electrochemistry and novel applications of chemically synthesized conductive polymer electrodes. Tartu, 2014, 206 p.
17. **Indrek Must.** Ionic and capacitive electroactive laminates with carbonaceous electrodes as sensors and energy harvesters. Tartu, 2014, 133 p.
18. **Veiko Voolaid.** Aquatic environment: primary reservoir, link, or sink of antibiotic resistance? Tartu, 2014, 79 p.
19. **Kristiina Laanemets.** The role of SLAC1 anion channel and its upstream regulators in stomatal opening and closure of *Arabidopsis thaliana.* Tartu, 2015, 115 p.

20. **Kalle Pärn**. Studies on inducible alphavirus-based antitumour strategy mediated by site-specific delivery with activatable cell-penetrating peptides. Tartu, 2015, 139 p.
21. **Anastasia Selyutina.** When biologist meets chemist: a search for HIV-1 inhibitors. Tartu, 2015, 172p.