DISSERTATIONES
TECHNOLOGIAE
UNIVERSITATIS
TARTUENSIS
72

LAURA SANDRA LELLO

Unraveling the intricate nature of the alphavirus RNA replicase

DISSERTATIONES TECHNOLOGIAE UNIVERSITATIS TARTUENSIS
72

LAURA SANDRA LELLO

Unraveling the intricate nature of the alphavirus RNA replicase

UNIVERSITY OF TARTU
Press
1632

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

This dissertation was accepted for the commencement of the degree of Doctor of Philosophy in Biomedical Technology on 19th of June 2023 by the Council of the 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
Reviewer:	Margit Mutso, PhD, Research Fellow, Institute of Technology, University of Tartu, Estonia
Opponent:	Scott Weaver, PhD Professor, Microbiology & Immunology, and Pathology, University of Texas Medical Branch, USA
Commencement:	Auditorium 121, Nooruse 1, Tartu, Estonia, at 12.15 on August 25th, 2023

Publication of this dissertation is granted by the Institute of Technology, Faculty of Science and Technology, University of Tartu

ISSN 2228-0855 (print)
ISBN 978-9916-27-263-3 (print)
ISSN 2806-2620 (pdf)
ISBN 978-9916-27-264-0 (pdf)

University of Tartu Press
www.tyk.ee

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS

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

I **Lello, L.S.**, Utt, A., Bartholomeeusen, K., Wang, S., Rausalu, K., Kendall, C., Coppens, S., Fragkoudis, R., Tuplin, A., Alphey, L., Ariën, K.K. and Merits, A. (2020). Cross-utilisation of template RNAs by alphavirus replicases. *Plos Pathogens* **16**(9): e1008825. https://doi.org/10.1371/journal.ppat.1008825.

II **Lello, L.S.,** Bartholomeeusen, K., Wang, S., Coppens, S., Fragkoudis, R., Alphey, L., Arien, K.K., Merits, A. and Utt, A. (2021). nsP4 is a major determinant of alphavirus replicase activity and template selectivity. *Journal of Virology*, https://doi.org/10.1128/JVI.00355-21.

III **Lello, L.S.,** Miilimäe, A., Cherkashchenko, L., Omler, A., Skilton, R., Ireland, R., Ulaeto, D. and Merits, A (2022). Activity, template preference and compatibility of components of RNA replicase of eastern equine encephalitis virus. *Journal of Virology*; e0136822. https://doi.org/10.1128/jvi.01368-22.

IV Tan, Y.B., **Lello, L.S.,** Liu, X., Law, Y.-S., Kang, C., Lescar, J., Zheng, J., Merits, A. and Luo, D. (2022). Crystal structures of alphavirus nonstructural protein 4 (nsP4) reveal an intrinsically dynamic RNA-dependent RNA polymerase fold. *Nucleic Acids Research.* https://doi.org/10.1093/nar/gkab1302.

The articles listed above are reprinted with the permission of the copyright owners. This thesis also contains unpublished data.

Author's contribution:

I. I participated in the experimental design, performed the optimization of the *trans*-replicase assays for different formats and executed all experiments except for the *in silico* analysis of the RNA structures, generation of the DNA constructs and the rescue of recombinant viruses.

II. I participated in the experimental design and performed all experiments except for generating the DNA constructs.

III. I participated in the experimental design, generated the DNA constructs for the two-component EEEV replicase, performed the component titration experiments for flow cytometry analysis, confirmed the results from the *trans*-replicase assay via northern blot analysis and designed, optimized and performed the transfection–infection experiments for the EEEV and CHIKV biosensors using CHIKV, SINV and KUNV infections.

IV. I performed the *trans*-replicase assays for structure-guided mutagenesis of the structurally important elements of the nsP4 proteins of SINV and RRV.

ABBREVIATONS

1C	one-component
2C	two-component
3'R	3' region (the 3' UTR of a relevant virus)
3C	three-component
5'R	5' region (the 5' UTR and start of nsP1 of a relevant virus)
ADPr	ADP ribose
AUD	alphavirus unique domain
BFV	Barmah Forest virus
BSL	biosafety level
CHIKV	chikungunya virus
CMV	human cytomegalovirus
CP	capsid protein
CPV-I	type-I cytopathic vacuoles
CPV-II	type-II cytopathic vacuoles
CSE	conserved sequence elements
ds	double-stranded
EEEV	Eastern equine encephalitis virus
EILV	Eilat virus
Fluc	firefly luciferase
ZsG	ZsGreen fluorescent reporter
Gluc	*Gaussia* luciferase
GTase	guanylyltransferase
HDV RZ	antisense strand ribozyme of hepatitis delta virus
HSPolI	*Homo sapiens* polymerase I
HVD	hypervariable domain
ICA	infectious centre assay
icDNA	infectious cDNA
IFN	interferon
ISV	insect specific virus
LDLRAD3	low density lipoprotein receptor class A domain-containing 3
MAR	mono-ADP-ribose
MAYV	Mayaro virus
MmTer	*Mus musculus* polymerase I terminator
MTase	methyltransferase
NC	nucleocapsid
NRAMP	natural resistance-associated macrophage protein
ns	nonstructural
nsP	nonstructural protein
NTD	N-terminal domain
NTPase	nucleoside triphosphatase
ONNV	O'nyong nyong virus
ORF	open reading frame

PARP	poly(ADPr)polymerase
PHB1	prohibitin I
PI3K	plasma membrane associated lipid kinases
RC	replication complex
RdRp	RNA-dependent RNA polymerase
RRV	Ross River virus
RTPase	RNA triphosphatase
SFV	Semliki Forest virus
SG	subgenomic
SHAPE	Selective 2’ Hydroxyl Acylation analyzed by Primer Extension
SINV	Sindbis virus
SL	stem-loop
ss	single-stranded
SV40Ter	Simian virus 40 terminator
TF	trans-frame
TIM	T-cell immunoglobulin mucin
ts	temperature-sensitive
ubi	ubiquitin
UTR	untranslated region
VEEV	Venezuelan equine encephalitis virus
WEEV	Western equine encephalitis virus
VLDLR	very low density lipoprotein receptor
VLP	virus-like particles
wt	wild-type

INTRODUCTION

Alphaviruses are important human and animal pathogens responsible for many outbreaks and epidemics in recent decades. The majority of alphaviruses use insect vectors to spread between vertebrate hosts. The most common vector species are *Aedes* and *Culex* mosquitoes, which are found on all continents except Antarctica. The increasing geographical distribution of alphaviruses is supported by growing intercontinental travel, trade, and climate change, which have led to the vector mosquitoes inhabiting new areas. Together with urbanization, alphaviruses are increasingly converting from natural enzootic cycles to urban cycles, resulting in urban epidemics and high attack rates.

One of the most medically important alphaviruses is chikungunya virus (CHIKV), known as the source of large-scale epidemics in Africa, the Indian Ocean islands, Southeast Asia, and the Americas. One of the main concerns regarding CHIKV is its ability to adapt to new vector species. This adaptation was the primary cause for the epidemic on Reunion Island in 2005–2006, where CHIKV mainly spread by the *Aedes albopictus* mosquitoes instead of its usual *Ae. aegypti* vector. Other alphaviruses of growing concern are the encephalitic New World viruses, which are geographically confined to the American continent. These viruses, for example, the Eastern equine encephalitic virus (EEEV), are usually maintained in an enzootic cycle accompanied by rare spillover events to dead-end hosts, such as humans and horses. However, in recent years, a growing number of EEEV infections in humans have been recorded in the United States of America. Furthermore, CHIKV and EEEV are only a fraction of the total number of highly pathogenic alphaviruses that cause distress around the world. It is only a matter of time as to when another alphavirus re-emerges, the ramifications of which could be disastrous. For example, Mayaro virus (MAYV) has the potential to become another major public health concern due to its increasing prevalence in the Amazon rural region caused by changes in the ecosystem. In addition, while the primary vector of MAYV is the *Haemagogus* species, MAYV is also spread by *Aedes* mosquitoes, which are continuously inhabiting new areas acting as the main cause of metropolitan outbreaks.

Whenever a virus outbreak occurs, three questions are inevitably raised. First, which counteractive measures can be used to stop the virus from spreading? Second, is there a vaccine to prevent illness? Third, are there any antivirals that can be used to treat patients? Often, no vaccines or antivirals exist that would allow for rapid and successful outbreak intervention. Developing vaccines and antivirals, however, requires extensive knowledge about the viruses, especially the key steps of their infection and transmission cycles. Performing such studies can be arduous and time-consuming, hence, knowledge should be acquired before an outbreak begins. Here, we performed extensive analysis on the properties of the RNA replicase using several alphaviruses that together represent most of the known diversity of alphaviruses. We evaluated the ability of the alphavirus replicases to cross-utilize heterologous template RNAs and analyzed which determinants are required by the alphavirus replicases to successfully replicate the

template RNAs. In addition, we reconstructed the alphavirus replicases from two functional components and performed an in-depth analysis on the compatibility of heterologous combinations of these components. We assessed the possibility of using template RNAs as biosensors for the detection of alphavirus infections in human cells. Finally, we participated in a study that resulted in acquiring the 3D structures of catalytic subunits of the RNA-dependent RNA polymerases of two medically important alphaviruses and performed functional analysis on these enzymes using structure-guided mutagenesis.

These specific studies are highly important for acquiring comprehensive information on the fundamental principles of RNA replication, while using alphaviruses from distinctively different origins. RNA replication is the center of existence of any virus with an RNA genome and directly or indirectly affects all other properties of the virus. The knowledge obtained here helps to understand the absolute requirements of alphavirus replicase formation and sheds light on how replicases recognize genomic RNAs. Identifying the special features typical of the individual alphaviruses accompanied by more general characteristics that apply to many, or maybe all pathogenic alphaviruses, can help in generating novel antiviral approaches as well as tools for the diagnosis and detection of multiple alphaviruses.

1. REVIEW OF LITERATURE

1.1. Alphaviruses

Alphaviruses are emerging human and animal pathogens. The genus *Alphavirus* is the sole member of the *Togaviridae* family and includes over 30 known virus species (1). Alphaviruses infecting birds and mammals are divided into seven antigenic complexes: Venezuelan equine encephalitis complex, Eastern equine encephalitis complex, Western equine encephalitis complex, Barmah Forest complex, Semliki Forest complex, Middelburg complex and Ndumu complex (1). Mostly spread via an arthropod vector, alphaviruses can infect diverse organisms, such as humans, birds, rodents, insects and fish. Infection by some alphaviruses is limited strictly to mosquitoes (insect-specific viruses, ISVs). The majority of known alphaviruses use mosquitoes from the *Aedes* and *Culex* families to spread between vertebrate hosts. The geographical distribution of alphaviruses is restricted by their vector preference and host range. However, factors such as climate change contribute to expanding the areas with suitable proliferation conditions for mosquito vectors. Additionally, increased travel and urbanization cause alphaviruses to emerge in new areas (2). CHIKV, for example, mainly spread via *Aedes aegypti* mosquitoes, was historically confined to regions such as sub-Saharan Africa and Southeast Asia. After 32 years of dormancy, CHIKV re-emerged in India in 2005 and gave rise to an outbreak in Southeast Asia (3). The global expansion of CHIKV began with the virus spreading into the southwestern Indian Ocean region in 2005, rapidly adapting to *Aedes albopictus* mosquitoes, and was followed by the introduction of CHIKV into Italy in 2006 via travelers from the Indian Ocean islands and India (4–6) (Figure 1). While the first cases of CHIKV were confirmed in France in 2010, the local transmission of the virus via *Aedes albopictus* mosquitoes was established in 2017 (7). Prior to 2013, CHIKV was constrained to Africa, Asia, Europe and the Indian and Pacific Oceans. However, in late 2013, local transmission of CHIKV was also confirmed in the Caribbean islands, which was followed by the virus spreading all across the Americas (8, 9) (Figure 1).

Based on geographical distribution and symptoms they cause, arbovirus members of the *Alphavirus* genus are divided into New World and Old World alphaviruses (10). New World alphaviruses, also called encephalitic alphaviruses, include Eastern equine encephalitis virus (EEEV), Western equine encephalitis virus (WEEV) and Venezuelan equine encephalitis virus (VEEV). New World alphaviruses are spread in North, Central and South America. The symptoms of New World alphavirus infection manifest in an encephalitic phenotype by causing various neurological diseases, such as encephalitis and meningitis (11).

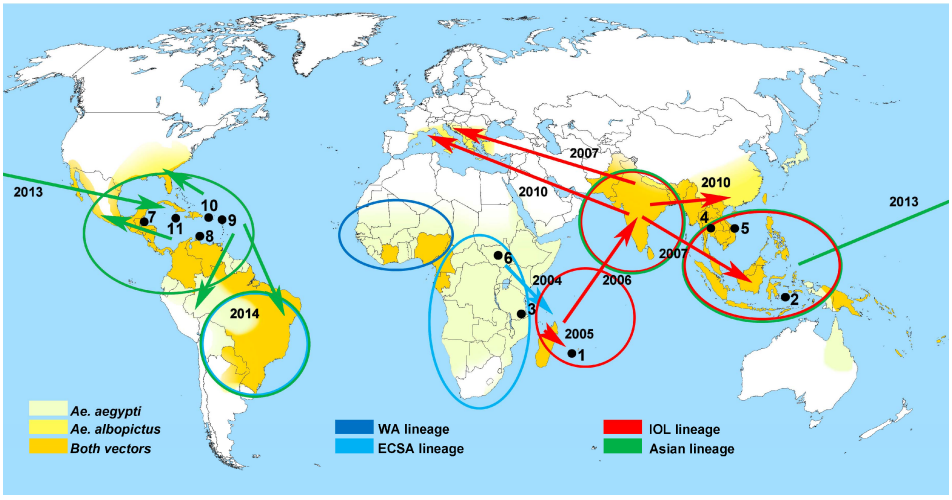

Figure 1. Schematic representation of the global expansion of CHIKV. Numbers 1–11 indicate several regions with endemic CHIKV and/or outbreaks in the order of their appearance. 1. Reunion; 2. Malay Archipelago; 3. Makonde/Tanzania; 4. Thailand; 5. Vietnam; 6. Southern Sudan; 7. Belize; 8. Curacao; 9. Saint Martin; 10. Puerto Rico; 11. Jamaica. Arrows and years indicate the global spread of CHIKV. Figure originally from (9).

EEEV, endemic to the northwestern region of the United States, is usually maintained in an enzootic cycle between birds and *Culiseta melanura* mosquitoes in forested swamp areas. However, abnormally humid and warm weather conditions may lead to an increased number of spillover events via bridge vectors to dead-end hosts, such as humans and horses (Figure 2). EEEV is subject to a growing concern in the United States, as an increased number of EEEV infections in humans have been reported in recent years. Most people infected with EEEV do not develop any symptoms or only present febrile illness defined by fever, myalgia and joint pain. In rare cases, however, EEEV infection can persist into neurological diseases with an approximately 30% fatality rate (12). VEEV is transmitted via *Culex* mosquitoes and remains in a constant enzootic cycle between the vectors and small mammals. One of the largest human outbreaks of VEEV originated in Colombia in 1962, affecting approximately 3,000 humans. The outbreak expanded into Venezuela, where it caused over 23,000 human cases, 960 of which included neurological symptoms (13). Additionally, in 1969, an outbreak that originated from El Salvador and Guatemala spread over Central America and Mexico and resulted in approximately 50,000 deaths in horses and approximately 52,000 infected humans (13).

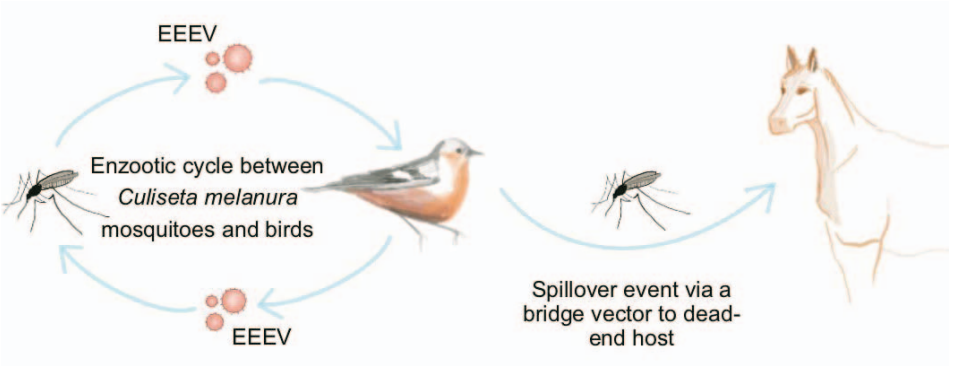

Figure 2. Schematic representation of a transmission cycle of EEEV. EEEV is amplified and maintained in an enzootic cycle between *Culiseta melanura* mosquitoes and birds in forested swamp areas. Unusually warm and humid weather may contribute to aberrancies in mosquito populations, which can lead to spillover events to dead-end hosts such as horses.

Old World alphaviruses, also called arthritogenic alphaviruses, are distributed in Eurasia, Australasia, Africa and the Indian and Pacific islands. The only exception is MAYV which is classified as an Old World alphavirus but causes outbreaks in the Americas (14). Upon infection, Old World alphaviruses can cause febrile illness that is characterized by fever, myalgia and joint pain, with the possibility of persisting into polyarthritis. Old World alphaviruses include, for example, CHIKV, MAYV, Semliki Forest virus (SFV), Barmah Forest virus (BFV), o'nyong'nyong virus (ONNV), Ross River virus (RRV) and Sindbis virus (SINV) (15). SINV, the type member of the genus *Alphavirus*, belongs to the Western equine encephalitis antigenic complex. SINV is spread by *Culex* mosquitoes and has been found in Europe, Africa, Asia and Australia. SINV infection in humans causes symptoms such as mild fever, rash and painful joints. Despite the wide distribution of SINV, the disease caused by its infection mostly occurs in Northern Europe, where SINV is reported to cause intermittent outbreaks (16). Clinical cases have also been reported in South Africa. Currently, SINV is the only alphavirus also found in Estonia, where the disease caused by the virus is called Karelian fever. The same disease in neighboring countries has different names: Ockelbo disease in Sweden and Pogosta disease in Finland. BFV and RRV are among the most common arboviruses in Australia (17). BFV is the sole member of the Barmah Forest antigenic complex. The main antigenic complex that includes important arthritogenic alphaviruses is the Semliki Forest antigenic complex, including CHIKV, SFV, ONNV, RRV and MAYV.

Currently, CHIKV is considered to be the most medically important alphavirus. Over 100,000 CHIKV cases worldwide have already been recorded this year alone as of March 2023 (18). CHIKV infection is almost always symptomatic, although rarely fatal. Symptoms of CHIKV infection include fever, chills, rash, muscle pain, fatigue, nausea and debilitating joint pain, which can be prolonged for months or even years (19). CHIKV is usually maintained in an enzootic cycle between mosquitoes and nonhuman primates. However, similar to other

arboviruses originating from enzootic cycles, it can be transmitted to humans via spillover events (Figure 3). These spillovers can happen accidentally when a mosquito vector, after biting an infected host followed by a period required for virus dissemination (i.e., viral replication in the midgut of the vector, infection of salivary glands and the release of progeny virions into saliva), feeds on a human. While these events can trigger outbreaks, they are not the main reason for massive outbreaks of CHIKV infection. CHIKV, similar to some other arboviruses, can also be amplified in humans and spread by certain mosquito species, such as *Aedes aegypti* and *Ae. albopictus*, to other people, bypassing the enzootic cycle entirely and resulting in so-called urban epidemic transmission (Figure 3) (20). This is possible due to the high potential of CHIKV to use humans as primary amplification hosts (i.e., to cause high-titer viremia) and the ability to infect the above-mentioned anthropophilic mosquitos. The ramifications of the urban CHIKV cycles have been painfully clear during recent outbreaks in Reunion and the Caribbean islands. The constant increase in urbanization, globalization and areas suitable for vector proliferation are likely to contribute to more frequent occurrences of CHIKV urban epidemics in the future.

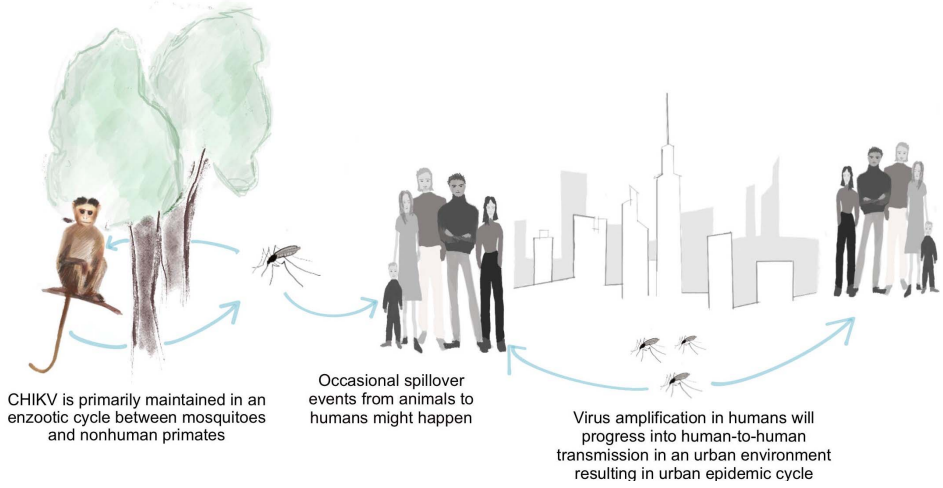

Figure 3. Different transmission cycles of CHIKV. CHIKV is primarily maintained in an enzootic cycle between mosquitoes and nonhuman primates. However, occasional spillovers to humans might happen. This can result in the amplification of the virus in humans with human–vector–human spreading potential. Interhuman transmission of CHIKV via *Ae. aegypti* or *Ae. albopictus* mosquitoes in an urban environment can progress into an urban epidemic cycle of the virus.

Despite being classified as an Old World alphavirus, SFV is neurovirulent in mice. Although it is known to infect humans, the infection is asymptomatic or results only in a mild febrile illness (21). These properties have facilitated the use of SFV as a model virus for studying alphavirus infection and as a tool for the development of alphavirus-based bio- and gene technology systems. MAYV is

endemic in South America and remains an important arthralgia-causing virus. MAYV cases are most likely underreported due to the symptomatic similarities with diseases caused by CHIKV and dengue virus (genus *Flavivirus*, family *Flaviviridae*). The main vectors for MAYV are *Haemagogus* mosquitoes, however, under experimental conditions, MAYV can also be transmitted by *Ae. aegypti* (22). Thus, the number of MAYV infections is likely to increase in the future with the globally widening distribution of the *Ae. aegypti* and *Ae. albopictus* vector mosquitoes (14). ONNV is the only known alphavirus that is transmitted via night-feeding *Anopheles* mosquitos, such as *Anopheles gambiae* and *Anopheles funestus*. It is suspected that the unique vector preference of ONNV is due to some specific properties of the nsP3 protein of the virus (23). Apart from the unique vector preferences, ONNV is very similar to CHIKV, and these two viruses form a monophyletic group inside the SFV clade.

1.1.1. The alphavirus virion

The alphavirus virion is approximately 70 nm in diameter and is arranged in a T=4 icosahedral symmetry. The virion comprises a nucleocapsid (NC) core, a lipid bilayer of host origin, and 80 trimeric spikes forming the outermost layer of the virion (24) (Figure 4). The NC core is formed of a single copy of genomic RNA assembled into the NC structure by 240 molecules of capsid protein (CP). Alphaviruses encode the glycoproteins E1 and E2, which form heterodimers that, in turn, are organized into trimers, also referred to as spikes. The lipid bilayer acquired during the budding of virions is penetrated by the transmembrane domains of each glycoprotein (10).

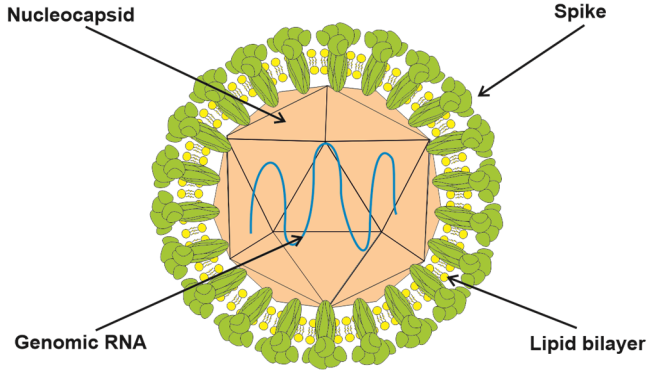

Figure 4. Schematic representation of the alphavirus virion. The genomic RNA (blue) is packed into a nucleocapsid (orange), which is surrounded by the lipid bilayer (yellow). E1-E2 heterotrimers are organized into spike structures on the surface of the virion (green).

1.1.2. Organization of the alphavirus genome

The genome of alphaviruses comprises a single positive-strand RNA molecule (Figure 5). The ssRNA genome has a cap-0 structure in the 5' terminus and a poly(A) tail in the 3' terminus. The alphavirus genome is approximately 12 kb in length and contains two open reading frames (ORFs, Figure 5) (10). The first ORF codes for precursor(s) of nonstructural proteins (nsPs), responsible for generating the alphavirus replication machinery. The second ORF codes for the precursor(s) of structural proteins, acting as building blocks for the assembly of new virions. ORF1 is translated directly from the virus genome. In contrast, ORF2 is translated from specific subgenomic (SG) RNA that is synthesized in infected cells.

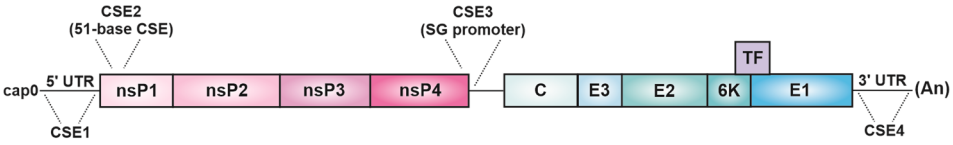

Figure 5. Schematic representation of the alphavirus genome. Explanations of the designations used are provided in the text. See also Figure 9.

The genome includes three noncoding regions: a 5' untranslated region (5' UTR) preceding the first ORF, an intergenic region between the two ORFs and a 3' UTR downstream of the second ORF. Noncoding regions of the alphavirus genome include important RNA regulatory elements (25). Two conserved sequence elements (CSEs) are present in the 5' region of the genome: CSE1 located in the 5' UTR and a 51-base-long CSE2 located in the region coding for the N-terminal region of nsP1 (26, 27). These elements include important *cis*-regulatory sequences required for positive- and negative-strand RNA synthesis (28). Structural analysis of the RNA secondary structures in the 5' region of the CHIKV genomic RNA *via* Selective 2′ Hydroxyl Acylation analyzed by Primer Extension (SHAPE) method revealed seven stem–loop (SL) structures in the first 303 nucleotides of the CHIKV genome (29) (Figure 6). The first two, named SL3 and SL47 after the nucleotide from which the SL structure starts in the genomic RNA sequence, are part of the 5' UTR. Disrupting SL47 results in reduced viral replication in both human and mosquito cells. Five of the SL structures were mapped to the region encoding the beginning of nsP1 and were named SL85, SL102, SL165, SL194 and SL246. SL165 and SL194 were confirmed to be included in CSE2. However, SL47, SL85, SL102 and SL246 are not localized in the conserved region and have not been described before (29). Analyzing the SL structures mapped downstream of the start of nsP1 (AUG, residues 77–79 of the CHIKV genome) revealed that SL85, SL102, SL165 and SL194 carry an important host-specific role during virus replication, as they are essential for replication in human cells but do not have this effect in mosquito cells. In contrast, mutations in SL246 significantly inhibited virus replication in mosquito cells but not in human cells (29).

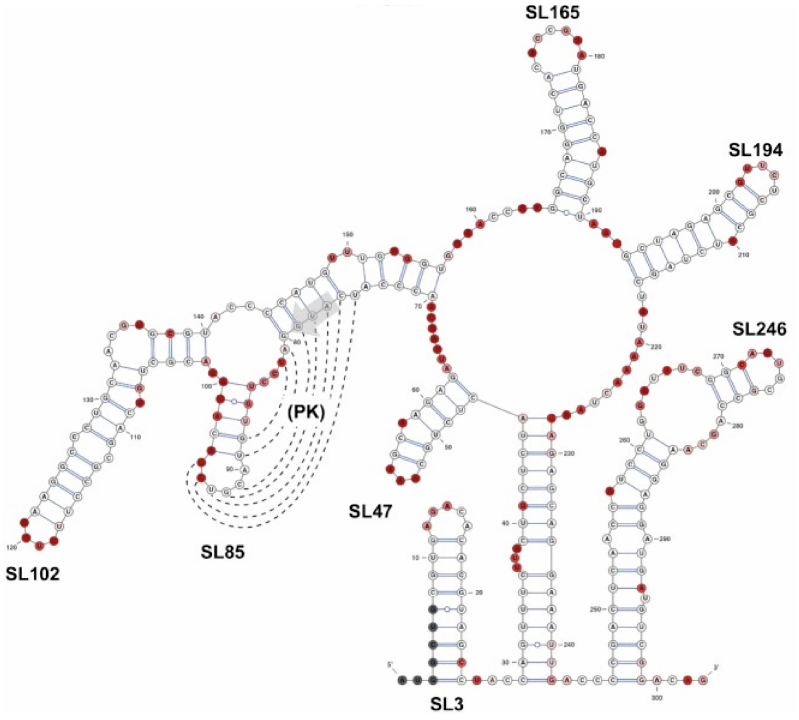

Figure 6. Structure of the 5' end of the CHIKV genome. SL – stem–loop structure; SL numbers indicate from which nucleotide the SL structure starts in the genomic RNA sequence. (PK) designates the potential pseudoknot structure. Figure is adapted from (29).

To continue, a 21-base-long CSE3 overlaps with the sequence encoding a few C-terminal amino acid residues of nsP4. This region contains sequences crucial for the synthesis of SG RNA, as the start site of SG RNA is located between residues 19 and 20 of CSE3 (30). Thus, CSE3 is a part of the SG promoter of alphaviruses. However, the functional SG promoter is somewhat longer than CSE3. For SINV, the minimal required length for the SG promoter is 19 residues located upstream and 5 residues located downstream from the start of the SG RNA. More precisely, the SG RNA of SINV can be synthesized when the SG promoter spans either from position –19 to +6 or from position –20 to +5 with respect to the start of the SG RNA. Shorter sequences spanning from –19 to +4 or from –17 to +6 were unable to trigger SG RNA synthesis (31). An SG promoter with minimal length is relatively weak, as its efficiency is approximately 3- to 6-fold lower than that of the SG promoter spanning from positions –98 to +14 (32). Thus, contrary to a rather common statement that the SG promoter is in the intergenic region of the alphavirus genome, most of the minimal SG promoter sequence as well as the sequences needed for its full activity are located in the region encoding the C-terminus of nsP4.

CSE4 is located in the 3' UTR immediately upstream of the poly(A) tail (33). It is 19 bases long, and together with a minimum of 11 or 12 A bases, it forms the

core promoter for negative-strand RNA synthesis (34). In contrast to CSE1, which roughly corresponds to the entire 5' UTR, CSE4 is much shorter than the 3' UTR. The nonconserved region of the alphavirus 3' UTR is nearly 1 kb in length and tends to contain repeated sequence elements and duplications. Using CHIKV infectious clones, it has been shown that introducing variations into this part of the 3' UTR results in host-specific effects. If the repeated sequence elements found in the 3' UTR are duplicated, virus replication is enhanced in mosquito cells. Conversely, when the 3' UTR is truncated, replication is inhibited. Such variation has no effects on virus replication in mammalian cells and does not affect CHIKV-induced pathology (footpad swelling) in mouse models (35).

1.2. The infection cycle of alphaviruses

The majority of alphaviruses are arboviruses infecting both insect vectors and vertebrate hosts. Hence, to perform replication, an alphavirus needs to be able to use mechanisms existing in both of these hosts. As insects and vertebrates are very different, it is not unexpected that some steps of alphavirus infection in vertebrates and insects are different as well. Furthermore, the goals of the virus are also different. In vertebrate cells, the virus needs to use robust replication, which in *in vivo* conditions is a prerequisite for rapid development of high-titer viremia. Robust replication is also required in mosquitoes, however, in this case, the damage caused by the virus must be minimal, as infected mosquitoes need to be able to fly and feed. Most studies on the alphavirus replication cycle have been performed using vertebrate cells, and the information provided below mostly originates from these studies.

1.2.1. Entry: the attachment and internalization of alphavirus virions

The infection cycle of alphaviruses begins with virus particle binding to the host cell via molecules present on the cellular surface (Figure 7). Binding of the cells is mediated via glycoprotein spikes – trimeric E1 and E2 heterodimers – on the surface of the virions (36). The ectodomain of the E2 glycoprotein is responsible for binding to the host cell surface (37). E1 mediates the fusion of cellular and viral membranes, however, evidence has suggested that E1 is also involved in host cell binding (37–40). The entry of alphaviruses occurs via clathrin-mediated endocytosis, which is followed by the fusion of viral and host membranes triggered by low pH (41). Exposure to low pH causes the E2/E1 heterodimers to disassociate, which leads to the formation of E1 homotrimers (40, 42, 43). These conformational changes allow the nucleocapsid to enter the cytoplasm followed by its disintegration and the release of genomic RNA.

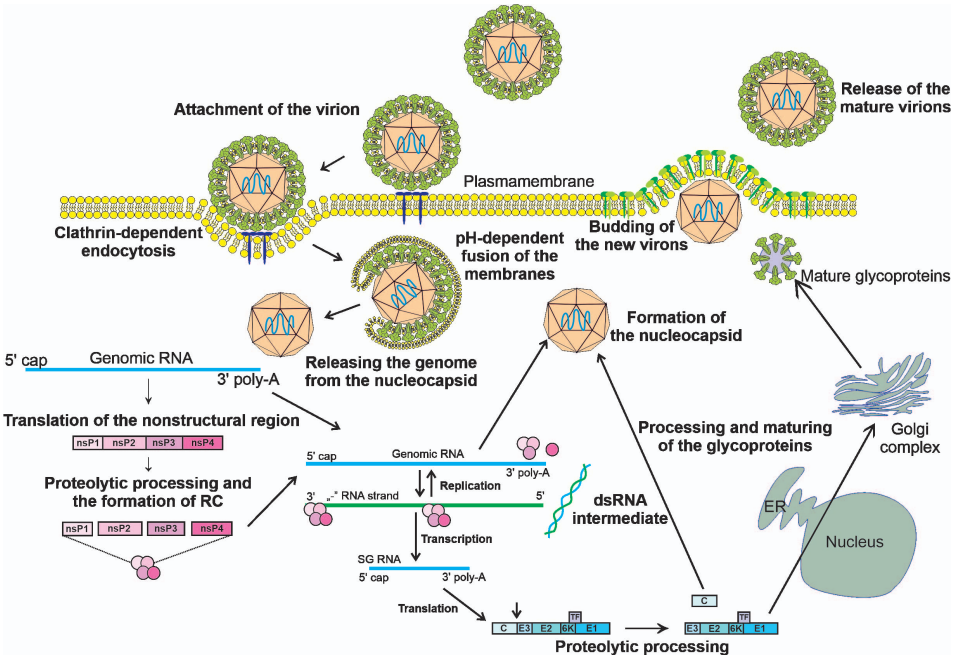

Figure 7. Infection cycle of the alphaviruses. Explanations are provided in the text.

The attachment of alphavirus virions to host cells occurs by binding various molecules on the cell surface. Part of these molecules, here referred to as the attachment factors, are relatively unspecific and include, for example, heparan sulfate, C-type lectins and phosphatidylserines (44, 45). Heparan sulfates have been shown to increase the replication of EEEV in the brain and promote neurovirulence in mice. Moreover, disrupting heparan sulfate binding by the E2 protein of EEEV decreases replication in lymphoid tissue and reduces the signs of febrile illness (46). Heparan sulfates also promote neurovirulence of arthritogenic SINV in adult mice (47). C-type lectins have been exploited by alphaviruses, as they usually act as pattern recognition receptors (PRRs) as part of cellular defense mechanisms. In, for example SINV, transfecting otherwise nonpermissive cells with genes expressing C-type lectins allows for the infection of those cells with SINV (48). T-cell immunoglobulin mucin (TIM) domain family proteins bind phosphatidylserines, and for various alphaviruses (and other enveloped viruses), phosphatidylserines have been shown to mediate the infection of cells that express TIM (49).

The attachment factors alone are not enough to promote entry of the viral particles into host cells. For this purpose, cellular receptors that interact with the E2 glycoprotein in a highly specific manner are critical (44, 45). For a long time, the specific receptors used by the alphaviruses remained unknown, and only recently a number of these receptors have been identified. Natural resistance-associated macrophage protein (NRAMP) is a divalent metal ion transporter that has been identified as a receptor for SINV in *Drosophila* cells (dNRAMP) (50).

NRAMP directly influences the entry of SINV into host cells, as viral RNA transfected into *Drosophila* cells lacking dNRAMP is still able to promote successful infection. NRAMP2, a vertebrate homolog of dNRAMP, mediates the binding and entry of SINV into mammalian cells, and NRAMP2-deficient mouse cells are nonresponsive to SINV infection (50). To continue, laminin receptors have also been proposed as receptors for SINV infection (51). In addition to SINV, laminin receptors are also thought to act as receptors for VEEV, as they appear to enhance VEEV infection in both mosquito and human cells (52). Recently, however, a specific receptor for VEEV, low density lipoprotein receptor class A domain containing 3 (LDLRAD3), was discovered using a genome-wide CRISPR/Cas9-based screen (53). In murine cells, the deletion of LDLRAD3 results in reduced infection by VEEV but not infection by EEEV or arthritogenic alphaviruses. Additionally, VEEV infection of human tumor cell lines lacking LDLRAD3 was shown to be unsuccessful, and ectopic expression of the receptor resulted in a gain-of-function phenotype (53). Recently, a very low density lipoprotein receptor (VLDLR) was shown to act as a receptor for SFV. Somewhat surprisingly, the same molecule also acts as a receptor for EEEV and contributes to the binding of SINV virions. The ligand-binding domains of VLDLR (and those of closely related apolipoprotein E receptor 2) bind the dimers of E2 and E1 glycoproteins of these viruses. This interaction is important, as ectopic expression of these proteins facilitates cellular attachment and internalization of virus-like particles (VLPs) of these viruses. It was also demonstrated that the receptor orthologs from mosquitoes can serve as functional alphavirus receptors, indicating that the entry of these viruses into the cells of the vectors is similar to that of the vertebrate cells (54). Interestingly, a more recent study demonstrated, that VLDLR binds specifically the E1, more commonly known as the mediator of the fusion of viral and cellular membranes, instead of the E2 (55).

To date, the most extensively studied receptor of alphaviruses is one that mediates the binding of CHIKV, ONNV, MAYV and RRV. Originally, prohibitin 1 (PHB1) was suggested as a possible receptor for CHIKV, as silencing of the PHB1 gene results in a small decrease in CHIKV infection, and PHB1 colocalizes with the CHIKV E2 glycoprotein at the plasma membrane (56). However, genome-wide CRISPR/Cas9-based screens identified Mxra8 as the main receptor for CHIKV, ONNV, MAYV and RRV (57). In NIH 3T3 and MEF cells, the knockout of Mxra8 either abolishes or reduces CHIKV infection, respectively. *Trans*-complementation of Mxra8 expression restores CHIKV infection in NIH 3T3 cells. The impact of Mxra8 on the replication of CHIKV was also studied by transfecting viral RNA into Mxra8 knockout cells (thus bypassing the attachment and internalization steps). In these experiments, no negative effects on replication were observed, suggesting that Mxra8 is important only for the binding and entry of viral particles (57). Mutant mice expressing truncated and soluble variants of Mxra8, which cannot be expressed on the cell surface, showed reduced susceptibility to CHIKV, MAYV, RRV and ONNV infection. Conversely, a recombinant variant of CHIKV unable to bind Mxra8 was shown to be attenuated in mice (58). The crystal structure of Mxra8 has been resolved at 2.2 Å, and structures have

been built for Mxra8 bound to the CHIKV VLP at a 4–5 Å resolution. Mxra8 binds the E2/E1 dimers in a 1:1 ratio, with the total number of Mxra8 receptors connected to the virion being up to 240. Interestingly, no significant conformational changes were observed in the structure of Mxra8-bound CHIKV VLPs compared to the unbound control VLPs (59).

1.2.2. Formation of the alphavirus RNA replicase and RNA replication

The first ORF of the alphavirus genome encodes the nsP1-4 proteins, which are responsible for the replication of the viral genome. Polyproteins P123 and P1234, precursors of these proteins, are translated directly from genomic RNA. The reason behind the translation of two polyproteins lies in an opal (UGA) stop codon in the region coding for the C-terminal end of nsP3 (60). As a result, P1234 is translated only in the case of a readthrough event, which has been shown to occur with an efficacy of 10–20% (61, 62). The opal codon is present in the genomes of the majority of alphaviruses. However, in some strains of SFV, ONNV and CHIKV, this opal codon is substituted by a sense codon (10, 63, 64). A study was carried out in pursuit of learning about the effects of replacing the UGA with an Arg codon in CHIKV. Although the *in vitro* replication of the virus was not affected by this switch, CHIKV-induced arthritis in mice was significantly reduced (65). To continue, for ONNV, the replacement of the Arg codon with an opal codon increased the fitness of the virus for infecting *Anopheles gambiae* mosquitos but reduced its infectivity in BHK21 and C6/36 *Aedes albopictus* (which is not a vector for ONNV) cells (66). These data suggest that the sense codons in the place of the opal stop codon might have a host-specific role during alphavirus infection.

P123 and P1234 have several important biological activities, however, they alone are not capable of replicating viral RNA (67, 68). To acquire this ability, P1234 needs to be processed in a strictly regulated and perfectly timed manner first into processing intermediates and finally into mature nsPs (69) (Figure 8).

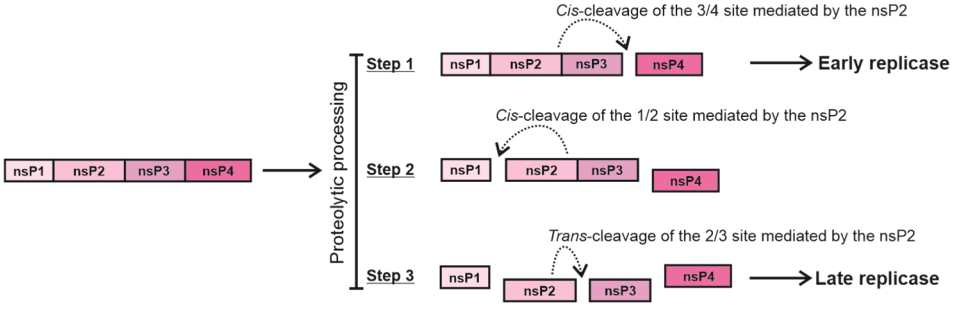

Figure 8. Proteolytic processing of P1234. Explanations are in the text.

The processing is performed by the nsP2 region of P1234 and begins with the release of nsP4 *via in cis* cleavage of the processing site between nsP3 and nsP4 (62, 70). This results in the formation of an early replicase composed of P123 and nsP4 and is responsible for synthesizing negative-strand RNAs (62, 71). The early replicase uses the genomic RNA as a template for synthesizing negative-strand RNAs that remain in a duplex with the genomic RNA forming double stranded RNA (dsRNA) replication intermediates. It is believed that each early replicase complex carries out a single synthesis event. This coincides with the formation of a specific structure called a spherule. This spherule contains a single dsRNA molecule and is attached to the plasma membrane of the cell. In alphavirus-infected cells, the number of spherules greatly exceeds the number of incoming virus genomes. The formation of additional spherules therefore requires positive-strand RNA synthesis and their conversion into dsRNA molecules using the above-described pathway (72). The synthesis of negative strands occurs only during the first 3–4 h of infection and is not detected in later stages (73). Thus, early in infection, most positive-strand genomes originating from incoming virions or synthesized by replicase complexes enter RNA replication. Requirements for negative-strand RNA synthesis are complex (Figure 9). It has been shown that the deletions of one or several SL structures located in the 5' region of the genomic RNA of SINV affect the synthesis of negative-strand RNAs (28). In addition, experiments conducted with SFV/SINV chimeric RNA templates and their respective replicases demonstrated that the 5' UTR of the alphavirus genomic RNA (CSE1) is an essential element in the promoters required for the initiation of both negative- and positive-strand RNA synthesis. Furthermore, it was postulated that a component of the replication machinery binds to the 5' end of the positive-strand RNA prior to initiating negative-strand RNA synthesis at the 3' end (28). In another study, the requirement of CSE4 for SINV negative-strand RNA synthesis was confirmed: the deletion of CSE4 completely abolished negative-strand RNA synthesis (34). The length of the poly(A) tail is also important for the synthesis of negative-strand RNA (34): a SINV template RNA with a poly(A) tail spanning in length from 25 to 34 adenylate residues can efficiently be used for negative-strand RNA synthesis. However, the synthesis of negative strands from a template with a poly(A) tail shorter than 20 residues is significantly reduced. The minimum requirement for the successful initiation of negative-strand RNA synthesis was found to be 11–12 A-residues immediately following CSE4 (34). To continue, the production of SINV negative-strand RNAs is severely hampered in the absence or replacement of a C residue at the –1 position (relative to the poly(A)) at the 3' end of the genomic RNA, suggesting the presence of a negative-strand RNA initiation site (74). Coupled with the fact that the poly(A) tail in alphavirus dsRNA is unpaired, alphavirus replicase is indicated to initiate the synthesis of negative-strand RNAs at the last non-A residue in the 3' end of the virus genome, and contrary to some early reports, the alphavirus negative strands lack the poly(U) sequence complementary to the poly(A) tail (75). It has been debated which nsP(s) recognize and bind genomic positive-strand RNA prior to the initiation of negative-strand RNA synthesis. A study by Thal and

colleagues confirmed that nsP4 binds CSE4 (more precisely, the A–U rich region within the CSE) when initiating negative-strand synthesis (76). However, as the promoter of negative-strand RNA synthesis also includes the 5' UTR, it is unclear whether nsP4 also recognizes this element. The importance of both ends of the genomic RNA has also been demonstrated in the example of SFV, where the requirements for negative-strand RNA synthesis were determined to be the 3' UTR of the genomic RNA and the 5' region of the genome, including the beginning of the nsP1 encoding sequence (77). These results suggest that both UTRs are required for negative-strand RNA synthesis. Furthermore, the UTRs most likely interact with each other to successfully initiate the production of negative-strand RNAs (28, 34, 76, 77). Additionally, studies have suggested that the interaction of nsP1 and nsP4 is also required for negative-strand RNA synthesis. In SINV, a suppressor mutation nsP1-T349K was found to induce the replication and production of new virions for viruses where the N-terminus of nsP4 contained a nonaromatic residue (see 1.3.4). Interestingly, this mutation was positioned immediately adjacent to the position of a temperature-sensitive nsP1-A348T mutation, which causes a defect in negative-strand RNA synthesis (78–80). These data emphasize the complex nature of negative-strand RNA synthesis and the existence of multiple important interactions between the viral RNA and replicase proteins of alphaviruses (and/or their precursors).

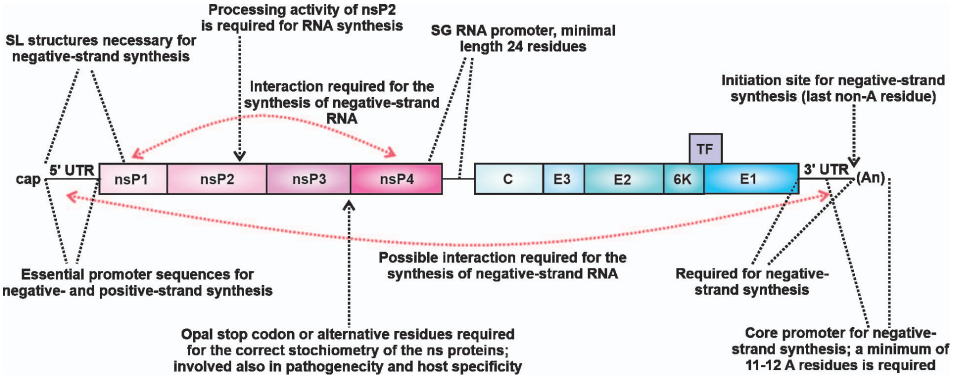

Figure 9. *Cis*-elements and some critical activities required for formation of alphavirus replicase and for RNA synthesis. Red dotted lines indicate possible interactions. Enzymatic activities other than the protease activity of nsP2 are not shown.

P123 is stable enough to allow the formation of an early replicase, binding of the template RNA and initiation of negative-strand RNA synthesis. It is not clear how P123 is stabilized and what signal triggers the following processing events. However, the processing of the polyprotein continues with another *cis* cleavage: nsP1 is cleaved from P123, and a nsP1 + P23 + nsP4 complex, sometimes called an intermediate replicase, is formed. This replicase can synthesize both positive- and negative-strand RNAs, however, it is extremely short-lived and therefore unlikely capable of synthesizing any detectable amount of viral RNAs (71). Instead, P23 is rapidly cleaved into nsP2 and nsP3 proteins. This is a *trans*

cleavage (Figure 8) and switches the replication complex from negative- to positive-sense RNA synthesis (71, 81). The replicase comprised of mature ns proteins is referred to as the late replicase. This complex is stable and capable of synthesizing multiple positive-strand RNAs. Full-length negative-strand RNA, almost certainly in the form of dsRNA, acts as a template for the synthesis of both genomic and SG RNAs (10). The latter occurs via the use of an internal SG promoter, as described in 1.1.2. The promoter of SG RNA synthesis interacts directly with nsP4, however, the presence of all nsPs is required for SG RNA synthesis to occur (82). nsP4 additionally binds the genomic promoter located at the 3' end of the negative strand. Interestingly, different regions of nsP4 are involved in the binding of genomic and SG promoters (83).

Evidence suggests that the pathway described above is the sole pathway used for the formation of functional replicase complexes. For example, expressing ns proteins individually results in a lack of RNA replication, suggesting that polyproteins and their processing intermediates are crucial for effective viral RNA replication (84). Cleavage-deficient mutants of SINV have been constructed to study the impact of polyprotein processing on RNA replication. Using these tools, an unprocessed P123 was shown to function as part of the RNA replicase and, along with P34 (which is, based on contemporary knowledge, processed into nsP3 and nsP4), is able to synthesize both positive- and negative-strand RNAs (85). In addition, if the cleavage of nsP4 from P1234 is entirely blocked, the synthesis of both positive- and negative-strand RNAs is abolished. This suggests that cleavage of the 3/4 site, resulting in the release of nsP4, is the absolute requirement for RNA replication (85). The speed of processing the polyproteins plays an additional role in RNA synthesis. Thus, an Asn614 to Asp substitution in the nsP2 of SINV has been shown to enhance the proteolytic processing of the P123 polyprotein in a way that no P123 can be detected in an *in vitro* assay. Originally, this mutation was reported to be completely lethal (86). However, more sensitive analysis performed later revealed that the mutant virus is viable, although severely attenuated (69). Despite the discrepancy, both studies confirmed that to support replication, P123 needs to be relatively stable. In another study, the Asn614 to Asp substitution was introduced into an uncleavable P123 construct (81). Predictably, the mutation had a negative effect on the cleavable wild-type (wt) P123 construct but showed no impact on the ability of the uncleavable P123 to participate in the synthesis of both positive- and negative-strand RNAs (81). These findings suggest that the processing of ns polyproteins is highly regulated, and the speed of processing plays an important role in the assembly of active replication complexes. The same is true for the order of the processing events: if P1234 is first cleaved at the 2/3 site, the resulting nsPs do not form a functional replicase complex. This so-called late processing pathway is activated by the accumulation of free nsP2 in infected cells and is involved in the cessation of negative-strand RNA synthesis. This most likely also plays a role in the establishment of superinfection exclusion, a phenomenon where the cell infected by an alphavirus rapidly becomes refractory to infection with the same or similar alphavirus (87, 88). To conclude, the RNA replication of alphaviruses is a highly

regulated and precise process that requires the presence of genomic RNA, polyproteins, processing intermediates and mature ns proteins (70, 71, 81, 84, 85).

The RCs of alphaviruses are assembled on the plasma membrane of infected cells. Of all alphavirus-encoded replicase proteins, only the nsP1 protein has an affinity toward membranes. Thus, it is plausible that the cleavage of the 1/2 site is delayed to allow the transport of nsP2 and nsP3 to the plasma membrane, where they can be incorporated into the RCs. If so, speeding up the processing can result in the premature release of free nsP2 and nsP3 that are not able to reach the plasma membrane and therefore be used for RC formation.

The replication of alphaviruses takes place in membranous invaginations called spherules (89–91). These structures are located in large cytoplasmic vacuoles that were first observed in WEEV- and SFV-infected chick embryo cells and were designated type-1 cytopathic vacuoles (CPV-I) (89, 91, 92). All four nsPs along with the newly synthesized viral RNAs are associated with CPV-Is (90). As mentioned above, the spherule structures that host the RCs of alphaviruses are first formed on the plasma membrane (72, 93). The nsP1 of alphaviruses possesses high affinity toward negatively charged phospholipids, an abundant component of the inner leaflet of the plasma membrane (94). The membrane affinity of nsP1 and the formation of spherules on the plasma membrane are common for all alphaviruses analyzed thus far. However, there are clear differences in the subsequent events: for some alphaviruses, such as CHIKV, the spherules stay on the plasma membrane, while for others, such as SFV, they become internalized (95). The significance of spherule internalization is not obvious, as inhibitors of this process do not affect the virus replication efficiency in cell culture (72, 93). Nevertheless, it is a prominent effect and relatively well studied. Time course experiments have shown that the RCs of SFV, which are formed at the plasma membrane, are later scattered in small cytoplasmic vesicles and finally localized in large perinuclear CPV-Is (93). The internalization of RCs is dependent on class I PI3Ks (plasma membrane associated lipid kinases), while the trafficking of the spherules is dependent on the cellular microtubule network. As mentioned above, RCs cannot be formed when nsPs are expressed as separate proteins (67, 81). The minimal requirement for the formation of spherule-like structures is the expression of nsP4 together with either P123 or together with nsP1 and an uncleavable P23. For efficient spherule formation, the presence of a replication-competent template RNA is also needed. In the presence of such RNA, nsP1, P23 and nsP4 can form active RCs that are able to synthesize negative- and positive-strand RNA (96).

While it has been clear that the spherules contain the RNA replication complexes, it has been unknown where exactly in the spherule the complex is positioned and how it is organized. Very recently, the molecular structure of the CHIKV RC was determined at a high resolution (97, 98) (Figure 10). Cryo-electron tomography confirmed that nsP1 acts as the base for the RC, anchoring the entire protein complex to the membrane, and that the RCs contain a single copy of the viral genome in the form of dsRNA (97). In another study, the core of CHIKV RC was reconstituted from individually expressed recombinant

proteins (98). As these studies are not only relevant for this thesis but also directly connected to Publications II and IV, these findings are discussed in detail in the Results and Discussion section. Briefly, a study by Tan and colleagues revealed that in infected cells, RCs are located on the cytosolic side of the neck of the spherule and consist of two structures: the membrane-bound replicase core and the amorphous cytosolic ring structure. The replicase core consists of nsP1, nsP2 and nsP4: 12 subunits of nsP1 form a dodecameric ring structure with nsP4 docking into the central pore and nsP2 extending above nsP4 toward the cytoplasm from the ring structure (98).

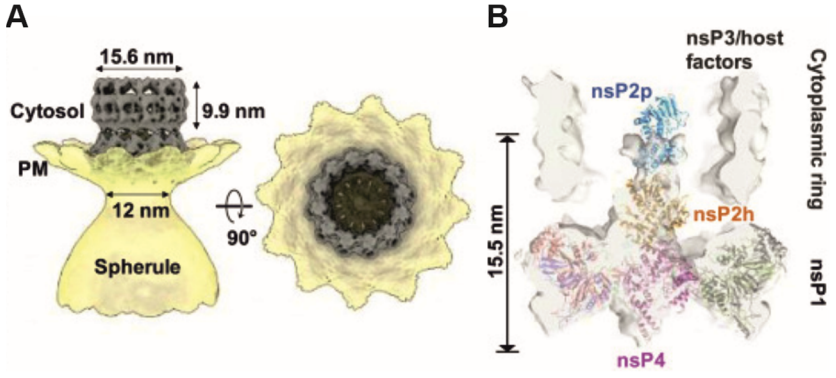

Figure 10. A. Subtomogram of the CHIKV spherule; B. The RC core complex (nsP1 + 2 + 4) fitted into the subtomogram. Figure is adapted from (98).

1.2.3. Expression of the structural proteins, virion formation and release

Following the synthesis of SG RNA, a structural CP-E3-E2-6K-E1 polyprotein is translated and subjected to co- and posttranslational processing (10). First, the capsid protein (CP) is cleaved from the polyprotein and released into the cytoplasm (99). The CP contains a C-terminal chymotrypsin-like protease domain performing this cleavage. Resulting from the cleavage, the enzymatic activity of the protease domain of the CP is inactivated. In addition to the protease domain, the CP also contains a disordered N-terminal domain (NTD). In VEEV, the protease domain and the extreme N-terminal region of the NTD of the CP have been shown to interact with the viral RNA and that this interaction is required for nucleocapsid (NC) assembly (100, 101). During assembly, only full-length genomic positive-sense RNAs, but not SG RNAs, become incorporated into the NC. This suggests the existence of a signal that directs the encapsidation of the viral RNA. Indeed, for SINV, the encapsidation signal has been mapped to the region coding for nsP1 corresponding to the sequences that have been shown to interact directly with the CP (102). Insertion of the packaging signal into SG RNAs increases packaging of mutant SG RNAs into the NC, supporting the role of this element as the encapsidation signal (103, 104). However, the localization

of the packaging signal in nsP1 is not conserved among alphaviruses. Thus, for viruses from the Semliki Forest complex, the packaging signal has been shown to reside in nsP2. Furthermore, the packaging signal of viruses, such as SINV, VEEV, EEEV and WEEV, was discovered to be recognized by the CPs of heterologous viruses. For example, the CP of CHIKV can utilize the packaging signal of VEEV. However, the CP of VEEV is not able to use the packaging signal of CHIKV. These results indicate that the viruses of the Semliki Forest complex have evolved in a way that ended in replacing the packaging signal in nsP1 with that in nsP2 (105).

Following the release of the CP, the E3 region of the structural polyprotein becomes exposed. E3 contains a signal sequence for transporting the polyprotein into the ER. The processing of the remaining structural polyprotein continues in the ER, where it is processed into the p62 polypeptide, a precursor for E2 and E3, and the 6K peptide and E1 glycoprotein. In the host cell ER, the p62 polyprotein and E1 interact to form heterodimers, which are then matured and organized into trimeric spikes (106). The trimers of the p62-E1 heterodimers are subsequently transported to the Golgi apparatus for glycosylation and are finally transported to the plasma membrane. During this transport, p62 is processed in the *trans*-Golgi network by furin-like cellular proteases, resulting in E2 remaining in a heterodimer–trimer organization with E1 and E3 generally being released (10, 107). The association of E1 with p62 rather than mature E2 until it reaches the surface of the cell has been the subject of multiple studies. E1 protein mediates the fusion of viral and cellular membranes and this fusion potential has been shown to strongly depend on the cleavage of p62 into E2 and E3 (108). Furthermore, the fusogenic properties of E1 are activated when exposed to low pH when this protein is alone. The p62–E1 complexes have been found to possess stronger resistance to acidic environments than the E2–E1 complexes (109, 110). This suggests that E3 in p62 has a stabilizing effect on the activities of E1, preventing premature formation of the fusogenic conformation. On the plasma membrane, the NC cores interact with the spikes. The most important interaction, essential for virus budding, takes place between the NC and the cytoplasmic domain of E2 (111). Mutating sequences in the cytoplasmic domain of E2 results in impaired budding, and competitive peptides designed based on the sequence of this region inhibit virion formation (112, 113).

6K is a transmembrane ion channel protein (viroporin). It is also incorporated into virions in small amounts (114, 115). 6K colocalizes with E2 in the Golgi apparatus and facilitates the formation of type-II cytopathic vacuoles (CPV-II). 6K is dispensable for viral entry and replication but is important for the transport of viral glycoproteins to the plasma membrane (116) and viral particle budding (117, 118). However, even for budding, the functions of 6K are not critical. Thus, the partial deletion of the 6K-encoding region in SINV and the complete deletion of the 6K region in salmonid alphavirus result in abnormal processing and transportation of glycoproteins but do not prevent virion formation and release (119, 120). The same has been observed for RRV and CHIKV (121, 122). Historically, 6K was suggested to be present in two forms: 4K, a partially acylated form,

and 6K (114, 123). This, however, was overturned by the identification of a frameshift protein TF, which is produced via a –1 ribosomal frameshift occurring on a slippery consensus motif located in the coding sequences of 6K (124). This frameshift resulting in TF, rather than 6K and E1 proteins, is conserved among alphaviruses (125). The 6K and TF proteins possess an identical N-terminal transmembrane domain and a cysteine-rich cytoplasmic loop followed by their unique C-terminal ends (124). The infection efficiency of a TF deletion mutant of SFV was approximately 56% lower than that of wt SFV. SDS–PAGE analysis of virions indicated that TF, rather than 6K, is being incorporated into new virions (124). TF is thought to regulate the interferon response in murine models, as 6K-only mutants (lacking TF) of SINV and VEEV have an attenuated phenotype in mice (126–128). In addition, two TF palmitoylation mutants either lacking palmitoylation or being hyperpalmitoylated demonstrated that only palmitoylated TF is incorporated into SINV virions (129).

As described above, in infected mammalian cells, the NCs formed in the cytosol and spike proteins matured and organized in the ER and Golgi apparatus are assembled into virions at the plasma membrane. The release of virions occurs by budding, and during this process, the lipid bilayer around the NC is acquired. The lipid bilayer is penetrated by the transmembrane helix of each glycoprotein, while the intracellular tails of E2 proteins are inserted into the hydrophobic pockets of CP, generating vertical links connecting the NC and spikes (10, 130, 131).

1.3. Nonstructural proteins

1.3.1. nsP1

nsP1 is the alphaviral N7-guanine-methyltransferase (MTase) and guanylyltransferase (GTase) and thus possesses the activities that are required for adding cap-0 structures to viral positive-strand RNAs (132–134). Capping of the viral RNAs promotes the replication and translation of the genome. In addition, with the cap-0 structure in the 5’ terminus and the poly(A) sequence in the 3’ terminus, the viral genome mimics eukaryotic mRNAs, which is important in bypassing cellular defense mechanisms (10). Alphaviruses are capable of synthesizing cap-0 structures identical to cellular cap-0, albeit in an unconventional manner (133). First, γ-phosphate is hydrolyzed from nascent viral RNA, resulting in a diphosphate RNA (this reaction is catalyzed by nsP2, see 1.3.2). Second, a GTP is methylated at the N7 position using S-adenosyl-L-methionine as a donor. GTase then binds N7-methyl-GTP, releasing pyrophosphate and forming m7-GMP. Finally, the enzyme transfers m7-GMP to viral diphosphate RNA, generating a m7GpppN cap-0 structure (133).

nsP1 is the only ns protein of alphaviruses with strong membrane binding properties. nsP1 is also posttranslationally palmitoylated. The palmitoylation site consists of one (SINV) or three (SFV, CHIKV) Cys residues located in the C-terminal part of nsP1 (135, 136). Palmitoylation is thought to be important in

the membrane binding of nsP1. Thus, transiently expressed palmitoylation-defective nsP1 of SFV has a reduced ability to bind membranes. However, at least for nsP1 of SFV, the lack of palmitoylation does not abolish the enzymatic activities of the protein (136). The effects of palmitoylation in the context of different alphaviruses have been studied by mutating the palmitoylation sequences in the nsP1 of SFV, SINV and CHIKV. Somewhat surprisingly, these studies revealed significant differences between these viruses. Palmitoylation-negative SINV and SFV grew to high titers in BHK cells, however, the SFV mutant presented slightly slower growth (137). In addition, the localization of the nonpalmitoylated nsP1s of SFV and SINV was different compared to that of wt viruses: in addition to the characteristic localization to membranes, the mutant nsP1 proteins were also found to be distributed in the cytoplasm (137). A subsequent study, however, revealed that SFV harboring mutations in the palmitoylation site of nsP1 was, in fact, severely attenuated. The growth of this mutant, observed in a previous study, was due to the accumulation of second-site mutations. These mutations did not restore the palmitoylation of nsP1, instead, they restored the interaction of nsP1 with nsP4 that was disrupted by the mutations in the palmitoylation site (138). In contrast to SFV and SINV, introducing mutations into the palmitoylation site of nsP1 of CHIKV completely abolishes viral RNA replication and the infectivity of the mutant genomes (139, 140). This phenotype is likely because the nsP1 of CHIKV targets cholesterol-rich plasma membrane microdomains. This localization depends on palmitoylation and is crucial for the RNA replication of CHIKV (141).

Recently, the structure of an active nsP1 complex of CHIKV was resolved at a high resolution (142, 143). The nsP1 of CHIKV was shown to be inactive as a monomer, only reaching its active form once assembled into oligomers. Active nsP1 is assembled into a single dodecameric ring, approximately 19 nm in diameter and 7 nm in height. The ring structure has a 7 nm wide opening in the center and is divided into three areas: the crown, the waist and the membrane-binding skirt. The crown area includes 12 capping domains of nsP1. The top of the crown, approximately 14 nm wide, and the waist region, approximately 7 nm wide, are positively charged. The inner part of the pore has a neutral charge, potentially allowing RNA, nucleotides and small globular proteins, approximately 70–90 kDa in size, to pass through (Figure 11). The membrane-binding skirt is organized into spikes that are positively charged on the surface. The tips of the spikes remain hydrophobic (142, 143). nsP1 folds into two domains: the MTase/ GTase catalytic domain positioned at the crown region and the membrane association and oligomerization domain positioned at the waist and skirt regions. The availability of the high-resolution 3D structure of active nsP1 of CHIKV has allowed for detailed analysis of the structural bases and dynamics of alphavirus RNA capping. The RNA capping reaction was found to be reversible, indicating that nsP1 can also function as an RNA decapping enzyme. Analysis of the molecular basis of RNA recognition by nsP1 revealed that the enzyme has clear specificity for the RNA substrate: it strongly prefers RNAs harboring AU dinucleotides at their 5’ ends. This is a sequence present at the native 5’ ends of CHIKV positive-

strand RNAs, thus, nsP1 can distinguish between viral and nonviral and positive- and negative-strand RNAs. Consistently, it was observed that mutating the first or second residue of the CHIKV genome strongly reduces or even abolishes the replication of such mutant templates (144, 145).

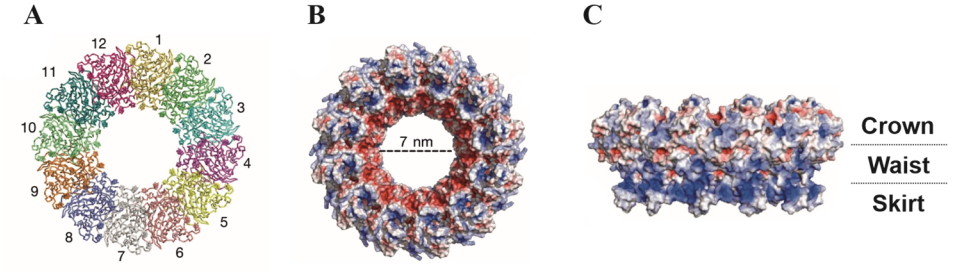

Figure 11. Cryo-EM structure of nsP1 of CHIKV. (A) Ribbon representation of the nsP1 complex formed from the 12 subunits viewed from the cytosolic side; (B) surface representation of the structure viewed from the cytosolic side; (C) lateral view of the nsP1 complex. Figure is adapted from (143).

nsP1 is important in the synthesis of negative-strand RNA, although its role in this process remains unclear. A temperature-sensitive (ts) mutant of SINV named ts11 contains a mutation in the nsP1 sequence. At nonpermissive temperatures, this mutant is unable to replicate because of the lack of negative-strand RNA synthesis. This phenotype is not caused by a defect in ns protein synthesis, suggesting that nsP1 has a direct function in negative-strand RNA synthesis (79). To continue, ts mutants of SFV (ts14 and ts10) containing mutations in nsP1 also have impaired production of negative-strand RNAs. The mutations present in ts14 and ts10 have also been shown to not affect other nsP1 activities, such as the enzymatic activities required for capping, membrane binding and interactions with other ns proteins (146, 147). These data suggest that these ts phenotypes are due to defects in the formation of RCs and/or their functioning. Interestingly, the analysis of ts10 and ts14 of SFV resulted in the hypothesis that nsP1 must interact with nsP4 and nsP2 (146). More than a decade later, this was confirmed by structural analysis. During the formation of functionally active RCs, the interactions between nsP1 and nsP4 are critical (98). It is possible that the interactions between nsP1 and nsP4 in the RCs of the ts mutants are conformationally divergent from those of wt viruses and thus do not permit the production of negative-strand RNAs. A conserved Arg183 residue in the nsP4 of SINV has been shown to be important in negative-strand RNA synthesis, as substitution with either Ser, Ala or Lys results in a ts phenotype characterized by defective production of negative-strand RNAs (148). However, the Arg183 to Ser mutant was found to recover, and the ability to synthesize negative-strand RNAs at the nonpermissive temperature was regained when an Asn374 to His or Ile mutation was introduced into nsP1 (149). This clearly indicates the presence of an interaction between these sites in nsP1 and nsP4. In addition, a mutation in nsP1 allows for the replication of SINV when containing a nonaromatic residue in place of a conserved Tyr at

the N-terminus of nsP4 (in the absence of compensatory mutations, these mutant viruses are defective, see Chapter 1.3.4.). Interestingly, this mutation (Thr349 to Lys) is located adjacent to a mutation (Ala348 to Thr) present in the ts11 mutant of SINV (79, 80). These findings further emphasize the importance of nsP1 in negative-strand RNA synthesis as well as the importance of interactions between nsP1 and nsP4.

1.3.2. nsP2

nsP2 is the largest ns protein of alphaviruses and possesses a wide array of important functions. The N-terminal part of nsP2 is known to contain motifs required for RNA helicase, NTPase and RNA triphosphatase activities, i.e., functions that are required for the RNA replication and capping of viral positive-strand RNAs (86, 150). The C-terminal part of nsP2 possesses proteolytic activity and is responsible for ns polyprotein processing (Figure 12) (70). The N-terminal part consists of three domains: the N-terminal domain (NTD) and two RecA-like RNA helicase superfamily 1 domains (151). A corresponding recombinant protein has been shown to possess two enzymatic activities: NTPase activity and RNA triphosphatase activity (150, 152). Although the N-terminal part of nsP2 is also clearly responsible for RNA helicase activity, this function can only be detected in the context of a full-length nsP2, indicating functional cross-talk between different parts of the protein (153). The C-terminal part of nsP2 folds into two domains: the N-terminal protease domain, which contains the catalytic core of the enzyme, and the C-terminal domain, which is similar to S-adenosyl-L-methionine methyltransferases (154). The protease active site of nsP2 of SINV contains Cys (C^{481}) and His (H^{558}) residues, and mutating these residues results in the loss of protease activity and consequently the loss of the ability of the replicase to produce viral RNAs (86). Similarly, corresponding mutations in SFV and CHIKV (in both cases $C^{478}A$) render the ns polyprotein unprocessed, confirming that nsP2 is the sole protease performing processing of the SFV polyprotein and that the protease activity of nsP2 of CHIKV is dependent on the catalytic Cys residue in the active site of the protease (68, 155). The N- and C-terminal parts of nsP2 are connected by a flexible linker that is crucial for the infectivity of alphaviruses (156).

nsP2 plays an important role in the replication of viral RNA. For instance, the helicase activity of the nsP2 that unwinds the RNA duplexes in the 5' to 3' direction is needed to access the template RNA to synthesize new strands. Thus, nsP2 acts coordinately with nsP4 (RdRp), which also has 5' to 3' activity. The helicase part of nsP2 has been shown to bind RNA by forming stacking interactions. Substituting the hydrophobic residues that mediate those interactions (Y161, F164 and F287) has a negative impact on the RNA replication and infectivity of CHIKV (151). In addition, nsP2 also demonstrates RNA annealing activity (153). Das and colleagues proposed a model in which nsP4 initiates the synthesis of either genomic or SG RNA, which is then capped quickly after the initiation of RNA synthesis. Capping requires nsP2 to be present, which positions nsP2 immediately following nsP4. This might create a situation where nsP4 synthesizes

a new RNA strain, and nsP4 is followed by nsP2, which separates the daughter and parental strains via its RNA helicase activity and then anneals the parental strains again into dsRNA intermediates (153).

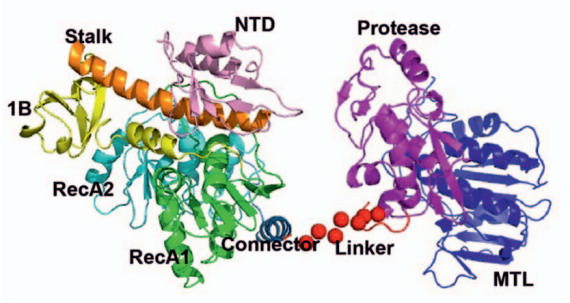

Figure 12. Structural overview of the domains of nsP2. NTD – N-terminal domain, MTL – methyltransferase-like domain. Figure is adapted from (156).

In addition to its multiple enzymatic activities and interactions with RNA, nsP2 is also a major factor involved in virus–host interactions, including suppression of cellular responses unfavorable for alphavirus infection. The nsP2 of Old World alphaviruses is important for downregulating cellular antiviral responses in vertebrate cells as nsP2 mediates transcriptional shut-down of the cell and induces cytopathic effects. In infected cells, nsP2 is found not only in the RCs but also diffusely in the cytoplasm and nucleus. Approximately half of the nsP2 found in alphavirus-infected cells is located in the nucleus (157). Certain mutations in nsP2 of SFV have been shown to keep the protein from moving from the cytoplasm into the nucleus while also reducing the cytotoxicity of the virus (158). Several mutations have been documented that result in a noncytotoxic phenotype and persistent SINV infection in mammalian cells (159, 160). Interestingly, those mutations localize at the interface of nsP2 and nsP3 of SINV (161). Unlike many other viral proteases that act as suppressors of cellular transcription, nsP2 of alphaviruses is not known to cleave any host transcription factors. In fact, the ability of nsP2 to suppress cellular transcription is independent of its protease activity. Instead, the nsP2s of several Old World alphaviruses have been shown to be capable of shutting down host cell mRNA transcription by degrading the catalytic subunit of DNA-dependent RNA polymerase (Rpb1) in vertebrate cells. Most likely, all Old World alphaviruses directly induce the degradation of Rpb1, while the nsP2s of New World alphaviruses, such as VEEV, do not give rise to such effects (162). The nsP2 of alphaviruses is also involved in fighting cellular defense mechanisms by regulating the interferon response in infected cells. For example, an SFV mutant in which nuclear translocation is blocked induces a significantly higher interferon response than wt SFV (163). The ability of nsP2 to interfere with the interferon response is also documented for CHIKV and, most likely, is a general property of the nsP2 of Old World alphaviruses (164).

1.3.3. nsP3

nsP3 remains the most enigmatic of all the ns proteins of alphaviruses. It consists of three domains, the first being the N-terminal macrodomain, also referred to as the X domain, which is conserved among the alphaviruses and also found in rubella virus, hepatitis E virus, coronaviruses and eukaryotic organisms (165). The second (central) domain of nsP3 is the zinc-binding domain that is unique to alphaviruses, for this reason, it is referred to as the alphavirus unique domain (AUD). The C-terminal domain is intrinsically disordered and highly variable both in length and sequence and is therefore called the hypervariable domain (HVD).

nsP3 is crucial for alphavirus replication and is thought to be a part of the RNA synthesis machinery. However, the recently resolved structure of the core of the alphavirus RNA replicase revealed that, unlike the other nsPs, nsP3 is not included in this complex (Figure 10). Instead, it is likely a major component of the cytosolic ring structure (98). The ring structures are characteristic of active RCs, confirming the role of nsP3 in alphavirus RNA replication. However, how exactly nsP3 functions is not obvious. In addition, nsP3 interacts with several host proteins that are also important for viral RNA synthesis (166).

Viral macrodomains are known to have both ADP-ribosylation and de-ADP-ribosylation activities (167). ADP-ribosylation is a posttranslational modification that is associated with gene regulation, protein expression, DNA damage repair and stress response (168, 169). ADP-ribosylation is catalyzed by poly(ADPr)-polymerases (PARPs), and it has been shown that the macrodomain of nsP3 shares structural similarities with a PARP domain (165). PARPs can be induced by interferons, suggesting that ADP-ribosylation has a function in cellular antiviral mechanisms. Indeed, PARP10 and PARP12 have been shown to inhibit the replication of CHIKV (170). Alphavirus nsP3 can bind free ADP-ribose (ADPr) as well as poly(ADPr)s (165). The nsP3 of CHIKV is also able to hydrolyze this modification from mono-ADP-ribosylated proteins, and mutants lacking this ability are unable to replicate in mammalian and mosquito cells (171). In addition, mutations in the ADPr binding region of nsP3 of SINV result in impaired viral replication in neurons (172). Accordingly, the effects caused by ADP-ribosylation have been suggested to be countered by the de-ADP-ribosylation activity of the viral macrodomains (173). Very recently, mono-ADP-ribosylation (MARylation) mediated by PARP10 was shown to reduce the processing of ns polyproteins. Reduced processing was also observed for mutant polyproteins harboring enzymatically inactive macrodomains. These findings led to the identification of nsP2 as a substrate for MARylation. More precisely, the MAR hydrolase activity of the macrodomain was found to remove MARylation from nsP2, reactivating its processing activities (170). These findings confirm that de-ADP-ribosylation is a counteractive mechanism of the alphaviruses required for combating the host cell's antiviral measures. Coupled with an older finding showing that the macrodomain is a crucial cofactor for the nsP2 needed for processing the bond between nsP2 and nsP3 (174), this new information places the macrodomain of nsP3 among the key factors required for alphavirus RC formation.

The AUD of nsP3 contains four highly conserved Cys residues involved in the binding of zinc ions. Site-directed mutagenesis revealed that each of these residues is crucial for SINV replication (161). Mutations in the AUD of CHIKV have presented both cell-type- and species-specific phenotypes, and it has been suggested that the AUD is important in alphavirus replication and transcription (175). However, it is still not known how AUD contributes to these activities.

The HVD of nsP3 is phosphorylated at Ser and Thr residues. This modification has been experimentally confirmed for the nsP3s of SINV, SFV, ONNV and CHIKV and is assumed to exist in the nsP3s of other alphaviruses as well (176–178). While the phenomenon of phosphorylation is conserved, its significance for different alphaviruses varies. For example, the substitution of all potential phosphorylation sites in the nsP3 of CHIKV results in the total abolition of viral RNA synthesis and infectivity (178). However, in the case of SFV, a similar modification has only a minute effect on virus replication in cell culture (179). Perhaps the most important function of the HVD of nsP3 is to act as a binding hub for cellular proteins (180). Multiple studies have shown that New World and Old World alphaviruses have evolved to interact with different host proteins. For example, the HVD of nsP3 of VEEV interacts with FXR protein family members, while the HVD of nsP3 of SINV and CHIKV binds G3BP protein family members (181). Interestingly, nsP3 of EEEV was shown to bind both FXR and G3BP proteins (182). FXR and G3BP protein family members are important for facilitating stress granule formation in cells. While counteracting stress granule formation is likely favorable for alphavirus infection, this is not the main reason why the interaction between the HVD and these proteins is needed (183, 184). A number of alphaviruses are unable to induce RNA replication in the absence of G3BPs or if their nsP3s contain mutations that prevent interaction with G3BPs (181). This suggests a more direct involvement of G3BPs in RNA replication, and indeed, in the absence of the G3BP/nsP3 interaction, the synthesis of alphavirus negative-strand RNAs has been shown to be abolished or attenuated (185). Interestingly, while the requirement for G3BPs (or FXR proteins) is conserved among alphaviruses, the impact of lacking these host factors on virus replication varies considerably. Thus, some Old World alphaviruses, such as SINV, are relatively resistant to G3BP deletion, while others (such as CHIKV) are very sensitive. Surprisingly, the high sensitivity is not caused by differences in the HVD of nsP3. Instead, this sensitivity is connected to the presence of an Arg residue in the P4 position (i.e., four residues upstream from the scissile bond) of the 1/2 cleavage site (185). As this residue is clearly involved in the regulation of ns polyprotein processing, this correlation again indicates a role of nsP3 in the activation, formation and/or functioning of alphavirus RCs (69, 186). In addition to the G3BPs and the FXR protein, the HVD interacts with a number of other host proteins, including NAP1, several SH3 domain-containing proteins (such as CD2AP) and FHL1. Knockout of the expression of any of these proteins typically has a rather limited impact on alphavirus replication. However, to achieve a detectable level of virus replication, the HVD needs to interact with at least one of these proteins (180).

1.3.4. nsP4

nsP4 is an RNA-dependent RNA polymerase (RdRp) and the most conserved ns protein of the alphaviruses. The core catalytic domain of nsP4 is responsible for *de novo* RNA synthesis and adding adenine residues to the 3' end of an acceptor RNA (187, 188). The nsP4 of alphaviruses is able to produce negative-strand RNAs using positive-strand RNA as a template and positive-strand RNAs using negative-strand RNA as a template even in the absence of other ns proteins (76). However, on its own, nsP4 is very inefficient and, especially in the form of a recombinant protein produced in bacteria, has very low processivity. Thus, the other nsPs are clearly required for reasonably efficient RNA synthesis. The interaction of nsP4 with other nsPs has been suggested to occur via its N-terminal domain, which is unique to alphaviruses. The C-terminal region of nsP4 is the RdRp domain and includes the highly conserved GDD (Gly-Asp-Asp) motif characteristic of viral RdRps (189).

nsP4s of all alphaviruses have a Tyr residue in their N-termini. This results from the processing of the 3/4 site in which Tyr occupies the P1' position (i.e., is located immediately downstream of the scissile bond). According to the N-end rule, Tyr is a destabilizing amino acid residue. Indeed, in rabbit reticulocyte lysates, nsP4 of SINV has been shown to be degraded by the ubiquitin-dependent pathway (190). Interestingly, the addition of a Met residue in the N-terminus of nsP4 results in a functionally compromised protein that, when expressed together with P123, is only capable of synthesizing negative-strand RNAs (84). Therefore, to study the requirements for the assembly of RCs, nsP4 fused with ubiquitin at its N-terminus was used (ubiquitin fused to the N-terminus of a protein is rapidly removed by cellular deubiquitinating enzymes). The ubi-nsP4 expression construct was shown to produce nsP4 with an authentic N-terminal Tyr residue and, when coexpressed with P123, was able to synthesize both positive- and negative-strand RNAs (81). The N-terminal Tyr residue is also necessary for the activity of nsP4 in the context of virus infection. Mutational analysis of SINV has revealed that Tyr can only be replaced with His or with an aromatic amino acid residue to preserve a near-wild-type growth and plaque phenotype (191). However, mutant viruses with nonaromatic N-terminal residues can function upon adding AU, AUA or AUU to the 5' terminus of the genome or when replacing the third nucleotide of the genome (A) with a U (192). In addition, second-site suppressor mutants mapped to nsP4 and nsP1 also allowed for the efficient replication and production of progeny virions (albeit in a temperature-sensitive manner), even if nsP4 starts with a nonaromatic residue (80). These findings suggest that the aromatic residue in the N-terminus of nsP4 has an important function. Furthermore, the instability of nsP4 is not a problem for alphaviruses. Instead, it is an essential property that is probably needed for removing the excess nsP4 and/or maintaining a close-to-optimal ratio of nsP4 to other ns proteins. This assumption has been supported by showing that bortezomib, an inhibitor of proteasomal degradation, acts as an inhibitor of CHIKV infection. In the cells treated with this compound, nsP4 is stabilized, and the levels of nsP4 (and, to an

extent, other ns proteins) are increased, resulting in the overproduction of viral RNAs. These effects are not, however, favorable for the virus, as both the expression of structural proteins and the formation/release of CHIKV virions are reduced upon bortezomib treatment (193).

Historically, numerous efforts have been made to analyze the activities of alphavirus nsP4s using recombinant proteins expressed in *E. coli*. However, obtaining soluble recombinant nsP4 protein with an authentic Tyr residue in the N-terminus has been extremely difficult due to the poor solubility of this protein as well as its tendency to aggregate. Despite this, Tomar and colleagues managed to express and purify the nsP4 of SINV via the removal of a portion of the N-terminus of the protein. The obtained protein was used to reveal the terminal adenylyl transferase activity of nsP4 (187). Rubach and colleagues were also able to express and purify the full-length nsP4 of SINV via its expression in the form of SUMO-nsP4 fusion protein. This full-length recombinant nsP4 of SINV was able to induce the synthesis of negative-strand RNAs, however, its activity was rather low (188). The recombinant nsP4 proteins of other alphaviruses have been even harder to obtain. Repeated efforts have been made in our laboratory to obtain soluble and enzymatically active nsP4s, which have consistently resulted in failures. To the best of our knowledge, the results (or the lack of them) obtained by other groups have been similar. The only alphavirus, aside from SINV, where some success has been obtained is CHIKV. Researchers in Singapore expressed and purified the core domain of the CHIKV nsP4 and demonstrated that in the presence of certain detergents, the recombinant protein had minimal (only slightly above background) primed template extension and terminal adenylyl transferase activities (194). Naturally, difficulties in producing soluble and active nsP4 have also hampered structural studies of this enzyme. However, finally, the structures of the nsP4 of SINV and RRV became available in 2022 (Publication IV). How this was achieved, what was discovered and how this information was subsequently used are described in the Results and Discussion section.

1.4. The alphavirus *trans*-replication systems and their use in the analysis of alphavirus RNA replication

Alphavirus-induced pathogenesis is a source of great concern around the globe, as there are no licensed antivirals available. One of the main targets for novel antivirals is RNA replication, a process unique to RNA viruses. However, to effectively counteract alphavirus RNA replication, thorough knowledge about the production, processing and properties of ns proteins, RC formation, requirements of template RNAs, the impact of structural elements in viral RNAs and the roles of host-derived factors on viral RNA synthesis is needed.

The genomic RNA of alphaviruses acts as an mRNA to produce ns proteins and as a template RNA for the synthesis of negative-strand RNA. Thus, the production of ns proteins and the replication of the genome are connected. This connection works in several ways. On the one hand, genomic RNA cannot serve

as mRNA for ns proteins and as template RNA for the synthesis of negative-strand RNA at the same time. Thus, there must be a switch from translation to replication. On the other hand, any defects in RNA replication inevitably result in a reduction in the amount of synthesized ns proteins, which in turn has an impact on RNA synthesis. Furthermore, studying the RNA replication of the alphaviruses in the context of a virus infection presents additional challenges. For example, mutations introduced into ns proteins are often reversed or compensated for. In addition, alphaviruses of high medical importance are dangerous pathogens, and biosafety level 3 laboratories are required for studying these viruses, complicating the research. These problems have been countered by generating *trans*-replication systems, in which the replication of the viral minigenome (template RNA) and the production of the ns proteins (replicase) are uncoupled. This is possible thanks to the ability of the alphavirus replicase proteins to find, bind and replicate suitable RNA templates. This transactivation activity of the alphavirus replicase proteins becomes evident from the generation of efficiently replicating defective interfering RNAs, a phenomenon that has been known for half a century (195). The same ability of the alphavirus replicase has been used for the generation of helper RNAs used for the packaging of alphavirus replicon vectors more than three decades ago (196). *Trans*-replicases were developed more recently. However, they also utilize the *trans*-activity of the alphavirus replicase proteins.

Trans-replicases constructed for SFV, SINV, CHIKV and other alphaviruses have been used to study the requirements of *cis*-elements for RNA replication, the importance of ns polyprotein processing for the efficiency of RNA replication in both mammalian and mosquito cells, the formation of RCs and the requirements for various host factors (72, 77, 96, 140, 185, 186, 197–199). Although all *trans*-replicases of alphaviruses rely on the same general principle, somewhat different designs have been used, and therefore, systems with different properties are available. Spuul and colleagues designed a *trans*-replication system for SFV, in which the template RNA was produced using bacteriophage T7 RNA polymerase and the replicase proteins were expressed from a separate plasmid. The transcription of the plasmids was performed using T7 RNA polymerase, which created noncapped transcripts. Poor translation of the noncapped RNAs is beneficial for the template RNA (see below) but is not acceptable for RNA encoding replicase proteins. Therefore, the translation of those transcripts was performed using the internal ribosomal entry site of the encephalomyocarditis virus (EMCV IRES) (198). The template RNA represented the genomic RNA of SFV, in which the CSEs absolutely required for the replication of the genome were preserved while the ORFs encoding the viral proteins were substituted with sequences encoding marker proteins such as *Renilla* luciferase (Rluc) and fluorescent Tomato protein.

The original system relied on using transgenic cell lines permanently expressing the DNA-dependent RNA polymerase of bacteriophage T7. The T7 RNA polymerase-based *trans*-replication system is highly efficient due to the high activity of T7 RNA polymerase. This system demonstrated high “boost values”, i.e., the expression of markers in the presence of an active replicase was much

higher than in the absence of the replicase or in the presence of a mutant replicase lacking RdRp activity. This is because the template RNAs generated by the T7 RNA polymerase lack the 5' cap structure and are therefore poorly translated, resulting in a low background of marker expression. In contrast, positive-strand RNAs (i.e., genomic and SG RNA) generated by the alphavirus replicase are capped by the alphavirus capping machinery and therefore much more efficiently translated. This system was shown to be a useful tool for quantifying the RNA replication activities of different templates and for analyzing the impact of certain replicase mutations on the activity of the alphavirus RNA replicase. The high efficiency of the system made the detection of spherules harboring the RCs relatively easy and allowed for the analysis of their morphology. This analysis revealed that the size of the spherules is dependent on the length of the template RNA provided (197). The T7 RNA polymerase-based SFV *trans*-replication system was also utilized by Hellström and colleagues in a study in which they focused on the requirements of *cis*-acting elements in genomic RNA during replication (77). The 5' UTR along with the beginning of the nsP1 coding region and the end of the 3' UTR were shown to be required for effective synthesis of negative-strand RNAs. These data are in good accordance with previously reported data from a replicon-based analysis, where SINV/SFV chimeras were used to study the requirement for *cis*-acting elements in alphavirus RNA replication (28). As mentioned above, spherule structures were effectively formed in the presence of a replication-competent template RNA accompanied by a replicase polyprotein expressed from a separate plasmid. The active spherule formation allowed to address several questions regarding the basics of this process. First, can the spherules also be formed from a template RNA and separately expressed components of the replicase? Second, is the template RNA expression plasmid strictly needed? Third, must the replicase (or its nsP4 component) be enzymatically active? None of these topics could be analyzed using mutant virus genomes or replicons. However, the *trans*-replicase assay made such an analysis possible. Using the *trans*-replicase of SFV, spherules were also found to be formed when expressing either P123 and nsP4 or nsP1, uncleavable P23 and nsP4 instead of P1234. Spherule-like structures were observed in the absence of the template RNA or in the absence of RdRp activity of nsP4 if the P123 or P23 used in the assay contained a mutation preventing the processing of the 2/3 site (96). Thus, spherules or spherule-like structures can be formed, albeit rather inefficiently, in the absence of RNA replication.

Although effective, a *trans*-replication system wherein the template RNA is expressed using the T7 RNA polymerase also possesses a significant limitation: the requirement for the presence of the bacteriophage T7 RNA polymerase. This makes the system easy to use in transgenic cells expressing the bacteriophage T7 polymerase, however, for other cells, cotransfection with the T7 RNA polymerase expression plasmid is needed. This represents a problem for cells that are difficult to transfect, as cotransfection with at least three plasmids may be challenging. Furthermore, if one aims to integrate the template RNA expressing construct into the host genome (i.e., generate stable reporter cell lines), using the

T7 promoter is a poor choice as the T7 RNA polymerase does not enter the nucleus of the cell. Therefore, *trans*-replicase systems in which the transcription of the replicase mRNA and the replication competent template RNA is performed *via* a host cell's DNA-dependent RNA polymerase would offer additional possibilities. The first of such systems utilized the human cytomegalovirus immediate early promoter (CMV promoter) for the transcription of both template RNA and replicase mRNA. In this system, ns polyprotein mRNA and template RNA are generated in the nucleus of the transfected cell using cellular RNA polymerase II, which produces capped transcripts that are transported into the cytoplasm. This system was first developed for SFV and used to study the host cell's innate immune response. Surprisingly, the replicase of SFV alone was found to be sufficient to induce IFN-β expression. This was the first description that an alphavirus can induce IFN-β expression in a viral RNA replication-independent manner (75).

Next, a CMV promoter-based CHIKV *trans*-replicase was constructed. The template RNA included all *cis*-acting elements required for the replication and transcription of the viral genome, i.e., the 5' UTR followed by the first 231 nucleotides of the nsP1 encoding sequence, the long version of the SG promoter (for maximal activity) and the last 110 nucleotides from the 3' UTR followed by a poly(A) tail of 60 nucleotides. A firefly luciferase (Fluc) marker gene was positioned under the genomic promoter, and the changes in Fluc activities in the presence of an active viral replicase were attributed to RNA replication (i.e., production of full-length positive-strand copies of the template RNA, which correspond to the genomic RNA) activities. The *Gaussia* luciferase (Gluc) marker was positioned under the control of the SG promoter, and accordingly, the changes in Gluc activities in the presence of the viral replicase represented transcription (i.e., production of SG RNA) activity (199). The new *trans*-replicase was compared to its counterpart with the T7 promoter. For both systems, the production of Gluc in the presence of an active CHIKV replicase was highly effective. The transcription boost (activity in the presence of an active replicase compared to the activity in the presence of a polymerase-inactive replicase) exceeded 1000-fold. However, due to the high background levels resulting from the translation of Fluc from the capped transcripts generated by RNA polymerase II, the replication boost (increase in Fluc activity) for the CMV promoter-based system was very low (199). Nevertheless, this CMV promoter-based *trans*-replicase system proved to be an effective, sensitive and versatile system for studying the replication of CHIKV. However, this system could not be used in mosquito cells. To overcome this limitation, an *Aedes aegypti* polyubiquitine promoter-based CHIKV *trans*-replicase was constructed and utilized to study the effects of various mutations in ns proteins in mosquito cells (186). As expected, mutations inhibiting the enzymatic functions of the ns proteins abolished RNA replication in both mammalian and mosquito cells, i.e., their impact was not host specific. In contrast, mutations designed to inhibit ns polyprotein processing typically had positive effects on RNA replication in mosquito cells while negatively impacting RNA replication in mammalian cells. The increase in RNA replication in mosquito cells might have resulted from a stabilizing effect of these mutations on the ns polyprotein

processing intermediates (i.e., components of alphavirus early replicase), allowing more time for the formation of RCs. This, however, does not explain why the same mutations negatively affected CHIKV replication in mammalian cells (186).

Transcripts produced by RNA polymerase II using CMV or polyubiquitinated promoters are capped. However, as mentioned above, the capped full-length (hereafter referred to as "genomic") RNAs generated by cellular RNA polymerase II are translated very effectively, causing a high background for Fluc activity, and therefore, only a small boost is generated in the presence of active replicase. Utt and colleagues proceeded to combat this restriction by generating cellular RNA polymerase I-based *trans*-replication systems for CHIKV (Figure 13). In these systems, the DNA copy of the template RNA was positioned under the control of a truncated RNA polymerase I promoter from *Homo sapiens, Aedes albopictus* or *Aedes aegypti* (140).

Transcripts produced by RNA polymerase I are uncapped and thus poorly translated in the cell, which greatly reduces the background levels of marker protein expression. The novel RNA polymerase I-based CHIKV *trans*-replicase system proved to be very effective, as it displayed high boosts in the activities of both replication and transcription. The correlation of the Fluc and Gluc boost values with actual positive-strand RNA synthesis was analyzed by detecting the genomic and SG RNAs by northern blot analysis. This analysis revealed excellent correlation between the boosts of marker activities and the abundance of corresponding RNAs generated by the replicase (140). As the RNA polymerase I-based CHIKV *trans*-replication system proved to be highly effective, the span of available systems was expanded significantly by generating similar *trans*-replication systems for SINV, BFV, VEEV, SFV, ONNV, MAYV and RRV. These systems were used to analyze the dependence of alphavirus replication on G3BP proteins (185), and the results originating from this study are described above (Section 1.3.3).

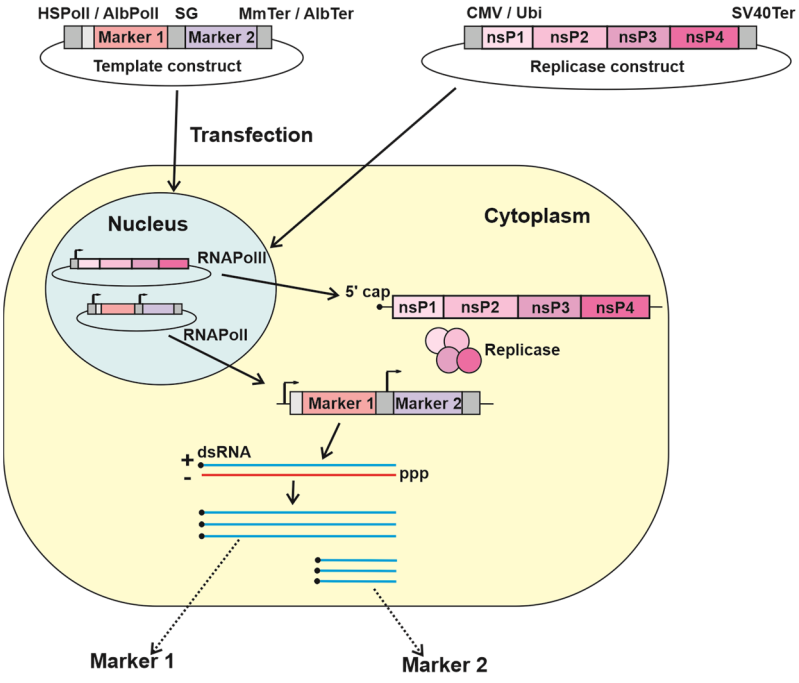

Figure 13. *Trans*-replication assay based on the use of constructs harboring promoters for cellular RNA polymerases. Cells are cotransfected with plasmids expressing alphavirus replicase proteins and a replication-competent RNA template. The expression of mRNA for the replicase proteins occurs in the nucleus from CMV (mammalian cells) or *Aedes aegypti* polyubiquitine (mosquito cells) promoters using cellular RNA polymerase II. The synthesized capped mRNAs are transported to the cytoplasm and translated. Expression of the replication-competent template RNA occurs in the nucleus using RNA polymerase I (resulting in noncapped transcripts). These template RNAs contain the *cis*-elements required for replication and are recognized by the alphavirus replicase. The replicase synthesizes negative-strand RNA and then starts producing capped full-length template RNAs from which Marker 1 is produced and shorter SG RNAs from which Marker 2 is produced. HSPolI and AlbPolI– truncated promoters for human and *Aedes albopictus* RNA polymerase I, respectively; MmTer and AlbTer – terminators for mouse and *Aedes albopictus* RNA polymerase I, respectively.

2. AIMS OF THE STUDY

The general aim of this study was to find new and exciting information on alphavirus replication, to develop new tools and to obtain knowledge that promotes scientific understanding while providing opportunities to put the acquired information into practical use. Specifically, the aim was to provide novel detailed information on the RNA replication of alphaviruses, how different components involved in this process interact with each other and how these interactions may shape alphavirus evolution.

Alphaviruses are abundant and medically important human pathogens. Increased travel, transport of goods and climate change favorable for mosquito vectors make it simple to anticipate the increasing medical importance of alphaviruses. Furthermore, constant evolution and adaptation makes alphaviruses susceptible to inhabiting new regions. To date, tools for counteracting alphaviruses remain very limited. Understanding the mechanisms behind the virus's spread, adaptations and pathogenicity as well as developing efficient antiviral strategies, such as vaccines and antiviral drugs, require detailed knowledge on the molecular biology of the viruses. In the world of RNA viruses, all viral properties are either directly or indirectly linked to RNA replication. Thus, in the case of alphaviruses, it is fitting to say that all roads lead to the replication of viral RNA, making it the most important (and interesting) aspect of alphavirus biology to study.

Studies included in this thesis were focused on the following objectives:

1. Analyzing the interactions between the two main components of the alphavirus RNA replication: the replicating RNA and the nonstructural (poly)-proteins (virus-encoded subunits of RNA replicase).
2. Analyzing the functional interactions between the RNA-dependent RNA polymerase (RdRp) of the alphaviruses and other virus-encoded subunits of the RNA replicase.
3. To perform a comprehensive analysis, the studies described in Objectives 1 and 2 were carried out using a panel of viruses that included representatives of all major groups of alphaviruses.
4. Using the obtained knowledge to partake in an analysis of the structure of the alphavirus RdRp and performing analysis on its activities using structure-guided mutagenesis.
5. Developing a set of tools and methods for performing the abovementioned research that are also applicable to a wide range of basic and applied studies.

3. MATERIALS AND METHODS

The methods used in this study can be divided into several groups. First, standard approaches in molecular biology and virology, such as DNA manipulation, cell transfection, virus rescue, titrations, etc., which are provided in the corresponding sections of Publications I, II, III and IV. Second, a significant part of this study was dedicated to the development of original approaches allowing in-depth analysis of the RNA replication of alphaviruses. These methods are considered original results of the studies presented in this thesis and are, if needed, described in the Results and Discussion section. Finally, there is a set of methods and approaches developed in our laboratory and adjusted to the studies presented in this thesis. These methods are overviewed in this section accompanied by a brief rationale as to why these methods were selected together with descriptions of their benefits and limitations.

The *trans*-replication assay (Publications I, II, III and IV)

Principles of the *trans*-replication assay are explained in detail in Section 1.4. In this study, the *trans*-replication system exploiting the template RNA expressed by RNA polymerase I was used (Figure 13). The selection was based on the high sensitivity of this system and the possibility of adapting the assays to a 96-well plate format. As described in Section 1.4, the *trans*-replication assay used in this study was based on the cotransfection of the cells with two basic components: a plasmid expressing the replication competent template RNA and another plasmid (or a set of plasmids) expressing the ns proteins and their precursors (Figure 14).

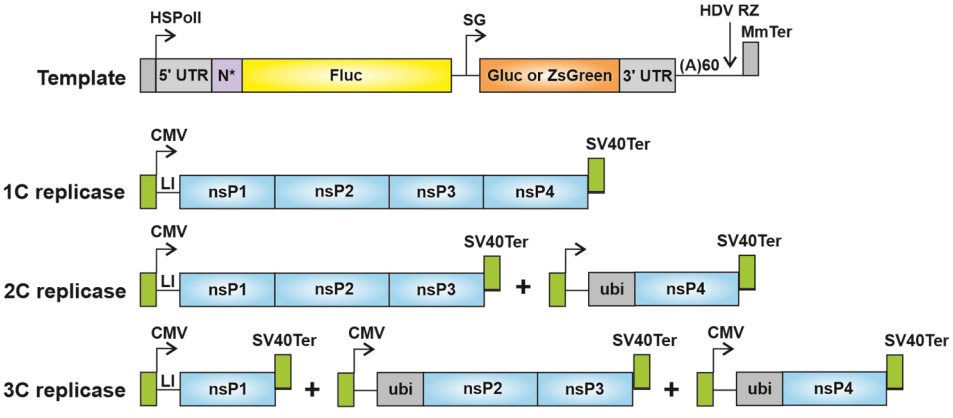

Figure 14. Schematic overview of the plasmids used to express components of the *trans*-replication system. The expression cassette for the template RNA (top) consisting of the following: HSPolI – a truncated promoter (residues −211 to −1) for human RNA polymerase I; 5′ UTR – the full length 5′ UTR of an alphavirus; nsP1 N* – region encoding the N-terminal 77 to 114 amino acid residues of nsP1, depending on the virus; SG – SG promoter spanning (with respect to termination codon of nsP4) from position –79 to the end of intergenic region; 3′ UTR – truncated (last 110 residues) 3′ UTR of an alphavirus; HDV RZ – antisense strand ribozyme of hepatitis delta virus; and MmTer – a terminator for RNA polymerase I in mice. In the ns (poly)protein expression constructs, the following components were used: CMV – the immediate early promoter of human cytomegalovirus; LI – leader sequence of the herpes simplex virus thymidine kinase gene with artificial intron; SV40Ter – simian virus 40 late polyadenylation region; and Ubi – sequence encoding for human ubiquitin. The vector backbones are not shown, and drawings are not to scale. 1C, 2C and 3C (C stands for component) replicases represent replicase proteins expressed from either one, two or three separate expression constructs (components), respectively.

Plasmids expressing the ns (poly)proteins contained the CMV promoter and were used in human cells. Based on the promoter used and the expressed viral ns (poly)proteins, the constructs were designated, for example, CMV-P1234-CHIKV, in which CMV refers to the promoter used, P1234 to the sequence encoding the full-length ns polyprotein and CHIKV to the virus from which the polyprotein coding sequence originates. In the presence of fusions or mutations, appropriate

specifications were incorporated into the name of the plasmid. In addition to the plasmids expressing a single replicase precursor (full-length P1234 polyprotein, also referred to as the one-component (1C) replicase), in Publications II, III and IV, the alphavirus replicase was also expressed as a combination of two (P123 + nsP4) or three (nsP1 + P23 + nsP4) ns (poly)protein expression plasmids, and these combinations are referred to as the two-component (2C) and three-component (3C) replicases, respectively (Figure 14).

Sequences corresponding to the template RNAs were positioned under the control of the *Homo sapiens* RNA polymerase I (HSPolI) promoter. Notably, a truncated promoter lacking the region located downstream of the transcription initiation site was used. This truncation most likely reduces the activity of the promoter used but is essential, as extra (nonviral) nucleotides in the 5' end of the RNA transcripts can reduce or even eliminate the ability of the replicases to replicate the template RNAs. Furthermore, alphaviruses have a strict requirement for a native AU dinucleotide to be present at the 5' end of the replicating RNA (144). The template RNAs included *cis*-elements from the viral genomes necessary for RNA replication and transcription. These elements included the 5' UTR, the sequences coding for the N-terminal region of nsP1 containing stem–loop structures involved in RNA replication (Figure 6), the end of the nsP4-encoding sequences together with the intergenic region comprising the SG promoter and the truncated version (last 110 residues) of the 3' UTR with the poly(A) tail having 69 residues. The length of the 3' UTR was selected because it has previously been shown to be sufficient for active RNA replication in mosquito and mammalian cells. In addition, the template RNAs were designed to include reporters instead of the native viral ORFs: Marker 1 (Fluc in all used templates) substituting the ns region and Marker 2 (either Gluc or a fluorescent ZsGreen protein) substituting the structural region of the genome. The template RNA expression plasmids were designated: for example, in HSPolI-FG-CHIKV, HSPolI refers to the promoter used, FG refers to Fluc and Gluc reporters and CHIKV refers to the virus from which the virus-specific elements originate.

The basic setup of the *trans*-replication assay is shown in Figure 15. Briefly, plasmids expressing the template RNA and replicase proteins were cotransfected into cells that were grown, depending on the experiment, on a 6, 12, 24, 48 or 96-well cell culture plate. The cells were incubated either at 37 °C and 5% CO_2 for 16–18 hours or at 28 °C and 5% CO_2 for 48 hours (human cells transfected with plasmids of EILV *trans*-replicase). Following incubation, the cells were lysed, and Fluc and Gluc activities were measured using the Dual Luciferase Reporter Assay System (Promega). Alternatively, if the cells were transfected with a plasmid expressing template RNA with the ZsGreen marker, the cells were harvested and analyzed using flow cytometry (Figure 15).

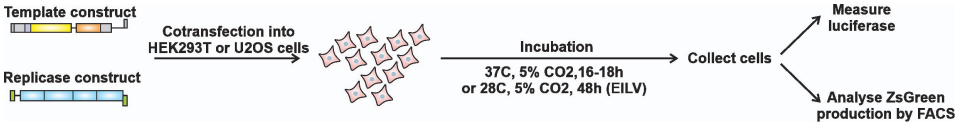

Figure 15. Experimental flow of the *trans*-replication assays. In the experiments where the template encoded the Fluc and Gluc reporters, the activities of these markers were measured using luminometers. In the experiments where the template contained the ZsGreen marker expressed from the SG RNA, the percentage of ZsGreen-positive cells and mean fluorescent intensity in these cells were determined using a cell sorter.

The transfection–infection assay (Publication III)

Transfection–infection assays were conducted to assess the potential of the template RNAs to serve as reporters in detecting viral infection (Figure 16). Cells were transfected with template RNA-encoding plasmids (no replicase-encoding plasmids were used in these experiments) containing the ZsGreen marker under the SG promoter to enable visual conformation of template activation by a relevant virus. Transfected cells were incubated at 37 °C, and at 18 hours post-transfection, the cells were infected with a relevant virus at an MOI of 10. At 24 hours post-infection (h.p.i.), the fluorescent reporter production was visually observed. Alternatively, time-lapse images of infected cells were recorded using an EVOS m7000 microscope (Thermo Fisher Scientific).

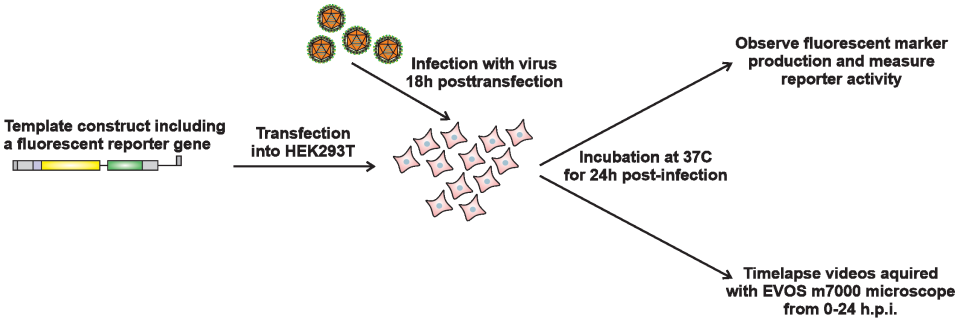

Figure 16. Experimental flow of the transfection–infection assays.

4. RESULTS AND DISCUSSION

This section summarizes data from Publications I–IV accompanied by some unpublished results. It should be noted that the data collected in Publication I were obtained prior to the PhD studies and were previously used in my MSc thesis. However, information in Publication I provides essential background for the rest of the studies included in this thesis and, most importantly, only together these four publications form a logically connected story. For these reasons, the data from Publication I are mostly used to explain the bases of other studies and are discussed mostly with the aim of demonstrating how the data are connected to the other publications and recent unpublished results.

4.1. The cross-utilization of template RNAs by replicases of heterologous alphaviruses

The *trans*-replication system uncouples the replication of the template RNA (a minigenome) from the synthesis of the ns proteins, which are the virus-encoded subunits of the RNA replicase. Furthermore, expressing the template RNAs and replicase proteins using separate expression constructs allows for the analysis of heterologous template and replicase combinations based on a "mix and match" approach. In this study, one of the topics of significant interest was the utilization of template RNAs by replicases originating from the same virus (hereafter referred to as matching replicases or combinations) as well as by replicases from different alphaviruses (hereafter referred to as heterologous replicases or combinations). In previous studies carried out in our laboratory, HSPolI-based template RNA expression plasmids of eight different alphaviruses (SINV, BFV, VEEV, CHIKV, ONNV, RRV, MAYV and SFV) were designed, and their functionality was confirmed in human cells (140, 185). Here, we also designed and constructed the HSPolI-based RNA template and the replicase precursor (P1234) expression plasmids for the insect-specific EILV (Publication I) and an encephalitic New World alphavirus EEEV (Publication III), generating a set of 10 alphavirus *trans*-replicases in total that were used in human cells. In Publications I and III, we aimed to determine whether alphavirus replicases are capable of cross-utilizing heterologous template RNAs. The basic principle of the HSPolI-based *trans*-replicase assay is explained in Section 1.4 and illustrated in Figure 13. The basic setup of these experiments is explained in the Methods and illustrated in Figure 15. Details of the experimental procedures are also available in the corresponding sections of Publications I and III and are therefore not repeated here.

Interestingly, in the first set of experiments, it was quickly observed that based on the ability of the replicases to use each other's template RNAs, the alphaviruses can be separated into two groups: first, the viruses belonging to the Semliki Forest complex (CHIKV, ONNV, MAYV, RRV and SFV) and second, viruses

that originate from various antigenic complexes outside the Semliki Forest complex and were collectively referred to as the "outgroup" alphaviruses (SINV, BFV, EILV, VEEV and EEEV). Interestingly, the *trans*-replicases of viruses from these groups exhibited different properties in several different assays, i.e., this division can also be applied to the studies described in Publications II and III. It may not, however, reflect the true evolutionary connections between the "outgroup viruses", as the observed similarities of the *trans*-replicases of outgroup alphaviruses might have collectively differed from the properties of the replicases of Semliki Forest complex viruses. Thus, our data most likely indicate that the viruses from the Semliki Forest complex stand out from other alphaviruses, i.e., they likely acquired properties that are not present in other alphavirus groups. If so, the situation is similar to what has also been observed for the genome packaging signals: the location, nature and required signals of which are different in the SFV clade compared to other viruses.

Prior to studying the cross-utilization of the template RNAs by various replicases, a few pilot experiments had to be executed. First, we needed to determine the most suitable time as an endpoint for the experiments, i.e., the incubation time after which the cells should be either lysed (Fluc-Gluc templates) or collected (Fluc-ZsGreen templates) for measuring the reporter signals (Figure 15). In a previous study, the expression kinetics of CMV and T7 promoter-based CHIKV *trans*-replicases were assessed (199). However, data characterizing the HSPolII-based system and the *trans*-replicases of viruses other than CHIKV were missing. Measuring the expression kinetics of the Gluc marker was performed by transfecting HEK293T cells with all matching combinations of the template and replicase expression constructs (EILV was excluded from this analysis). While the replicases of different alphaviruses displayed some differences in activities, the kinetics of marker accumulation were overall similar, excluding the need to use different timepoints for different alphaviruses (Publication I, Figure 2A; unpublished data). Based on the replication kinetics, a 16- to 18-hour posttransfection endpoint was used for all *trans*-replicase experiments performed in HEK293T cells except for the *trans*-replicase of EILV. Furthermore, the high similarity of the kinetics of reporter expression was an interesting finding on its own, as the alphaviruses on which the *trans*-replicases were based have rather different growth characteristics in cell culture. For example, a cell culture adapted SINV grows much faster than an ONNV strain based on a recent clinical isolate.

After selecting a suitable endpoint for the experiments in human cells, we revealed that the matching combinations of the template RNA and replicase expression constructs resulted in high boosts of Fluc (i.e., replication) and Gluc (i.e., transcription) activities, with the exceptions of EILV and ONNV (Publication I, Figure 2B). No replication or transcription was observed for EILV at 37 °C, which is not entirely surprising, as EILV is an insect-specific virus, and the cells and/or temperature used in this experiment were likely unsuitable for the activity of its RNA replicase. Interestingly, at 28 °C, the replication and transcription activities of EILV RNA replicase were clearly detectable, albeit at a lower level compared to the other *trans*-replicases (Publication I, Figure 2B). This

indicates that EILV is highly sensitive to temperature, but its RNA replication has no fundamental need for mosquito cells. The *trans*-replicase of VEEV demonstrated an interesting phenotype: while a low replication efficiency was observed, the transcription was highly active (Publication I, Figure 2B). Originally, this result was assumed to be caused by the noncytotoxic nature of the VEEV replicase (nsP2 of the New World alphaviruses does not cause shutoff of the host cell's transcription). However, subsequent analysis of the *trans*-replicase of EEEV, another New World alphavirus, revealed that it is highly active in human cells with transcription activities surpassing those of all other *trans*-replicases (Publication III, Figure 1B). Comparison of the replication activities of these two New World alphaviruses revealed that the replication activity of the *trans*-replicase of EEEV exceeded that of the *trans*-replicase of VEEV by approximately 120-fold (Publication III, Figure 1B). This difference is unlikely to have been caused by the differences in the EEEV and VEEV ns proteins in the host cell. Furthermore, in subsequent experiments performed using infectious viruses, VEEV was observed to be rather inefficient in activating a replication-competent template RNA (Publication III, Figure 7). The logical conclusion is that the low efficiency of the RNA replication observed for the VEEV *trans*-replicase is not an artifact of the system or a host cell-mediated effect but reflects an intrinsic property of the VEEV replicase. Most likely, the reason resides in an unusually low (for an alphavirus) ability of the VEEV replicase to recognize, bind and utilize the template RNAs provided *in trans*. In subsequent studies (see Section 4.2.2.), the determinant within the replicase of VEEV responsible for this effect was mapped.

Altogether, the preliminary experiments demonstrated that all the *trans*-replicases are active in both replication and transcription in human cells and can be used in the following cross-utilization experiments.

4.1.1. Alphavirus replicase proteins can utilize heterologous template RNAs

Having confirmed that all 10 *trans*-replicases are active, we proceeded to study the cross-utilization of the alphavirus template RNAs by heterologous replicases. As we needed to test all template RNA expression constructs with every replicase expression construct, including negative controls (inactive replicase instead of the wt replicase) and biologically relevant repeats, we adapted the *trans*-replicase assay to a 96-well format. The 96-well format *trans*-replicase assay provided us with a tool where many combinations could be tested at once while reducing the volume of plasmids, cells and reagents needed. Overall, testing the alphavirus RNA templates and replicases in all possible combinations revealed that all alphavirus replicases can replicate heterologous template RNAs at least to some extent (Publication I, Figures 4 and 5; Publication III, Figure 2). The only exception was the template RNA corresponding to the salmonid alphavirus, which was not used by any of the alphavirus replicases included in our study (data not shown). It was hypothesized that either the salmonid alphavirus is too distant from other alphaviruses and/or that the assay conditions (temperature of 28 °C or 37 °C, human

or mosquito cells) were unsuitable for the replication of this template RNA. In retrospective, however, it seems more likely that there was a mistake in the design of this template RNA. Namely, the sequences available in the database for the salmonid alphavirus genome either lacked the 5' region or were inconsistent regarding the 5' residues. Therefore, information from a previous publication describing the infectious clone of the virus was used as a guideline, and the 5' end of the salmonid alphavirus genome was deduced to be GAUAAAUC (119). However, nsP1 was recently described as possessing a strong preference toward the AU dinucleotide at the 5' end of template RNA, making it likely that the G residue at the 5' end of our template RNA was an artificial addition (144). Apparently, such an addition is tolerated in the infectious clone of the virus but not in the template used in the *trans*-replicase system. Thus, we cannot exclude that even the salmonid alphavirus template RNA, if present in the correct form, could be used by replicases of heterologous alphaviruses.

While the ability of the alphavirus replicases to use heterologous template RNAs was universal, the efficiency regarding the use of the template RNAs displayed huge variation. The replicases of viruses belonging to the Semliki Forest complex generally had the ability to use each other's template RNAs, and the efficiency of cross-utilization correlated well with the phylogenetic relationships between those viruses. For example, CHIKV and ONNV, which are closely related, used each other's templates efficiently. At the same time, clear differences were observed for less closely related viruses, such as RRV and CHIKV (Publication I, Figures 4A–4C). Furthermore, the replicases of viruses belonging to the Semliki Forest complex were also capable of utilizing the template RNAs of the outgroup viruses (Publication I, Figures 5A–5D; Publication III, Figure 2A). The SINV template RNA was used especially efficiently (except by the replicase proteins of ONNV, VEEV and EILV; Publication I, Figure 5A).

None of the replicases of the outgroup viruses were able to use the template RNAs of the Semliki Forest complex viruses effectively (Publication I, Figure 4A–4E, Publication III, Figure 2B). Interestingly, the use of the template RNAs belonging to the outgroup viruses displayed a certain pattern (Publication I, Figure 5A–5D). Again, the SINV template RNA was efficiently used by most of the replicases of the outgroup viruses except for the EILV and VEEV replicases (Publication I, Figures 5A; Publication III, Figure 2B). The replication of the template RNAs of VEEV and BFV was also similar to that of SINV (Publication I, compare Figures 5A–5C). Notably, replication of the VEEV template RNA by several heterologous replicases was more efficient than that observed for the matching replicase, indicating that the low replication activity observed for the *trans*-replicase of VEEV can indeed be attributed to the properties of its replicase proteins (Publication I, Figure 5C; Publication III, Figure 2A and 2B). The EEEV template RNA was generally replicated and transcribed effectively. The only replicase that was not able to use the EEEV template RNA belonged to EILV, and modest replication activities were also displayed by the VEEV and ONNV replicases (Publication III, Figure 2A). However, the VEEV and ONNV replicases also displayed modest activities on their matching template RNAs

(Publication III, Figure 1B). In addition, the EEEV template RNA was relatively ineffectively used by the SINV replicase (Publication III, Figure 2A). This was not surprising, as the SINV replicase demonstrated high boosts only for its matching template RNA and that of EILV (Publication II, Figures 5A and 5D). Thus, while the SINV template RNA was rather promiscuous, its replicase proteins were quite selective. The EEEV replicase did not replicate the SINV template RNA very effectively, however, it did generate a boost in transcription similar to what was observed in the presence of the matching replicase (compare Publication III, Figure 1B and Figure 2B). One of the most interesting template RNAs was that of EILV. While it was poorly replicated by its matching replicase (likely because of unfavorable conditions in human cells), its replication by the SINV, BFV, MAYV and SFV replicases was very efficient (Publication I, Figure 5D). The transcription of the EILV template was efficient by all the replicases, with the exceptions of the New World alphaviruses (Publication II, Figure 5D; Publication III, Figure 2B). Somewhat surprisingly, the transcription activity of the matching replicase was also high. Overall, the utilization of the EILV template RNA was similar to that of the SINV template RNA.

Taken together, the cross-utilization of the alphavirus template RNAs revealed clear differences between the viruses from the Semliki Forest group and the outgroup viruses. The template RNAs of several outgroup viruses (SINV, EILV and EEEV) had universal properties, meaning that they could be used by most of the alphavirus replicases at least with moderate efficiency. In this study, the template specificity/preference provided a topic for additional analysis. The promiscuous properties of the SINV template RNA were beneficial for subsequent analysis of the functional compatibility of replicase proteins derived from different alphaviruses. More generally, the cross-utilization of template RNAs by heterologous alphavirus replicases provides insights into the evolution of alphaviruses. For example, in cells that are coinfected with different alphaviruses, their replicase proteins can cross-utilize each other's template RNAs, which might provide a pathway for the viruses to generate recombinant RNA genomes. These genomes, in turn, can be used by the original or recombinant replicases of viruses involved in the coinfection event. Altogether, the cross-utilization of the alphavirus RNA genomes might present an important factor in the evolution of the alphaviruses.

4.1.2. Toward understanding the molecular bases of the template specificity of alphavirus replicases – from the point of view of the template RNA

Analysis of the cross-utilization of the alphavirus template RNAs indicated that some alphavirus template RNAs are quite promiscuous, whereas others can only be used by selected replicases. Furthermore, additional preferences were observed for both groups. For viruses from the Semliki Forest complex, the template preferences were often reciprocal. For example, the CHIKV replicase used the RRV template RNA less efficiently than its matching template RNA, and similarly, the

RRV replicase preferred its own template RNA over the CHIKV template (Publication I, Figure 4A and 4C). However, when both the Semliki Forest complex and outgroup viruses were involved, the usage of template RNAs was often not reciprocal. Furthermore, no clear correlation could be observed for the functional combinations of template RNAs and replicases. For example, while the SINV, EILV and EEEV template RNAs could be used by most of the heterologous replicases, their own replicases displayed clear specificity toward their matching RNAs and were rather inefficient in using heterologous template RNAs. Furthermore, the SINV template RNA was efficiently used by the CHIKV replicase, however, the SINV replicase was almost unable to use the CHIKV template RNA (Publication I, Figures 4A and 5A). To continue, the SINV template RNA was efficiently transcribed (and moderately replicated) by the EEEV replicase, while the SINV replicase was very poor in both replication and transcription of the EEEV template RNA (Publication III, Figures 2A and 2B). These nonreciprocal combinations of heterologous template RNAs and replicases presented an opportunity to study which region(s) of the alphavirus genome are responsible for the template specificity.

To address this question, we first constructed a set of chimeric SINV and CHIKV template RNAs by swapping the regions corresponding to the 5' UTR and the start of nsP1 (hereafter referred to as the 5' region), the SG region and the 3' UTR (the 3' region) of the viral genomes (Publication I, Figure 9A). Experiments with the SINV and CHIKV chimeric templates revealed that the template specificity of the SINV replicase is determined by the 5' region and the SG region of the genome (Publication I, Figure 9C). More detailed analysis confirmed that the SINV replicase requires the presence of the first SL structure (SL3) in the 5' UTR of the SINV genome to achieve successful replication (Publication I, Figure 10B). Next, we constructed similar chimeric templates by swapping the 5' region, SG region and 3' region of the SINV and EEEV template RNAs (Figure 17A). The results obtained using this set of constructs were similar to what we had seen with the SINV/CHIKV chimeric templates, demonstrating that the SINV replicase requires the 5' end of its genome for successful replication (Figure 17B). The same trend was also observed for the EEEV replicase (Figure 17C). We had seen earlier that the EEEV replicase transcribed the SINV template RNA very effectively, however, the replication boost was moderate at best (Publication III, Figure 2B). As the transcription of the SINV template RNA was initially effective by the EEEV replicase, we did not detect any prominent effects on the transcription of the chimeric templates using the EEEV replicase. However, for replication, the EEEV replicase clearly preferred its matching template RNA, suggesting that the region responsible for this preference could be mapped. Once again, the results of the corresponding experiment demonstrated that the EEEV replicase requires the 5' region of its matching template RNA to perform effective replication (Figure 17C).

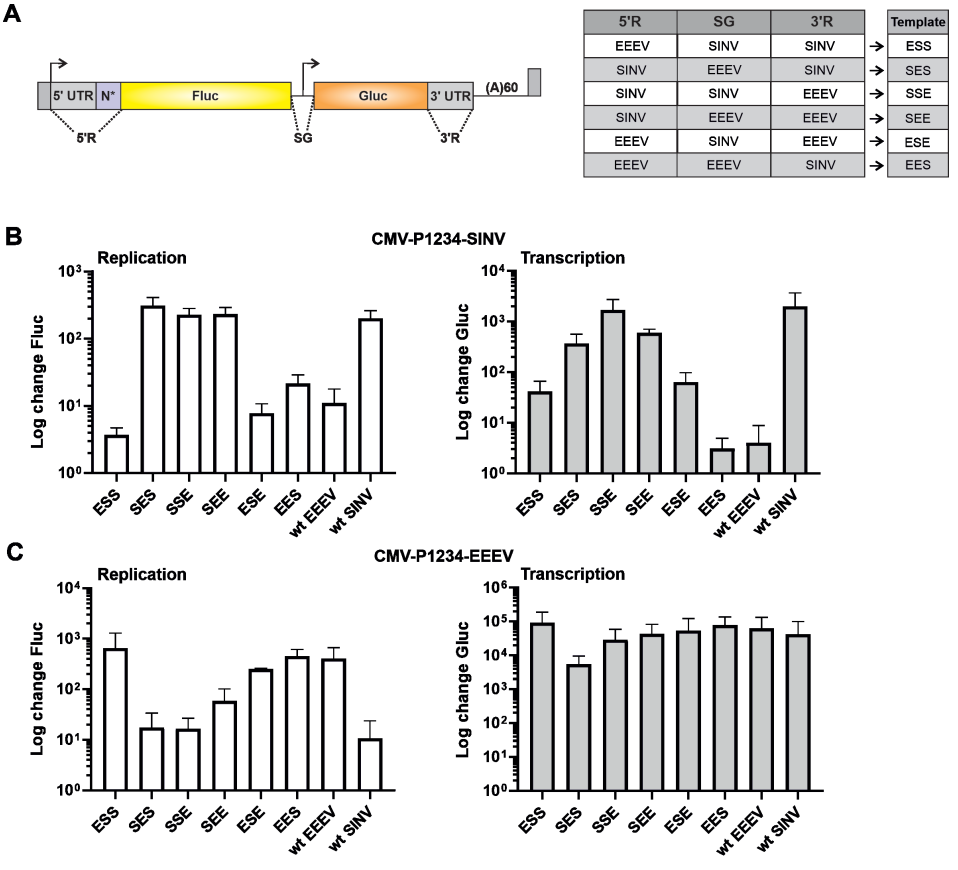

Figure 17. Analysis of the determinants of template preference of SINV and EEEV replicases (unpublished data). A. Schematic presentation of swapped templates used in the assay. B, C. Replication and transcription on the matching, heterologous and chimeric templates by SINV (B) and EEEV (C) replicases. HEK293T cells grown on a 48-well plate were cotransfected with 250 ng of template RNA and replicase expression plasmids. For control cells the plasmid expressing active replicase (CMV-P1234) was substituted with a corresponding CMV-P1234^{GAA} (polymerase inactive replicase) plasmid. Cells were lysed at 18 hours posttransfection. Data represent the luciferase activity from CMV-P1234 transfected cells normalized to the corresponding CMV-P1234^{GAA} control cells. Columns represent the means, and error bars represent the standard deviations of three independent experiments.

The studies (Publications I and III) on template RNA cross-utilization increased our understanding of the template RNA requirements for alphavirus replicase proteins. Perhaps the most important finding in the context of this thesis (Publications II and III) was the universality of the SINV template. Our data indicate that this is most likely because of the 5' terminal region of the SINV template RNA that is efficiently recognized by not only its matching replicase but also that of CHIKV and other alphaviruses. These data are in agreement with results from a previous study that was performed using replicons and minigenomes of SINV

and SFV (28). However, in both cases, one of the analyzed viruses (CHIKV or SFV) belongs to the Semliki Forest complex, while SINV is an outgroup virus. As the representatives of these two groups displayed several prominent differences in the cross-utilization of template RNAs, it was not obvious if template RNA preferences between the outgroup and the Semliki Forest complex viruses are determined in the same way. Nevertheless, our analysis revealed that for SINV and EEEV, which are both outgroup viruses, the region responsible for template specificity lies in the 5' regions of their genomes (Figure 17). Finally, the analysis performed using the swapped template RNAs of two viruses from the Semliki Forest complex, CHIKV and RRV, also confirmed the leading role of the 5' end of the genome in template preference (Publication I, Figure 9D and 9E). Collectively, these data strongly suggest that there must be a component(s) among the replicase proteins of alphaviruses that specifically recognize the 5' region of the template RNAs to successfully launch RNA replication. In the next chapters, we took advantage of the universality of the SINV template RNA and analyzed the compatibility of replicase proteins of different alphaviruses and their role in template RNA preference.

4.2. Two-component *trans*-replicases as tools in the analysis of interactions between the components of the alphavirus RNA replicase

During alphavirus infection, replicase proteins are produced by translating the P1234 (or P123 and P1234) polyprotein directly from genomic RNA. Next, these ns polyproteins undergo posttranslational (and possibly cotranslational) proteolytic processing. The processing events are mandatory, as no functional alphavirus replicase can be formed from separately expressed mature ns proteins. On the other hand, full-length P1234 cannot perform RNA replication, as the minimal requirement for the activation of replication is the cleavage of P1234, resulting in the release of the nsP4 (RdRp) protein. It has been demonstrated that an active RNA replicase can be formed from a mutant P1234 in which only the nsP4 component is released (200). Furthermore, the SINV replicase can be effectively formed by coexpressing the P123 and P34 polyproteins (84). The P34 polyprotein is cleaved by the protease activity of P123 and results, again, in the release of nsP4. These data imply that the alphavirus RNA replicase can be formed from separately produced P123 and nsP4 components that interact with suitable template RNA. However, the principles of this process have not been studied in detail.

In Publications I and III, we found that the determinants of the template RNA specificity of the alphavirus replicases are located in the 5' region of their respective genomes (the first SL structure of the 5' UTR for SINV) (Publication I, Figures 9 and 10; Figure 17). For effective replication, these sequences of the template RNA need to be matched (and interact) with one or several replicase proteins. Naturally, we questioned which part of the alphavirus replicase accounts

for this template recognition. To address this question, we either needed to construct chimeric viruses, produce replicases *via* the expression of chimeric P1234 polyproteins or generate replicases from separately expressed heterologous replicase components. Obviously, the latter option is much less laborious and allows for using a "mix and match" approach. Thus, we hypothesized that it would be possible to split the alphavirus replicases into two or more functional components, and such a system could be used in determining which of these components recognizes the template RNA. Equally important, splitting P1234 also allows us to analyze the functional compatibility of the replicase components originating from different alphaviruses.

As the 3/4 cleavage, resulting in the release of the nsP4 RdRp, is the only processing event absolutely required for the formation of the replication complexes, we decided to split the alphavirus P1234 into two: the P123 and the nsP4 components by expressing them from separate plasmids. The N-terminus of nsP4 corresponds to the conserved P1' tyrosine residue at the 3/4 cleavage site. Thus, the mature nsP4 has a Tyr residue (and not a Met) at its N-terminus. Introducing mutations to the N-terminus of nsP4 is known to compromise its activity and trigger the generation of adaptative mutations in nsP4 and nsP1 (see Chapter 1.3.1.). Hence, using an nsP4 with a nonnative N-terminus in the analysis of the compatibility of the replicase components and the template RNA was obviously not acceptable. To preserve the biologically relevant N-terminus of the protein, we expressed nsP4 in fusion with ubiquitin. Such expression plasmids were designated CMV-ubi-nsP4. After translation of the ubi-nsP4 fusion protein, ubiquitin is rapidly cleaved off, resulting in nsP4 with a natural N-terminus. The constructs expressing the P123 components of the alphavirus replicases were designated CMV-P123. The replicases expressed from two separate constructs, CMV-P123 and CMV-ubi-nsP4, are hereafter referred to as the two-component (2C) replicases (Figure 14) and were constructed for the same set of 10 alphaviruses covered in Chapter 4.1. (Publications I and III).

4.2.1. Separately expressed P123 and nsP4 components form highly efficient RNA replicases

The functionality of the 2C replicases was first confirmed by comparing their activities to those of the one-component replicases (1C). For this, HEK293T cells were cotransfected with plasmids expressing the template RNA and either CMV-P123 and CMV-ubi-nsP4 (2C replicase; in these experiments, the plasmids were used in an equimolar ratio) or with the P1234 expression plasmids (1C replicase). All 10 2C replicases were active, and, somewhat surprisingly, in most cases, their activities exceeded those of the corresponding 1C replicases (Publication II, Figure 2B; Publication III, Figure 3B). To confirm that the observed boost in Fluc and Gluc reporter activities accurately represented the synthesis of the viral RNAs, we purified the total RNA from the transfected cells and conducted a northern blot analysis (Publication II, Figure 2C; Publication III, Figure 5E). The production of all viral RNAs remained below the detection limit for EILV,

a finding that was in good accordance with the low boost value obtained from the Fluc and Gluc signals (Publication II, compare Figures 2B and 2C). In addition, the production of negative-strand RNAs remained below the detection limit for the 2C ONNV, MAYV, RRV and VEEV *trans*-replicases (Publication II, Figure 2C). Low levels of positive-strand genomic RNAs were detected for VEEV, ONNV and MAYV, again in correlation with the modest boosts in Fluc activities (Publication II, compare Figures 2B and 2C). It is also noteworthy that in all cases, the synthesis of the viral RNAs by the 2C replicases was in good accordance with the synthesis of viral RNAs by their 1C counterparts (compare Publication I, Figure 2C and Publication II, Figure 2C). Thus, it was concluded that boosts in reporter activities accurately reflect the synthesis of viral positive-strand RNAs. As the measurement of Fluc and Gluc activities is much easier to perform than a northern blot analysis, this method was used as the main approach to analyze the replication and transcription activities in all subsequent experiments. Furthermore, measuring the reporters' activities also turned out to be more informative, especially for the *trans*-replicases with low or modest activities that were not sensitive enough to be revealed by northern blot analysis.

Based on the preliminary analysis of the activities of the 2C replicases, we questioned why many of the 2C replicases displayed higher activities than their corresponding 1C replicases. Evidently, one important difference between these two systems is the production of nsP4. For 1C replicases, P1234 (source of nsP4) is produced by readthrough of an opal stop codon at the end of the sequences encoding nsP3. In our study, the stop codon was present in eight out of ten 1C replicase plasmids, the two exceptions being the CHIKV and SFV P1234 plasmids. Accordingly, the amount of nsP4 produced in cells from P1234 is typically only 10–20% of that of the other ns proteins. Obviously, the 2C replicase lacks this internal downregulation of nsP4 production. Another factor that influences the amount of nsP4 in alphavirus-infected cells is the instability of nsP4: unless stabilized by interactions with other ns proteins, nsP4 is rapidly degraded by proteasomes. The impact of this mechanism might also be different for the 1C and 2C replicases, as the abovementioned interaction may be more efficient for the former. Hence, we considered it necessary to analyze the impact of the P123:nsP4 ratio using component titration experiments in which the P123 and nsP4 expression plasmids were used at various ratios. For this analysis, we selected the 2C replicases of SINV, BFV, VEEV, EILV, CHIKV (Publication II, Figure 3), and EEEV (Publication III, Figures 3C–E), as these viruses belong to different antigenic complexes. Including these viruses in the analysis allowed us to obtain a broader picture, however, it also resulted in a technical limitation, as for most of these viruses, antibodies recognizing the ns proteins are not available. Therefore, we were unable to directly monitor the levels of these ns proteins in the transfected cells, and the accuracy of the changes resulting from the different ratios of the P123 and nsP4 components remains unknown.

For the nsP4 component titration experiments, we used nsP4:P123 expression plasmids in ratios from 1:10 to 8:1 (Publication II, Figure 3A). The results showed that the replication of alphaviruses is strongly dependent on the amount

of nsP4. At a nsP4:P123 expression plasmid ratio of 1:10, the activities of CHIKV, VEEV and EILV replicases remained close to the detection limit, and the activities of the SINV and BFV replicases were only approximately 10-fold over the background level (Publication II, Figure 3A). When the amount of nsP4 was increased by changing the nsP4:P123 expression plasmid ratio from 1:10 to 8:1, the replication activities improved approximately 10–100-fold. The impact of increasing the amount of nsP4 was similar in transcription, albeit slightly less pronounced (Publication II, Figure 3A). These results suggested that nsP4 is the component that limits alphavirus RNA replicase activity and that excess nsP4 results in increased RNA replication. However, similar analysis performed using the *trans*-replicase of EEEV revealed a somewhat different picture. Similar to the aforementioned alphaviruses, increasing the EEEV nsP4:P123 expression plasmid ratio from 1:10 to 1:1 improved replication (approximately 7-fold) and transcription (approximately 3-fold) (Publication III, Figure 3C). However, increasing the nsP4:P123 expression plasmid ratios from 1:1 to 8:1 resulted in a significant decrease in the activities of the EEEV replicase (Publication III, Figure 3C), in contrast to the replicases of CHIKV, SINV, BFV, EILV and VEEV (Publication II, Figure 3A). This indicates that for the EEEV replicase, there is an optimal ratio of P123 to nsP4, exceeding which results in a reduction in replicase activity. We confirmed this hypothesis by performing titration experiments with the EEEV P123 component. In this experiment, increasing the P123:nsP4 expression plasmid ratio from 1:10 to 1:1 was observed to have minimal or no effect on the activities of the EEEV replicase (Publication III, Figure 3D). However, an additional increase in the P123:nsP4 expression plasmid ratio from 1:1 to 8:1 resulted in a drastic, approximately 370-fold decrease in replication and a 125-fold decrease in transcription (Publication III, Figure 3D).

Naturally, we questioned if EEEV was an exception, and while this possibility cannot be excluded, we favor an alternative explanation. Recently, a study revealed that in the RNA replicase core, nsP1, nsP2 and nsP4 are present at a fixed 12:1:1 molar ratio (98). Most likely, the replication efficiency correlates with the number of active replication complexes formed in the transfected cells. It is obvious that an insufficient amount of nsP1, nsP2 or nsP4 hampers the formation of replication complexes. It is also conceivable that a large excess of either nsP4 or P123 processing products can be harmful as well. Furthermore, in addition to optimal amounts of P123 and nsP4, the components must also find and interact with each other in a timely manner. This interaction depends not only on the amount of the proteins but also on their stability (i.e., more stable proteins have more time to interact with each other). As described in Chapter 1.3.4., nsP4 is a rather unstable protein with the tendency to misfold and aggregate. It can be speculated that in cells, the nsP4 proteins that are more stable conformationally have an extended period of time to find the nsP1 and nsP2 proteins and interact with them, while the less stable nsP4 proteins need to interact more rapidly to acquire an active conformation and avoid degradation. Thus, nsP4 proteins could be more stable for viruses that require an optimal ratio of nsP4 to P123 components, such as EEEV. On the other hand, viruses with replication activities that are only

increased (no negative effect observed) by excess nsP4 could have a less stable nsP4 protein, such as all other viruses analyzed in the titration experiments. Taken together, it is possible that an optimal nsP4:P123 ratio is required for all alphaviruses and for most of them, this ratio was not achieved in the experiments described in Publication II.

Higher activities of the RNA replicase resulting from the increased amount of nsP4 could be attributed to a larger number of cells in which replication is initiated, to more efficient replication in such cells and to the combination of both of these things. To differentiate between these possibilities, plasmids expressing CHIKV and EEEV template RNAs were modified. Namely, the sequence encoding Gluc was replaced with a sequence encoding a fluorescent ZsGreen reporter. This modification allowed us to perform flow cytometry analysis of the transfected cells (Figure 15). To obtain more marker-positive cells required for this analysis, we used a 24-well plate format instead of the standard 96-well format and the LipoFectamine LTX and PLUS reagents (Thermo Fisher Scientific) as opposed to the FuGENE reagent (Promega) used in previous experiments. To account for differences that might result from the use of a modified template RNA and altered experimental conditions, we first performed nsP4:P123 titration experiments with the purpose of measuring replication (Fluc) activities using the new CHIKV-Fluc-ZsGreen template. In this experiment, the replication of the CHIKV template RNA reached maximal levels at a nsP4:P123 expression plasmid ratio of 2:1, remained at the same level of activity until a ratio of 4:1 and slightly decreased at ratios of 6:1 and 8:1 (Publication II, Figure 3B). The replication efficiencies of the templates harboring either Gluc or ZsGreen markers displayed slight differences (compare Publication II Figures 3A and 3B), which may be caused by several things. First, there is an unlikely option that the high levels of the ZsGreen reporter were harmful to the cells. Second, the differences may result from technical differences in the second experimental setup, i.e., from the combination of different template plasmids and the use of different (more efficient) transfection reagents. This might seem as a relatively unimportant technical issue, but, if this indeed is the case, it would align well with the differences observed in the EEEV replicase titration experiment (which was also done using the LipoFectamine instead of FuGENE *versus* the titration experiments for the replicases of other alphaviruses). Taken together, all changes in the second experiment setup (with the ZsGreen templates) were made with the aim of increasing the efficiency of RNA replication. Thus, it is plausible, and even very likely, that by doing so, we reached, and ultimately passed, the optimal nsP4:P123 ratio for the CHIKV replicase that was not reached in other experiments described in Publication II. If so, these findings support our hypothesis that viruses other than EEEV also have optimal nsP4:P123 ratios. Flow cytometry analysis also revealed that the percentage of ZsGreen-positive cells increased when the nsP4:P123 expression plasmid ratio was increased from 1:10 to 1:1, confirming that increasing the amount of nsP4 results in a higher number of cells in which RNA replication is initiated. However, from nsP4:P123 expression plasmid ratios of 1:1 to 8:1, the percentage of ZsGreen-positive cells declined (Publication II, Figure 3B). Thus, either large

volumes of plasmid DNA inhibited the transfection efficiency or/and the increased amounts of nsP4 were toxic to the cells. More importantly, however, the mean fluorescence intensity (MFI) was increased approximately 4.7-fold from the nsP4:P123 expression plasmid ratio of 1:10 to the ratio of 2:1. This finding confirms that the increased replication activities in the case of larger volumes of nsP4 result from both a higher number of cells in which replication is initiated and replication being more efficient in such cells. Interestingly, a slight decline in MFI was observed at nsP4:P123 expression plasmid ratios of 6:1 and 8:1. Although the effect was not statistically significant, it supports the above presented hypothesis, which is similar to that of the *trans*-replicase of EEEV: the activity of the CHIKV replicase (and, presumably, those of other alphaviruses) reaches a maximal level at a certain nsP4:P123 ratio and declines at ratios exceeding this level.

Analysis performed using plasmids expressing the EEEV template RNA and 2C replicase revealed identical trends. Again, the percentage of ZsGreen-positive cells increased when the nsP4:P123 expression plasmid ratio was increased from 1:10 to 4:10. Increasing the amount of the nsP4 expression plasmid up to a ratio of 1:1 had no further positive impact on the percentage of ZsGreen-positive cells. At the same time, increasing the nsP4:P123 expression plasmid ratio from 2:1 to 8:1 resulted in a severe decline in both the percentage of ZsGreen-positive cells and the MFI (Publication III, Figure 3E). Again, these data are consistent with the hypothesis that the maximal activity of the EEEV replicase requires an optimal nsP4:P123 ratio. We also transfected HEK293T cells with a constant amount of the template RNA and nsP4 expression plasmids while varying the amount of the P123 expression plasmid. We observed that the increase in the P123:nsP4 expression plasmid ratio from 1:10 to 1:1 resulted in a slight increase in the percentage of ZsGreen-positive cells while having minimal to no effect on the MFI (Publication III, Figure 3E). However, again, increasing the P123:nsP4 expression plasmid ratio from 1:1 to 8:1 had a debilitating effect on the percentage of ZsGreen-positive cells and the MFI, similar to what we observed in the experiment with the EEEV template RNA harboring the Gluc marker.

Altogether, the experiments performed with the 2C replicases and the ZsGreen-expressing template RNAs demonstrated that changes in the activities of the *trans*-replicases resulting from different nsP4:P123 ratios are causative of both the number of cells in which the replication is initiated and the efficiency of the RNA replication being higher in such cells. Finally, replicase component titration experiments confirmed that the critical factor for alphavirus RNA replication efficiency is the amount of nsP4 protein. However, it cannot be characterized as the-more-the-better principle. Instead, the maximal efficiency of replicase complex formation requires an optimal ratio of the P123 and nsP4 components.

4.2.2. Many heterologous combinations of P123 and nsP4 components can form a functional alphavirus RNA replicase

As described in the previous chapter, alphavirus RNA replicases can successfully be reconstructed from matching P123 and nsP4 components. The high activities of these 2C *trans*-replicases present a great opportunity for analyzing the compatibility of replicase proteins from different alphaviruses. Existing data obtained using chimeric alphavirus genomes suggested that the combinations of replicase proteins from closely related alphaviruses could be functional (23, 201). To confirm whether functional heterologous replicases can be formed of components originating from both closely and distantly related alphaviruses and to reveal the rules behind the compatibility/incompatibility of those components, we aimed to test the activities of all heterologous P123 and nsP4 combinations. However, testing all combinations of the P123, nsP4 and template RNA constructs from ten alphaviruses was technically impossible due to the large number of combinations. To simplify the initial analysis, an advantage was taken from our previous finding that the SINV template RNA is promiscuous and can be replicated by replicases of most alphaviruses (Publication I, Figure 5A). Thus, it was decided to perform component compatibility analysis in HEK293T cells using the SINV template RNA (with the Fluc and Gluc markers). This approach, however, was not ideal, as the replicases of some alphaviruses are less effective on the SINV template RNA compared to their own RNA. This created the possibility that the activities of some heterologous P123 and nsP4 combinations might be low specifically on the template RNA of SINV, essentially creating false-negative results. Therefore, some key findings were verified using template RNAs matching the P123 or nsP4 components used.

CHIKV P123 was found to be compatible with all nsP4s except for that of BFV (which was found to be active only with its matching P123). Interestingly, replicases formed from the CHIKV P123 and the SFV or ONNV nsP4s were even more active than the CHIKV 2C replicase (Publication II, Figure 4A). At the same time, the replicase formed from ONNV P123 and CHIKV nsP4 had very low activity (Publication II, Figure 4B). Furthermore, ONNV P123 was unable to form active replicases with any of the nsP4s, including its matching nsP4. Partly, this might have resulted from the fact that the SINV template RNA is not ideal for the ONNV RNA replicase (Publication I, Figure 5A). However, we had previously also observed that the ONNV 1C and 2C replicases have intrinsically low activity even on the matching template RNA, making them one of the least active *trans*-replicases studied here (Publication I, Figure 2B; Publication II, Figure 2B). At the same time, it was observed that the ONNV nsP4 was able to form very active replicases not only with the CHIKV P123 but also with the P123s of other viruses belonging to the Semliki Forest complex (Publication II, Figures 4A–4D). This clearly indicates that the low activity of the ONNV replicase is caused by its P123 component.

To continue, SFV and RRV P123s were generally able to form active replicases with all heterologous nsP4s other than that of BFV (Publication II, Figures 4C and 4D). In addition, the boost in transcription was also very low by

the replicase formed from the SFV P123 and the RRV nsP4 (Publication II, Figure 4C). Interestingly, the MAYV P123 displayed limited capacity to form active replicases with heterologous nsP4s, while the MAYV nsP4 formed very active replicases with the CHIKV, RRV and SFV P123s (compare Publication II, Figures 4A, 4C–E). Thus, the relatively low activity of the MAYV replicase, similar to ONNV, originates from its P123 component. However, the CHIKV, SFV and RRV P123 components (i.e., viruses in which the 2C replicases were highly active, Publication II, Figure 2B) were generally capable of forming active replication complexes with heterologous nsP4 components. Thus, the low activities of some heterologous replicases consisting of the P123 and nsP4 components of different Semliki Forest complex viruses mostly originated from intrinsically poor activities associated with the ONNV and MAYV P123s.

The situation for the outgroup viruses turned out to be different and more complex. BFV P123 formed active heterologous replicases with the nsP4 proteins from EILV, VEEV, EEEV, MAYV, RRV and SFV and was poorly compatible only with the CHIKV and ONNV nsP4s (Publication II, Figure 5A; Publication III, Figure 5D). In contrast, the P123 components of other outgroup viruses demonstrated a high rate of incompatibility with heterologous nsP4s (Publication II, Figure 5B–5D). In the presence of the SINV template RNA, EILV P123 was incompatible with all nsP4s (Publication II, Figure 5B; Publication III, Figure 5D). However, EILV nsP4 formed active replicases with several heterologous P123s (Publication II, Figures 4A–4E and 5A). The same was observed for VEEV, EEEV and SINV nsP4s. All these nsP4s could form active replicases with the P123 of BFV and several viruses from the Semliki Forest complex (Publication II, Figure 4A–4E, Figure 5A; Publication III, Figure 5D) but not with the P123 components of other members of this subgroup (Publication II, Figure 5B–5D; Publication III, Figure 5D). The P123s of SINV, VEEV and EEEV were highly selective and generally failed to form active replicases with any of the other heterologous nsP4s (Publication II, Figures 5C–D; Publication III, Figure 5B). We observed previously in Publication I that the VEEV replicase presented low compatibility with the SINV template RNA (Publication I, Figure 5A). Therefore, we questioned whether the incompatibility of VEEV P123 might be caused by using the SINV template RNA. To study this directly, the VEEV P123 was tested together with matching pairs of template RNAs and nsP4s from different viruses (Publication II, Figure 5E). Once again, only the matching 2C VEEV replicase displayed high replication and transcription activities. A similar experiment performed using the SINV P123 produced almost identical results (Publication II, Figure 5F). Thus, the incompatibility of the SINV and VEEV P123 components with heterologous nsP4s was considered to be independent of the template RNAs used. However, it should be mentioned that while the SINV P123 was also tested together with its matching template RNA and all heterologous nsP4s (Publication II, Figure 5D), such an experiment was not performed for the VEEV P123. The analysis performed later using the EEEV P123 suggested that this experiment might have been important. Namely, in the presence of the SINV template RNA, the EEEV P123 behaved very similarly to

that of VEEV P123 by displaying a limited ability to form active replicases with heterologous nsP4s (Publication III, Figure 5B). However, in the presence of its own template RNA, EEEV P123 formed active replicases not only with the matching nsP4 and VEEV nsP4 but also with SINV, MAYV, SFV and RRV nsP4s (Publication III, Figure 5A). Thus, the ability of P123 of this outgroup virus to form active heterologous replicases is, to an extent, dependent on the source of template RNA. These data also suggest that EEEV P123 might have an important role in the recognition of its template RNA.

4.2.3. The two-component trans-replicases consisting of heterologous components accurately portray the characteristics of corresponding chimeric viruses

Although *trans*-replicases are easy-to-use and sensitive tools, they are still artificial systems. Thus, we questioned whether the findings made using *trans*-replicases are also reproduced in the context of virus infections. Previous studies from our lab have revealed that this is indeed the case for mutations and insertions introduced into ns proteins (140, 199). Similarly, our data on the cross-utilization of the template RNAs (Publication I) are supported by previous findings made using alphavirus replicons (28). Here, we wanted to analyze whether the same applies to replicases containing the P123 and nsP4 components originating from different alphaviruses. Constructing the chimeric viruses for all 10 alphaviruses from our *trans*-replicase panel would have been impractical and, in some cases, ethically questionable (due to the need for using chimeras of highly pathogenic viruses). Therefore, we decided to limit this assay to two viruses, the P123 components of which displayed contrasting properties in our previous experiments. In the previous chapter, it was demonstrated that the SFV P123 can form active RNA replicases with the SINV and CHIKV nsP4s (and with those of several other viruses), while the replicases formed with the BFV and RRV nsP4s had rather low activities (Publication II, Figure 4C). In contrast, SINV P123 only formed active replicases with its matching nsP4 (Publication II, Figure 5D). Thus, we wondered if the same can be observed in the context of viruses with chimeric genomes.

The chimeras were constructed using the SFV and SINV icDNAs. Simply replacing the sequences encoding the SINV and SFV nsP4 proteins with their heterologous counterparts was, however, not possible, as this would have also resulted in swapping the SG promoters: corresponding sequences are mostly located in the region encoding the C-terminus of nsP4. As a solution, chimeric genomes were constructed by eliminating the heterologous SG promoters from the sequences of nsP4s via the introduction of synonymous mutations while preserving the native SFV and SINV SG promoter sequences. To account for any effects caused by this construction strategy, similar genomes were also constructed using matching SFV and SINV nsP4 regions and were designated SFV-SFV4 and SINV-SINV4, respectively. The construction of chimeric virus genomes and the design of subsequent experiments are also explained in Figure 18.

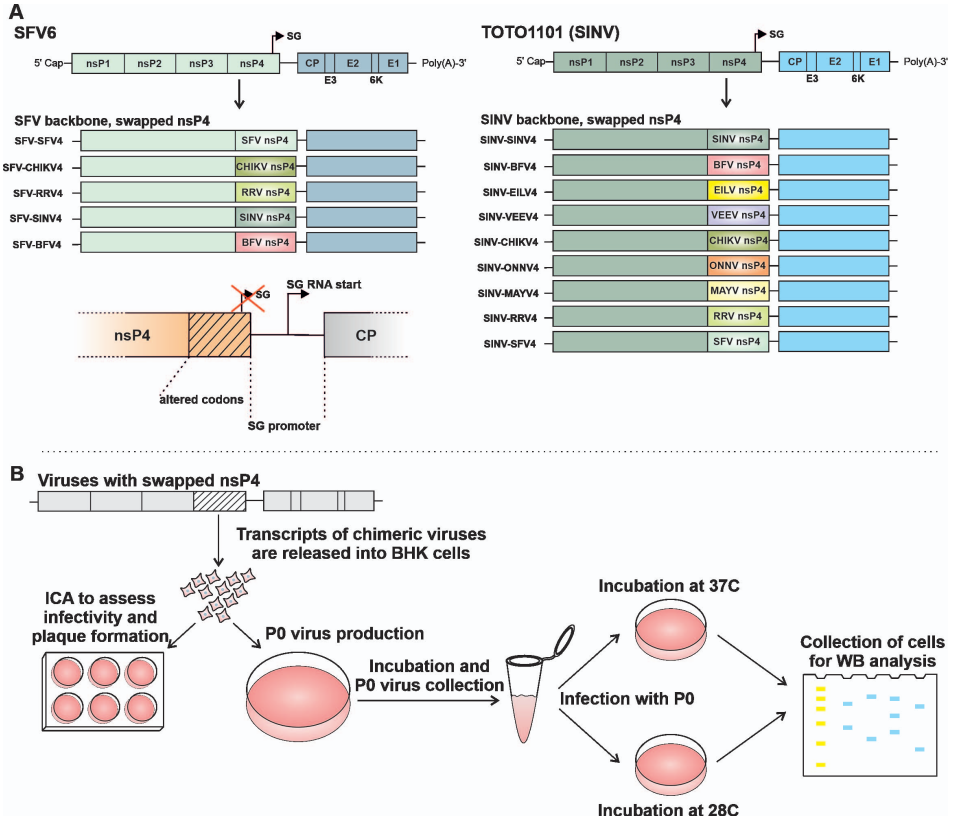

Figure 18. Analyses of the SFV and SINV genomes harboring the nsP4 regions of different alphaviruses. A. Cloning strategy used to obtain a panel of chimeric SFV and SINV genomes. B. Workflow of the chimeric virus rescue, propagation and infectivity experiments.

Using the approaches described and displayed above, chimeric SFV genomes in which the nsP4-encoding region was swapped with that of SINV, CHIKV, BFV or RRV were constructed. An infectious center assay (ICA) performed in BHK-21 cells confirmed that the infectivity of SFV-SFV4 was similar to that of wt SFV6, indicating that the developed design had no apparent flaws (Publication II, Table 2). However, none of the other chimeras were capable of plaque formation in the ICA. Nevertheless, development of cytopathic effects (CPEs) was observed in BHK-21 cell cultures transfected with these icDNAs. Importantly, CPE appeared earlier for the viruses harboring the CHIKV or SINV nsP4 and later for the viruses harboring the BFV or RRV nsP4. These data correlate well with the activities of corresponding heterologous *trans*-replicases. The replication of the virus genomes was analyzed by detecting the expression of the nsP2 protein, and virus rescue was confirmed by the detection of capsid protein by western blot analysis. Notably, the rescue performed at 28 °C was also successful (Publication II, Figure 6B). However, only minimal to modest CPE was observed (Publication II, Table 2). The chimeras based on the SINV genome displayed somewhat

different properties. The formation of plaques in the ICA was observed only for wt SINV and the SINV-SINV4 control construct (Publication II, Table 2). These two viruses were also the only viruses that could be rescued at 37 °C (Publication II, Figure 6B). Thus, similar to the *trans*-replicase assay (Publication II, Figure 5D), the chimeric viruses also confirmed that the SINV P123 can only form an active replicase with its matching nsP4. Interestingly, however, when the transfected cells were incubated at 28 °C, strong CPE was observed for the chimeras harboring BFV, EILV, CHIKV, ONNV or MAYV nsP4, and mild to moderate CPE was also observed for viruses harboring RRV or VEEV nsP4. The only construct for which the development of CPE was not detected was the chimera with the SFV nsP4 (Publication II, Table 2). However, even for this one, the infectious virus rescue was confirmed by western blot analysis (Publication II, Figure 6B).

Taken together, the results obtained using the chimeric genomes clearly correlated with the results of the *trans*-replicase assay. The fact that all chimeras based on the SINV genome could be rescued at reduced temperature indicates that the incompatibility of the SINV P123 with heterologous nsP4 components is a temperature-dependent effect. It can be speculated that the interaction between SINV nsP1 (and/or nsP2) and heterologous nsP4s is relatively weak, and at 37 °C, the formation of a stable core structure of the RNA replicase is hampered. At reduced temperatures, however, the interactions between these proteins are strong enough to allow for the formation of functional replication complexes. Structural studies using cryo-EM are required to reveal whether this is indeed the case.

4.2.4. Further splitting of the alphavirus replicase polyprotein drastically reduces the efficiency of RNA replication and transcription

The two-component (2C) *trans*-replicases proved to be sensitive, efficient and relevant tools for studying the compatibility of heterologous P123 and nsP4 components. To extend the analysis to the compatibility of the nsP1, nsP2 and nsP3 proteins, splitting of the P123 component would be needed. However, several studies reviewed in the Literature section have demonstrated that individual nsP1, nsP2, nsP3, and nsP4 cannot form a functional replicase. Furthermore, splitting of P123 into P12 and nsP3 would result in a similar outcome. Thus, the only option was to split P123 into nsP1 and P23 components. However, the P23 polyprotein is unstable and extremely short-lived. Therefore, we suspected that if we supplied the cells with nsP1, P23 and nsP4 components, the outcome would be similar to supplying the cells with four separate ns proteins. For these reasons, we decided to follow a strategy previously described for SFV. For SFV, it has been demonstrated that P123, in which the nsP2 protease is inactive ($P12^{CA}3$), is able to form active replication complexes in the presence of nsP4. In addition, the formation of the spherules is preserved even when the SFV $P12^{CA}3$ is split into nsP1 and $P2^{CA}3$ (96). Thus, we considered that the formation of active replicases from nsP1, $P2^{CA}3$ and nsP4 expressed from separate plasmids (Figure 14) should also be possible for other alphaviruses and constructed three-component (3C)

trans-replicases for SINV, RRV, CHIKV and EEEV. Analysis of their functionality in HEK293T cells, however, failed to reveal high activities similar to those reported for the SFV 3C replicase. Instead, the activities of the RRV 3C replicase remained close to the detection limit (Publication II, data not shown), and the SINV, CHIKV and EEEV 3C replicases had very low replication activities and strongly reduced (compared to the 2C replicases) transcription activities (Publication II, Figure 7D; Publication III, Figure 4B).

As seen previously (Chapter 4.2.1.), the activities of the 2C replicases can be drastically increased by optimizing the nsP4:P123 ratio (Publication II, Figure 3). Therefore, the same approach was used for the 3C replicases. It was observed that the activity of the 3C replicases also depends on the nsP4:(nsP1+$P2^{CA}3$) ratio. The trend was similar to the 2C replicases, however, the 3C replicases failed to achieve replication and transcription activities comparable with the levels obtained from the 2C replicases (Publication II, data not shown; Publication III Figures 4C and 4E). Flow cytometry analysis revealed that the low activities of the 3C replicases were due to both a small percentage of the cells in which replication was initiated, (ZsGreen-positive cells) and the low efficiency of RNA replication (low MFI) in these cells (Publication III, Figure 4D).

The activity of the 3C replicase may also depend on the nsP1:($P2^{CA}3$+nsP4) ratio or the $P2^{CA}3$:(nsP1+nsP4) ratio. In the core of the mature RNA replicase, the nsP1:nsP2:nsP4 ratio is 12:1:1 (97, 98). Therefore, we attempted to improve the activities of the 3C CHIKV replicase by changing the nsP1:($P2^{CA}3$+nsP4), $P2^{CA}3$:(nsP1+nsP4) or nsP4:(nsP1+$P2^{CA}3$) ratios. These experiments revealed that increasing the amounts of nsP1 and nsP4 had similar effects on the transcription efficiency of CHIKV. Increasing the amount of nsP4 in the nsP1:$P2^{CA}3$:nsP4 ratio from 10:10:2.5 to 1:1:2.5 also increased the transcription activity by approximately 175-fold. Further increases in the amount of nsP4 resulted in a debilitating effect on the transcription efficiency (Figure 19B). Increasing the amount of nsP1 had a similar effect on transcriptional activity. Thus, changing the nsP1:$P2^{CA}3$:nsP4 ratio in the expression plasmids from 1:10:10 to 2.5:1:1 increased the transcription activities approximately 875-fold. Again, increasing the amount of nsP1 further resulted in a decrease in transcription efficiency (Figure 19A). However, this decrease was more modest compared to what was observed for the nsP4 component. This difference might be explained by the fact that in the replication complexes, nsP1 is present in excess compared to the other nsPs. Finally, increasing the amount of the $P2^{CA}3$ component by changing the nsP1:$P2^{CA}3$:nsP4 ratio in the expression plasmids from 10:1:10 to 10:2.5:10 increased the transcription activities by approximately 69-fold. However, again, the activities dropped drastically after reaching a ratio of 1:1:1 (Figure 19C). This drop could, theoretically, be attributed to the cytotoxic properties of the nsP2 protein, however, no CPE in transfected cell cultures was observed visually.

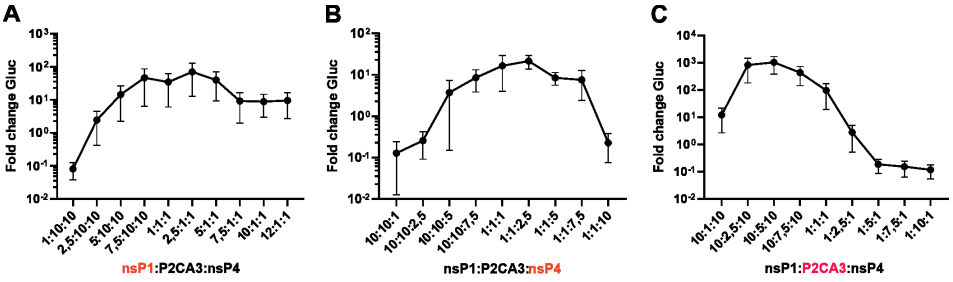

Figure 19. Impact of the different nsP1:P2^{CA}3:nsP4 ratios in the expression plasmids on the transcription activity of the CHIKV 3C replicase (unpublished data). A. The amount of nsP1 was changed by varying the nsP1:P2^{CA}3:nsP4 ratio from 1:10:10 to 12:1:1; B. The amount of nsP4 was changed by varying the nsP1:P2^{CA}3:nsP4 ratio from 10:10:1 to 1:1:10; C. The amount of P2^{CA}3 was changed by varying the nsP1:P2^{CA}3:nsP4 ratio from 10:1:10 to 1:10:1. For each panel, the component indicated in red was used in different amounts, while the amounts of the other two expression plasmids as well as the plasmid expressing template RNA were kept constant. Experiments were performed and data were analyzed as described for similar experiments in Publication III (Figure 4), however, a 48-well plate format was used as explained in the legend of Figure 17.

The new 3C replicase titration experiments revealed, using CHIKV as an example, that for each component, there is an optimal amount (relative to the other two components) that ensures maximal transcription efficiency. However, even if all three components were used in their optimal amounts, the transcription activity remained relatively low compared to the 2C replicase, and the replication efficiency was close to the background level. These properties eliminated the possibility of using the 3C replicases in analyzing the functional compatibility of the heterologous components. In addition, this result also triggered the question of why 3C replicases behave so differently from 2C replicases. We showed that high activities cannot be achieved by changing the ratios of the components of the 3C replicases. Thus, there must be a molecular reason(s) for the low activity. First, we previously observed that a replicase consisting of P12^{CA}3 and nsP4 demonstrated lower activities than wt P123 and nsP4 (Publication II). Thus, we wondered if the low activities might be caused by the lack of protease activity of nsP2. We refuted this hypothesis by making an additional P23 construct in which nsP2 regained its protease ability and harbored a short sequence from the end of nsP1 in its N-terminus to delay the processing of the 2/3 site. However, this 3C replicase demonstrated even lower activities (data not shown).

Another possibility is that there is nothing wrong with the functionalities of our three replicase components *per se*. Instead, by expressing the components separately, we might have compromised their localization in the transfected cells and forced them into a nonfunctional “long distance relationship”. Mutations in the 1/2 cleavage site have a strong negative impact on the activity of the *trans*-replicase and the infectivity of the corresponding transcript (69). Thus, it is likely that the nsP1 component of the 3C replicase localizes correctly to the plasma membrane. In addition, our data using 2C replicases indicate that nsP1 can inter-

act correctly with nsP4, which might occur at the plasma membrane or during the transport of nsP1 to the membrane. However, the transport of the $P2^{CA}3$ component might be hampered without efficient interactions with nsP1.

As of now, we have obtained evidence supporting this hypothesis. First, we discovered recently that a transiently expressed wt nsP1 can efficiently complement a nonfunctional P1234 harboring mutations in the MT/GT active sites (unpublished data). Thus, if the products of P1234 processing are located correctly, then the individual nsP1 supplied separately can find these products. As a result, wt nsP1 becomes incorporated into the functional replicase complexes. Second, we added a membrane attachment signal (palmitoylation–myristylation peptide) to the N-terminus of a construct expressing the P23 polyprotein (protease active) preceded by a fusion of the C-terminal end of the nsP1 encoding sequences (delnsP1-P23). This modification led to an approximately 10-fold increase in the activity of the corresponding 3C replicase (unpublished data). Therefore, we hypothesize that generating a construct in which the processing of P23 is slowed down and harbors a membrane attachment signal might result in a more active 3C replicase.

4.3. For EEEV and SINV, the P123 component of the RNA replicase has a leading role in template RNA recognition

We demonstrated previously (see Chapter 4.2.2.) that the P123 components of the outgroup alphaviruses are generally poorly compatible with heterologous nsP4s. These results were, however, mostly obtained using SINV template RNA. While this template was an excellent tool for determining the compatible P123 and nsP4 combinations for viruses from the Semliki Forest complex, this might not have been the case for the outgroup viruses. Therefore, it is possible that the P123 components of several outgroup viruses were not incompatible with the heterologous nsP4 components but with the SINV template RNA. Indeed, experiments with the 2C replicases containing either the SINV or EEEV P123 indicated that, at least for these viruses, P123 might be involved in template RNA recognition.

To study this directly, we tested the compatibility of EEEV P123 with all heterologous nsP4 components first on the EEEV template RNA and then on the SINV template RNA. On the EEEV template RNA, the matching 2C replicase was highly active in both replication and transcription. The RNA replicase consisting of EEEV P123 and VEEV nsP4 had similar or even higher activities (Publication III, Figure 5A). The activities were also high for heterologous replicases consisting of EEEV P123 and SINV, MAYV, and SFV nsP4s. Somewhat lower replication and transcription activities were observed for the combinations of EEEV P123 and EILV, RRV, or ONNV nsP4s. In this experiment, the only nsP4 proteins that were poorly compatible with EEEV P123 were derived from

CHIKV and BFV (Publication III, Figure 5A). When the same experiment was performed using the SINV template RNA, all heterologous replicases except for EEEV P123 of and VEEV nsP4 displayed distinctively lower activities (Publication III, Figure 5B). Altogether, the data obtained from these experiments showed that heterologous replicases containing the EEEV P123 are more active with its matching RNA template.

To confirm the importance of P123 and to study the possible role of EEEV nsP4 in the formation of active replicases, we performed a similar set of experiments as described above by combining EEEV nsP4 with all heterologous P123 components. This experiment revealed that although the boosts in Fluc and Gluc activities were slightly higher for the EEEV template RNA, the source of template RNA had very little impact on these 2C replicases (Publication III, Figures 5C and 5D). The only exception was the 2C replicase consisting of SINV P123 and EEEV nsP4. This combination was more efficient with the SINV template RNA, suggesting that the SINV P123 might also have an important role in template RNA recognition. Altogether, these results indicate that EEEV P123 plays a leading role in the recognition of the template RNA. In addition, data from Publication II also demonstrated that the 2C replicases containing the SINV P123 were more active on its matching template RNA (Publication II, Figure 5D and 5F). Therefore, it can be concluded that for EEEV and SINV, the P123 component has a leading role in template RNA recognition/utilization.

As of now, whether the findings made for EEEV and SINV also apply to other alphaviruses cannot be deduced with full certainty. We have demonstrated with the 2C replicases that VEEV P123 is highly incompatible with heterologous nsP4s on the SINV template RNA (Publication II, Figure 5C). However, as already noted before, the 2C replicases involving VEEV P123 were not tested on the VEEV template RNA. In addition, there are many discrepancies between different alphaviruses. For example, the BFV P123 component formed active heterologous replicases with different nsP4 components on the SINV template RNA (Publication II, Figure 5A). This was also a general trend for the 2C replicases containing the P123 components of Semliki Forest complex viruses (Publication II, Figure 4). As we did not perform extensive experiments using the matching templates for any of these viruses, the question of whether the match between the P123 and template RNAs contributes to the activities of these 2C replicases remains generally unanswered.

4.3.1. The SINV and EEEV P123 replicase components are responsible for the recognition/utilization of the 5′ region of the template RNA

The EEEV replicase modestly replicates and efficiently transcribes the SINV template RNA. At the same time, the SINV replicase shows low replication and transcription activities on the EEEV template RNA (Publication III, Figures 2A and 2B). Analysis performed using a set of SINV/EEEV chimeric templates

revealed that the SINV and EEEV replicases require the presence of the 5' regions of their respective genomes to launch efficient RNA synthesis (Figures 17B and 17C). The 2C replicases containing heterologous combinations of the P123 and nsP4 proteins of these viruses allowed us to extend this analysis. Thus, we decided to study which component of each replicase is responsible for the preference (i.e., recognition and/or utilization) of the 5' regions of their template RNAs.

The 2C replicase consisting of SINV P123 and EEEV nsP4 demonstrated low replication activity, moderate transcription activity and a clear preference toward the wt SINV template RNA (Publication III, Figure 5C and 5D). Consistently, the Fluc activities obtained using the chimeric templates remained close to the background level (Figure 20A). Therefore, no solid conclusion about the template preference could be made based on these data. However, the transcription of the chimeric templates was more efficient, and clear differences between the templates were observed. In all cases, the transcription signal was higher for the template RNAs harboring the 5' regions of the SINV genome (Figure 20A). Some importance of the SG promoter region was also detected in the transcription activities. Although transcription depends largely on the use of the SG promoter, replication of the template RNA appears to be an absolute prerequisite for SG RNA synthesis. Thus, it is logical that out of all the chimeric templates, the template RNA containing both the 5' and SG regions (SSE) demonstrated the highest activities for transcription. Taken together, our data suggest that SINV P123 most likely recognizes and/or utilizes both the 5' and SG regions in the template RNA.

The same analysis was also performed using the reciprocal variant of this 2C replicase, which further supported our previously stated conclusions. The high activity of the 2C replicase consisting of the EEEV P123 and the SINV nsP4 also allowed for analyzing the changes in Fluc activities (i.e., replication). The replication of the chimeric templates critically depended on the match between the P123 component and the 5' region of the template RNA (Figure 20B). Interestingly, the ESS and EES template RNAs were better than the ESE template RNA. This observation correlates well with the data obtained using the EEEV 1C replicase (Figure 17C). This effect might be caused by the incompatibility of the SG region of SINV with the 3' region of EEEV. Similar trends were also revealed for the transcription efficiencies: templates harboring the 5' region of the EEEV genome clearly outperformed templates where the region originated from the SINV genome (Figure 20B). As suggested above, this likely reflects the fact that more effective replication is a prerequisite for more efficient transcription. Thus, the EEEV P123 component of the replicase has a leading role in the recognition/utilization of the 5' region of the template RNA.

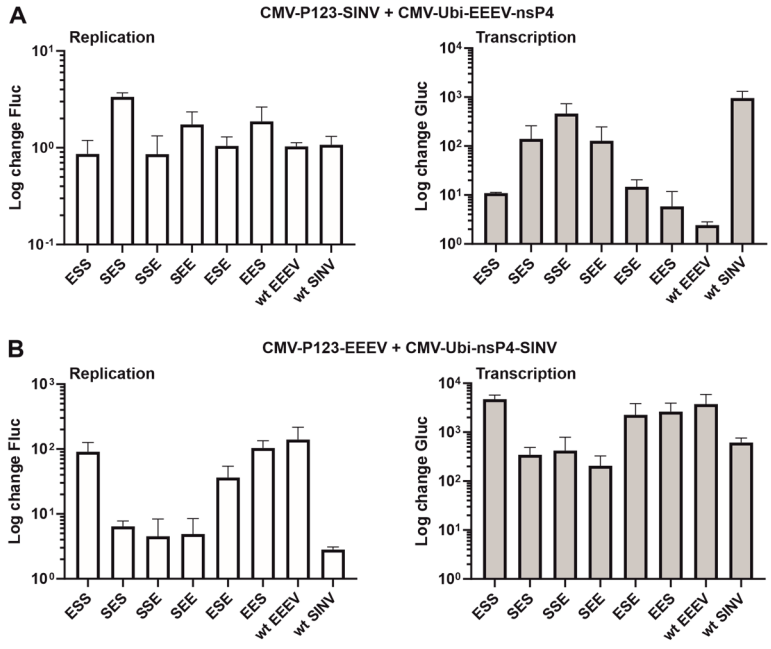

Figure 20. Mapping the regions responsible for the template preferences using the 2C replicases formed from EEEV and SINV P123 and nsP4 components (unpublished data). The experiment was performed, and the data were analyzed and presented as described for Figure 17 except that 2C replicases were used instead of 1C replicases. wt – wild-type EEEV or SINV template RNA.

4.3.2. For viruses belonging to the Semliki Forest complex, the nsP4 component has a leading role in the recognition/utilization of the template RNAs

The analysis of the determinants responsible for the template RNA specificity performed using the EEEV and SINV 2C replicases suffers from one important drawback: the heterologous 2C replicase consisting of SINV P123 and EEEV nsP4 had very low replication activity. Hence, in replication, the template preference of this replicase cannot be directly analyzed (Figure 20A). This might not be the sole pair of P123 and nsP4 components that works "unidirectionally" (meaning that the combinations of P123 and nsP4 are not active reciprocally). An ideal pair of viruses for the template specificity analysis would be one where the activities of all four 2C replicases – two matching and two heterologous combinations – are similar to each other. In this case, the differences in template RNA replication and transcription could be solely attributed to the template RNA preference of the replicase.

Thus, we decided that when analyzing the factors that are responsible for the template RNA preference, two conditions should be met: i) replicases of the analyzed alphaviruses should recognize and use each other's templates with different efficiencies (the larger the difference, the better), and ii) the P123 of

virus 1 should form an active replicase with the nsP4 of virus 2 and *vice versa* (preferably, these heterologous combinations should have activities similar to the homologous combinations). Based on a finding from Publication I that revealed that the SINV template RNA is universal (or at least acceptable) for the replicases of different alphaviruses, we proceeded by identifying which P123 and nsP4 combinations match the abovementioned requirements. The rationale of this experiment is also described in Figure 21.

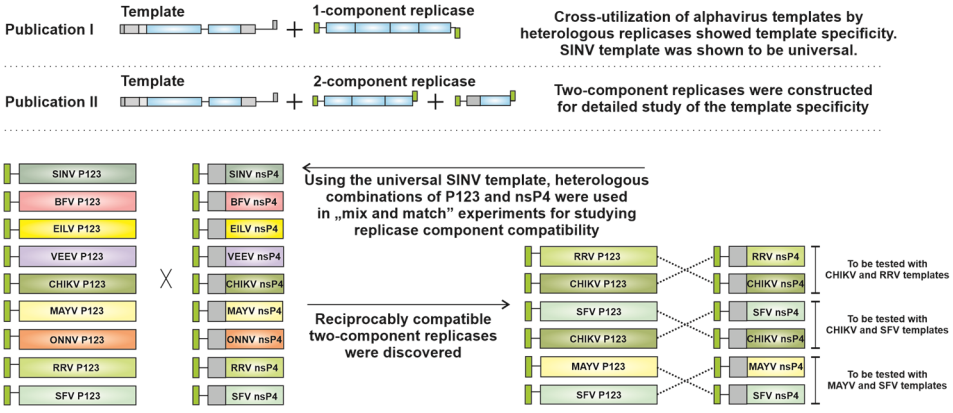

Figure 21. Schematic representation of the workflow of "mix-and-match" experiments resulting in the selection of viruses for which the determinants responsible for the template RNA preference were analyzed. In Publication I, heterologous combinations of replicases and templates in which the replicase proteins of one virus had a clear preference for their homologous template RNA were identified (not shown). It was also found that SINV template RNA is relatively universal and can be used by 1C and 2C replicases of different alphaviruses. Using SINV template RNA to screen for heterologous combinations of P123 and nsP4 components uncovered combinations that are capable of forming active replicase complexes in a reciprocal manner (i.e., both complexes had comparable activity). Three suitable pairs, CHIKV and RRV, CHIKV and SFV, and MAYV and SFV, were identified and used in the analysis of the template RNA preferences of these viruses.

Somewhat unexpectedly, the analysis performed using the SINV template RNA and the P123 and nsP4 components of ten different alphaviruses revealed that pairs of replicases that work reciprocally are rare. In fact, no such pair was found for alphaviruses belonging to different antigenic complexes. However, from the Semliki Forest complex, three combinations of P123 and nsP4 expression plasmids, largely corresponding to the criteria described above, were found: CHIKV P123, RRV nsP4 and *vice versa*, CHIKV P123, SFV nsP4 and *vice versa*, and (albeit with some reservations) MAYV P123, SFV nsP4 and *vice versa* (Figure 21).

On the SINV template RNA, the heterologous 2C replicases formed by the CHIKV and RRV P123 and nsP4 components were active reciprocally. However, both of the replicases displayed lower activities than the matching 2C replicases. The CHIKV template RNA was effectively used by its matching 2C replicase, while for the RRV 2C replicase, the activities were significantly lower (Publication II, Figure 8A). The replicase consisting of CHIKV P123 and RRV nsP4

could not replicate or transcribe the CHIKV template RNA effectively. In contrast, the activities of the 2C replicase consisting of RRV P123 and CHIKV nsP4 were high. Combined, these results indicate that the CHIKV nsP4 has a leading role in the recognition/utilization of both the genomic and SG promoters in the CHIKV template RNA. The analysis performed using the RRV template RNA resulted in similar data: the 2C replicase formed from CHIKV P123 and RRV nsP4 was significantly more active than the 2C replicase consisting of RRV P123 and CHIKV nsP4. Similar results were observed for both replication and transcription activities, again suggesting the leading role of the RRV nsP4 in the recognition and use of the genomic and SG promoters of the matching template RNA (Publication II, Figure 8A). Similar trends were also observed for the SFV and CHIKV replicase components (Publication II, Figure 8B) as well as for the MAYV and SFV replicase component (Publication II, Figure 8C). For the latter, the leading role of the nsP4 component could only be confirmed for the recognition and usage of the SG promoter. The problem was that the 2C replicase of SFV is significantly more active than its counterpart from MAYV. This difference was too prominent in the replication activities of the 2C replicases and masked the possible effects resulting from the preference of the genomic promoter (i.e., no matter what template was used, the 2C replicase of SFV was always more active). However, for transcription, the template preference was clearly determined by the nsP4 component (Publication II, Figure 8C). Altogether, these results indicate that for the Semliki Forest complex virus replicases, the nsP4 component has the leading role in recognizing/utilizing the SG promoter. The results were less clear for the recognition/utilization of the genomic promoter. However, it is very likely that the leading role in genomic promoter preference also belongs to the nsP4 component.

The most intriguing question resulting from our studies is whether Semliki Forest complex viruses use a different mechanism for template RNA recognition than other alphaviruses. At glance, the data seem to be very clear: for outgroup viruses the template RNA preference associates with the P123 component and for Semliki Forest complex viruses with the nsP4 protein. Indeed, it is possible that different mechanisms are used. The same has also been described for nucleocapsid formation, in which case the localization of the packaging signal (and possibly even its very nature) is different for outgroup and Semliki Forest complex viruses. However, it is also possible that for all alphaviruses, the primary recognition/utilization of the template RNA is determined by the P123 part of the replicase, and the viruses from the Semliki Forest complex have simply acquired an additional mechanism allowing them to differentiate between the template RNAs of viruses belonging to the same complex. To verify whether this is the case, one needs to perform experiments using heterologous 2C replicases consisting of the P123 and nsP4 components of viruses belonging to different serocomplexes. Unfortunately, these combinations always demonstrated unidirectional compatibility, and such experiments were therefore not performed. Alternatively, one could analyze the specificity of the interactions between template RNAs of alphaviruses and their replicase components. Unfortunately, thus far, such studies

have failed to produce any results, as in cell-free systems, the alphavirus replicase proteins display no selectivity in their interactions with RNAs (our unpublished data).

4.4. EEEV and CHIKV minigenomes represent tools that can be used in molecular biotechnology

Alphavirus-specific *cis*-elements in the template RNAs are recognized by their matching as well as multiple heterologous replicases. This allows for the replication and transcription of the template RNA and thus the expression of the reporters encoded by the template. A SINV minigenome, including a fluorescent reporter under the control of the SG promoter, has been shown to be inducible by viral infection (202). The activation of reporter expression, however, was approximately 10-fold, a modest effect that most likely can be attributed to either suboptimal design of the template RNA or to the low efficiency of the *trans*-replicase system in mosquito cells, a property that we also noted (Publication I). In contrast, a CHIKV minigenome identical to the template RNAs used in this study was shown to be activated by viral infection. Furthermore, the activation allowed for an approximately 30,000-fold boost in the expression of the reporter gene from the SG promoter of the template RNA (140). As many template RNAs developed in this study were highly active, we hypothesized that they could also serve as detection tools – or at least as prototypes of such tools – for the infection of unmodified wt alphaviruses.

The potential of using minigenomes as biosensors was evaluated using the CHIKV and EEEV template RNAs encoding the ZsGreen reporter. A set of experiments performed by cotransfecting HEK293T cells confirmed that these templates can be activated by replicase expression plasmids. Not surprisingly, the matching combination of the EEEV template and replicase resulted in the largest number of ZsGreen-positive cells (Publication III, Figure 6A). In addition, green cells were also observed for the combination of the EEEV template and SFV replicase. Importantly, ZsGreen expression was also activated in cells that were first transfected with the EEEV template RNA expression plasmid and subsequently infected with SFV (Publication III, Figure 6A). However, cells transfected with the EEEV template RNA expression plasmid and subsequently infected with Kunjin virus (genus *Flavivirus*) did not activate the production of the ZsGreen marker. Thus, the EEEV template RNA can only be activated by alphavirus infection the replicases of which can cross-utilize the EEEV template. Similar analysis was also performed with the CHIKV template RNA. Again, the production of the ZsGreen reporter was apparent when cotransfecting the cells with the template RNA and its matching replicase expression plasmids (Publication III, Figure 6B). Consistently, infecting the template-transfected cells with CHIKV resulted in the production of the ZsGreen reporter, while SINV infection did not induce reporter expression (Publication III, Figure 6B). This was exactly the outcome that was expected based on the inability of the SINV replicase to use

CHIKV template RNA (Publication I, Figure 4A). Taken together, the transfection–infection experiments with plasmids expressing the CHIKV and EEEV template RNAs revealed that these constructs serve as accurate and sensitive tools for detecting SFV and CHIKV infection. These, however, are both Old World alphaviruses. We were curious to also assess the ability of several New World alphaviruses to activate CHIKV and EEEV template RNAs. For biosafety reasons, these experiments were performed in the BSL3 facilities of the Defense Science and Technology Laboratory (Dstl, UK) by colleagues who have extensive experience in working with encephalitic New World alphaviruses. These experiments revealed that EEEV, VEEV and WEEV infections activate the ZsGreen reporter in cells transfected with the expression construct of the EEEV template. Interestingly, the EEEV RNA template was highly activated upon infection with EEEV and WEEV and to a much lesser extent by VEEV (Publication III, Figure 7A). Again, this is consistent with the data obtained in the cross-utilization experiments using the EEEV and VEEV replicases (Publication III, Figure 2A). Predictably, the CHIKV template RNA was used rather inefficiently (Publication III, Figure 7B). However, the trend of EEEV and WEEV infection resulting in higher expression of the ZsGreen reporter compared to VEEV infection remained. This is again well correlated with what we saw previously using the CHIKV template RNA and VEEV and EEEV replicases (Publication I, Figure 4A; Publication III, Figure 2B).

Taken together, these results indicate that the EEEV and CHIKV template RNAs are good biosensors for detecting infections with matching alphaviruses as well as infections with several related alphaviruses. The detection of virus infection via expression of fluorescent ZsGreen protein was in perfect correlation with the ability of the corresponding replicases to replicate and transcribe the used template RNAs. Thus, these biosensors have an opportunity to significantly simplify the characterization of these viruses in tissue culture. These biosensors could be used for real-time virus enumeration and quantification while limiting the need for virus manipulation. Perhaps more importantly, integration of such sensors into the genomes of vertebrate hosts or insect vectors would result in transgenic animals allowing for monitoring alphavirus infections using noninvasive methods. Such highly sensitive models would be especially useful for monitoring early stages of virus infection when the amount of virus is low and in the chronic stages of infection when virus replication, if it exists at all, is likely limited to a small number of cells. However, it is also clear that the great performance of the minigenomes in *in vitro* settings does not necessarily mean that the same will occur in *in vivo* models. There are several challenges one needs to overcome before truly effective *in vivo* models can be obtained. This includes (but, unfortunately, is not limited to) host and tissue specificity of alphaviruses, different replication rates in different cell types, low numbers of cells in which the replication of the minigenome is activated and possible adverse effects of the sensor (essentially defective interfering RNA) on virus replication. Thus, there is a long way to go before such advanced models can become available for studies of alphaviruses.

4.5. nsP4 as the central part of the alphavirus RNA replication machinery

It has long been well known that nsP4 is the catalytic subunit of the alphavirus RNA replicase. However, as an individual protein, it is nearly nonfunctional. Thus, nsP4 can be characterized as more of a team player than a solo artist, and to reveal how it is organized and how it functions, one needs to understand the context in which it performs its biological activities. Our studies (Publications I, II and III) represent important steps toward this end. We have emphasized the important roles of nsP4, whether regarding its role as the key factor in template RNA preference, its role as the rate limiting factor of RNA replicase activity, or the compatibility of nsP4 with other virus-encoded replicase subunits to define the formation and efficiency of an active RNA replicase. Furthermore, these studies revealed a number of small and seemingly unconnected details about the properties and functioning of the nsP4 proteins of different alphaviruses. However, when all these were combined with the rest of the functional data, they provided essential clues that ultimately allowed us to solve a long-standing problem: the crystallization of recombinant nsP4 and the determination of its 3D structure at high resolution. This achievement is important for understanding the biology of alphaviruses. In addition, this work provides rational bases for designing compounds that target the activities of nsP4 and subsequently allowed for solving the core structure of a functional alphavirus RNA replicase.

Over several decades, many attempts have been made to express an active recombinant nsP4 protein. This has also been a topic of study carried out in our research group since its establishment more than 20 years ago. The efforts of researchers have been, however, hampered by the intrinsic properties of nsP4, most notably the poor solubility of the recombinant proteins accompanied by their tendency to aggregate. In hindsight, these properties did, in fact, accurately reflect how nsP4 functions in cells and highlighted its need for interaction partners. Originally, however, they were regarded as hard-to-overcome obstacles. Our studies on the alphavirus RNA replicases revealed several facts that were helpful for our collaborators at Nanyang Technological University (Singapore) in their efforts to obtain soluble recombinant nsP4. Thus, the finding that the SINV nsP4 could form functional replicases with heterologous P123 components (Publication II, Figures 4 and 5) allowed for them to concentrate their efforts on this protein, as it is easier to work with than most other nsP4 proteins. The finding that RRV nsP4 can form functional replicases with CHIKV P123 allowed us to switch focus from the nearly insoluble CHIKV nsP4 to a more manageable RRV nsP4. Sure, on their own, these pieces of information would not have changed much. However, in combination with the skills and dedication of the research team of Dr Luo Dahai, they became important. Consequently, these studies resulted in the crystal structures of the RdRp domains for both SINV and RRV that were determined at 2.6 Å and 1.9 Å resolutions, respectively (Publication IV).

The N-terminal domains of these proteins could not be crystallized and were assumed (partly incorrectly) to be unfolded. The more conserved RdRp domains did crystallize, however, their structures revealed the presence of several regions that were not shown in the crystals. RRV RdRp was shown to be structurally very dynamic and contain several disordered segments, while SINV RdRp showed more order. More precisely, the crystal structure of RRV RdRp was missing a total of 83 residues, while only 39 residues were missing from SINV RdRp. Both RdRps demonstrated a right-hand fold characteristic of viral RdRps, with the finger-palm-thumb domains spanning sequentially from the N- to C-terminus of the protein (Figure 22; Publication IV, Figures 1B and 1C). The core catalytic domains of RRV and SINV RdRps were shown to be structurally well conserved. Nevertheless, some differences in the thumb and finger subdomains were identified: in the case of RRV RdRp, the fingertips extended from the fingers to the thumb, forming an encircled ring structure, while for SINV RdRp, the index fingertip remained detached from the core domain and extended over to the neighboring RdRp, forming a dimeric structure. However, this dimeric form of SINV RdRp was considered to be a crystal structure artifact, as alphavirus nsP4s are known to be monomeric (a fact that was later confirmed by resolving the structure of the RNA replicase core). The recombinant proteins corresponding to the RdRp domain and full-length recombinant nsP4 proteins were shown to be enzymatically active, although their activities were extremely low and very difficult to detect.

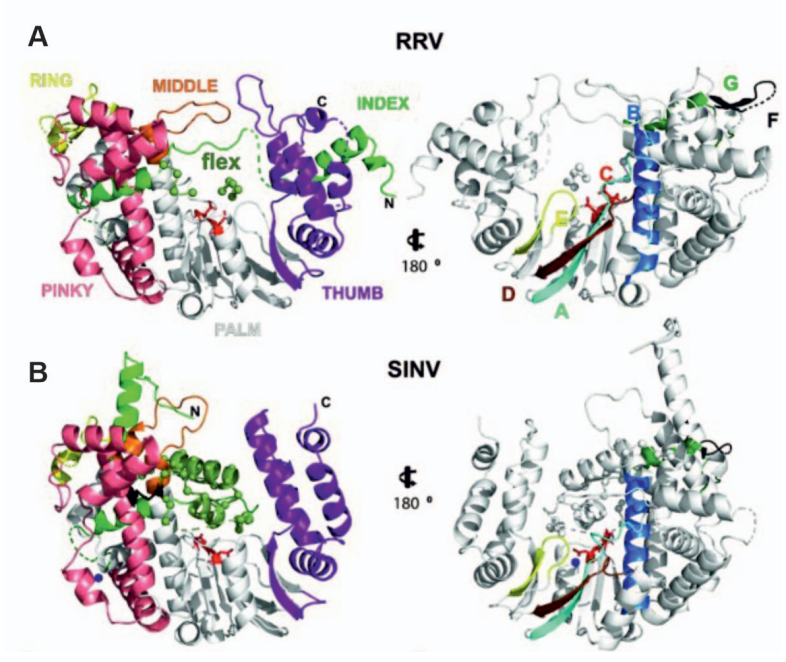

Figure 22. Crystal structures of the RdRp domains of RRV and SINV nsP4. Figure is adapted from (Publication IV, Figure 1).

The availability of the structures of the RRV and SINV nsP4 proteins resulted in a number of questions, for example: are the flexible regions of the RdRps essential for RNA replication (Publication IV, Supplemental Figure S8)? The importance of the revealed structural elements was therefore analyzed using structure-guided mutagenesis. As mentioned above, the enzymatic activities of the recombinant nsP4 proteins were very low, and therefore, using the recombinant proteins in such assays was impractical. On the other hand, introducing mutations that potentially have a negative impact on the replication to the infectious clones very likely would have resulted in the appearance of reversions or second-site compensatory mutations. However, this is not the case for the *trans*-replication system, which is why we decided to study the effects of these mutations using the SINV and RRV 2C *trans*-replicases. The mutations were introduced into nsP4 expression constructs and tested by cotransfecting U2OS cells with plasmids expressing nsP4 (including the mutations), matching P123 and the template RNA.

Overall, the results obtained for SINV and RRV were rather similar, which was not surprising considering the similar structures of their nsP4 proteins (Publication IV, Figures 5B and 5C). Substituting Arg63 and Arg67 (both located in the NTD of nsP4) with Ala residues was observed to have no significant effect on the replication activity of the RRV RNA replicase. However, this substitution did result in a minor but significant decrease in transcription (Publication IV, Figure 5B). Corresponding mutations in the nsP4 of SINV resulted in similar effects (Publication IV, Figure 5C). Inserting a Strep-FLAG tag after residue 92 of RRV nsP4 had a debilitating effect on both the replication and transcription of RRV. However, adding the same tag after residue 109 was tolerated better as the activities of the transcription decreased but not to the same extent as with the previous mutant. Thus, these results indicate that while the replacement of charged residues with alanines in the NTD of RRV nsP4 is well tolerated, inserting a Strep-FLAG tag results in impaired activities of the RRV replicase. Inserting a Strep-FLAG tag after residues 91 and 108 in SINV nsP4 also resulted in decreased replication and transcription activities, however, the decrease was less prominent than in the case of RRV (Publication IV, compare Figures 5B and 5C).

To continue, the interaction between the index finger and thumb regions of RRV nsP4 was tested by substituting the Cys125 residue with Val or Asp residues. Interestingly, while the RRV C125V mutant had replication and transcription activities similar to those of the wt replicase, the C125D mutant was characterized by an approximately 60-fold decrease in replication and an approximately 14-fold decrease in transcription (Publication IV, Figure 5B). Again, the same mutations introduced to the SINV nsP4 resulted in similar effects, albeit the decrease in replication and transcription observed for C124D (mutation that corresponds to C125D in RRV nsP4) was less prominent (Publication IV, Figure 5C). Next, the importance of the flexible loop structures mapped to residues 136–185 of RRV nsP4 and residues 135–184 of SINV nsP4 was studied by replacing these regions with a Strep-FLAG tag. This substitution had a debilitating effect on the activities of both SINV and RRV replicases, indicating the

importance of these regions in RNA replication (Publication IV, Figure 5B and 5C). The significance of this region in RRV was further confirmed by substituting the highly conserved Asp143 and Asp150 residues with alanines, which resulted in prominently decreased replication and transcription activities compared to the wt replicase. In contrast, substituting the Asp146 and Asp153 residues of RRV nsP4 with Ala residues had a more modest effect on the replication and transcription activities. Once again, similar results were obtained for SINV, where the respective Asp142 and Asp149 mutations also resulted in a decrease in the replicase activities, while the mutations of Asp145 and Asp152 to alanines had a less prominent effect (Publication IV, Figure 5C). This information clearly suggests that the flex region of the nsP4 protein plays a crucial role in the RNA replication in RRV and SINV (Publication IV, Figure 5B and 5C). Indeed, data that were subsequently obtained for the structure of the RNA replicase core revealed that in the presence of nsP1, this region is fully folded (98).

A charged helix tip in the thumb subdomain of RRV nsP4 was predicted to interact directly with RNA. The functional importance of this region was studied by substituting Asp543 and Asp545 with Ala residues. This substitution resulted in an approximately 25-fold decrease in replication and a 35-fold decrease in transcription activities (Publication IV, Figure 5B). To continue, the substitution of Arg546 and Arg548 residues with Ala also had a negative impact on replication and transcription. Interestingly, the substitution of Asp542 and Asp544 with Ala residues had only a minimal effect on the replication and transcription activities of SINV RNA replicase (Publication IV, Figure 5C). Finally, the extreme C-terminus of nsP4 (10 amino acid residues that remained unresolved in the crystal structures) was removed by adding a stop codon after codon 601 (RRV) or 600 (SINV). For RRV, this mutation resulted in the total loss of the activity of the RNA replicase, while some low-level transcription activity was detected for SINV replicase. Similar results were observed when the C-terminal tail was substituted by a Strep-FLAG tag. Collectively, these data indicated that the last 10 residues in the C-terminus of nsP4 are essential for the replication in RRV and SINV (Publication IV, Figures 5B and 5C).

As seen from our data, the unfolded (or flexible) regions, not resolved in the crystal structures of RdRps, are functionally highly important. The logical explanation is that these regions must be involved in the interactions with other components of the RNA replicase or, at the very least, stabilized and folded upon such interactions. This was subsequently confirmed by the analysis of the RNA replicase core structure, which was found to be a complex consisting of 12 molecules of nsP1, one molecule of nsP2 and one molecule of nsP4 (98). In this complex, the dynamically flexible nsP4 is stabilized by interactions with other replicase proteins, primarily with nsP1, in such a way that all its regions (including the N-terminal domain) display a clear fold (Figure 23; 98).

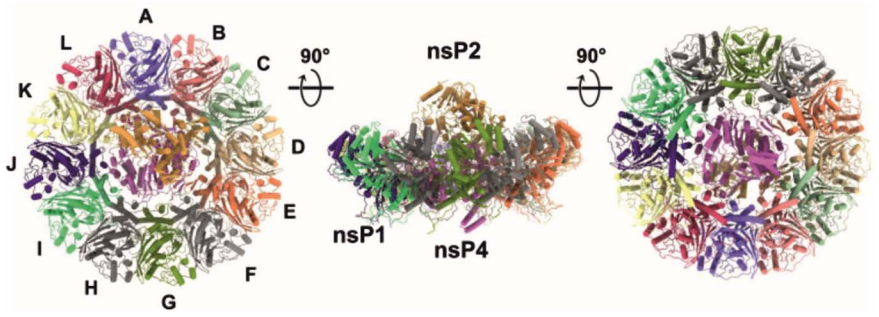

Figure 23. Graphical overview of the structure of the alphavirus replication core. Letters A to L represent each nsP1 subunit. Figure is adapted from (98).

The organization of the RNA replicase core is also highly logical and consistent with the needs of viral RNA synthesis and capping. nsP4 is quite literally the central part of the RNA replication of these viruses, as it occupies the central space in the nsP1 ring and interacts with much of everything else, including nsP2 and viral RNAs. Evidently, all elements of the RNA replicase core work in a coordinated manner. Furthermore, to achieve full enzymatic activity, functional interactions are required with elements located outside the RNA replicase core structure. The availability of these 3D structures and our 2C replicases generates possibilities to study the molecular bases of the compatibility and incompatibility of nsP4 and other replicase components. However, these studies are rather complex, as nsP1 and nsP4 form 10 contact surfaces that are different from each other. Accordingly, mutations in nsP1 can affect different interaction surfaces in a different manner. Altogether, the 2C replicases and mutant forms of P123 might represent tools for studying the biogenesis of the RNA replicase core. This kind of analysis is needed to shed light on the structure of an immature (early) replicase that may or may not be organized in a manner similar to that of the mature (late) replicase.

5. SUMMARY

In the studies that are included in this thesis, I performed extensive analysis on the RNA replication of 10 alphaviruses that differ in their geographical distribution, host and vector range and symptoms. By their most basic properties, the alphaviruses analyzed in these studies formed two clear groups, the viruses from the Semliki Forest antigenic complex and outgroup viruses originating from several other complexes. I have described the ability of the RNA replicases to cross-utilize heterologous template RNAs and revealed that the replicases of Semliki Forest group viruses can cross-utilize template RNAs of both Semliki Forest group and outgroup viruses. In contrast, the template RNAs of Semliki Forest group viruses were only used by the viral replicases belonging to the same group. The replicases of outgroup viruses cross-utilized the heterologous templates of the outgroup viruses, however, were incapable of using the template RNAs of the Semliki Forest group viruses. In contrast, the template RNAs of outgroup viruses were widely used by replicases originating from several complexes. Using incompatible combinations of replicases and templates, I demonstrated that for SINV and EEEV replicases, the 5' ends of their respective genomes are required for successful RNA replication.

I also reconstructed the RNA replicases of 10 alphaviruses from two functional components, P123 and nsP4, and demonstrated that such two-component replicases are highly active. In addition, the rate-limiting component of the two-component replicases was shown to be nsP4. Analysis of the requirements of the formation of alphavirus replication complexes by studying the compatibility of heterologous P123 and nsP4 components revealed that some replicases formed of the heterologous components are highly active. Interestingly, the P123 components of several alphaviruses were found to be highly specific toward matching nsP4s and template RNAs and were not compatible with components originating from different alphaviruses. Taking advantage of the split replicases and the ability of the alphavirus replicases to cross-utilize heterologous templates, I determined that the nsP4 component is a major determinant for the template RNA preferences of the Semliki Forest group viruses. On the other hand, EEEV and SINV P123 clearly had a leading role in template RNA specificity, suggesting that alphaviruses from various complexes might possess different mechanisms for recognizing template RNAs. I also demonstrated that the template RNA expression constructs used here can be exploited as a novel type of sensor for detecting alphavirus infections. It was found that the EEEV RNA template can be effectively used for detecting the infection of several New World alphaviruses. In addition, the CHIKV template was effectively used for detecting CHIKV infection.

Finally, I participated in a study that ended a decades-long string of unsuccessful attempts to resolve the 3D structure of the central part of the alphavirus RNA replicase: the RNA-dependent RNA polymerase (nsP4 RdRp). Crystal structures of these enzymes from SINV and RRV were determined, and the

importance of some key structural elements was confirmed via structure-guided mutagenesis.

All in all, I consider it highly important to conduct such comparative studies on the RNA replication of alphaviruses of distinct origin. The functional analysis of the properties of the RNA replicases performed here coupled with novel information on the fundamental principles of alphavirus RNA replicase formation provides insights into the possible recombination and evolution of alphaviruses. It also provides tools and knowledge that can be used for the development of advanced anti-alphavirus approaches.

6. SUMMARY IN ESTONIAN

Alfaviiruste RNA replikaasi keerukas loomus ning selle käsitlus

Selle väitekirja raames uurisin ma kümne erineva alfaviiruse RNA replikatsiooni. Uuringutesse kaasatud alfaviirused erinesid nii leviku, peremeesorganismide kui ka vektorite poolest ning jagunesid omaduste poolest kaheks grupiks. Esimene grupp koosnes viirustest, mis kuuluvad Semliki Forest viiruse (SFV) antigeensesse kompleksi ning teise grupi moodustasid viirused, mis kuuluvad erinevatesse kompleksidesse SFV grupi väliselt. Alustuseks uurisin alfaviiruste RNA replikaaside võimet replitseerida heteroloogseid RNA matriitse ning selgus, et SFV gruppi kuuluvate viiruste replikaasid kasutavad efektiivselt kõiki uuringusse kaasatud viiruste matriitse. SFV kompleksi RNA matriitse suutsid replitseerida siiski ainult samasse kompleksi kuuluvate viiruste replikaasid. SFV kompleksi väliste viiruste RNA matriitse suutsid kasutada replikaasid erinevatest kompleksidest, samas kui nende viiruste replikaasid ei olnud võimalised replitseerima SFV gruppi kuuluvate viiruste matriitse. Kasutades mittesobivaid matriitside ja replikaaside kombinatsioone, näitasin, et SINVi and EEEVi replikaasid vajavad matriits RNA replitseerimiseks enda genoomide 5’ otsa.

Järgnevalt rekonstrueerisin kümne erineva alfaviiruse RNA replikaasid kahest funktsionaalsest komponendist. Sellised replikaasid, mis moodustati eraldi ekspresseeritud P123 ning nsP4 komponentidest olid väga aktiivsed. Ma näitasin, et alfaviiruste RNA replikatsiooni efektiivsus sõltub nsP4 komponendi hulgast. Kasutades P123 ning nsP4 komponente, mis pärinevad erinevatelt alfaviirustelt, näitasin, et aktiivseid alfaviiruste replikaase on võimalik moodustada heteroloogsetest komponentidest. Lisaks tuvastasin, et mõnede viiruste P123 komponendid on väga spetsiifilised enda homoloogse nsP4 suhtes ning ei moodusta aktiivseid heteroloogseid replikaase mitte ühegi võõra nsP4 komponendiga.

Kahest komponendist rekonstrueeritavad replikaasid võimaldasid mul uurida, milline komponent replikaasist tunneb ära RNA matriitsi ning määrab, kas seda on võimalik replitseerida või mitte. Ma näitasin, et SFV grupi viiruste jaoks on kõige olulisem nsP4 komponent, samas kui EEEVi ning SINVi jaoks on määrav hoopis P123 komponent. Sellest võib järeldada, et erinevad alfaviirused on arenenud kasutama matriitside äratundmisel erinevaid mehhanisme.

Lisaks näitasin, et uuringute käigus konstrueeritud EEEVi ning CHIKVi matriitse on võimalik kasutada biosensoritena viirusnakkuste tuvastamiseks. EEEVi RNA matriits aktiveerub kui rakke nakatada erinevate Uue Maailma viirustega ning CHIKVi RNA matriitsi aktiveerib edukalt CHIKV nakkus.

Ma osalesin uuringus, mille käigus määrati esmakordselt SINV ning RRV viiruste RNA-sõltuva RNA polümeraasi 3D struktuur ning analüüsisin mitmete struktuuriliselt oluliste elementide rolli RNA replikatsioonil.

Sellised mitmekülgsed uuringud, mis keskenduvad erinevate alfaviiruste RNA replikatsiooni uurimisele on väga olulised. Käesoleva väitekirja raames läbiviidud RNA replikaaside omaduste funktsionaalne analüüs võimaldas koguda informatsiooni RNA replikaaside moodustumise, võimaliku rekombinatsiooni ning üldisemalt alfaviiruste evolutsiooni kohta. Lisaks, töös kogutud informatsiooni ning loodud töövahendeid on tulevikus võimalik kasutada uudsete alfaviirustevastaste vahendite väljatöötamiseks.

7. REFERENCES

1. Chen R, Mukhopadhyay S, Merits A, Bolling B, Nasar F, Coffey LL, Powers A, Weaver SC, ICTV Report Consortium. 2018. ICTV Virus Taxonomy Profile: Togaviridae. Journal of General Virology 99:761–762.
2. Lwande OW, Obanda V, Bucht G, Mosomtai G, Otieno V, Ahlm C, Evander M. 2015. Global emergence of Alphaviruses that cause arthritis in humans. Infect Ecol Epidemiol 5:10.3402/iee.v5.29853.
3. Sunil S. 2021. Current Status of Chikungunya in India. Frontiers in Microbiology 12.
4. Beltrame A, Angheben A, Bisoffi Z, Monteiro G, Marocco S, Calleri G, Lipani F, Gobbi F, Canta F, Castelli F, Gulletta M, Bigoni S, Del Punta V, Iacovazzi T, Romi R, Nicoletti L, Ciufolini MG, Rorato G, Negri C, Viale P. 2007. Imported Chikungunya Infection, Italy. Emerg Infect Dis 13:1264–1266.
5. Reiter P, Fontenille D, Paupy C. 2006. Aedes albopictus as an epidemic vector of chikungunya virus: another emerging problem? Lancet Infect Dis 6:463–464.
6. Enserink M. 2006. Massive Outbreak Draws Fresh Attention to Little-Known Virus. Science 311:1085–1085.
7. Chikungunya – France. https://www.who.int/emergencies/disease-outbreak-news/item/25-august-2017-chikungunya-france-en. Retrieved 1 December 2022.
8. 2022. Areas at Risk for Chikungunya | Chikungunya virus | CDC. https://www.cdc.gov/chikungunya/geo/index.html. Retrieved 1 December 2022.
9. Frickmann H, Herchenröder O. 2019. Chikungunya Virus Infections in Military Deployments in Tropical Settings – A Narrative Minireview. 6. Viruses 11:550.
10. Strauss JH, Strauss EG. 1994. The alphaviruses: gene expression, replication, and evolution. Microbiol Rev 58:491–562.
11. Baxter VK, Heise MT. 2020. Immunopathogenesis of alphaviruses. Adv Virus Res 107:315–382.
12. 2023. Eastern Equine Encephalitis | Eastern Equine Encephalitis | CDC. https://www.cdc.gov/easternequineencephalitis/index.html. Retrieved 12 June 2023.
13. Aguilar PV, Estrada-Franco JG, Navarro-Lopez R, Ferro C, Haddow AD, Weaver SC. 2011. Endemic Venezuelan equine encephalitis in the Americas: hidden under the dengue umbrella. Future Virol 6:721–740.
14. Diagne CT, Bengue M, Choumet V, Hamel R, Pompon J, Missé D. 2020. Mayaro Virus Pathogenesis and Transmission Mechanisms. Pathogens 9:738.
15. Powers AM, Brault AC, Shirako Y, Strauss EG, Kang W, Strauss JH, Weaver SC. 2001. Evolutionary Relationships and Systematics of the Alphaviruses. J Virol 75:10118–10131.
16. Facts about Sindbis fever. European Centre for Disease Prevention and Control. https://www.ecdc.europa.eu/en/sindbis-fever/facts. Retrieved 1 December 2022.
17. Ross River virus and Barmah Forest virus. https://www.healthywa.wa.gov.au/Articles/N_R/Ross-River-virus-and-Barmah-Forest-virus. Retrieved 13 June 2023.
18. Chikungunya worldwide overview [Internet]. European Centre for Disease Prevention and Control. [cited 2023 June 11]. Available from: https://www.ecdc.europa.eu/en/chikungunya-monthly
19. Chikungunya fact sheet. https://www.who.int/news-room/fact-sheets/detail/chikungunya. Retrieved 1 December 2022.
20. Weaver SC. 2018. Prediction and prevention of urban arbovirus epidemics: A challenge for the global virology community. Antiviral Res 156:80–84.

21. Atkins GJ, Sheahan BJ, Liljeström PY 1999. The molecular pathogenesis of Semliki Forest virus: a model virus made useful? Journal of General Virology 80:2287–2297.
22. Long KC, Ziegler SA, Thangamani S, Hausser NL, Kochel TJ, Higgs S, Tesh RB. 2011. Experimental transmission of Mayaro virus by Aedes aegypti. Am J Trop Med Hyg 85:750–757.
23. Saxton-Shaw KD, Ledermann JP, Borland EM, Stovall JL, Mossel EC, Singh AJ, Wilusz J, Powers AM. 2013. O'nyong nyong Virus Molecular Determinants of Unique Vector Specificity Reside in Non-Structural Protein 3. PLoS Negl Trop Dis 7:e1931.
24. 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.
25. 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.
26. Niesters HG, Strauss JH. 1990. Defined mutations in the 5' nontranslated sequence of Sindbis virus RNA. J Virol 64:4162–4168.
27. Niesters HG, Strauss JH. 1990. Mutagenesis of the conserved 51-nucleotide region of Sindbis virus. J Virol 64:1639–1647.
28. 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.
29. Kendall C, Khalid H, Müller M, Banda DH, Kohl A, Merits A, Stonehouse NJ, Tuplin A. 2019. Structural and phenotypic analysis of Chikungunya virus RNA replication elements. Nucleic Acids Res 47:9296–9312.
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. Wielgosz MM, Raju R, Huang HV. 2001. Sequence Requirements for Sindbis Virus Subgenomic mRNA Promoter Function in Cultured Cells. Journal of Virology 75:3509–3519.
33. Kuhn RJ, Hong Z, Strauss JH. 1990. Mutagenesis of the 3' nontranslated region of Sindbis virus RNA. J Virol 64:1465–1476.
34. Hardy RW, Rice CM. 2005. Requirements at the 3′ End of the Sindbis Virus Genome for Efficient Synthesis of Minus-Strand RNA. J Virol 79:4630–4639.
35. Madden EA, Plante KS, Morrison CR, Kutchko KM, Sanders W, Long KM, Taft-Benz S, Cisneros MCC, White AM, Sarkar S, Reynolds G, Vincent HA, Laederach A, Moorman NJ, Heise MT. 2020. Using SHAPE-MaP To Model RNA Secondary Structure and Identify 3′UTR Variation in Chikungunya Virus. Journal of Virology 94.
36. Rice CM, Strauss JH. 1982. Association of Sindbis virion glycoproteins and their precursors. Journal of Molecular Biology 154:325–348.
37. Lescar J, Roussel A, Wien MW, Navaza J, Fuller SD, Wengler G, Wengler G, Rey FA. 2001. The Fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenic activation at endosomal pH. Cell 105:137–148.

38. Voss JE, Vaney M-C, 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.
39. 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.
40. Wahlberg JM, Garoff H. 1992. Membrane fusion process of Semliki Forest virus. I: Low pH-induced rearrangement in spike protein quaternary structure precedes virus penetration into cells. J Cell Biol 116:339–348.
41. 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.
42. Wahlberg JM, Boere WA, Garoff H. 1989. The heterodimeric association between the membrane proteins of Semliki Forest virus changes its sensitivity to low pH during virus maturation. J Virol 63:4991–4997.
43. 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.
44. Zimmerman O, Holmes AC, Kafai NM, Adams LJ, Diamond MS. Entry receptors – the gateway to alphavirus infection. J Clin Invest 133:e165307.
45. Holmes AC, Basore K, Fremont DH, Diamond MS. 2020. A molecular understanding of alphavirus entry. PLoS Pathog 16:e1008876.
46. 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.
47. 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.
48. Klimstra WB, Nangle EM, Smith MS, Yurochko AD, Ryman KD. 2003. DC-SIGN and L-SIGN Can Act as Attachment Receptors for Alphaviruses and Distinguish between Mosquito Cell- and Mammalian Cell-Derived Viruses. J Virol 77:12022–12032.
49. Moller-Tank S, Kondratowicz AS, Davey RA, Rennert PD, Maury W. 2013. Role of the Phosphatidylserine Receptor TIM-1 in Enveloped-Virus Entry. J Virol 87:8327–8341.
50. 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 (NRAMP) is a cellular receptor for Sindbis virus in both insect and mammalian hosts. Cell Host Microbe 10:97–104.
51. 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.
52. Ludwig GV, Kondig JP, Smith JF. 1996. A putative receptor for Venezuelan equine encephalitis virus from mosquito cells. J Virol 70:5592–5599.
53. Ma H, Kim AS, Kafai NM, Earnest JT, Shah AP, Case JB, Basore K, Gilliland TC, Sun C, Nelson CA, Thackray LB, Klimstra WB, Fremont DH, Diamond MS. 2020. LDLRAD3 is a receptor for Venezuelan equine encephalitis virus. Nature 588:308–314.
54. Clark LE, Clark SA, Lin C, Liu J, Coscia A, Nabel KG, Yang P, Neel DV, Lee H, Brusic V, Stryapunina I, Plante KS, Ahmed AA, Catteruccia F, Young-Pearse TL, Chiu IM, Llopis PM, Weaver SC, Abraham J. 2022. VLDLR and ApoER2 are receptors for multiple alphaviruses. 7897. Nature 602:475–480.

55. Cao D, Ma B, Cao Z, Zhang X, Xiang Y. 2023. Structure of Semliki Forest virus in complex with its receptor VLDLR. Cell 186:2208-2218.e15.
56. Wintachai P, Wikan N, Kuadkitkan A, Jaimipuk T, Ubol S, Pulmanausahakul R, Auewarakul P, Kasinrerk W, Weng W-Y, Panyasrivanit M, Paemanee A, Kittisenachai S, Roytrakul S, Smith DR. 2012. Identification of prohibitin as a Chikungunya virus receptor protein. J Med Virol 84:1757–1770.
57. Zhang R, Kim AS, Fox JM, Nair S, Basore K, Klimstra WB, Rimkunas R, Fong RH, Lin H, Poddar S, Crowe JE, Doranz BJ, Fremont DH, Diamond MS. 2018. Mxra8 is a receptor for multiple arthritogenic alphaviruses. Nature 557:570–574.
58. Zhang R, Earnest JT, Kim AS, Winkler ES, Desai P, Adams LJ, Hu G, Bullock C, Gold B, Cherry S, Diamond MS. 2019. Expression of the Mxra8 Receptor Promotes Alphavirus Infection and Pathogenesis in Mice and Drosophila. Cell Rep 28:2647–2658.e5.
59. Basore K, Kim AS, Nelson CA, Zhang R, Smith BK, Uranga C, Vang L, Cheng M, Gross ML, Smith J, Diamond MS, Fremont DH. 2019. Cryo-EM structure of Chikungunya virus in complex with the Mxra8 receptor. Cell 177:1725–1737.e16.
60. 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.
61. 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.
62. de Groot RJ, Hardy WR, Shirako Y, Strauss JH. 1990. Cleavage-site preferences of Sindbis virus polyproteins containing the non-structural proteinase. Evidence for temporal regulation of polyprotein processing in vivo. EMBO J 9:2631–2638.
63. Strauss EG, Levinson R, Rice CM, Dalrymple J, Strauss JH. 1988. Nonstructural proteins nsP3 and nsP4 of Ross River and O'Nyong-nyong viruses: Sequence and comparison with those of other alphaviruses. Virology 164:265–274.
64. Chen KC, Kam Y-W, Lin RTP, Ng MM-L, Ng LF, Chu JJH. 2013. Comparative analysis of the genome sequences and replication profiles of chikungunya virus isolates within the East, Central and South African (ECSA) lineage. Virol J 10:169.
65. Jones JE, Long KM, Whitmore AC, Sanders W, Thurlow LR, Brown JA, Morrison CR, Vincent H, Peck KM, Browning C, Moorman N, Lim JK, Heise MT. 2017. Disruption of the Opal Stop Codon Attenuates Chikungunya Virus-Induced Arthritis and Pathology. mBio 8:e01456–17.
66. Myles KM, Kelly CLH, Ledermann JP, Powers AM. 2006. Effects of an Opal Termination Codon Preceding the nsP4 Gene Sequence in the O'Nyong-Nyong Virus Genome on Anopheles gambiae Infectivity. J Virol 80:4992–4997.
67. Salonen A, Vasiljeva L, Merits A, Magden J, Jokitalo E, Kääriäinen L. 2003. Properly Folded Nonstructural Polyprotein Directs the Semliki Forest Virus Replication Complex to the Endosomal Compartment. J Virol 77:1691–1702.
68. Rausalu K, Utt A, Quirin T, Varghese FS, Žusinaite E, Das PK, Ahola T, Merits A. 2016. Chikungunya virus infectivity, RNA replication and non-structural polyprotein processing depend on the nsP2 protease's active site cysteine residue. 1. Sci Rep 6:37124.
69. Lulla V, Karo-Astover L, Rausalu K, Saul S, Merits A, Lulla A. 2018. Timeliness of Proteolytic Events Is Prerequisite for Efficient Functioning of the Alphaviral Replicase. Journal of Virology 92:e00151–18.

70. 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.
71. 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.
72. Frolova EI, Gorchakov R, Pereboeva L, Atasheva S, Frolov I. 2010. Functional Sindbis Virus Replicative Complexes Are Formed at the Plasma Membrane. Journal of Virology 84:11679–11695.
73. Sawicki DL, Sawicki SG. 1980. Short-lived minus-strand polymerase for Semliki Forest virus. J Virol 34:108–118.
74. Hardy RW. 2006. The role of the 3′ terminus of the Sindbis virus genome in minus-strand initiation site selection. Virology 345:520–531.
75. Nikonov A, Mölder T, Sikut R, Kiiver K, Männik 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.
76. Thal MA, Wasik BR, Posto J, Hardy RW. 2007. Template requirements for recognition and copying by Sindbis virus RNA-dependent RNA polymerase. Virology 358:221–232.
77. Hellström K, Kallio K, Meriläinen H-M, Jokitalo E, Ahola T. 2016. Ability of minus strands and modified plus strands to act as templates in Semliki Forest virus RNA replication. J Gen Virol 97:1395–1407.
78. Hahn YS, Strauss EG, Strauss JH. 1989. Mapping of RNA- temperature-sensitive mutants of Sindbis virus: assignment of complementation groups A, B, and G to nonstructural proteins. J Virol 63:3142–3150.
79. Wang YF, Sawicki SG, Sawicki DL. 1991. Sindbis virus nsP1 functions in negative-strand RNA synthesis. J Virol 65:985–988.
80. 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.
81. Lemm JA, Rümenapf 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.
82. Li M-L, 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.
83. Li M-L, 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.
84. 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.
85. Lemm JA, Rice CM. 1993. Roles of nonstructural polyproteins and cleavage products in regulating Sindbis virus RNA replication and transcription. J Virol 67: 1916–1926.

86. Strauss EG, De Groot RJ, Levinson R, Strauss JH. 1992. Identification of the active site residues in the nsP2 proteinase of Sindbis virus. Virology 191:932–940.
87. Vasiljeva L, Merits A, Golubtsov A, Sizemskaja V, Kääriäinen L, Ahola T. 2003. Regulation of the sequential processing of Semliki Forest virus replicase polyprotein. J Biol Chem 278:41636–41645.
88. Cherkashchenko L, Rausalu K, Basu S, Alphey L, Merits A. 2022. Expression of Alphavirus Nonstructural Protein 2 (nsP2) in Mosquito Cells Inhibits Viral RNA Replication in Both a Protease Activity-Dependent and -Independent Manner. 6. Viruses 14:1327.
89. 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.
90. Kujala P, Ikäheimonen A, Ehsani N, Vihinen H, Auvinen P, Kääriäinen L. 2001. Biogenesis of the Semliki Forest Virus RNA Replication Complex. J Virol 75:3873–3884.
91. 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.
92. Morgan C, Howe C, Rose HM. 1961. Structure and development of viruses as observed in the electron microscope. V. Western equine encephalomyelitis virus. J Exp Med 113:219–234.
93. Spuul P, Balistreri G, Kääriäinen 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.
94. Ahola T, Lampio A, Auvinen P, Kääriäinen L. 1999. Semliki Forest virus mRNA capping enzyme requires association with anionic membrane phospholipids for activity. EMBO J 18:3164–3172.
95. Thaa B, Biasiotto R, Eng K, Neuvonen M, Götte 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.
96. Hellström K, Kallio K, Utt A, Quirin T, Jokitalo E, Merits A, Ahola T. 2017. Partially Uncleaved Alphavirus Replicase Forms Spherule Structures in the Presence and Absence of RNA Template. J Virol 91:e00787–17.
97. Laurent T, Kumar P, Liese S, Zare F, Jonasson M, Carlson A, Carlson L-A. Architecture of the chikungunya virus replication organelle. eLife 11:e83042.
98. Tan YB, Chmielewski D, Law MCY, Zhang K, He Y, Chen M, Jin J, Luo D. 2022. Molecular architecture of the Chikungunya virus replication complex. Science Advances 8:eadd2536.
99. Melancon P, Garoff H. 1987. Processing of the Semliki Forest virus structural polyprotein: role of the capsid protease. J Virol 61:1301–1309.
100. Lulla V, Kim DY, Frolova EI, Frolov I. 2013. The Amino-Terminal Domain of Alphavirus Capsid Protein Is Dispensable for Viral Particle Assembly but Regulates RNA Encapsidation through Cooperative Functions of Its Subdomains. J Virol 87:12003–12019.

101. 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.
102. Bredenbeek PJ, Frolov I, Rice CM, Schlesinger S. 1993. Sindbis virus expression vectors: packaging of RNA replicons by using defective helper RNAs. J Virol 67:6439–6446.
103. Weiss B, Nitschko H, Ghattas I, Wright R, Schlesinger S. 1989. Evidence for specificity in the encapsidation of Sindbis virus RNAs. J Virol 63:5310–5318.
104. Frolova E, Frolov I, Schlesinger S. 1997. Packaging signals in alphaviruses. J Virol 71:248–258.
105. 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. Journal of Virology 85:8022–8036.
106. Barth BU, Wahlberg JM, Garoff H. 1995. The oligomerization reaction of the Semliki Forest virus membrane protein subunits. J Cell Biol 128:283–291.
107. Yap ML, Klose T, Urakami A, Hasan SS, Akahata W, Rossmann MG. 2017. Structural studies of Chikungunya virus maturation. Proc Natl Acad Sci U S A 114: 13703–13707.
108. Lobigs M, Garoff H. 1990. Fusion function of the Semliki Forest virus spike is activated by proteolytic cleavage of the envelope glycoprotein precursor p62. J Virol 64:1233–1240.
109. Sanz MA, Rejas MT, Carrasco L. 2003. Individual expression of sindbis virus glycoproteins. E1 alone promotes cell fusion. Virology 305:463–472.
110. Andersson H, Barth BU, Ekström M, Garoff H. 1997. Oligomerization-dependent folding of the membrane fusion protein of Semliki Forest virus. J Virol 71:9654–9663.
111. Metsikkö K, Garoff H. 1990. Oligomers of the cytoplasmic domain of the p62/E2 membrane protein of Semliki Forest virus bind to the nucleocapsid in vitro. J Virol 64:4678–4683.
112. Gaedigk-Nitschko K, Schlesinger MJ. 1991. Site-directed mutations in Sindbis virus E2 glycoprotein's cytoplasmic domain and the 6K protein lead to similar defects in virus assembly and budding. Virology 183:206–214.
113. Lee S, Owen KE, Choi H-K, Lee H, Lu G, Wengler G, Brown DT, Rossmann MG, Kuhn RJ. 1996. Identification of a protein binding site on the surface of the alphavirus nucleocapsid and its implication in virus assembly. Structure 4:531–541.
114. Gaedigk-Nitschko K, Schlesinger MJ. 1990. The Sindbis virus 6K protein can be detected in virions and is acylated with fatty acids. Virology 175:274–281.
115. Lusa S, Garoff H, Liljeström P. 1991. Fate of the 6K membrane protein of Semliki Forest virus during virus assembly. Virology 185:843–846.
116. Elmasri Z, Negi V, Kuhn RJ, Jose J. 2022. Requirement of a functional ion channel for Sindbis virus glycoprotein transport, CPV-II formation, and efficient virus budding. PLoS Pathog 18:e1010892.
117. Loewy A, Smyth J, von Bonsdorff CH, Liljeström P, Schlesinger MJ. 1995. The 6-kilodalton membrane protein of Semliki Forest virus is involved in the budding process. J Virol 69:469–475.
118. McInerney GM, Smit JM, Liljeström 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.

119. Guo T-C, Johansson DX, Haugland Ø, Liljeström P, Evensen Ø. 2014. A 6K-Deletion Variant of Salmonid Alphavirus Is Non-Viable but Can Be Rescued through RNA Recombination. PLoS One 9:e100184.
120. Sanz MA, Carrasco L. 2001. Sindbis Virus Variant with a Deletion in the 6K Gene Shows Defects in Glycoprotein Processing and Trafficking: Lack of Complementation by a Wild-Type 6K Gene in trans. J Virol 75:7778–7784.
121. Taylor A, Melton JV, Herrero LJ, Thaa B, Karo-Astover L, Gage PW, Nelson MA, Sheng K-C, Lidbury BA, Ewart GD, McInerney GM, Merits A, Mahalingam S. 2016. Effects of an In-Frame Deletion of the 6k Gene Locus from the Genome of Ross River Virus. J Virol 90:4150–4159.
122. Hallengärd D, Kakoulidou M, Lulla A, Kümmerer BM, Johansson DX, Mutso M, Lulla V, Fazakerley JK, Roques P, Le Grand R, Merits A, Liljeström P. 2014. Novel attenuated Chikungunya vaccine candidates elicit protective immunity in C57BL/6 mice. J Virol 88:2858–2866.
123. Gaedigk-Nitschko K, Ding MX, Levy MA, Schlesinger MJ. 1990. Site-directed mutations in the Sindbis virus 6K protein reveal sites for fatty acylation and the underacylated protein affects virus release and virion structure. Virology 175:282–291.
124. Firth AE, Chung BY, Fleeton MN, Atkins JF. 2008. Discovery of frameshifting in Alphavirus 6K resolves a 20-year enigma. Virol J 5:108.
125. Ramsey J, Mukhopadhyay S. 2017. Disentangling the Frames, the State of Research on the Alphavirus 6K and TF Proteins. Viruses 9:228.
126. Snyder JE, Kulcsar KA, Schultz KLW, 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.
127. Kendra JA, de la Fuente C, Brahms A, Woodson C, Bell TM, Chen B, Khan YA, Jacobs JL, Kehn-Hall K, Dinman JD. 2017. Ablation of Programmed −1 Ribosomal Frameshifting in Venezuelan Equine Encephalitis Virus Results in Attenuated Neuropathogenicity. J Virol 91:e01766–16.
128. Rogers KJ, Jones-Burrage S, Maury W, Mukhopadhyay S. 2020. TF protein of Sindbis virus antagonizes host type I interferon responses in a palmitoylation-dependent manner. Virology 542:63–70.
129. Ramsey J, Chavez M, Mukhopadhyay S. 2019. Domains of the TF protein important in regulating its own palmitoylation. Virology 531:31–39.
130. Tang J, Jose J, Chipman P, Zhang W, Kuhn RJ, Baker TS. 2011. Molecular Links between the E2 Envelope Glycoprotein and Nucleocapsid Core in Sindbis Virus. J Mol Biol 414:442–459.
131. Zhang R, Hryc CF, Cong Y, Liu X, Jakana J, Gorchakov R, Baker ML, Weaver SC, Chiu W. 2011. 4.4 Å cryo-EM structure of an enveloped alphavirus Venezuelan equine encephalitis virus. EMBO J 30:3854–3863.
132. Ahola T, Laakkonen P, Vihinen H, Kääriäinen L. 1997. Critical residues of Semliki Forest virus RNA capping enzyme involved in methyltransferase and guanylyltransferase-like activities. J Virol 71:392–397.
133. Ahola T, Kääriäinen 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.
134. Laakkonen P, Hyvönen M, Peränen J, Kääriäinen L. 1994. Expression of Semliki Forest virus nsP1-specific methyltransferase in insect cells and in Escherichia coli. J Virol 68:7418–7425.

135. Peränen J, Laakkonen P, Hyvönen M, Kääriäinen L. 1995. The alphavirus replicase protein nsP1 is membrane-associated and has affinity to endocytic organelles. Virology 208:610–620.
136. Laakkonen P, Ahola T, Kääriäinen L. 1996. The effects of palmitoylation on membrane association of Semliki forest virus RNA capping enzyme. J Biol Chem 271: 28567–28571.
137. 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.
138. Žusinaite 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.
139. Zhang N, Zhao H, Zhang L. 2019. Fatty Acid Synthase Promotes the Palmitoylation of Chikungunya Virus nsP1. Journal of Virology 93:e01747–18.
140. Utt A, Rausalu K, Jakobson M, Männik A, Alphey L, Fragkoudis R, Merits A. 2019. Design and Use of Chikungunya Virus Replication Templates Utilizing Mammalian and Mosquito RNA Polymerase I-Mediated Transcription. J Virol 93:e00794–19.
141. Bakhache W, Neyret A, Bernard E, Merits A, Briant L. 2020. Palmitoylated Cysteines in Chikungunya Virus nsP1 Are Critical for Targeting to Cholesterol-Rich Plasma Membrane Microdomains with Functional Consequences for Viral Genome Replication. J Virol 94.
142. Zhang K, Law Y-S, Law MCY, Tan YB, Wirawan M, Luo D. 2021. Structural insights into viral RNA capping and plasma membrane targeting by Chikungunya virus nonstructural protein 1. Cell Host & Microbe 29:757–764.e3.
143. Jones R, Bragagnolo G, Arranz R, Reguera J. 2021. Capping pores of alphavirus nsP1 gate membranous viral replication factories. Nature 589:615–619.
144. Zhang K, Law MCY, Nguyen TM, Tan YB, Wirawan M, Law Y-S, Jeong LS, Luo D. 2022. Molecular basis of specific viral RNA recognition and 5'-end capping by the Chikungunya virus nsP1. Cell Rep 40:111133.
145. Jones R, Hons M, Rabah N, Zamarreño N, Arranz R, Reguera J. 2023. Structural basis and dynamics of Chikungunya alphavirus RNA capping by nsP1 capping pores. Proceedings of the National Academy of Sciences 120:e2213934120.
146. 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. Journal of Virology 82:9236–9244.
147. Lulla V, Merits A, Sarin P, Kääriäinen L, Keränen S, Ahola T. 2006. Identification of mutations causing temperature-sensitive defects in Semliki Forest virus RNA synthesis. J Virol 80:3108–3111.
148. Fata CL, Sawicki SG, Sawicki DL. 2002. Alphavirus minus-strand RNA synthesis: identification of a role for Arg183 of the nsP4 polymerase. J Virol 76:8632–8640.
149. Fata CL, Sawicki SG, Sawicki DL. 2002. Modification of Asn374 of nsP1 Suppresses a Sindbis Virus nsP4 Minus-Strand Polymerase Mutant. Journal of Virology 76:8641–8649.
150. Vasiljeva L, Merits A, Auvinen P, Kääriäinen L. 2000. Identification of a novel function of the alphavirus capping apparatus. RNA 5'-triphosphatase activity of Nsp2. J Biol Chem 275:17281–17287.

151. Law Y-S, Utt A, Tan YB, Zheng J, Wang S, Chen MW, Griffin PR, Merits A, Luo D. 2019. Structural insights into RNA recognition by the Chikungunya virus nsP2 helicase. Proc Natl Acad Sci U S A 116:9558–9567.
152. Rikkonen M, Peränen J, Kääriäinen L. 1994. ATPase and GTPase activities associated with Semliki Forest virus nonstructural protein nsP2. Journal of Virology 68:5804–5810.
153. 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.
154. Russo AT, White MA, Watowich SJ. 2006. The Crystal Structure of the Venezuelan Equine Encephalitis Alphavirus nsP2 Protease. Structure 14:1449–1458.
155. Merits A, Vasiljeva L, Ahola T, Kääriäinen L, Auvinen P. 2001. Proteolytic processing of Semliki Forest virus-specific non-structural polyprotein by nsP2 protease. Journal of General Virology 82:765–773.
156. Law Y-S, Wang S, Tan YB, Shih O, Utt A, Goh WY, Lian B-J, Chen MW, Jeng U S, Merits A, Luo D. 2021. Interdomain Flexibility of Chikungunya Virus nsP2 Helicase-Protease Differentially Influences Viral RNA Replication and Infectivity. J Virol 95:e01470–20.
157. Peränen J, Rikkonen M, Liljeström P, Kääriäinen L. 1990. Nuclear localization of Semliki Forest virus-specific nonstructural protein nsP2. J Virol 64:1888–1896.
158. 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.
159. Frolov I, Agapov E, Hoffman TA, Prágai BM, Lippa M, Schlesinger S, Rice CM. 1999. Selection of RNA Replicons Capable of Persistent Noncytopathic Replication in Mammalian Cells. J Virol 73:3854–3865.
160. Dryga SA, Dryga OA, Schlesinger S. 1997. Identification of Mutations in a Sindbis Virus Variant Able to Establish Persistent Infection in BHK Cells: The Importance of a Mutation in the nsP2 Gene. Virology 228:74–83.
161. 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.
162. 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.
163. Breakwell L, Dosenovic P, Karlsson Hedestam GB, D'Amato M, Liljeström 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.
164. 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. Journal of Virology 84:10877–10887.
165. 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.
166. Götte B, Liu L, McInerney GM. 2018. The Enigmatic Alphavirus Non-Structural Protein 3 (nsP3) Revealing Its Secrets at Last. Viruses 10:105.

167. Feijs KLH, Forst AH, Verheugd P, Lüscher B. 2013. Macrodomain-containing proteins: regulating new intracellular functions of mono(ADP-ribosyl)ation. Nat Rev Mol Cell Biol 14:443–451.
168. Laing S, Unger M, Koch-Nolte F, Haag F. 2011. ADP-ribosylation of arginine. Amino Acids 41:257–269.
169. Fehr AR, Jankevicius G, Ahel I, Perlman S. 2018. Viral Macrodomains: Unique Mediators of Viral Replication and Pathogenesis. Trends Microbiol 26:598–610.
170. Krieg S, Pott F, Potthoff L, Verheirstraeten M, Bütepage M, Golzmann A, Lippok B, Goffinet C, Lüscher B, Korn P. 2023. Mono-ADP-ribosylation by PARP10 inhibits Chikungunya virus nsP2 proteolytic activity and viral replication. Cell Mol Life Sci 80:72.
171. McPherson RL, Abraham R, Sreekumar E, Ong S-E, Cheng S-J, Baxter VK, Kistemaker HAV, Filippov DV, Griffin DE, Leung AKL. 2017. ADP-ribosylhydrolase activity of Chikungunya virus macrodomain is critical for virus replication and virulence. Proc Natl Acad Sci U S A 114:1666–1671.
172. Park E, Griffin DE. 2009. The nsP3 macro domain is important for Sindbis virus replication in neurons and neurovirulence in mice. Virology 388:305–314.
173. Atasheva S, Akhrymuk M, Frolova EI, Frolov I. 2012. New PARP Gene with an Anti-Alphavirus Function. J Virol 86:8147–8160.
174. Lulla A, Lulla V, Merits A. 2012. Macromolecular Assembly-Driven Processing of the 2/3 Cleavage Site in the Alphavirus Replicase Polyprotein. Journal of Virology 86:553–565.
175. Gao Y, Goonawardane N, Ward J, Tuplin A, Harris M. 2019. Multiple roles of the non-structural protein 3 (nsP3) alphavirus unique domain (AUD) during Chikungunya virus genome replication and transcription. PLoS Pathog 15:e1007239.
176. Li GP, La Starza MW, Hardy WR, Strauss JH, Rice CM. 1990. Phosphorylation of Sindbis virus nsP3 in vivo and in vitro. Virology 179:416–427.
177. Peränen J, Takkinen K, Kalkkinen N, Kääriäinen L. 1988. Semliki Forest virus-specific non-structural protein nsP3 is a phosphoprotein. J Gen Virol 69 (Pt 9): 2165–2178.
178. Teppor M, Žusinaite E, Merits A. 2021. Phosphorylation Sites in the Hypervariable Domain in Chikungunya Virus nsP3 Are Crucial for Viral Replication. J Virol 95:e02276–20.
179. Vihinen H, Ahola T, Tuittila M, Merits A, Kääriäinen L. 2001. Elimination of phosphorylation sites of Semliki Forest virus replicase protein nsP3. J Biol Chem 276:5745–5752.
180. Meshram CD, Agback P, Shiliaev N, Urakova N, Mobley JA, Agback T, Frolova EI, Frolov I. 2018. Multiple Host Factors Interact with the Hypervariable Domain of Chikungunya Virus nsP3 and Determine Viral Replication in Cell-Specific Mode. J Virol 92.
181. Kim DY, Reynaud JM, Rasalouskaya A, Akhrymuk I, Mobley JA, Frolov I, Frolova EI. 2016. New World and Old World Alphaviruses Have Evolved to Exploit Different Components of Stress Granules, FXR and G3BP Proteins, for Assembly of Viral Replication Complexes. PLoS Pathog 12:e1005810.
182. Frolov I, Kim DY, Akhrymuk M, Mobley JA, Frolova EI. 2017. Hypervariable Domain of Eastern Equine Encephalitis Virus nsP3 Redundantly Utilizes Multiple Cellular Proteins for Replication Complex Assembly. J Virol 91.

183. 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. Journal of Virology 86:10873–10879.
184. Panas MD, Varjak M, Lulla A, Er Eng K, Merits A, Karlsson Hedestam GB, McInerney GM. 2012. Sequestration of G3BP coupled with efficient translation inhibits stress granules in Semliki Forest virus infection. MBoC 23:4701–4712.
185. Götte B, Utt A, Fragkoudis R, Merits A, McInerney GM. 2020. Sensitivity of Alphaviruses to G3BP Deletion Correlates with Efficiency of Replicase Polyprotein Processing. J Virol 94.
186. Bartholomeeusen K, Utt A, Coppens S, Rausalu K, Vereecken K, Ariën KK, Merits A. 2018. A Chikungunya Virus trans-Replicase System Reveals the Importance of Delayed Nonstructural Polyprotein Processing for Efficient Replication Complex Formation in Mosquito Cells. J Virol 92:e00152–18.
187. 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.
188. 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.
189. Kamer G, Argos P. 1984. Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. Nucleic Acids Res 12:7269–7282.
190. de Groot RJ, Rümenapf 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.
191. 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.
192. Shirako Y, Strauss EG, Strauss JH. 2003. Modification of the 5′ Terminus of Sindbis Virus Genomic RNA Allows nsP4 RNA Polymerases with Nonaromatic Amino Acids at the N Terminus To Function in RNA Replication. J Virol 77:2301–2309.
193. Kaur P, Lello LS, Utt A, Dutta SK, Merits A, Chu JJH. 2020. Bortezomib inhibits chikungunya virus replication by interfering with viral protein synthesis. PLoS Negl Trop Dis 14:e0008336.
194. Chen MW, Tan YB, Zheng J, Zhao Y, Lim BT, Cornvik T, Lescar J, Ng LFP, Luo D. 2017. Chikungunya virus nsP4 RNA-dependent RNA polymerase core domain displays detergent-sensitive primer extension and terminal adenylyltransferase activities. Antiviral Res 143:38–47.
195. Shenk TE, Stollar V. 1973. Defective-interfering particles of Sindbis virus: I. Isolation and some chemical and biological properties. Virology 53:162–173.
196. Liljeström P, Garoff H. 1991. A New Generation of Animal Cell Expression Vectors Based on the Semliki Forest Virus Replicon. 12. Nat Biotechnol 9:1356–1361.
197. Kallio K, Hellström K, Balistreri G, Spuul P, Jokitalo E, Ahola T. 2013. Template RNA length determines the size of replication complex spherules for Semliki Forest virus. J Virol 87:9125–9134.
198. Spuul P, Balistreri G, Hellström 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.

199. Utt A, Quirin T, Saul S, Hellström K, Ahola T, Merits A. 2016. Versatile Trans-Replication Systems for Chikungunya Virus Allow Functional Analysis and Tagging of Every Replicase Protein. PLoS One 11:e0151616.
200. Gorchakov R, Frolova E, Sawicki S, Atasheva S, Sawicki D, Frolov I. 2008. A New Role for ns Polyprotein Cleavage in Sindbis Virus Replication. J Virol 82:6218–6231.
201. Teppor M, Žusinaite E, Karo-Astover L, Omler A, Rausalu K, Lulla V, Lulla A, Merits A. 2021. Semliki Forest Virus Chimeras with Functional Replicase Modules from Related Alphaviruses Survive by Adaptive Mutations in Functionally Important Hot Spots. Journal of Virology 95:e00973–21.
202. Steel JJ, Franz AWE, Sanchez-Vargas I, Olson KE, Geiss BJ. 2013. Subgenomic Reporter RNA System for Detection of Alphavirus Infection in Mosquitoes. PLOS ONE 8:e84930.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my supervisor professor Andres Merits. Thank you for all the invaluable advice, discussions and guidance during my years in MeritsLab. There are no words big enough to describe how much I have appreciated being a part of your team, Andres! I am very thankful to Age Utt who has been nothing short of amazing. Thank you for your kindness, patience and all the voice messages to guide me through the seemingly endless amounts of northern blots. Special thanks go to Margit Mutso for the extremely valuable advice, feedback and questions.

I would like to thank all my colleagues in our lab, TÜTI and ABSL3 for creating a supportive and inspiring workspace. My sincerest thanks go to Ailar, Sandra, Sainan and Kristel. Even the toughest days are fun with you guys around! You should all be rewarded a medal for your exceptional problem-solving skills (and perhaps an additional medal for Ailar for all the impeccable jokes).

I am beyond grateful to my family. I want to thank my mom Sirle and dad Priit who have given me everything (including two slightly annoying, however, incredibly sweet and awesome sisters). Thank you for always supporting me. Special thanks go to a young architect Tuule who so valiantly took on the quest to help me illustrate some alphavirus transmission cycles.

Finally, I would like to express my most sincere gratitude to Siim. Thank you for always keeping my feet on the ground, for standing by me through all the stressful times and making me laugh every single day. My life with you is all I ever could have asked for.

PUBLICATIONS

CURRICULUM VITAE

Name Laura Sandra Lello
Date of birth 26th of July 1996
Citizenship Estonian
E-mail laura.lello@ut.ee

Education and professional employment

2004–2015	Tallinn School No. 21
2015–2018	University of Tartu, BSc in Gene Technology
2018–2020	University of Tartu, MSc in Gene Technology
2019–2020	University of Tartu, Laboratory Assistant in Virology
2020–	University of Tartu, doctoral studies in Biomedical Technology
2021–	University of Tartu, Junior Research Fellow in Virology

Publications

Lello, L.S.; Miilimäe, A,; Cherkashchenko, L.; Omler, A., Skilton, R., Ireland, R.; Ulaeto, D.; Merits, A. (2023). Activity, template preference and compatibility of components of RNA replicase of Eastern equine encephalitis virus. Journal of Virology, 97 (1), e01368–22. DOI: 10.1128/jvi.01368-22.

Tan, Yaw Bia; Lello, Laura Sandra; Liu, Xin; Law, Yee-Song; Kang, Congbao; Lescar, Julien; Zheng, Jie; Merits, Andres; Luo, Dahai (2022). Crystal structures of alphavirus nonstructural protein 4 (nsP4) reveal an intrinsically dynamic RNA-dependent RNA polymerase fold. Nucleic Acids Research, 50 (2), 1000–1016. DOI: 10.1093/nar/gkab1302.

Kangro, Kadri; Kurašin, Mihhail; Gildemann, Kiira; Sankovski, Eve; Žusinaite, Eva; Lello, Laura Sandra; Pert, Raini; Kavak, Ants; Poikalainen, Väino; Lepasalu, Lembit; Kuusk, Marilin; Pau, Robin; Piiskop, Sander; Rom, Siimu; Oltjer, Ruth; Tiirik, Kairi; Kogermann, Karin; Plaas, Mario; Tiirats, Toomas; Aasmäe, Birgit ... Ustav, Mart (2021). Bovine Colostrum Derived Antibodies Against SARS-CoV-2 Show Great Potential to Serve as a Prophylactic Agent. Medrxiv. DOI: 10.1101/2021.06.08.21258069.

Lello, Laura Sandra; Bartholomeeusen, Koen; Wang, Sainan; Coppens, Sandra; Fragkoudis, Rennos; Alphey, Luke; Ariën, Kevin K.; Merits, Andres; Utt, Age (2021). nsP4 is a major determinant of alphavirus replicase activity and template selectivity. Journal of Virology. DOI: 10.1128/JVI.00355-21.

Rihn, Suzannah J.; Merits, Andres; Bakshi, Siddharth; Turnbull, Matthew L.; Wickenhagen, Arthur; Alexander, Akira J. T.; Baillie, Carla; Brennan, Benjamin; Brown, Fiona; Brunker, Kirstyn; Bryden, Steven R.; Burness, Kerry A.; Carmichael, Stephen; Cole, Sarah J.; Cowton, Vanessa M.; Davies, Paul; Davis, Chris; De Lorenzo, Giuditta; Donald, Claire L.; Dorward, Mark ... Mahalingam, Suresh (2021). A plasmid DNA-launched SARS-CoV-2 reverse genetics system and coronavirus toolkit for COVID-19 research. PLoS Biology, 19 (2), ARTN e3001091. DOI: 10.1371/journal.pbio.3001091.

Lello, Laura Sandra; Utt, Age; Bartholomeeusen, Koen; Wang, Sainan; Rausalu, Kai; Kendall, Catherine; Coppens, Sandra; Fragkoudis, Rennos; Tuplin, Andrew; Alphey, Luke; Ariën, Kevin K.; Merits, Andres (2020). Cross-utilisation of template RNAs by alphavirus replicases. PLoS Pathogens, 16 (9), e1008825. DOI: 10.1371/journal.ppat.1008825.

Kaur, Parveen; Lello, Laura Sandra; Utt, Age; Dutta, Sujit Krishna; Merits, Andres; Chu, Justin Jang Hann (2020). Bortezomib inhibits chikungunya virus replication by interfering with viral protein synthesis. PLoS Neglected Tropical Diseases, 14 (5), e0008336. DOI: 10.1371/journal.pntd.0008336.

ELULOOKIRJELDUS

Nimi Laura Sandra Lello
Sünniaeg 26.07.1996
Kodakondsus eesti
E-post laura.lello@ut.ee

Akadeemiline elukäik

2004–2015	Tallinna 21. Kool
2015–2018	Tartu Ülikool, BSc geenitehnoloogia erialal
2018–2020	Tartu Ülikool, MSc geenitehnoloogia erialal
2019–2020	Tartu Ülikool, viroloogia spetsialist
2020–	Tartu Ülikool, doktoriõpe biomeditsiini tehnoloogia erialal
2021–	Tartu Ülikool, viroloogia nooremteadur

Publikatsioonid

Lello, L.S.; Miilimäe, A,; Cherkashchenko, L.; Omler, A., Skilton, R., Ireland, R.; Ulaeto, D.; Merits, A. (2023). Activity, template preference and compatibility of components of RNA replicase of Eastern equine encephalitis virus. Journal of Virology, 97 (1), e01368–22. DOI: 10.1128/jvi.01368-22.

Tan, Yaw Bia; Lello, Laura Sandra; Liu, Xin; Law, Yee-Song; Kang, Congbao; Lescar, Julien; Zheng, Jie; Merits, Andres; Luo, Dahai (2022). Crystal structures of alphavirus nonstructural protein 4 (nsP4) reveal an intrinsically dynamic RNA-dependent RNA polymerase fold. Nucleic Acids Research, 50 (2), 1000−1016. DOI: 10.1093/nar/gkab1302.

Kangro, Kadri; Kurašin, Mihhail; Gildemann, Kiira; Sankovski, Eve; Žusinaite, Eva; Lello, Laura Sandra; Pert, Raini; Kavak, Ants; Poikalainen, Väino; Lepasalu, Lembit; Kuusk, Marilin; Pau, Robin; Piiskop, Sander; Rom, Siimu; Oltjer, Ruth; Tiirik, Kairi; Kogermann, Karin; Plaas, Mario; Tiirats, Toomas; Aasmäe, Birgit ... Ustav, Mart (2021). Bovine Colostrum Derived Antibodies Against SARS-CoV-2 Show Great Potential to Serve as a Prophylactic Agent. Medrxiv. DOI: 10.1101/2021.06.08.21258069.

Lello, Laura Sandra; Bartholomeeusen, Koen; Wang, Sainan; Coppens, Sandra; Fragkoudis, Rennos; Alphey, Luke; Ariën, Kevin K.; Merits, Andres; Utt, Age (2021). nsP4 is a major determinant of alphavirus replicase activity and template selectivity. Journal of Virology. DOI: 10.1128/JVI.00355-21.

Rihn, Suzannah J.; Merits, Andres; Bakshi, Siddharth; Turnbull, Matthew L.; Wickenhagen, Arthur; Alexander, Akira J. T.; Baillie, Carla; Brennan, Benjamin; Brown, Fiona; Brunker, Kirstyn; Bryden, Steven R.; Burness, Kerry A.; Carmichael, Stephen; Cole, Sarah J.; Cowton, Vanessa M.; Davies, Paul; Davis, Chris; De Lorenzo, Giuditta; Donald, Claire L.; Dorward, Mark ... Mahalingam, Suresh (2021). A plasmid DNA-launched SARS-CoV-2 reverse genetics system and coronavirus toolkit for COVID-19 research. PLoS Biology, 19 (2), ARTN e3001091. DOI: 10.1371/journal.pbio.3001091.

Lello, Laura Sandra; Utt, Age; Bartholomeeusen, Koen; Wang, Sainan; Rausalu, Kai; Kendall, Catherine; Coppens, Sandra; Fragkoudis, Rennos; Tuplin, Andrew; Alphey, Luke; Ariën, Kevin K.; Merits, Andres (2020). Cross-utilisation of template RNAs by alphavirus replicases. PLoS Pathogens, 16 (9), e1008825. DOI: 10.1371/journal.ppat.1008825.

Kaur, Parveen; Lello, Laura Sandra; Utt, Age; Dutta, Sujit Krishna; Merits, Andres; Chu, Justin Jang Hann (2020). Bortezomib inhibits chikungunya virus replication by interfering with viral protein synthesis. PLoS Neglected Tropical Diseases, 14 (5), e0008336. DOI: 10.1371/journal.pntd.0008336.

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, 172 p.
22. **Sirle Saul**. Towards understanding the neurovirulence of Semliki Forest virus. Tartu, 2015, 136 p.
23. **Marit Orav**. Study of the initial amplification of the human papillomavirus genome. Tartu, 2015, 132 p.
24. **Tormi Reinson**. Studies on the Genome Replication of Human Papillomaviruses. Tartu, 2016, 110 p.
25. **Mart Ustav Jr**. Molecular Studies of HPV-18 Genome Segregation and Stable Replication. Tartu, 2016, 152 p.
26. **Margit Mutso**. Different Approaches to Counteracting Hepatitis C Virus and Chikungunya Virus Infections. Tartu, 2016, 184 p.
27. **Jelizaveta Geimanen**. Study of the Papillomavirus Genome Replication and Segregation. Tartu, 2016, 168 p.
28. **Mart Toots**. Novel Means to Target Human Papillomavirus Infection. Tartu, 2016, 173 p.
29. **Kadi-Liis Veiman**. Development of cell-penetrating peptides for gene delivery: from transfection in cell cultures to induction of gene expression *in vivo*. Tartu, 2016, 136 p.
30. **Ly Pärnaste**. How, why, what and where: Mechanisms behind CPP/cargo nanocomplexes. Tartu, 2016, 147 p.
31. **Age Utt**. Role of alphavirus replicase in viral RNA synthesis, virus-induced cytotoxicity and recognition of viral infections in host cells. Tartu, 2016, 183 p.
32. **Veiko Vunder**. Modeling and characterization of back-relaxation of ionic electroactive polymer actuators. Tartu, 2016, 154 p.
33. **Piia Kivipõld**. Studies on the Role of Papillomavirus E2 Proteins in Virus DNA Replication. Tartu, 2016, 118 p.
34. **Liina Jakobson**. The roles of abscisic acid, CO_2, and the cuticle in the regulation of plant transpiration. Tartu, 2017, 162 p.
35. **Helen Isok-Paas**. Viral-host interactions in the life cycle of human papillomaviruses. Tartu, 2017, 158 p.
36. **Hanna Hõrak**. Identification of key regulators of stomatal CO_2 signalling via O_3-sensitivity. Tartu, 2017, 260 p.
37. **Jekaterina Jevtuševskaja**. Application of isothermal amplification methods for detection of *Chlamydia trachomatis* directly from biological samples. Tartu, 2017, 96 p.
38. **Ülar Allas**. Ribosome-targeting antibiotics and mechanisms of antibiotic resistance. Tartu, 2017, 152 p.
39. **Anton Paier**. Ribosome Degradation in Living Bacteria. Tartu, 2017, 108 p.
40. **Vallo Varik**. Stringent Response in Bacterial Growth and Survival. Tartu, 2017, 101 p.

41. **Pavel Kudrin**. In search for the inhibitors of *Escherichia coli* stringent response factor RelA. Tartu, 2017, 138 p.
42. **Liisi Henno**. Study of the human papillomavirus genome replication and oligomer generation. Tartu, 2017, 144 p.
43. **Katrin Krõlov**. Nucleic acid amplification from crude clinical samples exemplified by *Chlamydia trachomatis* detection in urine. Tartu, 2018, 118 p.
44. **Eve Sankovski**. Studies on papillomavirus transcription and regulatory protein E2. Tartu, 2018, 113 p.
45. **Morteza Daneshmand**. Realistic 3D Virtual Fitting Room. Tartu, 2018, 233 p.
46. **Fatemeh Noroozi**. Multimodal Emotion Recognition Based Human-Robot Interaction Enhancement. Tartu, 2018, 113 p.
47. **Krista Freimann**. Design of peptide-based vector for nucleic acid delivery in vivo. Tartu, 2018, 103 p.
48. **Rainis Venta**. Studies on signal processing by multisite phosphorylation pathways of the *S. cerevisiae* cyclin-dependent kinase inhibitor Sic1. Tartu, 2018, 155 p.
49. **Inga Põldsalu**. Soft actuators with ink-jet printed electrodes. Tartu, 2018, 85 p.
50. **Kadri Künnapuu**. Modification of the cell-penetrating peptide PepFect14 for targeted tumor gene delivery and reduced toxicity. Tartu, 2018, 114 p.
51. **Toomas Mets**. RNA fragmentation by MazF and MqsR toxins of *Escherichia coli.* Tartu, 2019, 119 p.
52. **Kadri Tõldsepp**. The role of mitogen-activated protein kinases MPK4 and MPK12 in CO_2-induced stomatal movements. Tartu, 2019, 259 p.
53. **Pirko Jalakas**. Unravelling signalling pathways contributing to stomatal conductance and responsiveness. Tartu, 2019, 120 p.
54. **S. Sunjai Nakshatharan**. Electromechanical modelling and control of ionic electroactive polymer actuators. Tartu, 2019, 165 p.
55. **Eva-Maria Tombak**. Molecular studies of the initial amplification of the oncogenic human papillomavirus and closely related nonhuman primate papillomavirus genomes. Tartu, 2019, 150 p.
56. **Meeri Visnapuu**. Design and physico-chemical characterization of metal-containing nanoparticles for antimicrobial coatings. Tartu, 2019, 138 p.
57. **Jelena Beljantseva**. Small fine-tuners of the bacterial stringent response – a glimpse into the working principles of Small Alarmone Synthetases. Tartu, 2020, 104 p.
58. **Egon Urgard**. Potential therapeutic approaches for modulation of inflammatory response pathways. Tartu, 2020, 120 p.
59. **Sofia Raquel Alves Oliveira**. HPLC analysis of bacterial alarmone nucleotide (p)ppGpp and its toxic analogue ppApp. Tartu, 2020, 122 p.
60. **Mihkel Örd**. Ordering the phosphorylation of cyclin-dependent kinase Cdk1 substrates in the cell cycle. Tartu, 2021, 228 p.
61. **Fred Elhi**. Biocompatible ionic electromechanically active polymer actuator based on biopolymers and non-toxic ionic liquids. Tartu, 2021, 140 p.

62. **Liisi Talas**. Reconstructing paleo-diversity, dynamics and response of eukaryotes to environmental change over the Late-Glacial and Holocene period in lake Lielais Svētiņu using sedaDNA. Tartu, 2021, 118 p.
63. **Livia Matt**. Novel isosorbide-based polymers. Tartu, 2021, 118 p.
64. **Koit Aasumets**. The dynamics of human mitochondrial nucleoids within the mitochondrial network. Tartu, 2021, 104 p.
65. **Faiza Summer**. Development and optimization of flow electrode capacitor technology. Tartu, 2022, 109 p.
66. **Olavi Reinsalu**. Cancer-testis antigen MAGE-A4 is incorporated into extracellular vesiclesand is exposed to the surface. Tartu, 2022, 130 p.
67. **Tetiana Brodiazhenko**. RelA-SpoT Homolog enzymes as effectors of Toxin-Antitoxin systems. Tartu, 2022, 132 p.
68. **Georg-Marten Lanno**. Development of novel antibacterial drug delivery systems as wound scaffolds using electrospinning technology. Tartu, 2022, 175 p.
69. **Liubov Cherkashchenko**. New insights into alphaviral nsP2 functions. Tartu, 2023, 171 p.
70. **Kristina Kiisholts**. Peptide-based drug carriers and preclinical nanomedicine applications for endometriosis treatment. Tartu, 2023, 138 p.
71. **Kai Rausalu**. Alphaviral nsP2 protease: From requirements for functionality to inhibition. Tartu, 2023, 175 p.