Universiteit
Leiden
The Netherlands

Functions and requirements of conserved RNA structures in the 3' untranslated region of Flaviviruses

Agostinho Gonçalves Costa da Silva, P.

Citation
Agostinho Gonçalves Costa da Silva, P. (2011, June 27). *Functions and requirements of conserved RNA structures in the 3' untranslated region of Flaviviruses*. Retrieved from https://hdl.handle.net/1887/17775

Version:	Corrected Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/17775

Note: To cite this publication please use the final published version (if applicable).

Functions and requirements of conserved RNA structures in the 3′ untranslated region of Flaviviruses

Patrícia Agostinho Gonçalves Costa da Silva

ISBN: 978-94-6169-098-2

Layout and printing: Optima Grafische Communicatie, Rotterdam, The Netherlands

Images on the cover are taken from or based on experiments described in this thesis.
All rights reserved. No part of this book may be reproduced or transmitted, in any form or by any means, without permission of the author.

Functions and requirements of conserved RNA structures in the 3' untranslated region of Flaviviruses

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op maandag 27 juni 2011
klokke 11.15 uur

door

Patrícia Agostinho Gonçalves Costa da Silva

geboren te Leiria, Portugal
in 1983

PROMOTIECOMISSIE

Promotor: Prof. Dr. W.J.M. Spaan

Copromotor: Dr. P.J. Bredenbeek

Overige leden: Prof. Dr. E.J. Snijder
Prof. Dr. B. Berkhout (Universiteit van Amsterdam)
Prof. Dr. J.H. Neyts (Katholieke Universiteit Leuven)
Dr. G.P. Pijlman (Universiteit Wageningen)

"Truth in science can be defined as the working hypothesis best suited to open the way to the next better one."

Konrad Lorenz

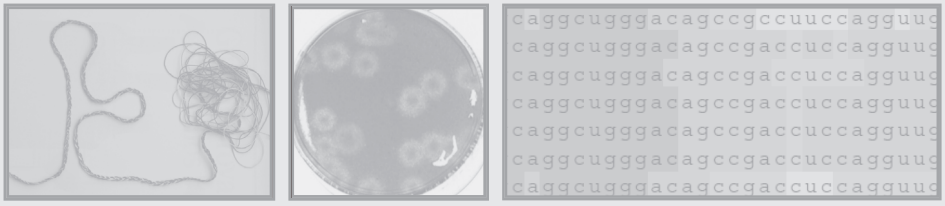
Aos meus pais
E todos os meus "anjinhos"

Contents

Chapter 1	General introduction	9
Chapter 2	Conservation of the pentanucleotide motif at the top of the yellow fever virus 17D 3′ stem-loop structure is not required for replication	35
Chapter 3	An RNA pseudoknot is required for production of yellow fever virus subgenomic RNA by the host nuclease XRN1	55
Chapter 4	Characterization of the sfRNAs that are produced in cells infected with flaviviruses with no known vector and cell fusing agent	83
Chapter 5	Characterization of a stable full-length cDNA clone for the transcription of infectious RNA of a Flavivirus with no known vector	103
Chapter 6	Functions and requirements of conserved RNA structures in the 3′ untranslated region of flaviviruses	115
Chapter 7	Epilogue	145
	Summary	153
	Samenvatting	157
	Curriculum Vitae	161

Chapter 1

General introduction



GENERAL INTRODUCTION

Infectious diseases are estimated to be the direct cause of more than 25% of all annual human deaths worldwide [1]. In fact, infection itself exerts a tremendous selective pressure that has driven the evolution of resistance mechanisms and, to a large extent, has shaped the human genome [2]. Infectious diseases are caused by many different organisms from several biological taxa. There are more than 1.400 species known to be pathogenic to humans [3]. Many of these species are associated with emerging diseases and they mainly consist of zoonotic pathogens, with 44 % of these agents being viruses and prions [3]. RNA viruses in particular, are the cause of many of the emerging and re-emerging diseases of the last decades [4]. RNA viruses have the highest mutation rate among species (estimated at 10^{-3} to 10^{-5} misincorporations per nucleotide and replication cycle) due to the lack (or low efficiency) of proof-reading activity by their RNA-dependent RNA polymerase (RdRp). As a consequence, RNA viruses replicate as complex and dynamic swarms of virus mutants known as virus quasispecies. In combination with short replication times and extremely large populations, this explains why RNA viruses can efficiently adapt to new selective pressures in the environment and are able to exploit new ecological niches and to jump between host species (reviewed in [5]). There are currently 95 virus families and unassigned genera approved by the International Committee on Taxonomy of Viruses [6]; among these, positive-stranded RNA viruses undoubtedly comprise the biggest fraction and it is within this group that we encounter the Flaviviruses.

Flavivirus genus

The family *Flaviviridae* currently consists of three genera: *Flavivirus* (from the Latin *flavus*, "yellow"), *Pestivirus* (from the Latin *pestis*, "plague"), and *Hepacivirus* (from the Greek *hepar, hepatos*, "liver"). Besides these genera, two distinct groups of viruses have tentatively been assigned to the family, GBV-A and GBV-C [7]. All members of the *Flaviviridae* family share similar characteristics in virion morphology, genome organization, and replication strategy. In contrast, members of the three genera are antigenically unrelated and exhibit different biological properties, such as host range and transmission [7]. The *Flavivirus* genus contains nearly 80 viruses and its members show a worldwide distribution. The majority of Flaviviruses is arthropod-borne and many of them are important human pathogens that can cause a variety of diseases including encephalitis and hemorrhagic fevers. Flaviviruses of major global concern include dengue virus (DENV), yellow fever virus (YFV), Japanese encephalitis virus (JEV), West Nile virus (WNV), and tick-borne encephalitis virus (TBEV). Flaviviral infections have dramatically increased in frequency [8] and the reasons underlying this phenomenon are complex. The decrease in mosquito

control measures during the last decades together with social and environmental factors such as the unprecedented population growth, increased urbanization, travel, trade, and deforestation are believed to be the main reasons for the re-emergence of flaviviruses. DENV, for instance, has spread into new areas and is now endemic in more than 100 countries where 2.5 billion people (40% of the world's population) are at risk of infection and an estimated 50 million people are infected every year [9]. YFV appeared to be under control in the middle of the previous century due to mass vaccination campaigns and eradication of the principal urban vector, *Aedes aegypti*. However, YFV is re-emerging as numerous outbreaks have been registered during the last decades in both Africa and South America, due to the declining vaccination coverage and mosquito reinfestation [10-12]. Furthermore, the vectors used by flaviviruses have the ability to infest alternative favourable habitats where the viruses can eventually cause an epidemic. WNV had never been isolated in the Americas until 1999, when it emerged in New York City [13]. By the end of 2003, it was present in almost every state of the United States, Mexico, and the Caribbean. It has now been detected as far south as Argentina (reviewed in [14]). Climate changes can also affect vector distribution implying that global warming, for instance, could significantly increase the potential for flavivirus dispersal (reviewed in [14,15]). As an example, *Ixodes ricinus*, the main vector of TBEV, used to be found in Europe up to 700-750 m above sea level in the early 1980s, but in 2001 it was found up to 1.000 m [16], and in 2009 ticks infected with TBEV were detected up to 1.140 m [17]. Global warming was suggested to be responsible for this shift in the habitat of ticks.

Vectors of members of the Flavivirus genus

Phylogenetically, the Flavivirus genus is grouped into three clusters based on the vector involved in transmission: (i) mosquito-borne, (ii) tick-borne, and (iii) no known vector (NKV) viruses [18,19]. The evolutionary relationship between these three clusters is not clear. Initial phylogenetic analysis using the amino acid sequences of the envelope gene established that mosquito- and tick-borne viruses represent two different evolutionary lineages [20]. In another study, the flaviviruses NS5 amino acid sequence was used for phylogenetic analysis [18]. This study included the NKV flaviviruses and postulated that the NKV and then the vector-borne flaviviruses have emerged from an ancestral insect-borne flavivirus. These vector-borne viruses later diverged into the tick-borne and then the mosquito-borne virus clusters, suggesting that arthropod-mediated transmission is a derived trait within the genus. Other studies supported this topology based on the NS5 gene [21]. However, an alternative phylogenetic tree based on the amino acid sequence encoded by either the NS3 gene or the entire open reading frame (ORF) demonstrated that the tick-borne and NKV viruses have diverged together and independently from the mosquito-borne flaviviruses, suggesting a common ancestor for the tick-borne and NKV viruses [19,22,23]. The mosquito-borne cluster can be further subdivided into two

epidemiologically distinct groups: the *Culex* and the *Aedes* mosquito clades. A strong correlation was observed between the mosquito clade, the mammalian host, and the type of disease caused by the virus. In general, flaviviruses transmitted by the *Culex* mosquito are neurotropic viruses associated with neurological diseases in both humans and livestock. These viruses usually cycle between mosquitoes and birds. In contrast, flaviviruses transmitted by the *Aedes* mosquito are associated with hemorrhagic fevers and are non-neurotropic in humans. These viruses usually cycle between mosquitoes and primate hosts [21].

Besides the viruses assigned to one of the three clusters within the Flavivirus genus, there are a few viruses that are currently considered tentative species of the flavivirus genus. This group includes viruses like cell fusing agent virus (CFAV), Kamiti River virus (KRV) and *Culex* flavivirus (CxFV), which have all been exclusively isolated from mosquitoes or insect cell lines [24-26]. There is no evidence that these viruses are able to infect a vertebrate host and therefore they are also referred to as insect flaviviruses. Interestingly, they have been suggested to represent the primordial forms of the Flavivirus genus [18-20,27], and CFAV- and KRV-related genomic fragments have been found integrated in a DNA form in the genome of Aedes mosquitoes [28,29].

Flaviviruses transmission by arthropods, such as mosquitoes or ticks, requires infection of the vector's midgut tissue after the ingestion of an infected blood meal and subsequent release of progeny virus. Subsequently, the virus escapes from the midgut and spreads to the hemocele, from where the virus is able to infect several other tissues, including the salivary glands. After infection of the salivary glands, the arthropod vector is able to transmit the virus via the infected saliva while feeding on a susceptible host (reviewed in [30]). The arthropod vector usually becomes persistently infected [31]. Vertebrate hosts that survive a flavivirus infection usually develop lifelong immunity, implying that flaviviruses need a constant supply of immunologically naïve hosts. As a result, the majority of the flaviviruses is enzootic and infects vertebrate hosts with a high reproductive rate like birds or small mammals (reviewed in [32]). In the case of DENV, which has four distinct serotypes, recovery from infection by one of the serotypes provides lifelong immunity against that particular serotype, but only partial and transient protection against the other three serotypes [9]. Except for dengue, humans are only infected accidentally when they intrude in the natural cycle of flaviviruses. For most flaviviruses, humans are in fact considered "dead end" hosts as these infections do not result in a significant viremia; humans are therefore unable to transmit the virus when bitten by arthropod vectors. Legendary exceptions are the dengue virus, for which humans actually seem to be the natural vertebrate host [9], and yellow fever virus (reviewed in [32]).

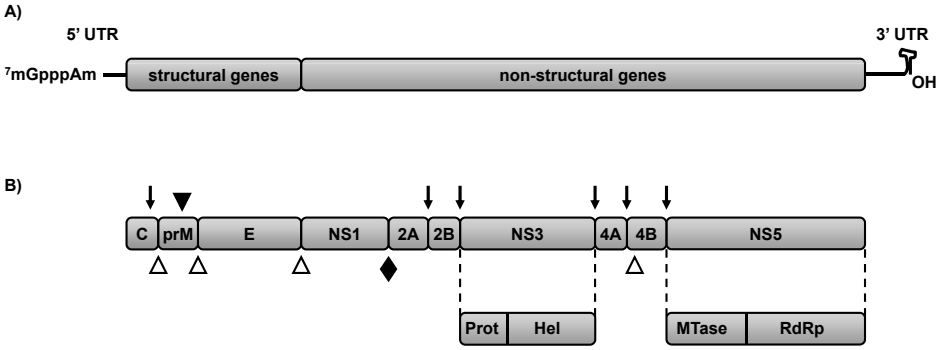


Fig. 1. Flavivirus genome organization.
A) Schematic representation of the flavivirus genome structure. The viral RNA encodes one large open reading frame (ORF). The 5'-terminal region of the ORF encodes three viral structural proteins whereas the remaining region encodes seven non-structural (NS) proteins. The ORF is flanked by 5' and 3' untranslated regions (UTR). **B)** Polyprotein processing and cleavage products. The cleavage sites for the host signal peptidase (Δ), the viral serine protease (↓), the furin-like protease (▾), as well as a yet unknown protease responsible for the NS1-NS2A cleavage (♦) are indicated. Prot and Hel in NS3 refer to the serine-like protease and helicase domains respectively. MTase and RdRp in NS5 reflect the position of the methyltransferase/ RNA capping enzyme activity and RNA-dependent RNA polymerase domains respectively.

Flavivirus RNA genome and life cycle

Flaviviruses are small (~50 nm), enveloped animal viruses containing a single positive-strand RNA genome of approximately 11 kb with a 5'-cap structure and a 3' non-polyadenylated terminus. The genomic RNA serves as the messenger RNA for translation of a single open reading frame (ORF) into a large polyprotein that is subsequently co- and post-translationally processed into the functional viral proteins by cellular and viral proteases (fig. 1). The flavivirus ORF is flanked by 5' and 3' untranslated regions (UTRs) of approximately 100 nts and 400 to 700 nts, respectively. The N-terminal region of the polyprotein encodes the viral structural proteins core (C), membrane (prM/M), and envelope (E), which are involved in the formation of the virus particle (reviewed in [33]). The core or capsid protein is a small (≈11 kDa) basic protein that forms the icosahedral nucleocapsid in which the virus genome is packaged. Nascent C (or anchored C) protein contains a COOH-terminal hydrophobic anchor that serves as a signal sequence for ER translocation of the prM protein. This hydrophobic domain is cleaved from anchored C protein by the viral protease to produce C protein for capsid assembly (reviewed in [33]). The prM protein is a glycoprotein precursor of the viral M protein. It serves as chaperone for the E protein and forms prM-E heterodimers at the envelope of the newly formed, immature virions. This prM-E interaction prevents acid-induced conformational changes in the E protein during transit through the secretory pathway [34,35]. The conversion of immature to mature virions requires the cleavage of the prM protein into pr and M

fragments by the Golgi-resident protease furin [36]. The E protein (≈ 53 kDa) is the most prominent protein on the flavivirus surface. It mediates receptor binding and membrane fusion and is an important target for the humoral immune response. The E protein structure as present in the mature virion as well as that of the post-fusion form have been determined and the combination of these structural data with cryo-EM studies have resulted in fairly detailed models for flavivirus maturation and entry (reviewed in [37]). The furin-mediated maturation of the virion (see above) catalyzes a major rearrangement of the interactions and structure of the E protein. During entry, upon exposure to low pH, the E protein homodimers dissociate into monomers which then form trimers. This reconfiguration of the E proteins exposes the previously buried fusion peptide that is subsequently inserted into the host endosomal membrane to mediate fusion between the viral envelope and the endosomal membrane; after fusion the virion RNA is released into the cytoplasm (reviewed in [37]).

The C-terminal two-thirds of the polyprotein include seven nonstructural (NS) proteins (fig. 1) that are primarily involved in viral RNA replication. NS1 is a glycoprotein of approximately 46 kDa that can be excreted from infected cells. The role of this protein in the viral life cycle is poorly understood although there is compelling evidence that it is required for RNA replication [38-41]. NS2A is a small (22 kDa) hydrophobic transmembrane protein that is important for assembly and/or release of the newly formed virus particles [42-44]. NS2B serves as an essential cofactor for the viral serine protease activity that is associated with the N-terminal region of NS3. This protease activity mediates the cleavage of the viral polyprotein at the C-terminal side of two highly conserved basic residues located at the C-terminal of the capsid protein and at the junctions NS2A/NS2B, NS2B/NS3, NS3/NS4A, and NS4B/NS5 (reviewed in [45,33]) (fig. 1). In addition, the C-terminal half of this protein functions as RNA helicase during viral RNA transcription. NS4A is a small (16 kDa) integral membrane protein which has been shown to induce membrane rearrangements [46,47]. An interaction between NS4A and NS1 was reported to be important for RNA replication [41]. NS4B is a small (27 kDa) hydrophobic protein that colocalizes with NS3 and viral double-stranded RNA in membrane-associated replication complexes [48]. NS5 is the largest viral protein (approximately 103 kDa). The N-terminal region of this protein has methyltransferase (MTase) activity and is required for the capping of the newly synthesized genomic RNA. The C-terminal part of NS5 contains the viral RdRp (reviewed in [33]) (fig. 1). The protein structure has been determined for the full-length NS3 [49,50] and for both MTase and RdRp NS5 domains [51-56] of several flaviviruses. A model for the full-length WNV NS5 structure has been proposed based on an *in silico* docking approach [52]. These structures are currently used for the rational design of inhibitors to block the essential function of these proteins in the virus life cycle. Moreover, NS2A, NS4A, NS4B and NS5 were found to be able to inhibit the host-antiviral interferon response [57-68].

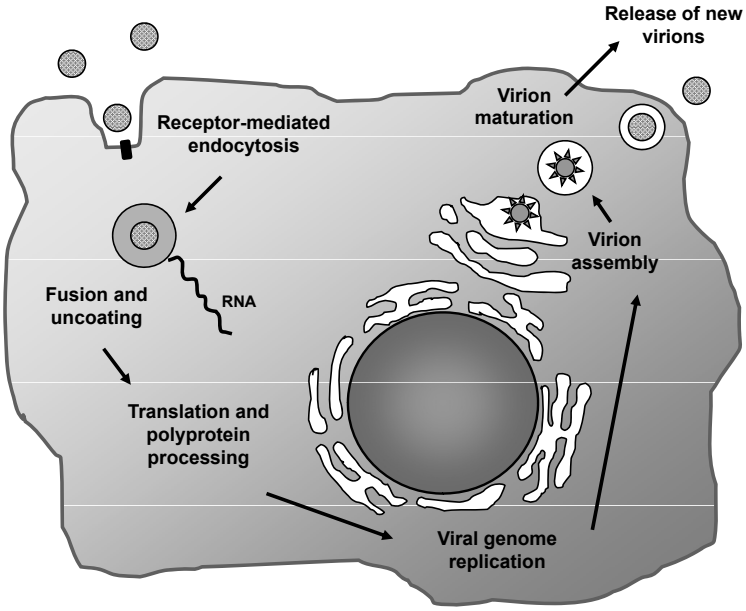


Fig. 2. Schematic representation of the Flavivirus life cycle.
See the text for more details.

Flaviviruses enter the cells by receptor-mediated endocytosis via clathrin-coated vesicles (fig. 2). They are then transported to a prelysosomal endocytic compartment. As explained in more detail above, acidification of this compartment induces a conformational change in the envelope protein that activates the fusion domain resulting in the fusion between the virus and the endosomal membrane, and ultimately resulting in the release of the viral genomic RNA into the cytoplasm. Once in the cytoplasm, the positive-strand viral RNA serves as mRNA and is translated by the host ribosomes (reviewed in [33]). Subsequent replication of the viral genome takes place in close association with virus-induced intracellular membrane structures. These membranes appear to be wrapped around the RNA amplification machinery. This replication complex (RC) is associated with unique perinuclear structures termed "vesicle packets" (VP) [69]. VPs are enriched in viral NS proteins (NS1, NS2A, NS3, NS4A, and NS5), dsRNA, and presumably some host factors [69-74]. While RNA replication takes place in these vesicle packets, translation and processing of the flavivirus polyprotein is thought to occur in association with different membrane structures designated convoluted membranes/paracrystalline arrays [72,70]. The suggested shielding of the RC by membranes is thought to prevent or reduce the exposure to cytoplasmic sensors like RIG-I or MDA5, and to dsRNA-induced host defence mechanisms, like protein kinase R, RNase L or RNA interference [75]. Additionally, it could also provide a stable and confined surface area for the RC to assemble and function [76].

A role for NS4A was demonstrated in the induction of these membrane alterations [46,47]. More recently, NS2A was also proposed to be involved in the induction of virus-specific membrane structures [44]. After the formation of the replication complex, negative-sense genome-length RNA is synthesized, which serves as a template for new positive-strand genomic molecules. Flavivirus RNA replication is an asymmetric process in which the positive-strand RNA is synthesized in 10- to 100-fold excess over the negative-strand RNA [77,78]. The newly synthesized positive-strand is subsequently used for (i) translation into new viral proteins, (ii) synthesis of negative-strand RNA, and/or (iii) encapsidation into new viral particles. Virus assembly is thought to occur by budding into the endoplasmatic reticulum (ER). The immature viral particles transit through the trans-Golgi network. Upon prM cleavage by the Golgi-resident protease furin [36], the immature viral particles turn into mature virions, which are released from the cell by the host secretory pathway (reviewed in [37], [33]) (fig. 2).

Three major viral RNA species have been detected in cells infected with flaviviruses: the genomic positive-strand RNA, a double-stranded replicative form (RF), and a heterogeneous population of replicative intermediate (RI) RNAs [77,78]. Surprisingly, an additional positive-sense small viral RNA species was reported to accumulate in both mammalian and insect cells and also in mouse brains infected with flaviviruses [79-82]. This small RNA was found to correspond to the 3′ terminal region of the viral genome [80] and to be generated by a mechanism independent of the endoribonuclease RNase L [81]. Recently, it was shown that this small flavivirus (sf) RNA is actually a product of incomplete degradation of the viral genomic RNA by the host 5′-3′ exoribonuclease XRN1 [83], the main mediator of the 5′ to 3′ mRNA decay that takes place in cytoplasmic processing bodies [84,] (reviewed in [85-87]). Interestingly, production of this sfRNA was shown to be an important parameter for viral pathogenicity [83]. The molecular basis for the role of the sfRNA in pathogenicity has not yet been elucidated.

Flavivirus genomic 3′ UTR

The 3′ UTR of flavivirus genomes is predicted to fold into a complex structure in which, despite the generally large sequence variability, a number of small, but well conserved RNA sequence elements as well as secondary and tertiary RNA structures have been identified (reviewed in [88]) (fig. 3.A). Some of these have been identified in all flaviviruses studied thus far, whereas others are characteristic for a particular cluster of the genus. The flavivirus 3′ UTR can be divided into a proximal part, immediately following the stop codon of the NS5 protein, which exhibits extensive heterogeneity in both length and sequence, and a more conserved distal part that has been defined as the core element of the 3′ UTR as it contains the majority of the elements involved in viral translation, replication, and assembly [89-96]. The 3′ end of the flavivirus genome is not polyadenylated; instead, all flavivirus genomes analyzed to date terminate with a large, stable stem-loop

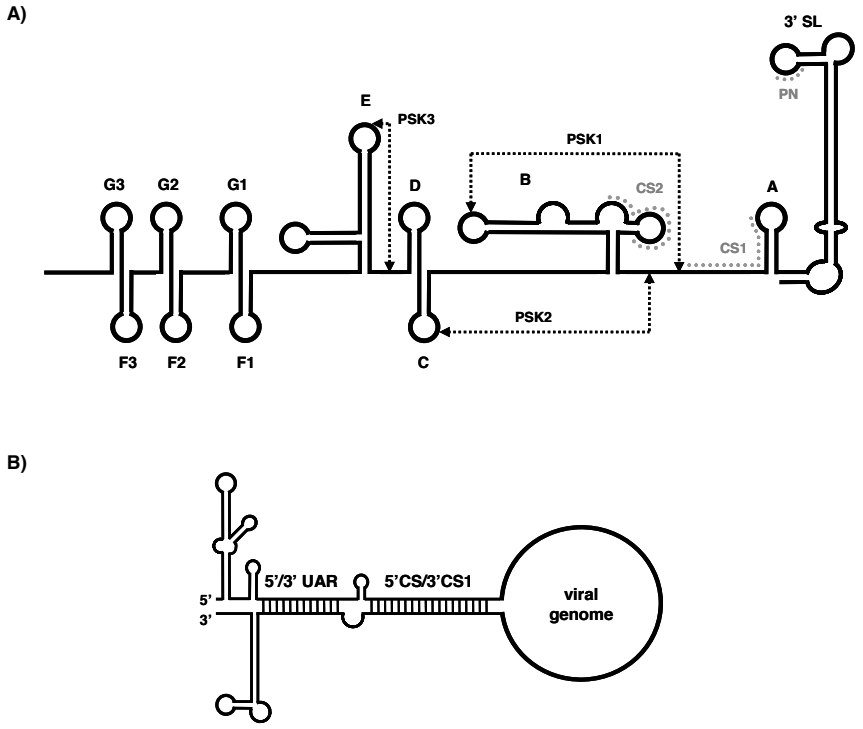


Fig. 3. Schematic model of the predicted RNA structure of the complete 3′ untranslated region (UTR) of the prototype flavivirus YFV and of the genome circularization.
A) The large 3′ terminal stem-loop structure is termed 3′ SL, other secondary RNA structures are indicated A through G as in Olsthoorn and Bol [122]. Sequences that are predicted to base pair and form an RNA pseudoknot (PSK) are connected by a dashed line. The location of conserved RNA sequences within the 3′ UTR is indicated by dots. PN, pentanucleotide motif; CS1, conserved sequence 1; CS2, conserved sequence 2. **B)** The circular conformation promoted by the long-range RNA interactions between the YFV 5′ and 3′ ends is schematically shown. The 5′/3′ UAR and 5′CS/3′CS1 interactions are indicated.

structure (3′ SL) involving 90 to 120 nts. Within the 3′ SL, two small conserved sequence motifs were identified and found to be required for viral RNA synthesis. One of these motifs is the dinucleotide 5′-CU-3′ at the 3′ end of the genome [97-99]. The other conserved sequence is the pentanucleotide motif 5′-CACAG-3′ in the top loop of the 3′ SL [97,100,101]. Upstream of the 3′ SL, there are two conserved sequence elements designated CS1 and CS2 that are well conserved among mosquito-borne flaviviruses (fig. 3.A). CS1 is found immediately adjacent to the 3′ SL and is involved in a long range RNA-RNA interaction with a complementary conserved sequence (5′ CS) near the 5′ end of the genome, downstream of the translation initiation codon in the capsid gene [102]. Base pairing of these two sequence elements allows the formation of a panhandle-like structure that mediates circularization of the viral genome (fig. 3.B). The base pairing between 3′ CS1

and 5′ CS has been shown to be critical for viral RNA synthesis [103-109]. Recent studies have demonstrated that, apart from the 5′CS – 3′CS1 interaction, another long-range RNA interaction also plays a role in promoting genome cyclization. This interaction involves complementary sequences at the 5′ end, located immediately upstream the AUG start codon region (UAR) of the ORF, and at the 3′ end within the bottom part of the 3′ SL (3′ UAR) [109] (fig. 3.B). This pair of complementary sequences has been shown to be important for viral replication [108,110-114]. In DENV and WNV yet a third interaction important for genome circularization and RNA replication was recently identified and involves nucleotides downstream of the AUG region (5′ DAR) and nucleotides downstream CS1 (3′ DAR) [115-117]. Similar long-distance RNA interactions involving elements in the terminal regions of the genome, but at a different location from 5′CS and CS1, have also been shown for the tick-borne [118,104,110,119] and NKV flaviviruses [120,121].

CS2 is approximately 24 nts in length and is located upstream of CS1. YFV contains only one copy of CS2 but the sequence is duplicated (RCS2) in members of the JEV and DENV subgroups [102]. A sequence motif with high sequence identity and a position that resembles the one of CS2 in mosquito-borne flaviviruses has also been identified in the 3′ UTR of NKV flaviviruses [120,121]; while it is apparently absent in tick-borne flaviviruses. Deletion of CS2 has little effect on viral RNA synthesis but seems to affect pathogenicity of at least YFV and DENV, as mutants lacking CS2 form turbid plaques [107,103]. Dengue viruses lacking this sequence are attenuated in rhesus monkeys [103]. The viral 3′ UTR region encompassing CS2 in mosquito-borne flaviviruses is predicted to fold into dumbbell-like structures of which a loop is thought to be involved in the formation of an RNA pseudoknot with downstream sequences [122] (fig. 3.A).

Yellow fever

Yellow fever was the first human disease shown to be caused by a virus and the third viral infection (after smallpox and rabies) to be controlled by vaccination [123]. Yellow fever is a mosquito-borne, viral hemorrhagic fever that is endemic in tropical regions of Africa and South America where it affects 45 countries with a combined population of over 900 million people [124]. WHO estimated 200,000 cases of yellow fever worldwide every year, resulting in 30.000 deaths. However, as with many diseases in rural Africa, underreporting of the actual number of infections is likely. More than 90% of the YFV cases occur in Africa and, according to the WHO, one single confirmed case of YFV in an unvaccinated population should be considered an outbreak [124].

YFV probably evolved from ancestral mosquito-borne viruses over 3.000 years ago [125]. It is postulated that the virus originated in Africa and was subsequently introduced from the Old World into the Americas during the slave trade period in the 16th century

(reviewed in [10]). In 1900, an American commission headed by Walter Reed proved that yellow fever was caused by a filterable agent and transmitted to humans by mosquitoes [126,] (reviewed in [127]). In 1927 the Rockefeller Foundation's West Africa Yellow Fever Commission isolated the virus by inoculation of a rhesus monkey with blood of an YFV-infected Ghanian male named Asibi [128]. Theiler and Smith [129] attenuated the YFV Asibi strain by serial passage in cultures of mouse brain and modified chick embryo tissues, and demonstrated the use of the resulting attenuated YFV-17D strain as a vaccine to protect humans from yellow fever infection. In 1951, Theiler was awarded with the Nobel Prize in medicine for this groundbreaking work. In 1985, the complete genome sequence of YFV-17D was reported [130]. Shortly thereafter, the nucleotide sequence of YFV-Asibi was determined and it was shown that the Asibi and 17D strains differ at 68 nucleotide positions resulting in 32 amino acid changes [131]. Despite the fact that infectious cDNA clones for both YFV-17D [132,107] and YFV-Asibi [133] (Bredenbeek, Dorner, Ploss and Rice, unpublished results) are available, the precise genetic determinants for attenuation of YFV-17D are still unknown. Several studies have shown that the flavivirus envelope protein contains important determinants for cell tropism, virulence, as well as immunity (reviewed in [134]). Therefore, it has been suggested that either one or several of the eight amino acid differences between the Asibi and 17D E proteins are critical for the attenuation of the YFV-Asibi.

Unfortunately, despite the availability of a very successful vaccine, yellow fever is still a major public health concern. Because of the variable clinical presentation of the infected individuals, yellow fever can be difficult to differentiate from other hemorrhagic fevers (e.g. Lassa fever, Ebola) and diseases like malaria, influenza, and typhoid fever, which often also occur in areas where YFV is endemic [12]. The liver is the target organ in humans and liver dysfunction results in patient's skin turning yellow, a characteristic from which the name of the disease was derived (reviewed in [135]). YFV infection presents a broad clinical spectrum varying from mild symptoms to a fatal hemorrhagic fever, with a biphasic pattern. The onset of the disease is typically within 3 to 6 days after a bite from an infected mosquito. The symptoms during this first phase of the infection include fever, headache, backache, myalgia (muscle pain), chills, malaise, nausea, dizziness and vomiting. During this phase, patients are viremic and infectious to mosquitoes. This period will last for several days and may be followed by a "period of remission", with mitigation of symptoms lasting up to 24 h. In this phase, the virus is cleared by antibodies and the cellular immune response of the individual. Most patients recover at this point without further signs or symptoms. However, in approximately 15 to 25% of those infected, symptoms reappear in a more severe form with high fever, vomiting, epigastric pain, jaundice, renal failure, hemorrhagic diathesis ("black vomit") and coma. Bleeding can occur from the mouth, nose, eyes, or stomach. This is the "period of intoxication". Viremia is generally absent, and anti-YFV antibodies appear during this phase. Depending on

the virulence of the YFV strain, up to 50% of the patients do not survive this second phase and die within 7 to 10 days after the onset of symptoms. Patients that survive the infection usually recover without significant organ damage or other lasting effects (reviewed in [10,12,124]).

There is no cure for yellow fever; vaccination is the single most important preventive measure against yellow fever. The YFV-17D vaccine is effective against all African and South American YFV genotypes and has an unmatched safety record [12]. This excellent safety record might be in part explained by the fact that the virus is quite stable as it accumulates mutations at a very low frequency in healthy vaccinees [136]. On the other hand, it has been well documented that YFV-17D vaccine preparations consist of a heterogeneous population containing a mixture of variants with distinct biological properties, such as plaque size in Vero cells, virulence for mice [137,138], and antigenicity [139-141].

The YFV-17D vaccine is an affordable, highly effective vaccine that is thought to provide protection for 30-35 years or more. Severe adverse reactions to vaccination have been reported but are extremely rare (reviewed in [135,124,142]). Despite its success, the mechanisms by which YF-17D induces protective immunity are not completely understood. Vaccination is followed by a rapid activation of both the cellular and humoral arms of the adaptive immune response. Long term protection against YFV infection appears to be exclusively mediated by neutralizing antibodies (protective levels of neutralizing antibody are found in 90% of vaccines within 10 days and in 99% within 30 days) [12]. In trying to understand the molecular basis of the efficacy of the YFV-17D vaccine, recent research has focused on the innate immune response upon infection of dendritic cells [143] and in vaccines [144] using gene expression profiling. From these studies it can be concluded that YFV-17D activates multiple Toll-like receptors (TLRs) which are likely to activate several arms of the innate immune response. A robust activation and upregulation of a complex network of genes involving innate sensing receptors (e.g. TLR7, RIG-I, MDA5), IFN-β stimulated transcription factors, and pro- and anti-inflammatory cytokines, ultimately culminates in a protective immune response [143,144] (reviewed in [142]).

YFV-17D as a platform for developing recombinant vaccines

Despite our expanding knowledge of the molecular biology, immunology and pathology of flaviviruses, relatively little progress has been made with respect to treatment of infected individuals. Currently, only a limited number of licensed vaccines to protect humans against flavivirus infections is available. These include the already discussed YFV-17D vaccine, a live attenuated as well as an inactivated JEV vaccine, and a TBEV vaccine based on inactivated virus. An inactivated WNV vaccine is also available but only licensed for use in livestock [145,146]. As for DENV, despite numerous efforts, there is no vaccine available; the situation is especially aggravated by the fact that a DENV vaccine should be tetravalent, inducing protection against the four dengue serotypes (reviewed

in [145]). Recent data indicate an even bigger challenge. DENV-infected cells apparently secrete high levels of particles containing prM instead of M due to inefficient cleavage. Antibodies against prM are generated and were found to be highly cross-reactive and able to promote the antibody-dependent enhancement (ADE) that is often associated with severe secondary infections by a different serotype [147,148]. These observations imply that an effective DENV vaccine should not only be able to offer protection from the four different serotypes but also minimize the anti-prM response.

The ability for genetic manipulation of the flavivirus genome by using available infectious cDNAs allowed the construction of live attenuated chimeric viruses. These novel recombinant vaccines make use of the fact that prM and E proteins of a particular flavivirus can be exchanged for the corresponding genes of another flavivirus, without significantly affecting the replication of the new recombinant virus. The expressed prM and E of the donor virus will drive the efficient assembly and budding of an enveloped virion in which the recombinant RNA is packaged, and will trigger an immune response against the donor virus upon vaccination. YFV-17D has been used as a vector backbone for the construction of such chimeric viruses due to its unique safety record and efficacy as a vaccine for humans. YFV-based chimeric candidate vaccine marketed as Chimerivax™ have been constructed for DENV, JEV, and WNV and have now been extensively tested in clinical trials with results that demonstrate their immunogenicity and excellent safety profile in humans [149] (reviewed in [150]).

Apart from being used as a vector for the construction of chimeric flavivirus vaccine candidates, YFV-17D has also been exploited as a vector for the expression of heterologous genes to develop recombinant vaccines against pathogens like malaria [151-153], tumours [154], Lassa virus [155,156], and HIV [157,158]. Although most of these recombinants show promising results in small scale animal experiments, genetic stability is often an issue especially with longer inserts [156,159].

SCOPE AND OUTLINE OF THIS THESIS

It is generally accepted that the 3' UTR of positive-strand RNA viruses has an important role in several steps of the virus life cycle. RNA sequences and/or structures have been implicated in the regulation of translation and replication, as well as encapsidation (reviewed in [160,161]). Flaviviruses are no exception regarding the importance of the 3' UTR for the virus cycle. Several motifs and RNA structures have been identified in the flavivirus 3' UTR (see fig. 3); most have been predicted based on phylogenetic analysis and RNA folding algorithms. Unfortunately, RNA probing data to support the predicted structures is scarce. In addition, our knowledge of the biological function of most of these predicted RNA elements is still rather vague and often limited to the biological

effect of deleting the predicted RNA structures (e.g., [107]). The only exception are the RNA sequences involved in genome circularization that have been studied in detail using a variety of techniques (e.g. *in vitro* RdRp assays, mutagenesis, atomic force microscopy, RNA structure probing) [103-110,112,113,118,119,162-167].

The major aim of the research described in this thesis was to characterize and further understand the sequence and structural requirements as well as the biological function of some of these predicted RNA elements in the flavivirus 3′ UTR. The work described in chapters 2 and 3 has been performed with the mosquito-borne YFV. The experiments in chapter 4 and 5 were carried out with several of the NKV flaviviruses that lack an arthropod vector and with the tentative flavivirus CFAV, which is thought to be an insect virus unable to infect vertebrate hosts. CFAV and the NKV flaviviruses are particularly interesting from the perspective that, as far as it is currently known, they do not cycle between different hosts. As a consequence of this apparently simpler life cycle, their 3′ UTR and the conserved RNA elements within it, are predicted to have evolved towards an optimal function in only one type of host, e.g. mosquitoes, bats, or rodents. This could potentially have resulted in a less complex 3′ UTR. These relatively unknown and poorly studied NKV- and insect flaviviruses may therefore be excellent tools to provide a better understanding of the function of conserved RNA structures and could yield valuable insight into virus-host interactions, host range restrictions or specific requirements for replication in different hosts.

Chapter 2 describes the importance and the sequence requirements of the pentanucleotide motif in the 3′ SL of YFV (see fig. 3). Of the five nucleotides (5′-CACAG-3′), only the G nucleotide at the 5th position was indispensable for viral replication. Mutations at the other positions were tolerated, although the nucleotide at the 1st position had to be able to base pair with the nucleotide four positions downstream of the PN sequence (9th position). This result provided experimental support for the predicted structure at the top of 3′ SL. Strikingly, YFV replication was found to be less dependent on sequence conservation of the pentanucleotide motif than reported for West Nile virus. Nonetheless, despite the fact that the majority of the mutations in the YFV PN motif did not seem to affect viral RNA synthesis, a clear preference for the wild-type sequence was observed when the fitness of these mutant viruses was analyzed in a competition experiment against the parental YFV-17D.

Chapter 3 presents a detailed description of the characteristics of the YFV sfRNA and the RNA structure within the viral 3′ UTR that is required for its production. Complementary *in vitro* and cell culture experiments confirmed the 5′ – 3′ RNase XRN1 as the host protein responsible for sfRNA generation. A predicted RNA pseudoknot with hitherto unknown function was shown to be essential for the production of the YFV sfRNA. Evidence to support the formation of this particular pseudoknot was obtained by RNA structure probing and mutagenesis studies.

From published data [79-81,83] and our work it became evident that all arthropod-borne flaviviruses produce at least one sfRNA in infected mammalian as well as insect cell lines. In addition, sfRNA production was shown to be an important determinant of virulence, as viruses that are unable to produce the sfRNA are less pathogenic in mice [83]. These data provide evidence that the sfRNA has at least a function in the vertebrate host, but do not necessarily exclude a function in the arthropod host. If the sfRNA has no function in the arthropod host it could be hypothesized that the (tentative) insect flavivirus CFAV would not produce an sfRNA whereas the bat- and rodent-infecting NKV flaviviruses would produce an sfRNA. Alternatively, sfRNA synthesis could be a unique hallmark of arthropod-borne flaviviruses or a characteristic feature of all flaviviruses irrespective of their host range. **Chapter 4** describes the experiments that were done to verify the various possibilities concerning sfRNA production in flaviviruses without a vector. From the data presented, it was concluded that all flaviviruses, including the tentative species CFAV, produce an sfRNA, suggesting that sfRNA generation is indeed a feature of the Flavivirus genus. The mechanism by which these sfRNAs are produced was shown to be similar to that of the arthropod-borne flaviviruses.

In contrast to most of the arthropod-borne flaviviruses, studies with NKV flaviviruses are hampered by the lack of infectious cDNA clones. **Chapter 5** describes the construction and characterization of a MODV full-length infectious cDNA clone. The clone was constructed in the low copy number vector pACNR that had been used before as a stable acceptor for the often "toxic" sequences of the *Flaviviridae* in *Escherichia coli*. MODV genome-length transcripts were shown to be highly "infectious" when transfected into BHK cells. The virus obtained from the transfected cells showed similar characteristics as the parental virus in terms of growth kinetics and plaque morphology. This clone can be used to study the function of predicted 3' UTR elements putatively important for NKV flaviviruses. Furthermore, the infectious MODV clone offers the possibility to construct chimeras with arthropod-borne flaviviruses in order to understand the molecular determinants required for a virus to be able to replicate in insect cells.

Chapter 6 presents an extensive literature review describing the characteristics and function of the RNA structures that were predicted within the Flavivirus 3' UTR. Data available for the well-studied arthropod-borne flaviviruses, as well as for the poorly studied NKV flaviviruses, is summarized and discussed. Emphasis was given to structures that were shown to be involved in viral replication and pathogenicity.

Chapter 7 is an epilogue in which the results of the experimental work are briefly summarized and discussed in a broader context. Potential functions and future research directions for the RNA structures that were studied in the work presented in this thesis are suggested.

REFERENCE LIST

1. **Morens, D. M., G. K. Folkers, and A. S. Fauci**. 2004. The challenge of emerging and re-emerging infectious diseases. Nature **430**:242-249.
2. **Beutler, B., C. Eidenschenk, K. Crozat, J. L. Imler, O. Takeuchi, J. A. Hoffmann, and S. Akira**. 2007. Genetic analysis of resistance to viral infection. Nat. Rev. Immunol. **7**:753-766.
3. **Taylor, L. H., S. M. Latham, and M. E. Woolhouse**. 2001. Risk factors for human disease emergence. Philos. Trans. R. Soc. Lond B Biol. Sci. **356**:983-989.
4. **Nichol, S.** 1996. RNA viruses. Life on the edge of catastrophe. Nature **384**:218-219.
5. **Domingo, E. and J. J. Holland**. 1997. RNA virus mutations and fitness for survival. Annu. Rev. Microbiol. **51**:151-178.
6. **ICTV**. 2005. Virus Taxonomy - Eight Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego.
7. **Thiel, H. J., M. S. Collet, E. A. Gould, F. X. Heinz, G. Meyers, R. H. Purcell, C. M. Rice, and M. Houghton**. 2005. Flaviridae, p. 981-998. *In* L. A. Ball (ed.), Virus Taxonomy - Eight Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego.
8. **Kuno, G. and G. J. Chang**. 2005. Biological transmission of arboviruses: reexamination of and new insights into components, mechanisms, and unique traits as well as their evolutionary trends. Clin. Microbiol. Rev. **18**:608-637.
9. WHO. World Health Organization. Fact sheet, no. 117. Dengue and dengue haemorrhagic fever. http://www.who.int/mediacentre/factsheets/fs117/en/. 2009.
10. **Monath, T. P.** 2001. Yellow fever: an update. Lancet Infect. Dis. **1**:11-20.
11. **Gubler, D. J.** 2004. The changing epidemiology of yellow fever and dengue, 1900 to 2003: full circle? Comp Immunol. Microbiol. Infect. Dis. **27**:319-330.
12. **WHO**. 2003. Yellow fever vaccine. WHO position paper. Wkly. Epidemiol. Rec. **78**:349-359.
13. **Briese, T., X. Y. Jia, C. Huang, L. J. Grady, and W. I. Lipkin**. 1999. Identification of a Kunjin/West Nile-like flavivirus in brains of patients with New York encephalitis. Lancet **354**:1261-1262.
14. **Gould, E. A. and S. Higgs**. 2009. Impact of climate change and other factors on emerging arbovirus diseases. Trans. R. Soc. Trop. Med. Hyg. **103**:109-121.
15. **Weissenbock, H., Z. Hubalek, T. Bakonyi, and N. Nowotny**. 2010. Zoonotic mosquito-borne flaviviruses: worldwide presence of agents with proven pathogenicity and potential candidates of future emerging diseases. Vet. Microbiol. **140**:271-280.
16. **Daniel, M., V. Danielova, B. Kriz, A. Jirsa, and J. Nozicka**. 2003. Shift of the tick Ixodes ricinus and tick-borne encephalitis to higher altitudes in central Europe. Eur. J. Clin. Microbiol. Infect. Dis. **22**:327-328.
17. **Danielova, V., M. Daniel, L. Schwarzova, J. Materna, N. Rudenko, M. Golovchenko, J. Holubova, L. Grubhoffer, and P. Kilian**. 2010. Integration of a tick-borne encephalitis virus and Borrelia burgdorferi sensu lato into mountain ecosystems, following a shift in the altitudinal limit of distribution of their vector, Ixodes ricinus (Krkonose mountains, Czech Republic). Vector. Borne. Zoonotic. Dis. **10**:223-230.
18. **Kuno, G., G. J. Chang, K. R. Tsuchiya, N. Karabatsos, and C. B. Cropp**. 1998. Phylogeny of the genus Flavivirus. J. Virol. **72**:73-83.
19. **Cook, S. and E. C. Holmes**. 2006. A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: Flaviviridae) and the evolution of vector transmission. Arch. Virol. **151**:309-325.

20. **Marin, M. S., P. M. Zanotto, T. S. Gritsun, and E. A. Gould**. 1995. Phylogeny of TYU, SRE, and CFA virus: different evolutionary rates in the genus Flavivirus. Virology **206**:1133-1139.
21. **Gaunt, M. W., A. A. Sall, L. de, X, A. K. Falconar, T. I. Dzhivanian, and E. A. Gould**. 2001. Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. J. Gen. Virol. **82**:1867-1876.
22. **Billoir, F., C. R. de, H. Tolou, M. P. de, E. A. Gould, and L. de, X**. 2000. Phylogeny of the genus flavivirus using complete coding sequences of arthropod-borne viruses and viruses with no known vector. J. Gen. Virol. **81 Pt 9**:2339.
23. **Grard, G., G. Moureau, R. N. Charrel, E. C. Holmes, E. A. Gould, and L. de, X**. 2010. Genomics and evolution of Aedes-borne flaviviruses. J. Gen. Virol. **91**:87-94.
24. **Stollar, V. and V. L. Thomas**. 1975. An agent in the Aedes aegypti cell line (Peleg) which causes fusion of Aedes albopictus cells. Virology **64**:367-377.
25. **Crabtree, M. B., R. C. Sang, V. Stollar, L. M. Dunster, and B. R. Miller**. 2003. Genetic and phenotypic characterization of the newly described insect flavivirus, Kamiti River virus. Arch. Virol. **148**:1095-1118.
26. **Hoshino, K., H. Isawa, Y. Tsuda, K. Yano, T. Sasaki, M. Yuda, T. Takasaki, M. Kobayashi, and K. Sawabe**. 2007. Genetic characterization of a new insect flavivirus isolated from Culex pipiens mosquito in Japan. Virology **359**:405-414.
27. **Cammisa-Parks, H., L. A. Cisar, A. Kane, and V. Stollar**. 1992. The complete nucleotide sequence of cell fusing agent (CFA): homology between the nonstructural proteins encoded by CFA and the nonstructural proteins encoded by arthropod-borne flaviviruses. Virology **189**:511-524.
28. **Crochu, S., S. Cook, H. Attoui, R. N. Charrel, C. R. de, M. Belhouchet, J. J. Lemasson, M. P. de, and L. de, X**. 2004. Sequences of flavivirus-related RNA viruses persist in DNA form integrated in the genome of Aedes spp. mosquitoes. J. Gen. Virol. **85**:1971-1980.
29. **Roiz, D., A. Vazquez, M. P. Seco, A. Tenorio, and A. Rizzoli**. 2009. Detection of novel insect flavivirus sequences integrated in Aedes albopictus (Diptera: Culicidae) in Northern Italy. Virol. J. **6**:93.
30. **Gubler, D. J., G. Kuno, and L. Markoff**. 2007. Flaviviruses *In* D. M. Knipe and P. M. Howley (eds.), Fields Virology. Lippincott Williams & Wilkins, Philadelphia.
31. **Blair, C. D., Z. N. Adelman, and K. E. Olson**. 2000. Molecular strategies for interrupting arthropod-borne virus transmission by mosquitoes. Clin. Microbiol. Rev. **13**:651-661.
32. **Solomon, T. and M. Mallewa**. 2001. Dengue and other emerging flaviviruses. J. Infect. **42**:104-115.
33. **Lindenbach, B. D., H. J. Thiel, and C. M. Rice**. 2007. Flaviviridae: The Viruses and Their Replication *In* D. M. Knipe and P. M. Howley (eds.), Fields Virology. Lippincott Williams & Wilkins, Philadelphia.
34. **Heinz, F. X., K. Stiasny, G. Puschner-Auer, H. Holzmann, S. L. Allison, C. W. Mandl, and C. Kunz**. 1994. Structural changes and functional control of the tick-borne encephalitis virus glycoprotein E by the heterodimeric association with protein prM. Virology **198**:109-117.
35. **Guirakhoo, F., R. A. Bolin, and J. T. Roehrig**. 1992. The Murray Valley encephalitis virus prM protein confers acid resistance to virus particles and alters the expression of epitopes within the R2 domain of E glycoprotein. Virology **191**:921-931.
36. **Stadler, K., S. L. Allison, J. Schalich, and F. X. Heinz**. 1997. Proteolytic activation of tick-borne encephalitis virus by furin. J. Virol. **71**:8475-8481.
37. **Mukhopadhyay, S., R. J. Kuhn, and M. G. Rossmann**. 2005. A structural perspective of the flavivirus life cycle. Nat. Rev. Microbiol. **3**:13-22.

38. **Muylaert, I. R., T. J. Chambers, R. Galler, and C. M. Rice**. 1996. Mutagenesis of the N-linked glycosylation sites of the yellow fever virus NS1 protein: effects on virus replication and mouse neurovirulence. Virology **222**:159-168.
39. **Muylaert, I. R., R. Galler, and C. M. Rice**. 1997. Genetic analysis of the yellow fever virus NS1 protein: identification of a temperature-sensitive mutation which blocks RNA accumulation. J. Virol. **71**:291-298.
40. **Lindenbach, B. D. and C. M. Rice**. 1997. trans-Complementation of yellow fever virus NS1 reveals a role in early RNA replication. J. Virol. **71**:9608-9617.
41. **Lindenbach, B. D. and C. M. Rice**. 1999. Genetic interaction of flavivirus nonstructural proteins NS1 and NS4A as a determinant of replicase function. J. Virol. **73**:4611-4621.
42. **Kummerer, B. M. and C. M. Rice**. 2002. Mutations in the yellow fever virus nonstructural protein NS2A selectively block production of infectious particles. J. Virol. **76**:4773-4784.
43. **Liu, W. J., H. B. Chen, and A. A. Khromykh**. 2003. Molecular and functional analyses of Kunjin virus infectious cDNA clones demonstrate the essential roles for NS2A in virus assembly and for a nonconservative residue in NS3 in RNA replication. J. Virol. **77**:7804-7813.
44. **Leung, J. Y., G. P. Pijlman, N. Kondratieva, J. Hyde, J. M. Mackenzie, and A. A. Khromykh**. 2008. Role of nonstructural protein NS2A in flavivirus assembly. J. Virol. **82**:4731-4741.
45. **Chambers, T. J., C. S. Hahn, R. Galler, and C. M. Rice**. 1990. Flavivirus genome organization, expression, and replication. Annu. Rev. Microbiol. **44**:649-688.
46. **Roosendaal, J., E. G. Westaway, A. Khromykh, and J. M. Mackenzie**. 2006. Regulated cleavages at the West Nile virus NS4A-2K-NS4B junctions play a major role in rearranging cytoplasmic membranes and Golgi trafficking of the NS4A protein. J. Virol. **80**:4623-4632.
47. **Miller, S., S. Kastner, J. Krijnse-Locker, S. Buhler, and R. Bartenschlager**. 2007. The non-structural protein 4A of dengue virus is an integral membrane protein inducing membrane alterations in a 2K-regulated manner. J. Biol. Chem. **282**:8873-8882.
48. **Miller, S., S. Sparacio, and R. Bartenschlager**. 2006. Subcellular localization and membrane topology of the Dengue virus type 2 Non-structural protein 4B. J. Biol. Chem. **281**:8854-8863.
49. **Luo, D., T. Xu, C. Hunke, G. Gruber, S. G. Vasudevan, and J. Lescar**. 2008. Crystal structure of the NS3 protease-helicase from dengue virus. J. Virol. **82**:173-183.
50. **Assenberg, R., E. Mastrangelo, T. S. Walter, A. Verma, M. Milani, R. J. Owens, D. I. Stuart, J. M. Grimes, and E. J. Mancini**. 2009. Crystal structure of a novel conformational state of the flavivirus NS3 protein: implications for polyprotein processing and viral replication. J. Virol. **83**:12895-12906.
51. **Egloff, M. P., D. Benarroch, B. Selisko, J. L. Romette, and B. Canard**. 2002. An RNA cap (nucleoside-2′-O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. EMBO J. **21**:2757-2768.
52. **Malet, H., M. P. Egloff, B. Selisko, R. E. Butcher, P. J. Wright, M. Roberts, A. Gruez, G. Sulzenbacher, C. Vonrhein, G. Bricogne, J. M. Mackenzie, A. A. Khromykh, A. D. Davidson, and B. Canard**. 2007. Crystal structure of the RNA polymerase domain of the West Nile virus nonstructural protein 5. J. Biol. Chem. **282**:10678-10689.
53. **Yap, T. L., Y. L. Chen, T. Xu, D. Wen, S. G. Vasudevan, and J. Lescar**. 2007. A multi-step strategy to obtain crystals of the dengue virus RNA-dependent RNA polymerase that diffract to high resolution. Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun. **63**:78-83.
54. **Yap, T. L., T. Xu, Y. L. Chen, H. Malet, M. P. Egloff, B. Canard, S. G. Vasudevan, and J. Lescar**. 2007. Crystal structure of the dengue virus RNA-dependent RNA polymerase catalytic domain at 1.85-angstrom resolution. J. Virol. **81**:4753-4765.

55. **Zhou, Y., D. Ray, Y. Zhao, H. Dong, S. Ren, Z. Li, Y. Guo, K. A. Bernard, P. Y. Shi, and H. Li**. 2007. Structure and function of flavivirus NS5 methyltransferase. J. Virol. **81**:3891-3903.
56. **Assenberg, R., J. Ren, A. Verma, T. S. Walter, D. Alderton, R. J. Hurrelbrink, S. D. Fuller, S. Bressanelli, R. J. Owens, D. I. Stuart, and J. M. Grimes**. 2007. Crystal structure of the Murray Valley encephalitis virus NS5 methyltransferase domain in complex with cap analogues. J. Gen. Virol. **88**:2228-2236.
57. **Munoz-Jordan, J. L., G. G. Sanchez-Burgos, M. Laurent-Rolle, and A. Garcia-Sastre**. 2003. Inhibition of interferon signaling by dengue virus. Proc. Natl. Acad. Sci. U. S. A **100**:14333-14338.
58. **Liu, W. J., H. B. Chen, X. J. Wang, H. Huang, and A. A. Khromykh**. 2004. Analysis of adaptive mutations in Kunjin virus replicon RNA reveals a novel role for the flavivirus nonstructural protein NS2A in inhibition of beta interferon promoter-driven transcription. J. Virol. **78**:12225-12235.
59. **Liu, W. J., X. J. Wang, V. V. Mokhonov, P. Y. Shi, R. Randall, and A. A. Khromykh**. 2005. Inhibition of interferon signaling by the New York 99 strain and Kunjin subtype of West Nile virus involves blockage of STAT1 and STAT2 activation by nonstructural proteins. J. Virol. **79**:1934-1942.
60. **Munoz-Jordan, J. L., M. Laurent-Rolle, J. Ashour, L. Martinez-Sobrido, M. Ashok, W. I. Lipkin, and A. Garcia-Sastre**. 2005. Inhibition of alpha/beta interferon signaling by the NS4B protein of flaviviruses. J. Virol. **79**:8004-8013.
61. **Best, S. M., K. L. Morris, J. G. Shannon, S. J. Robertson, D. N. Mitzel, G. S. Park, E. Boer, J. B. Wolfinbarger, and M. E. Bloom**. 2005. Inhibition of interferon-stimulated JAK-STAT signaling by a tick-borne flavivirus and identification of NS5 as an interferon antagonist. J. Virol. **79**:12828-12839.
62. **Liu, W. J., X. J. Wang, D. C. Clark, M. Lobigs, R. A. Hall, and A. A. Khromykh**. 2006. A single amino acid substitution in the West Nile virus nonstructural protein NS2A disables its ability to inhibit alpha/beta interferon induction and attenuates virus virulence in mice. J. Virol. **80**:2396-2404.
63. **Lin, R. J., B. L. Chang, H. P. Yu, C. L. Liao, and Y. L. Lin**. 2006. Blocking of interferon-induced Jak-Stat signaling by Japanese encephalitis virus NS5 through a protein tyrosine phosphatase-mediated mechanism. J. Virol. **80**:5908-5918.
64. **Park, G. S., K. L. Morris, R. G. Hallett, M. E. Bloom, and S. M. Best**. 2007. Identification of residues critical for the interferon antagonist function of Langat virus NS5 reveals a role for the RNA-dependent RNA polymerase domain. J. Virol. **81**:6936-6946.
65. **Werme, K., M. Wigerius, and M. Johansson**. 2008. Tick-borne encephalitis virus NS5 associates with membrane protein scribble and impairs interferon-stimulated JAK-STAT signalling. Cell Microbiol. **10**:696-712.
66. **Mazzon, M., M. Jones, A. Davidson, B. Chain, and M. Jacobs**. 2009. Dengue virus NS5 inhibits interferon-alpha signaling by blocking signal transducer and activator of transcription 2 phosphorylation. J. Infect. Dis. **200**:1261-1270.
67. **Ashour, J., M. Laurent-Rolle, P. Y. Shi, and A. Garcia-Sastre**. 2009. NS5 of dengue virus mediates STAT2 binding and degradation. J. Virol. **83**:5408-5418.
68. **Laurent-Rolle, M., E. F. Boer, K. J. Lubick, J. B. Wolfinbarger, A. B. Carmody, B. Rockx, W. Liu, J. Ashour, W. L. Shupert, M. R. Holbrook, A. D. Barrett, P. W. Mason, M. E. Bloom, A. Garcia-Sastre, A. A. Khromykh, and S. M. Best**. 2010. The NS5 protein of the virulent West Nile virus NY99 strain is a potent antagonist of type I interferon-mediated JAK-STAT signaling. J. Virol. **84**:3503-3515.
69. **Mackenzie, J. M., M. K. Jones, and P. R. Young**. 1996. Immunolocalization of the dengue virus nonstructural glycoprotein NS1 suggests a role in viral RNA replication. Virology **220**:232-240.

70. **Mackenzie, J. M., A. A. Khromykh, M. K. Jones, and E. G. Westaway**. 1998. Subcellular localization and some biochemical properties of the flavivirus Kunjin nonstructural proteins NS2A and NS4A. Virology **245**:203-215.
71. **Mackenzie, J. M., M. T. Kenney, and E. G. Westaway**. 2007. West Nile virus strain Kunjin NS5 polymerase is a phosphoprotein localized at the cytoplasmic site of viral RNA synthesis. J. Gen. Virol. **88**:1163-1168.
72. **Westaway, E. G., J. M. Mackenzie, M. T. Kenney, M. K. Jones, and A. A. Khromykh**. 1997. Ultrastructure of Kunjin virus-infected cells: colocalization of NS1 and NS3 with double-stranded RNA, and of NS2B with NS3, in virus-induced membrane structures. J. Virol. **71**:6650-6661.
73. **Westaway, E. G., A. A. Khromykh, and J. M. Mackenzie**. 1999. Nascent flavivirus RNA colocalized in situ with double-stranded RNA in stable replication complexes. Virology **258**:108-117.
74. **Welsch, S., S. Miller, I. Romero-Brey, A. Merz, C. K. Bleck, P. Walther, S. D. Fuller, C. Antony, J. Krijnse-Locker, and R. Bartenschlager**. 2009. Composition and three-dimensional architecture of the dengue virus replication and assembly sites. Cell Host. Microbe **5**:365-375.
75. **Uchil, P. D. and V. Satchidanandam**. 2003. Architecture of the flaviviral replication complex. Protease, nuclease, and detergents reveal encasement within double-layered membrane compartments. J. Biol. Chem. **278**:24388-24398.
76. **Mackenzie, J.** 2005. Wrapping things up about virus RNA replication. Traffic. **6**:967-977.
77. **Chu, P. W. and E. G. Westaway**. 1985. Replication strategy of Kunjin virus: evidence for recycling role of replicative form RNA as template in semiconservative and asymmetric replication. Virology **140**:68-79.
78. **Cleaves, G. R., T. E. Ryan, and R. W. Schlesinger**. 1981. Identification and characterization of type 2 dengue virus replicative intermediate and replicative form RNAs. Virology **111**:73-83.
79. **Urosevic, N., M. M. van, J. P. Mansfield, J. S. Mackenzie, and G. R. Shellam**. 1997. Molecular characterization of virus-specific RNA produced in the brains of flavivirus-susceptible and -resistant mice after challenge with Murray Valley encephalitis virus. J. Gen. Virol. **78 (Pt 1)**:23-29.
80. **Lin, K. C., H. L. Chang, and R. Y. Chang**. 2004. Accumulation of a 3′-terminal genome fragment in Japanese encephalitis virus-infected mammalian and mosquito cells. J. Virol. **78**:5133-5138.
81. **Scherbik, S. V., J. M. Paranjape, B. M. Stockman, R. H. Silverman, and M. A. Brinton**. 2006. RNase L plays a role in the antiviral response to West Nile virus. J. Virol. **80**:2987-2999.
82. **Liu, R., L. Yue, X. Li, X. Yu, H. Zhao, Z. Jiang, E. Qin, and C. Qin**. 2010. Identification and characterization of small sub-genomic RNAs in dengue 1-4 virus-infected cell cultures and tissues. Biochem. Biophys. Res. Commun. **391**:1099-1103.
83. **Pijlman, G. P., A. Funk, N. Kondratieva, J. Leung, S. Torres, L. van der Aa, W. J. Liu, A. C. Palmenberg, P. Y. Shi, R. A. Hall, and A. A. Khromykh**. 2008. A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity. Cell Host. Microbe **4**:579-591.
84. **Sheth, U. and R. Parker**. 2003. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science **300**:805-808.
85. **Anderson, P. and N. Kedersha**. 2006. RNA granules. J. Cell Biol. **172**:803-808.
86. **Eulalio, A., I. Behm-Ansmant, and E. Izaurralde**. 2007. P bodies: at the crossroads of post-transcriptional pathways. Nat. Rev. Mol. Cell Biol. **8**:9-22.
87. **Garneau, N. L., J. Wilusz, and C. J. Wilusz**. 2007. The highways and byways of mRNA decay. Nat. Rev. Mol. Cell Biol. **8**:113-126.
88. **Markoff, L.** 2003. 5′- and 3′-noncoding regions in flavivirus RNA. Adv. Virus Res. **59**:177-228.

89. **Mandl, C. W., C. Kunz, and F. X. Heinz**. 1991. Presence of poly(A) in a flavivirus: significant differences between the 3′ noncoding regions of the genomic RNAs of tick-borne encephalitis virus strains. J. Virol. **65**:4070-4077.
90. **Wallner, G., C. W. Mandl, C. Kunz, and F. X. Heinz**. 1995. The flavivirus 3′-noncoding region: extensive size heterogeneity independent of evolutionary relationships among strains of tick-borne encephalitis virus. Virology **213**:169-178.
91. **Poidinger, M., R. A. Hall, and J. S. Mackenzie**. 1996. Molecular characterization of the Japanese encephalitis serocomplex of the flavivirus genus. Virology **218**:417-421.
92. **Wang, E., S. C. Weaver, R. E. Shope, R. B. Tesh, D. M. Watts, and A. D. Barrett**. 1996. Genetic variation in yellow fever virus: duplication in the 3′ noncoding region of strains from Africa. Virology **225**:274-281.
93. **Gritsun, T. S., K. Venugopal, P. M. Zanotto, M. V. Mikhailov, A. A. Sall, E. C. Holmes, I. Polkinghorne, T. V. Frolova, V. V. Pogodina, V. A. Lashkevich, and E. A. Gould**. 1997. Complete sequence of two tick-borne flaviviruses isolated from Siberia and the UK: analysis and significance of the 5′ and 3′-UTRs. Virus Res. **49**:27-39.
94. **Proutski, V., E. A. Gould, and E. C. Holmes**. 1997. Secondary structure of the 3′ untranslated region of flaviviruses: similarities and differences. Nucleic Acids Res. **25**:1194-1202.
95. **Rauscher, S., C. Flamm, C. W. Mandl, F. X. Heinz, and P. F. Stadler**. 1997. Secondary structure of the 3′-noncoding region of flavivirus genomes: comparative analysis of base pairing probabilities. RNA. **3**:779-791.
96. **Mandl, C. W., H. Holzmann, T. Meixner, S. Rauscher, P. F. Stadler, S. L. Allison, and F. X. Heinz**. 1998. Spontaneous and engineered deletions in the 3′ noncoding region of tick-borne encephalitis virus: construction of highly attenuated mutants of a flavivirus. J. Virol. **72**:2132-2140.
97. **Khromykh, A. A., N. Kondratieva, J. Y. Sgro, A. Palmenberg, and E. G. Westaway**. 2003. Significance in replication of the terminal nucleotides of the flavivirus genome. J. Virol. **77**:10623-10629.
98. **Nomaguchi, M., M. Ackermann, C. Yon, S. You, and R. Padmanabhan**. 2003. De novo synthesis of negative-strand RNA by Dengue virus RNA-dependent RNA polymerase in vitro: nucleotide, primer, and template parameters. J. Virol. **77**:8831-8842.
99. **Tilgner, M. and P. Y. Shi**. 2004. Structure and function of the 3′ terminal six nucleotides of the west nile virus genome in viral replication. J. Virol. **78**:8159-8171.
100. **Tilgner, M., T. S. Deas, and P. Y. Shi**. 2005. The flavivirus-conserved penta-nucleotide in the 3′ stem-loop of the West Nile virus genome requires a specific sequence and structure for RNA synthesis, but not for viral translation. Virology **331**:375-386.
101. **Elghonemy, S., W. G. Davis, and M. A. Brinton**. 2005. The majority of the nucleotides in the top loop of the genomic 3′ terminal stem loop structure are cis-acting in a West Nile virus infectious clone. Virology **331**:238-246.
102. **Hahn, C. S., Y. S. Hahn, C. M. Rice, E. Lee, L. Dalgarno, E. G. Strauss, and J. H. Strauss**. 1987. Conserved elements in the 3′ untranslated region of flavivirus RNAs and potential cyclization sequences. J. Mol. Biol. **198**:33-41.
103. **Men, R., M. Bray, D. Clark, R. M. Chanock, and C. J. Lai**. 1996. Dengue type 4 virus mutants containing deletions in the 3′ noncoding region of the RNA genome: analysis of growth restriction in cell culture and altered viremia pattern and immunogenicity in rhesus monkeys. J. Virol. **70**:3930-3937.
104. **Khromykh, A. A., H. Meka, K. J. Guyatt, and E. G. Westaway**. 2001. Essential role of cyclization sequences in flavivirus RNA replication. J. Virol. **75**:6719-6728.

105. **Lo, M. K., M. Tilgner, K. A. Bernard, and P. Y. Shi**. 2003. Functional analysis of mosquito-borne flavivirus conserved sequence elements within 3′ untranslated region of West Nile virus by use of a reporting replicon that differentiates between viral translation and RNA replication. J. Virol. **77**:10004-10014.
106. **Corver, J., E. Lenches, K. Smith, R. A. Robison, T. Sando, E. G. Strauss, and J. H. Strauss**. 2003. Fine mapping of a cis-acting sequence element in yellow fever virus RNA that is required for RNA replication and cyclization. J. Virol. **77**:2265-2270.
107. **Bredenbeek, P. J., E. A. Kooi, B. Lindenbach, N. Huijkman, C. M. Rice, and W. J. Spaan**. 2003. A stable full-length yellow fever virus cDNA clone and the role of conserved RNA elements in flavivirus replication. J. Gen. Virol. **84**:1261-1268.
108. **Alvarez, D. E., A. L. De Lella Ezcurra, S. Fucito, and A. V. Gamarnik**. 2005. Role of RNA structures present at the 3′UTR of dengue virus on translation, RNA synthesis, and viral replication. Virology **339**:200-212.
109. **Alvarez, D. E., M. F. Lodeiro, S. J. Luduena, L. I. Pietrasanta, and A. V. Gamarnik**. 2005. Long-range RNA-RNA interactions circularize the dengue virus genome. J. Virol. **79**:6631-6643.
110. **Thurner, C., C. Witwer, I. L. Hofacker, and P. F. Stadler**. 2004. Conserved RNA secondary structures in Flaviviridae genomes. J. Gen. Virol. **85**:1113-1124.
111. **Filomatori, C. V., M. F. Lodeiro, D. E. Alvarez, M. M. Samsa, L. Pietrasanta, and A. V. Gamarnik**. 2006. A 5′ RNA element promotes dengue virus RNA synthesis on a circular genome. Genes Dev. **20**:2238-2249.
112. **Zhang, B., H. Dong, D. A. Stein, P. L. Iversen, and P. Y. Shi**. 2008. West Nile virus genome cyclization and RNA replication require two pairs of long-distance RNA interactions. Virology **373**:1-13.
113. **Alvarez, D. E., C. V. Filomatori, and A. V. Gamarnik**. 2008. Functional analysis of dengue virus cyclization sequences located at the 5′ and 3′UTRs. Virology **375**:223-235.
114. **Song, B. H., S. I. Yun, Y. J. Choi, J. M. Kim, C. H. Lee, and Y. M. Lee**. 2008. A complex RNA motif defined by three discontinuous 5-nucleotide-long strands is essential for Flavivirus RNA replication. RNA. **14**:1791-1813.
115. **Friebe, P. and E. Harris**. 2010. Interplay of RNA elements in the dengue virus 5′ and 3′ ends required for viral RNA replication. J. Virol. **84**:6103-6118.
116. **Friebe, P., P. Y. Shi, and E. Harris**. 2010. The 5′ and 3′ Downstream of AUG region (DAR) elements are required for mosquito-borne flavivirus RNA replication. J. Virol.
117. **Zhang, B., H. Dong, H. Ye, M. Tilgner, and P. Y. Shi**. 2010. Genetic analysis of West Nile virus containing a complete 3′CSI RNA deletion. Virology **408**:138-145.
118. **Mandl, C. W., H. Holzmann, C. Kunz, and F. X. Heinz**. 1993. Complete genomic sequence of Powassan virus: evaluation of genetic elements in tick-borne versus mosquito-borne flaviviruses. Virology **194**:173-184.
119. **Kofler, R. M., V. M. Hoenninger, C. Thurner, and C. W. Mandl**. 2006. Functional analysis of the tick-borne encephalitis virus cyclization elements indicates major differences between mosquito-borne and tick-borne flaviviruses. J. Virol. **80**:4099-4113.
120. **Leyssen, P., N. Charlier, P. Lemey, F. Billoir, A. M. Vandamme, C. E. De, L. de, X, and J. Neyts**. 2002. Complete genome sequence, taxonomic assignment, and comparative analysis of the untranslated regions of the Modoc virus, a flavivirus with no known vector. Virology **293**:125-140.
121. **Charlier, N., P. Leyssen, C. W. Pleij, P. Lemey, F. Billoir, L. K. Van, A. M. Vandamme, C. E. De, L. de, X, and J. Neyts**. 2002. Complete genome sequence of Montana Myotis leukoencephalitis virus, phylogenetic analysis and comparative study of the 3′ untranslated region of flaviviruses with no known vector. J. Gen. Virol. **83**:1875-1885.

122. **Olsthoorn, R. C. and J. F. Bol**. 2001. Sequence comparison and secondary structure analysis of the 3′ noncoding region of flavivirus genomes reveals multiple pseudoknots. RNA. **7**:1370-1377.
123. **Monath, T. P., R. M. Kinney, J. J. Schlesinger, M. W. Brandriss, and P. Bres**. 1983. Ontogeny of yellow fever 17D vaccine: RNA oligonucleotide fingerprint and monoclonal antibody analyses of vaccines produced world-wide. J. Gen. Virol. **64 Pt 3**:627-637.
124. WHO. World Health Organization. Fact sheet, no. 100. Yellow Fever. http://www.who.int/media-centre/factsheets/fs100/en/. 2009.
125. **Zanotto, P. M., E. A. Gould, G. F. Gao, P. H. Harvey, and E. C. Holmes**. 1996. Population dynamics of flaviviruses revealed by molecular phylogenies. Proc. Natl. Acad. Sci. U. S. A **93**:548-553.
126. **Reed, W., J. Carroll, A. Agramonte, and J. W. Lazear**. 1900. The Etiology of Yellow Fever-A Preliminary Note. Public Health Pap. Rep. **26**:37-53.
127. **Monath, T. P.** 1985. Glad tidings from yellow fever research. Science **229**:734-735.
128. **Stokes, A., J. H. Bauer, and N. P. Hudson**. 2001. The transmission of yellow fever to Macacus rhesus. 1928. Rev. Med. Virol. **11**:141-148.
129. **Theiler, M. and H. H. Smith**. 2000. The use of yellow fever virus modified by in vitro cultivation for human immunization. J. Exp. Med. 65, 787-800 (1937). Rev. Med. Virol. **10**:6-16.
130. **Rice, C. M., E. M. Lenches, S. R. Eddy, S. J. Shin, R. L. Sheets, and J. H. Strauss**. 1985. Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. Science **229**:726-733.
131. **Hahn, C. S., J. M. Dalrymple, J. H. Strauss, and C. M. Rice**. 1987. Comparison of the virulent Asibi strain of yellow fever virus with the 17D vaccine strain derived from it. Proc. Natl. Acad. Sci. U. S. A **84**:2019-2023.
132. **Rice, C. M., A. Grakoui, R. Galler, and T. J. Chambers**. 1989. Transcription of infectious yellow fever RNA from full-length cDNA templates produced by in vitro ligation. New Biol. **1**:285-296.
133. **McElroy, K. L., K. A. Tsetsarkin, D. L. Vanlandingham, and S. Higgs**. 2005. Characterization of an infectious clone of the wild-type yellow fever virus Asibi strain that is able to infect and disseminate in mosquitoes. J. Gen. Virol. **86**:1747-1751.
134. **McMinn, P. C.** 1997. The molecular basis of virulence of the encephalitogenic flaviviruses. J. Gen. Virol. **78 (Pt 11)**:2711-2722.
135. **Barrett, A. D. and D. E. Teuwen**. 2009. Yellow fever vaccine - how does it work and why do rare cases of serious adverse events take place? Curr. Opin. Immunol. **21**:308-313.
136. **Xie, H., A. R. Cass, and A. D. Barrett**. 1998. Yellow fever 17D vaccine virus isolated from healthy vaccinees accumulates very few mutations. Virus Res. **55**:93-99.
137. **Liprandi, F.** 1981. Isolation of plaque variants differing in virulence from the 17D strain of yellow fever virus. J. Gen. Virol. **56**:363-370.
138. **Barrett, A. D. and E. A. Gould**. 1986. Comparison of neurovirulence of different strains of yellow fever virus in mice. J. Gen. Virol. **67 (Pt 4)**:631-637.
139. **Gould, E. A., A. Buckley, N. Cammack, A. D. Barrett, J. C. Clegg, R. Ishak, and M. G. Varma**. 1985. Examination of the immunological relationships between flaviviruses using yellow fever virus monoclonal antibodies. J. Gen. Virol. **66 (Pt 7)**:1369-1382.
140. **Buckley, A. and E. A. Gould**. 1985. Neutralization of yellow fever virus studied using monoclonal and polyclonal antibodies. J. Gen. Virol. **66 (Pt 12)**:2523-2531.
141. **Gould, E. A., A. Buckley, P. A. Cane, S. Higgs, and N. Cammack**. 1989. Use of a monoclonal antibody specific for wild-type yellow fever virus to identify a wild-type antigenic variant in 17D vaccine pools. J. Gen. Virol. **70 (Pt 7)**:1889-1894.

142. **Pulendran, B.** 2009. Learning immunology from the yellow fever vaccine: innate immunity to systems vaccinology. Nat. Rev. Immunol. **9**:741-747.
143. **Querec, T., S. Bennouna, S. Alkan, Y. Laouar, K. Gorden, R. Flavell, S. Akira, R. Ahmed, and B. Pulendran**. 2006. Yellow fever vaccine YF-17D activates multiple dendritic cell subsets via TLR2, 7, 8, and 9 to stimulate polyvalent immunity. J. Exp. Med. **203**:413-424.
144. **Silva, M. L., M. A. Martins, L. R. Espirito-Santo, A. C. Campi-Azevedo, D. Silveira-Lemos, J. G. Ribeiro, A. Homma, E. G. Kroon, A. Teixeira-Carvalho, S. M. Eloi-Santos, and O. A. Martins-Filho**. 2010. Characterization of main cytokines source from the innate and adaptive immune responses following primary 17DD yellow fever vaccination in adults. Vaccine.
145. **Pugachev, K. V., F. Guirakhoo, and T. P. Monath**. 2005. New developments in flavivirus vaccines with special attention to yellow fever. Curr. Opin. Infect. Dis. **18**:387-394.
146. **Halstead, S. B. and S. J. Thomas**. 2010. Japanese encephalitis: new options for active immunization. Clin. Infect. Dis. **50**:1155-1164.
147. **Rodenhuis-Zybert, I. A., H. M. van der Schaar, J. M. Silva Voorham, H. Ende-Metselaar, H. Y. Lei, J. Wilschut, and J. M. Smit**. 2010. Immature dengue virus: a veiled pathogen? PLoS. Pathog. **6**:e1000718.
148. **Dejnirattisai, W., A. Jumnainsong, N. Onsirisakul, P. Fitton, S. Vasanawathana, W. Limpitikul, C. Puttikhunt, C. Edwards, T. Duangchinda, S. Supasa, K. Chawansuntati, P. Malasit, J. Mongkolsapaya, and G. Screaton**. 2010. Cross-reacting antibodies enhance dengue virus infection in humans. Science **328**:745-748.
149. **Morrison, D., T. J. Legg, C. W. Billings, R. Forrat, S. Yoksan, and J. Lang**. 2010. A novel tetravalent dengue vaccine is well tolerated and immunogenic against all 4 serotypes in flavivirus-naive adults. J. Infect. Dis. **201**:370-377.
150. **Guy, B., F. Guirakhoo, V. Barban, S. Higgs, T. P. Monath, and J. Lang**. 2010. Preclinical and clinical development of YFV 17D-based chimeric vaccines against dengue, West Nile and Japanese encephalitis viruses. Vaccine **28**:632-649.
151. **Bonaldo, M. C., R. C. Garratt, P. S. Caufour, M. S. Freire, M. M. Rodrigues, R. S. Nussenzweig, and R. Galler**. 2002. Surface expression of an immunodominant malaria protein B cell epitope by yellow fever virus. J. Mol. Biol. **315**:873-885.
152. **Tao, D., G. Barba-Spaeth, U. Rai, V. Nussenzweig, C. M. Rice, and R. S. Nussenzweig**. 2005. Yellow fever 17D as a vaccine vector for microbial CTL epitopes: protection in a rodent malaria model. J. Exp. Med. **201**:201-209.
153. **Stoyanov, C. T., S. B. Boscardin, S. Deroubaix, G. Barba-Spaeth, D. Franco, R. S. Nussenzweig, M. Nussenzweig, and C. M. Rice**. 2010. Immunogenicity and protective efficacy of a recombinant yellow fever vaccine against the murine malarial parasite Plasmodium yoelii. Vaccine **28**:4644-4652.
154. **McAllister, A., A. E. Arbetman, S. Mandl, C. Pena-Rossi, and R. Andino**. 2000. Recombinant yellow fever viruses are effective therapeutic vaccines for treatment of murine experimental solid tumors and pulmonary metastases. J. Virol. **74**:9197-9205.
155. **Bredenbeek, P. J., R. Molenkamp, W. J. Spaan, V. Deubel, P. Marianneau, M. S. Salvato, D. Moshkoff, J. Zapata, I. Tikhonov, J. Patterson, R. Carrion, A. Ticer, K. Brasky, and I. S. Lukashevich**. 2006. A recombinant Yellow Fever 17D vaccine expressing Lassa virus glycoproteins. Virology **345**:299-304.
156. **Jiang, X., T. J. Dalebout, P. J. Bredenbeek, R. Carrion, Jr., K. Brasky, J. Patterson, M. Goicochea, J. Bryant, M. S. Salvato, and I. S. Lukashevich**. 2010. Yellow fever 17D-vectored vaccines

expressing Lassa virus GP1 and GP2 glycoproteins provide protection against fatal disease in guinea pigs. Vaccine.

157. **Van Epps, H. L.** 2005. Broadening the horizons for yellow fever: new uses for an old vaccine. J. Exp. Med. **201**:165-168.

158. **Franco, D., W. Li, F. Qing, C. T. Stoyanov, T. Moran, C. M. Rice, and D. D. Ho**. 2010. Evaluation of yellow fever virus 17D strain as a new vector for HIV-1 vaccine development. Vaccine **28**:5676-5685.

159. **Bonaldo, M. C., S. M. Mello, G. F. Trindade, A. A. Rangel, A. S. Duarte, P. J. Oliveira, M. S. Freire, C. F. Kubelka, and R. Galler**. 2007. Construction and characterization of recombinant flaviviruses bearing insertions between E and NS1 genes. Virol. J. **4**:115.

160. **Liu, Y., E. Wimmer, and A. V. Paul**. 2009. Cis-acting RNA elements in human and animal plus-strand RNA viruses. Biochim. Biophys. Acta **1789**:495-517.

161. **Zoll, J., H. A. Heus, F. J. van Kuppeveld, and W. J. Melchers**. 2009. The structure-function relationship of the enterovirus 3′-UTR. Virus Res. **139**:209-216.

162. **You, S. and R. Padmanabhan**. 1999. A novel in vitro replication system for Dengue virus. Initiation of RNA synthesis at the 3′-end of exogenous viral RNA templates requires 5′- and 3′-terminal complementary sequence motifs of the viral RNA. J. Biol. Chem. **274**:33714-33722.

163. **Ackermann, M. and R. Padmanabhan**. 2001. De novo synthesis of RNA by the dengue virus RNA-dependent RNA polymerase exhibits temperature dependence at the initiation but not elongation phase. J. Biol. Chem. **276**:39926-39937.

164. **You, S., B. Falgout, L. Markoff, and R. Padmanabhan**. 2001. In vitro RNA synthesis from exogenous dengue viral RNA templates requires long range interactions between 5′- and 3′-terminal regions that influence RNA structure. J. Biol. Chem. **276**:15581-15591.

165. **Nomaguchi, M., T. Teramoto, L. Yu, L. Markoff, and R. Padmanabhan**. 2004. Requirements for West Nile virus (-)- and (+)-strand subgenomic RNA synthesis in vitro by the viral RNA-dependent RNA polymerase expressed in Escherichia coli. J. Biol. Chem. **279**:12141-12151.

166. **Dong, H., B. Zhang, and P. Y. Shi**. 2008. Terminal structures of West Nile virus genomic RNA and their interactions with viral NS5 protein. Virology **381**:123-135.

167. **Polacek, C., J. E. Foley, and E. Harris**. 2009. Conformational changes in the solution structure of the dengue virus 5′ end in the presence and absence of the 3′ untranslated region. J. Virol. **83**:1161-1166.

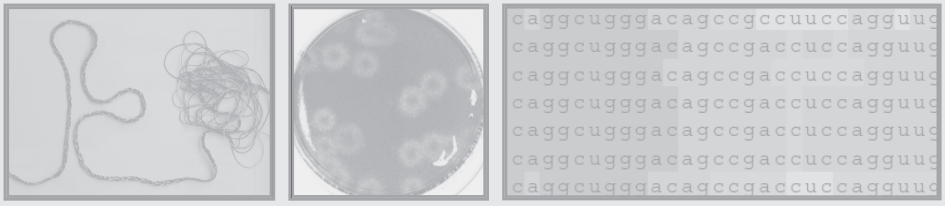
Chapter 2

Conservation of the pentanucleotide motif at the top of the yellow fever virus 17D 3′ stem-loop structure is not required for replication

Patrícia A. G. C. Silva
Richard Molenkamp
Tim. J. Dalebout
Nathalie Charlier
Johan Neyts
Willy J. M. Spaan
Peter J. Bredenbeek

Journal of General Virology (2007), 88, 1738-1747

Reprinted with permission



ABSTRACT

The pentanucleotide sequence (PN) 5′-CACAG-3′ at the top of the 3′ stem-loop structure of the flavivirus genome is well conserved in the arthropod-borne viruses but is more variable in flaviviruses with no known vector. In this study, the sequence requirements of the PN motif for yellow fever virus 17D (YFV) replication were determined. In general, individual mutations at either the 2nd, 3rd or 4th positions were tolerated and resulted in replication-competent virus. Mutations at the 5th position were lethal. Base pairing of the nucleotide at the 1st position of the PN motif and a nucleotide four positions downstream of the PN (9th position) was a major determinant for replication. Despite the fact that the majority of the PN mutants were able to replicate efficiently, they were outcompeted by parental YFV-17D virus following repeated passages in double-infected cell cultures. Surprisingly, some of the virus mutants at the 1st and/or the 9th position that maintained the possibility of forming a base pair were found to have a similar fitness to YFV-17D under these conditions. Overall, these experiments suggest that YFV is less dependent on sequence conservation of the PN motif for replication in animal cells than West Nile virus. However, in animal cell culture, YFV has a preference for the wt CACAG PN sequence. The molecular mechanisms behind this preference remain to be elucidated.

INTRODUCTION

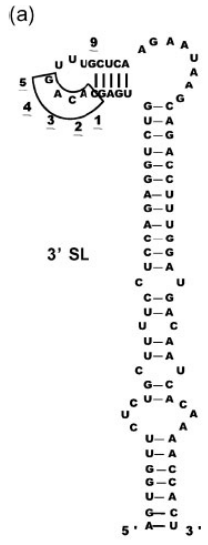
The genus *Flavivirus* consists of nearly 80 RNA viruses that are distributed worldwide. Many of these viruses are transmitted by mosquito or tick species to their vertebrate hosts. However, there are also flaviviruses for which no arthropod vector has been identified [1,2]. Phylogenetic analysis of the genus *Flavivirus* has grouped these viruses into three major clusters: (i) the mosquito-borne viruses; (ii) the tick-borne viruses; and (iii) the no known vector (NKV) viruses [1,3]. It is unknown whether the inability of NKV flaviviruses to infect arthropod vectors is due to a block at the level of entry, replication or assembly [4].

Flaviviruses are small, enveloped viruses containing a positive, single-stranded RNA genome of approximately 11 kb in length with a 5' cap structure and a 3' non-polyadenylated terminus. The genomic RNA encodes a single large open reading frame flanked by 5' and 3' untranslated regions (UTRs) of approximately 100 and 400 – 600 nt, respectively. Translation of the genome results in the synthesis of a large polyprotein, which is co- and post-translationally processed by viral and cellular proteases into three structural proteins (C, prM and E) and seven non-structural proteins that are primarily involved in the replication of the viral RNA (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) [5].

The 3' UTR of the mosquito-borne flaviviruses contains several conserved sequences and is predicted to fold into a complex structure including well-conserved secondary and tertiary RNA elements that are involved in the initiation and regulation of genome amplification and translation (reviewed by [6]). This conservation of RNA structure is especially obvious in the stem-loop (SL) that is predicted to be formed at the 3' end of every flavivirus genome. This structure involves 80 - 90 nucleotides that are not well conserved in primary sequence, except for the pentanucleotide (PN) CACAG [7-9] in the bulge at the top of the SL and the dinucleotide CU at the end of the genome (fig. 1.A). Deletion of the SL is lethal for flavivirus RNA synthesis [10-13]. The SL structure is also required for efficient translation of the flavivirus genome [14,15].

Specific binding of the viral polymerase (NS5) of Japanese encephalitis virus (JEV) to the SL has been demonstrated [16]. In addition, several host proteins such as translation elongation factor-1α [17,18], Mov34 [19] La and PTB [18,20] have been shown to interact with the SL of several flaviviruses.

Sequence comparison within the genus *Flavivirus* reveals that the CACAG sequence is only well conserved when the vector-borne viruses are aligned (fig. 1.B**,** numbering according to fig. 1.A). When the NKV flaviviruses are included in this comparison the PN sequence was shown to be more variable. Rio Bravo virus (RBV) contains a C residue at the 2nd position, whereas Montana *myotis* leukoencephalitis virus (MMLV) and Modoc virus (MODV) have a U [21]. In addition to the sequence variation at the 2nd position, Apoi virus (APOIV) and Yokose virus (YOKV) also contain different nucleotides at the 3rd and 4th positions (C**CU**AG and C**G**C**C**G, respectively) [21,22].



(b)

Cluster	Virus	Sequence	GenBank accession no.
Tick-borne viruses	TBEV	5'-CCCAGA**CACAG**ACA**G**UCUG-3'	AF069066
	LGTV	5'-CCUAGA**CACAG**AUA**G**UCUG-3'	AF253419
	POWV	5'-UCCAGG**CACAG**AUA**G**CCUG-3'	L06436
Mosquito-borne viruses	YFV-17D	5'-UGUGAG**CACAG**UUU**G**CUCA-3'	X03700
	DENV-1	5'-UCCAGG**CACAG**AAC**G**CCAG-3'	M87512
	WNV	5'-ACACGG**CACAG**UGC**G**CCGA-3'	M12294
	KUNV	5'-ACACGG**CACAG**UGC**G**CCGA-3'	L24512
	JEV	5'-ACUAGG**CACAG**AGC**G**CCGA-3'	M18370
	MVEV	5'-ACAAGG**CAC**C**G**AGC**G**CCGA-3'	AF161266
No-known-vector viruses	RBV	5'-GCAUUG**C**C**CAG**AUU**G**CAUG-3'	AF452049
	MODV	5'-AUCAAG**C**U**CAG**AUU**G**CUUA-3'	AJ242984
	MMLV	5'-AAAAGA**C**U**CAG**AUU**G**UCUC-3'	AJ299445
	APOIV	5'-GUAAUC**C**CU**AG**GUU**G**GAUU-3'	AF452050
	YOKV	5'-CCAGGG**C**G**C**C**G**UUU**G**CCCG-3'	AB114858

Fig. 1.

A) Secondary structure model for the YFV 3'terminal SL. The pentanucleotide motif is boxed. The numbering of the nucleotides used throughout this study is indicated. **B)** Sequence comparison of the PN motif and surrounding nucleotides of representatives of the genus *Flavivirus*. The PN motif and the 9th position are shaded. Conserved nucleotide residues are indicated in bold. TBEV, Tick-borne encephalitis virus; LGTV, Langat virus; POWV, Powassan virus; YFV-17D, Yellow fever virus; DENV-1, Dengue virus; WNV, West Nile virus; KUNV, Kunjin virus; JEV, Japanese encephalitis virus; MVEV, Murray Valley encephalitis virus; RBV, Rio Bravo virus; MODV, Modoc virus; MMLV, Montana *myotis* leukoencephalitis virus; APOIV, Apoi virus; YOKV, Yokose virus.

The conservation of the PN motif suggests that it is an important element for the replication of arthropod-borne flaviviruses. Mutagenesis of this CACAG sequence in replicons of West Nile virus (WNV) revealed that only the A at the 4th position could be replaced by another nucleotide without affecting virus replication [13]. These data were partially confirmed in a study using a full-length WNV cDNA instead of a replicon [11]. However, in contrast to the results obtained using the WNV replicon, mutagenesis of the A residue at the 2nd position of the PN motif did not impair RNA synthesis in the background of WNV full-length genomic RNA.

In view of the observed sequence variation in the PN motif of the NKV flaviviruses and the contradicting results concerning the PN sequence requirements in WNV replicon RNA versus genomic RNA, we performed an extensive mutagenesis of the PN motif of yellow fever virus (YFV) using an infectious full-length YFV and replicon RNA to determine the PN sequence requirements for replication in animal cells.

MATERIAL AND METHODS

Cell culture

Vero E6 cells were kindly provided by Professor A. Osterhaus (Rotterdam, The Netherlands). BHK-21J [23] and Vero E6 cells were grown at 37 ºC in 5 % CO_2 in Dulbecco's modified Eagle's medium (DMEM; Cambrex) supplemented with 8 % fetal calf serum (FCS; Bodinco).

Recombinant DNA techniques and plasmid constructions

Standard nucleic acid methodologies were used [24,25]. Chemically competent *Escherichia coli* DH5α cells [26] were used for cloning. Plasmid pACNR-MODV/YFV-pnMODV, a derivative of pACNR-MODV/YFV [27] in which the YFV PN motif CACAG was mutated to C**U**CAG, mimicking the sequence found in MODV, was digested with *Sfi I* and *Xho I*. The 644 bp fragment corresponding to the MODV/YFV cDNA 3' end was cloned into pBluescript-$YFV_{9845\text{-}10861}$ to yield pBlscrpt-3'YFV-pnC**U**CAG. This plasmid was used as template for mutagenesis of the PN sequence using the QuickChange Site-directed Mutagenesis strategy (Stratagene). The inserts were sequenced to verify the mutations and to exclude unintended mutations. The mutant pBlscrpt-3'YFV derivatives were digested with *Xba I* and *Xho I* and the DNA fragment containing the mutated PN motif was cloned into pACNR-FLYF17Dx [10].

Renilla luciferase-expressing YFV replicons containing a mutated PN motif were created by cloning YFV 3' UTR from the full–length YFV cDNA harboring the mutated PN motif with *Sfi I* and *Xho I* into pYF-R.luc2A-RP [28].

RNA secondary structure prediction

RNA secondary structure was predicted using MFOLD version 3.1 [29,30].

In vitro RNA transcription

Plasmid DNA for *in vitro* run-off RNA transcription was purified with a Qiagen Plasmid Midi kit. YF-R.luc2A-RP or pACNR-FLYF17Da [31] and their derivatives containing the mutated PN sequence were linearized with *Afl II* and purified by proteinase K digestion and phenol/chloroform extraction. Approximately 1 – 2 µg DNA was used as a template for

in vitro transcription using the mMESSAGE mMACHINE SP6 kit (Ambion). Trace amounts of [^{3}H]UTP were added to the reaction mixture to determine the yield [10]. Genomic full-length transcripts were used for transfection without any additional purification. *In vitro*-transcribed replicon RNA was purified according to the protocol supplied with the mMessage mMACHINE kit and the yield was quantified using a Nanodrop photospectrometer.

RNA transfections

BHK-21J cells were prepared for electroporation as described by Lindenbach and Rice [23]. Immediately after preparation, 5 µg of *in vitro*-transcribed RNA was mixed with 600 µl cell suspension and electroporated using an Easyject electroporator (Eurogentec) [32].

Labelling and analysis of viral RNAs

Viral RNA synthesis was analyzed by *in vivo* labeling of the transfected cells with [^{3}H] uridine at 18 to 24 hours post-electroporation (p.e.) [10]. At 24 h, total RNA was isolated, denatured with glyoxal and analyzed on a 0.8 % MOPS/agarose gel [25].

Virus stocks, infections and plaque assays

Medium was harvested from transfected cells to obtain virus stocks when complete cytopathogenic effect (CPE) was observed. For infections, the cells were washed once with PBS and infected with virus using the m.o.i. indicated in the relevant figure legends. After 1 h, the inoculum was replaced by DMEM containing 2 % FCS. For analysis of virus growth kinetics, the medium was collected and replaced by the same volume of fresh medium at the indicate times. Virus titers were determined as described previously [10] except that Vero cells were used instead of SW13 cells in the plaque assays.

RT-PCR

Total RNA was isolated using Trizol at 30 h p.i. from a 10 cm^2 dish containing Vero or BHK-21 cells infected with either YFV-17D or the mutant viruses. RNA was dissolved in 30 µl H_2O and 5 µl was used for RT-PCR to amplify the 3' UTR of the YFV genome using the ThermoScript RT-PCR system (Invitrogen). Primer sequences are available on request.

Renilla luciferase activity

Eletroporated BHK cells (800 μl) were seeded per well of a 12-well plate. At 2 and 18 h post-transfection, the cells were lysed in 200 μl passive lysis buffer (Promega). Luciferase activity was determined using the *Renilla* luciferase assay system (Promega) and a LB9507 luminometer (Berthold). Protein concentrations of the lysates were determined using the Bradford method (Bio-Rad Laboratories).

Virus competition experiments

Vero E6 cells were infected simultaneously with the mutant virus and YFV-17D at an m.o.i. of 5 and 0.5, respectively (ratio 10:1). After 72 h, 200 μl medium was used to infect fresh Vero cells. Intracellular RNA was isolated from the infected Vero cells at the 10th passage and used for RT-PCR. The RT-PCR products were cloned using the TA Cloning kit (Invitrogen). Plasmid DNA was isolated from bacteria cultures and sequenced to determine the nucleotide sequence of the PN motif.

RESULTS

The PN motif is required for YFV replication

Two YFV-17D mutants were constructed to determine whether the PN motif was essential for virus replication. In YFV-ΔpnCACAG, the complete PN sequence was deleted. Computer-aided RNA folding indicated that this deletion could significantly change the RNA structure at the top of the 3′ SL. Therefore, an additional mutant was constructed in which the CACAG sequence was changed to **UGUGA**. RNA modeling predicted that the 3′ SL structure of this mutant would adopt a similar structure to the wt YFV-17D (fig. 2.A). Viral RNA synthesis was detected only in BHK-21J cells transfected with YFV-17D RNA (fig. 2.B). No viral RNA was detected in cells electroporated with either YFV-ΔpnCACAG or YFV-pn**UGUGA** RNA. Even after prolonged incubation (96 – 120 hrs), no virus could be detected in the medium of cells transfected with YFV-pnΔCACAG or YFV-pn**UGUGA** RNA by plaque assay (data not shown). These results demonstrated that the PN motif is absolutely required for YFV replication.

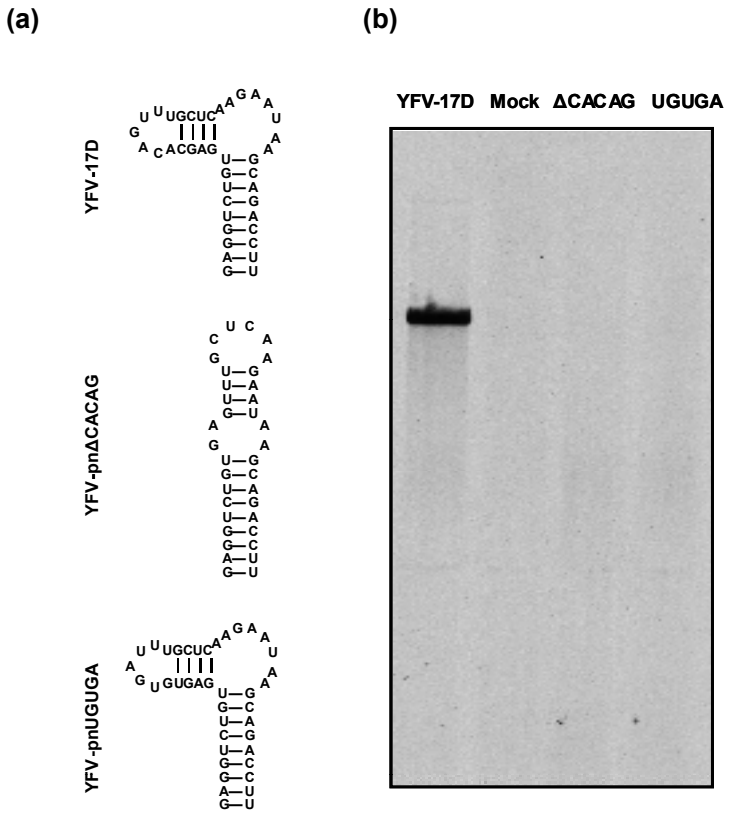


Fig. 2.
A) Secondary RNA structure model for the top of the 3' SL structure of YFV-17D, YFV-ΔpnCACAG and YFV-pn**UGUGA** as predicted by MFOLD. **B)** Viral RNA synthesis in BHK-21J cells transfected with *in vitro*-transcribed RNA of YFV-17D, YFV-ΔpnCACAG and YFV-pn**UGUGA**.

Mutations at the 2nd, 3rd and 4th position of the PN motif are tolerated

As illustrated in fig. 1.B, the PN sequence is not absolutely conserved in flaviviruses. Variation is observed in the 2nd, 3rd and 4th positions. To verify whether other nucleotides were tolerated at these positions in the YFV PN motif, a set of mutants was created in which the A at the 2nd position, the C at the 3rd position or the A at the 4th position was replaced by the alternative nucleotides. Some of these mutations resulted in PN motifs mimicking the sequence of NKV flaviviruses such as MODV and MMLV (C**U**CAG) or RBV (C**C**CAG). In addition to these YFV mutants containing a single mutation, a mutant was created in which the nucleotides at the 2nd and 3rd position were mutated, thereby mimicking the PN motif of the NKV APOIV (C**CU**AG).

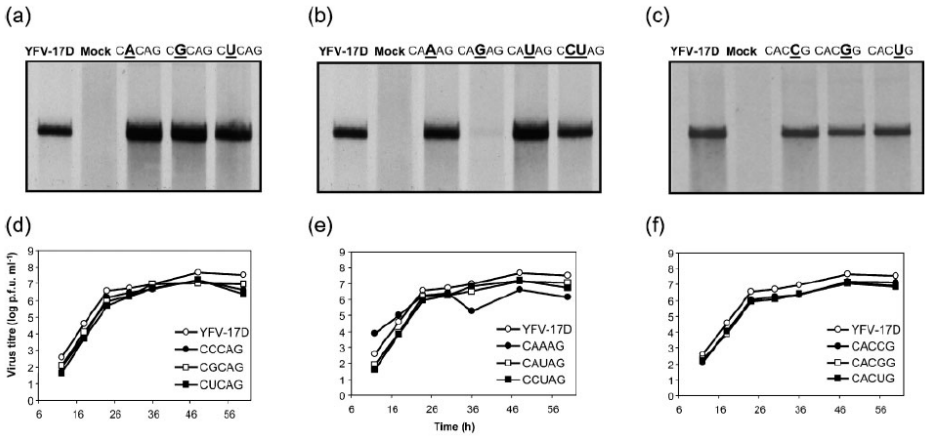


Fig. 3.
Effect of mutations at the 2nd, 3rd and 4thpositions of the PN motif on YFV replication. (**A** – **C**) Analysis of RNA synthesis in BHK-21J cells transfected with RNA of mutants at the 2nd, 2nd and 3rd, and 4th PN positions. The mutants tested are indicated above each lane. (**D** – **F**) Viral growth kinetics of the indicated YFV mutants. BHK cells were infected at an m.o.i. 1; the medium of the infected cells was sampled at the indicated times post-infection. Titers were determined by plaque assays on Vero cells.

As shown in fig. 3.A, the mutants in which the A residue at the 2nd position was replaced by either a C, G or U synthesized RNA at a similar rate to YFV-17D. In addition, the viral growth curves of these viruses showed similar growth kinetics when compared with YFV-17D (fig. 3.D).

A slightly different picture arose when the C at the 3rd position was mutated. Changing it to either an A or a U had no significant effect on viral RNA synthesis. However, a significant decrease in RNA synthesis was observed when this C was replaced by a G (fig. 3.B). The titer in the medium of cells transfected with YFV-pnCA**G**AG RNA was approximately 10^5 p.f.u./ml when CPE was complete. In addition, heterogeneity in plaque size was observed. Most of the plaques were small and turbid and therefore hardly visible, but larger plaques were also observed. RT-PCR on RNA isolated from Vero cells infected with this virus revealed that the introduced G at the 3rd position was replaced by a U. The original PN sequence contains a C residue at this position. However, YFV-pnCA**U**AG was also shown to replicate efficiently (fig. 3.B and E). Given the limited genetic stability of YFV-pnCA**G**AG, this mutant was excluded from the growth curves. The growth kinetics of YFV-pnCA**U**AG did not differ significantly from the parental virus, whereas the growth of YFV-pnCA**A**AG was slightly delayed (fig. 3.E).

In agreement with the above results, the mutant YFV-pnC**CU**AG mimicking the PN motif of APOIV was able to synthesize viral RNA efficiently and showed similar growth kinetics to YFV-17D, despite the fact that it contained two mutations within the PN motif.

Mutagenesis of the A residue at the 4th position had no significant effect on virus replication. The mutants YFV-pnCAC**C**G, YFV-pnCAC**G**G and YFV-pnCAC**U**G all synthesized RNA at comparable levels (fig. 3.C) and showed similar growth kinetics (fig. 3.F) when compared with YFV-17D.

To analyze whether reversion of the introduced mutations to the original YFV-17D PN sequence could have influenced the outcome of these experiments, mutant viruses from the 60 h time point of the growth curves were used to infect Vero cells. At 30 h p.i., total RNA was isolated and used for RT-PCR. All viruses had maintained the original mutation. However, second-site reversions in other regions of the genome could not be excluded.

Mutational analysis of the 1st position of the PN motif reveals the importance of base pairing

The C residue at the 1st position of the PN motif appears to be truly conserved in all flaviviruses. This C residue is predicted to base pair with an equally well conserved G four positions downstream the PN motif (fig. 1.B). This position will be referred to as the 9th position. To determine the importance of this C residue and the potential role of the C-G base pair in YFV replication, the C was replaced by each of the other three nucleotides. Mutagenesis to either an A (YFV-pn**A**ACAG9G) or G (YFV-pn**G**ACAG9G) was predicted to disrupt the base pairing, whereas this base pair was predicted to be maintained when the C was replaced by a U (YFV-pn**U**ACAG9G). As shown in fig. 4, no RNA synthesis was detected in cells transfected with YFV-pn**A**ACAG9G, whereas viral RNA synthesis was significantly impaired in cells transfected with YFV-pn**G**AGAC9G. YFV-pn**U**ACAG9G synthesized RNA with an efficiency that was similar to the parental virus. These data suggested that the formation of the base pair between the 1st and 9th positions is more important than the actual nucleotides. To verify this hypothesis, additional mutants were created by introducing mutations at the 9th position in combination with the 1st position mutants described above. This resulted in YFV-pn**A**ACAG9**A**, YFV-pnCACAG9**C**, YFV-pn**U**ACAG9**U**, YFV-pn**G**ACAG9**C**, YFV-pn**A**ACAG9**U** and YFV-pn**U**ACAG9**A**. In the first three mutants, the G residue at the 9th position was changed to the same nucleotide as in the 1st position of the mutated PN motif, thus impairing base pair formation. In the last three mutants, the potential for base pairing was restored, albeit it with different nucleotides compared with the parental virus. No YFV RNA was detected in cells transfected with either YFV-pn**A**ACAG9**A** or YFV-pnCACAG9**C**, and viral RNA synthesis was significantly impaired in cells transfected with YFV-pn**U**ACAG9**U** (fig. 4.A). In contrast to the above mutants, the mutant viruses YFV-pn**G**ACAG9**C**, YFV-pn**A**ACAG9**U** and YFV-pn**U**ACAG9**A** in which base pairing was restored, showed efficient RNA synthesis and viral growth kinetics similar to that of the parental virus (fig. 4). Taken together,

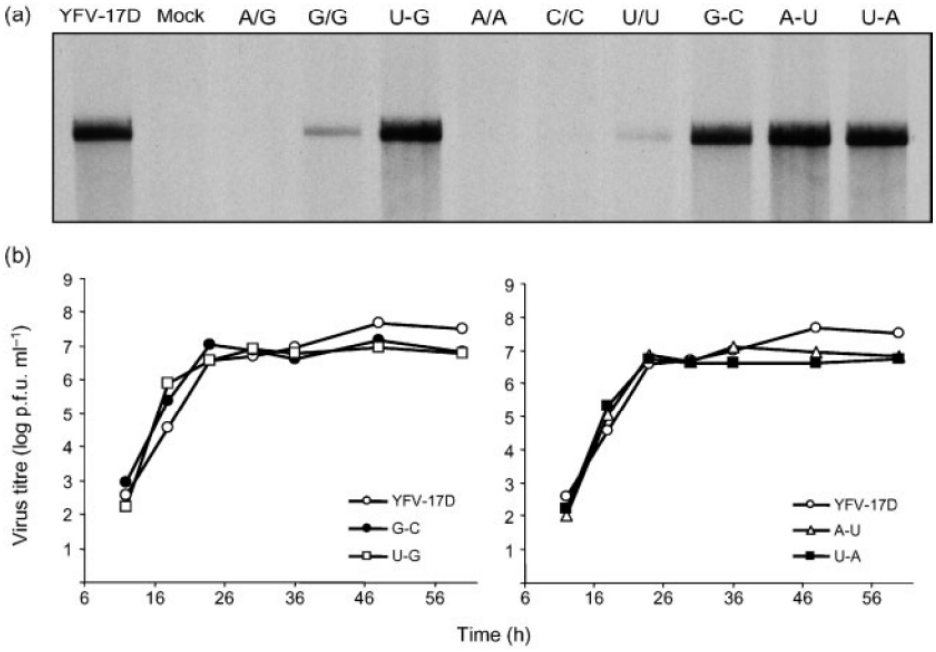


Fig. 4.

Characterization of YFV with mutations at the 1st PN position and the 9th position. **A)** RNA synthesis in BHK-21J cells transfected with RNA of mutants: YFV-pn**A**ACAG9G, YFV-pn**G**ACAG9G, YFV-pn**U**ACAG9G, YFV-pn**A**ACAG9**A**, YFV-pnCACAG9**C**, YFV-pn**U**ACAG9**U**, YFV-pn**G**ACAG9**C**, YFV-pn**A**ACAG9**U** and YFV-pn**U**ACAG9**A**. **B)** Viral growth kinetics of the indicated YFV mutants. BHK cells were infected at an m.o.i. of 1. The medium of the infected cells was harvested at the indicated times post-infection. Plaque assays on Vero cells were used to determine the virus titer.

these data clearly demonstrated that base pair formation between the 1st nucleotide of the PN motif and the nucleotide at the 9th position is more important for efficient virus replication than the nature of the nucleotides at these positions.

Analysis of the 3'UTR of viruses with mutations at either the 1st and/or the 9th position obtained at the 60 h time point of the growth curves showed no evidence for primary site reversion. Despite the fact that no RNA synthesis could be detected in cells transfected with YFV-pn**A**ACAG9G and YFV-pn**A**ACAG9**A**, these cells eventually developed CPE. Sequencing of the 3' UTR of these virus stocks revealed reversion to the wt virus and the mutant YFV-pn**A**ACAG9**U**. Interestingly, the latter was actually shown to replicate efficiently in this study (fig. 4). Second-site reversions in other regions of the genome could not be excluded.

The G residue at the 5th position is essential for virus replication

The well-conserved G residue at the 5th position of the PN motif was replaced by one of the other nucleotides. As shown in fig. 5, viral RNA synthesis was only detected in cells transfected with the parental YFV-17D transcript; no RNA was detected in cells electroporated with YFV-pnCACA**A**, YFV-pnCACA**C** or YFV-pnCACA**U** RNA, and no virus could be detected by plaque assay in the medium of the transfected cells (data not shown). These data demonstrated that the G at the 5th position of the PN motif is absolutely required for YFV replication.

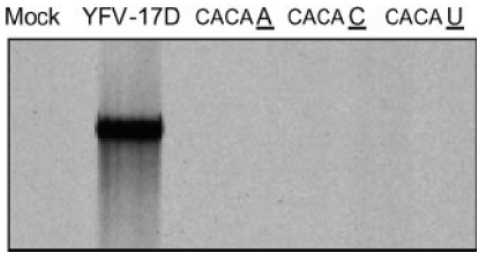


Fig. 5.
RNA synthesis in BHK-21J cells transfected with *in vitro*-transcribed YFV RNA containing mutations at the 5th position of the PN motif.

Competition between mutant virus and YFV-17D

The results presented so far in this study have demonstrated that, except for the 5th position, point mutations are generally well tolerated within the PN motif. Many of these PN mutant viruses replicated with an efficiency that was comparable to YFV-17D, indicating that they were just as fit as the parental virus in animal cells. This suggested that conservation of the wt CACAG sequence is not that important in an animal cell culture system. To test this hypothesis, Vero cells were simultaneously infected with efficiently replicating representatives of the PN mutant viruses and YFV-17D at a ratio of 10:1. Intracellular RNA was isolated after ten passages and used to determine the ratio of mutant to parental virus by sequencing the PN motif.

Viruses with a mutation at the 2nd position of the PN motif, such as YFV-pnC**C**CAG and YFV-pnC**U**CAG, were clearly outcompeted by YFV-17D within ten passages (Table 1). A similar result was also obtained for the 3rd position mutant YFV-pnCA**U**AG. These results were supported by the fact that the virus mimicking the APOIV PN motif (YFV-pnC**CU**AG) was also not detected after ten passages. Compared with viruses with mutations at the 2nd and 3rd positions, YFV-pnCAC**C**G replicated relatively well and was still the dominant virus after ten passages. However, the ratio of 12:7 for YFV-pnCAC**C**G versus YFV-17D at the 10th passage indicated that the parental virus was slowly outcompeting the mutant virus. These data demonstrated that viruses with a mutation at the 2nd, 3rd or 4th position

Table 1. Competition experiment in Vero E6 cells simultaneously infected with efficiently replicating representatives of the PN mutant viruses and YFV-17D at a ratio of 10:1. Intracellular RNA was isolated after ten passages and used to determine the ratio of PN mutant virus to parental virus by sequencing.

	Mutant	Number of sequences			Final ratio mutant/wt	Dominant virus
		Total	Mutant	YFV-17D		
1	pnC**C**CAG	20	3	17	3:17	YFV
2	pnC**U**CAG	18	1	17	1:17	YFV
3	pnCA**U**AG	22	0	22	0:22	YFV
4	pnC**CU**AG	19	0	19	0:19	YFV
5	pnCAC**C**G	19	12	7	12:7	CAC**C**G
6	pn**G**ACAG9**C**	26	20	6	10:3	**G**ACAG9**C**
7	pn**U**ACAG9**G**	20	0	20	0:20	YFV
8	pn**U**ACAG9**U**	19	0	19	0:19	YFV
9	pn**A**ACAG9**U**	29	27	2	27:2	**A**ACAG9**U**
10	pn**U**ACAG9**A**	26	18	8	9:4	**U**ACAG9**A**

of the PN motif were less fit than the parental YFV-17D in Vero cells, despite the fact that these mutant viruses showed similar replication efficiency and growth kinetics in individual infection experiments.

Mutant viruses that contained an alternative base pair at the 1st position of the PN motif and the 9th position were also analysed. After ten passages, the mutant YFV-pn**U**ACAG9G was completely outcompeted by the parental virus (Table 1). The result obtained for YFV-pn**U**ACAG9**A** was essentially similar to that of the mutant virus YFV-pnCAC**C**G. At the 10th passage, the YFV-pn**U**ACAG9**A** was still the dominant virus, but the ratio indicated that the mutant would eventually be outcompeted. A more interesting picture was observed for both YFV-pn**G**ACAG9**C** and YFV-pn**A**ACAG9**U**. At the 10th passage, they were clearly the dominant viruses, suggesting that equilibrium was possible between the mutant and the wt virus. As expected, a poorly replicating mutant such as YFV-pn**U**ACAG9**U**, in which the base pairing was disrupted, was easily outcompeted by YFV-17D.

Although the results obtained with YFV-pn**G**ACAG9**C** and YFV-pn**A**ACAG9**U** indicated that some of the mutants were as fit as the wt virus, it was obvious that the wt PN sequence still had an as yet undefined advantage over most of the mutant PN sequences when analyzed in animal cells.

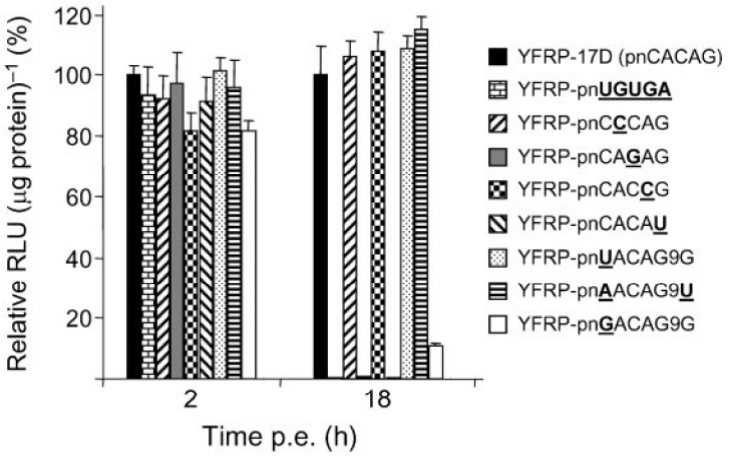


Fig. 6.

Luciferase expression of YFV replicons containing mutations in the PN motif at 2 and 18 h post-electroporation. The following replicons were analyzed: YFRP-17D (pnCACAG), YFRP-pn**UGUGA**, YFRP-pnC**C**CAG, YFRP-pnCA**G**AG, YFRP-pnCAC**C**G, YFRP-pnCACA**U**, YFRP-pn**U**ACAG9G, YFRP-pn**A**ACAG9**U** and YFRP-pn**G**ACAG9G. Results are given as relative luciferase units (RLU) relative to the activity measured in mock-transfected cells. No significant luciferase activity was detected in cells transfected with YFRP-pn**UGUGA**, YFRP-pnCA**G**AG or YFRP-pnCACA**U** at 18 h p.e.

Mutations in the PN motif do not affect translation

To analyze whether the effect of the PN mutations was due to a direct effect on RNA synthesis or an indirect effect by influencing virus RNA translation, a selected set of mutations was cloned into pYF-R.luc2A-RP. RNA transcribed from these plasmids was transfected into BHK cells, which were subsequently analysed for *Renilla* luciferase expression at 2 h (the peak time for translation of input RNA) and 18 h p.e. (when only virus synthesized RNA is translated). At 2 h p.e., all of the replicons expressed luciferase at a comparable level (fig. 6) including replicons such as YFRP-pn**UGUGA**, YFRP-pnCA**G**AG, YFRP-pnCACA**U** and YFRP-pn**G**ACAG9G for which no or hardly any RNA synthesis could be detected in the background of the full-length YFV RNA (fig. 2, 3, 4 and 5). From these data, it was concluded that the mutations in the PN motif had at best a relatively minor effect on translation. No luciferase was detected in cells transfected with YFRP-pn**UGUGA**, YFRP-pnCA**G**AG, or YFRP-pnCACA**U** at 18 h p.e., whereas YFRP-pn**G**ACAG9G showed low luciferase activity. The replicons YFRP-pnC**C**CAG, YFRP-pnCAC**C**G, YFRP-pn**U**ACAG9G and YFRP-pn**A**ACAG9**U** showed a high level of luciferase activity. These mutations also allowed efficient replication in the background of the full-length clone, demonstrating that mutagenesis of the PN motif in either the replicon or the full-length YFV RNA yielded identical results.

DISCUSSION

Sequence comparison of the PN motif and the surrounding nucleotides has shown that the PN sequence 5'-CACAG-3' is well conserved within the vector-borne flaviviruses (fig. 1.B). However, when the NKV flaviviruses are included in this comparison, the PN sequence is far less conserved. None of the NKV viruses sequenced to date contains the sequence CACAG as the PN motif. Other nucleotides are observed at either the 2nd or 3rd position and some viruses even contain substitutions at the 2nd and 3rd or the 2nd and 4th positions [21,22]. These nucleotide substitutions in the NKV viruses are rather surprising and suggest that mutations of some PN positions may be tolerated in arthropod-borne flaviviruses. These observations prompted us to analyze the requirement for the YFV PN sequence.

YFV mutants in which the PN sequence CACAG is either deleted or completely changed (YFV-pn**UGUGA**) were unable to replicate, demonstrating that at least part of this sequence is absolutely required for YFV replication. These data are in agreement with the observation for WNV in which deletion of the PN motif is also lethal [13].

To determine whether other nucleotides are tolerated in the YFV PN motif, we performed an extensive mutagenesis study using *in vitro*-transcribed YFV genomic and replicon RNA. Our data showed that mutations at either the 2nd, 3rd or 4th positions of the PN motif had no significant effect on YFV replication, except when the C at the 3rd position was replaced by a G, which severely impaired viral RNA synthesis and growth. Except for the 4th position, our data are clearly different from the published results on the WNV PN motif [13]. Using WNV replicon RNA, it was shown that the A and C residues at the 2nd and 3rd positions were absolutely required for WNV RNA synthesis. Another study using full-length WNV genome RNA transcripts instead of a replicon RNA confirmed the data of Tilgner *et al.* [13] concerning the 3rd position of the PN motif. However, in contrast to the WNV replicon, replacement of the A at the 2nd position by a U residue was tolerated in the complete WNV genome [11].

In addition to these point mutations, a YFV mutant mimicking the NKV APOIV PN motif (C**CU**AG) was constructed. This mutant replicated efficiently and showed only slightly slower growth kinetics compared to the YFV-17D. As a control for these results on the 2nd, 3rd and 4th positions, the mutant YFV-pnC**GUG**G was created. As expected from our previous observations this mutant was able to replicate, although less efficiently than YFV-17D (data not shown). These results confirm our initial finding that point mutations at the 2nd, 3rd or 4th position of the YFV PN motif are tolerated, although the almost undetectable effects of individual mutations become more obvious when mutations were combined.

Only the 1st and 5th positions of the PN motif appear to be truly conserved among all flaviviruses. Replacing the G at the 5th position of the YFV PN motif for another nucleo-

tide was lethal. A similar result was obtained for the WNV PN motif [11,13], suggesting that this G residue has a critical role in the replication of all flaviviruses.

Analysis of the sequence surrounding the PN motif revealed that the G residue at the 9th position is also strictly conserved. RNA folding of the flavivirus 3' SL structure predicted that this G will base pair with the well-conserved C residue at the 1st position of the PN motif, suggesting that the formation of this base pair is essential. Mutants at the 1st and/or the 9th position that disrupted this predicted base pair were either unable to replicate or were significantly impaired in virus replication, whereas mutants that allowed the formation of this base pair showed RNA synthesis and viral growth with similar kinetics to the parental virus. From these data, we concluded that the formation of a base pair between the 1st nucleotide of the PN motif and the nucleotide at the 9th position is a critical determinant for efficient virus replication. The importance of this base pair was also recognized for WNV [13]. However, WNV replicons with an alternative base pair showed only 10 – 20% of the luciferase activity of the wt replicon, whereas the comparable YFV mutants in this study were virtually indistinguishable from the wt virus or replicon. Using these luciferase-expressing YF replicons, we also demonstrated that the PN mutations only had a direct effect on viral RNA synthesis and did not affect virus RNA translation.

Flavivirus RNA replicons have been used extensively to study virus replication [31,33-37] and so far no significant differences have been observed when analyzing the effect of 3' UTR mutations on virus replication using flavivirus RNA replicons versus full-length genomic RNA. This is also true for the data presented in this study and we have no explanation as to why the sequence requirements of the YFV PN motif are so different from those determined using either a WNV replicon or a full-length RNA.

It has been stated that the C at the 8th position in WNV replication [11] is critical for replication. Substitution of the U at this position in YFV for a C yielded a virus with similar characteristics to wt YFV-17D, indicating that in YFV the nucleotide at this position is not critical for replication (data not shown).

The variability that can be introduced in the YFV PN sequence is somewhat surprising when we take into account the fact that the PN CACAG is well conserved within the arthropod-borne flaviviruses. To evaluate the importance of the wt sequence for replication in animal cells, competition experiments between YFV-17D and a set of mutants were performed. Although the results showed that some of the base pair mutants appeared to be as fit as the parental virus, the wt PN sequence had an advantage over most of the mutant PN sequences in animal cells. These results are to some extend similar to what has been observed for tick-borne encephalitis virus mutants where point mutations that seem to have little or no effect in animal cell culture were shown to have a clearly different phenotype in a relevant small animal model [38].

Taken together, our data support the fact that the PN CACAG is quite variable in sequence when analyzed in animal cell culture systems. Individual point mutations at the 2nd, 3rd and 4th positions are generally well tolerated in the YFV PN motif, whereas the G residue at the 5th position is truly conserved. In addition to this G, base pairing between the nucleotides at the 1st and 9th positions is also essential for efficient replication. Despite this sequence variability that can be introduced, there appears to be a preference for the parental CACAG sequence in animal cell culture. The reason for this is currently unclear. The PN motif may be part of either a host or viral protein RNA binding site. The G at the 5th position would then be crucial for protein binding, whilst the formation of the base pair might be required to form the proper RNA structure. Given the mutations that can be introduced into the PN sequence, it is unlikely that the PN motif is involved in an RNA-RNA interaction.

ACKNOWLEDGEMENTS

We thank Professor R. J. Kuhn for providing the YF-R.luc2A-rep DNA.

REFERENCE LIST

1. **Kuno, G., G. J. Chang, K. R. Tsuchiya, N. Karabatsos, and C. B. Cropp**. 1998. Phylogeny of the genus Flavivirus. J. Virol. **72**:73-83.
2. **Thiel, H. J., M. S. Collet, E. A. Gould, F. X. Heinz, G. Meyers, R. H. Purcell, C. M. Rice, and M. Houghton**. 2005. Flaviridae, p. 981-998. *In* L. A. Ball (ed.), Virus Taxonomy - Eight Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego.
3. **Cook, S. and E. C. Holmes**. 2006. A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: Flaviviridae) and the evolution of vector transmission. Arch. Virol. **151**:309-325.
4. **Lawrie, C. H., N. Y. Uzcategui, M. Armesto, L. Bell-Sakyi, and E. A. Gould**. 2004. Susceptibility of mosquito and tick cell lines to infection with various flaviviruses. Med. Vet. Entomol. **18**:268-274.
5. **Lindenbach, B. and C. M. Rice**. 2001. *Flaviviridae*: the viruses and their replication, p. 991-1041. *In* D. M. Knipe and P. M. Howley (eds.), Fields Virology. Lippincott, Williams, Wilkins, Philadelphia.
6. **Markoff, L.** 2003. 5'- and 3'-noncoding regions in flavivirus RNA. Adv. Virus Res. **59**:177-228.
7. **Brinton, M. A., A. V. Fernandez, and J. H. Dispoto**. 1986. The 3'-nucleotides of flavivirus genomic RNA form a conserved secondary structure. Virology **153**:113-121.
8. **Hahn, C. S., Y. S. Hahn, C. M. Rice, E. Lee, L. Dalgarno, E. G. Strauss, and J. H. Strauss**. 1987. Conserved elements in the 3' untranslated region of flavivirus RNAs and potential cyclization sequences. J. Mol. Biol. **198**:33-41.
9. **Wengler, G. and E. Castle**. 1986. Analysis of structural properties which possibly are characteristic for the 3'-terminal sequence of the genome RNA of flaviviruses. J. Gen. Virol. **67 (Pt 6)**:1183-1188.
10. **Bredenbeek, P. J., E. A. Kooi, B. Lindenbach, N. Huijkman, C. M. Rice, and W. J. Spaan**. 2003. A stable full-length yellow fever virus cDNA clone and the role of conserved RNA elements in flavivirus replication. J. Gen. Virol. **84**:1261-1268.
11. **Elghonemy, S., W. G. Davis, and M. A. Brinton**. 2005. The majority of the nucleotides in the top loop of the genomic 3' terminal stem loop structure are cis-acting in a West Nile virus infectious clone. Virology **331**:238-246.
12. **Nomaguchi, M., T. Teramoto, L. Yu, L. Markoff, and R. Padmanabhan**. 2004. Requirements for West Nile virus (-)- and (+)-strand subgenomic RNA synthesis in vitro by the viral RNA-dependent RNA polymerase expressed in Escherichia coli. J. Biol. Chem. **279**:12141-12151.
13. **Tilgner, M., T. S. Deas, and P. Y. Shi**. 2005. The flavivirus-conserved penta-nucleotide in the 3' stem-loop of the West Nile virus genome requires a specific sequence and structure for RNA synthesis, but not for viral translation. Virology **331**:375-386.
14. **Chiu, W. W., R. M. Kinney, and T. W. Dreher**. 2005. Control of translation by the 5'- and 3'-terminal regions of the dengue virus genome. J. Virol. **79**:8303-8315.
15. **Holden, K. L. and E. Harris**. 2004. Enhancement of dengue virus translation: role of the 3' untranslated region and the terminal 3' stem-loop domain. Virology **329**:119-133.
16. **Chen, C. J., M. D. Kuo, L. J. Chien, S. L. Hsu, Y. M. Wang, and J. H. Lin**. 1997. RNA-protein interactions: involvement of NS3, NS5, and 3' noncoding regions of Japanese encephalitis virus genomic RNA. J. Virol. **71**:3466-3473.
17. **Blackwell, J. L. and M. A. Brinton**. 1997. Translation elongation factor-1 alpha interacts with the 3' stem-loop region of West Nile virus genomic RNA. J. Virol. **71**:6433-6444.

18. **De Nova-Ocampo, M., N. Villegas-Sepulveda, and R. M. del Angel**. 2002. Translation elongation factor-1alpha, La, and PTB interact with the 3′ untranslated region of dengue 4 virus RNA. Virology **295**:337-347.
19. **Ta, M. and S. Vrati**. 2000. Mov34 protein from mouse brain interacts with the 3′ noncoding region of Japanese encephalitis virus. J. Virol. **74**:5108-5115.
20. **Kim, S. M. and Y. S. Jeong**. 2006. Polypyrimidine tract-binding protein interacts with the 3′ stem-loop region of Japanese encephalitis virus negative-strand RNA. Virus Res. **115**:131-140.
21. **Charlier, N., P. Leyssen, C. W. Pleij, P. Lemey, F. Billoir, L. K. Van, A. M. Vandamme, C. E. De, L. de, X, and J. Neyts**. 2002. Complete genome sequence of Montana Myotis leukoencephalitis virus, phylogenetic analysis and comparative study of the 3′ untranslated region of flaviviruses with no known vector. J. Gen. Virol. **83**:1875-1885.
22. **Tajima, S., T. Takasaki, S. Matsuno, M. Nakayama, and I. Kurane**. 2005. Genetic characterization of Yokose virus, a flavivirus isolated from the bat in Japan. Virology **332**:38-44.
23. **Lindenbach, B. D. and C. M. Rice**. 1997. trans-Complementation of yellow fever virus NS1 reveals a role in early RNA replication. J. Virol. **71**:9608-9617.
24. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl**. 2000. Current Protocols in Molecular Biology. Wiley Interscience, New York.
25. **Sambrook, J., T. Fritsch, and T. Maniatis**. 1989. Molecular Cloning: a Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
26. **Inoue, H., H. Nojima, and H. Okayama**. 1990. High efficiency transformation of Escherichia coli with plasmids. Gene **96**:23-28.
27. **Charlier, N., R. Molenkamp, P. Leyssen, J. Paeshuyse, C. Drosten, M. Panning, C. E. De, P. J. Bredenbeek, and J. Neyts**. 2004. Exchanging the yellow fever virus envelope proteins with Modoc virus prM and E proteins results in a chimeric virus that is neuroinvasive in SCID mice. J. Virol. **78**:7418-7426.
28. **Jones, C. T., C. G. Patkar, and R. J. Kuhn**. 2005. Construction and applications of yellow fever virus replicons. Virology **331**:247-259.
29. **Mathews, D. H., J. Sabina, M. Zuker, and D. H. Turner**. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. **288**:911-940.
30. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31**:3406-3415.
31. **Molenkamp, R., E. A. Kooi, M. A. Lucassen, S. Greve, J. C. Thijssen, W. J. Spaan, and P. J. Bredenbeek**. 2003. Yellow fever virus replicons as an expression system for hepatitis C virus structural proteins. J. Virol. **77**:1644-1648.
32. **van Dinten, L. C., J. A. den Boon, A. L. Wassenaar, W. J. Spaan, and E. J. Snijder**. 1997. An infectious arterivirus cDNA clone: identification of a replicase point mutation that abolishes discontinuous mRNA transcription. Proc. Natl. Acad. Sci. U. S. A **94**:991-996.
33. **Alvarez, D. E., A. L. De Lella Ezcurra, S. Fucito, and A. V. Gamarnik**. 2005. Role of RNA structures present at the 3′UTR of dengue virus on translation, RNA synthesis, and viral replication. Virology **339**:200-212.
34. **Holden, K. L., D. A. Stein, T. C. Pierson, A. A. Ahmed, K. Clyde, P. L. Iversen, and E. Harris**. 2006. Inhibition of dengue virus translation and RNA synthesis by a morpholino oligomer targeted to the top of the terminal 3′ stem-loop structure. Virology **344**:439-452.
35. **Khromykh, A. A. and E. G. Westaway**. 1997. Subgenomic replicons of the flavivirus Kunjin: construction and applications. J. Virol. **71**:1497-1505.

36. **Lo, M. K., M. Tilgner, K. A. Bernard, and P. Y. Shi**. 2003. Functional analysis of mosquito-borne flavivirus conserved sequence elements within 3' untranslated region of West Nile virus by use of a reporting replicon that differentiates between viral translation and RNA replication. J. Virol. **77**:10004-10014.
37. **Shi, P. Y., M. Tilgner, and M. K. Lo**. 2002. Construction and characterization of subgenomic replicons of New York strain of West Nile virus. Virology **296**:219-233.
38. **Gritsun, T. S., A. Desai, and E. A. Gould**. 2001. The degree of attenuation of tick-borne encephalitis virus depends on the cumulative effects of point mutations. J. Gen. Virol. **82**:1667-1675.

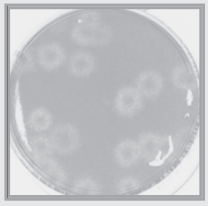
Chapter 3

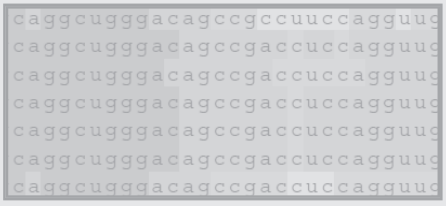
An RNA pseudoknot is required for production of yellow fever virus subgenomic RNA by the host nuclease XRN1

Patrícia A. G. C. Silva
Carina F. Pereira
Tim. J. Dalebout
Willy J. M. Spaan
Peter J. Bredenbeek

Journal of Virology (2010), 84, 11395-11406

Reprinted with permission





ABSTRACT

Cells and mice infected with arthropod-borne flaviviruses produce a small subgenomic RNA that is collinear with the distal part of the viral 3′ untranslated region (UTR). This small subgenomic flavivirus RNA (sfRNA) results from the incomplete degradation of the viral genome by the host 5′-3′ exonuclease XRN1. Production of the sfRNA is important for the pathogenicity of the virus. This study not only presents a detailed description of the yellow fever virus (YFV) sfRNA but, more importantly, describes for the first time the molecular characteristics of the stalling site for XRN1 in the flavivirus genome. Similar to the case for West Nile virus, the YFV sfRNA was produced by XRN1. However, in contrast to the case for other arthropod-borne flaviviruses, not one but two sfRNAs were detected in YFV-infected mammalian cells. The smaller of these two sfRNAs was not observed in infected mosquito cells. The larger sfRNA could also be produced *in vitro* by incubation with purified XRN1. These two YFV sfRNAs formed a 5′ nested set. The 5′ ends of the YFV sfRNAs were found to be just upstream of the previously predicted RNA pseudoknot PSK3. RNA structure probing and mutagenesis studies provided strong evidence that this pseudoknot structure was formed and served as the molecular signal to stall XRN1. The sequence involved in PSK3 formation was cloned into the Sinrep5 expression vector and shown to direct the production of a sfRNA-like RNA. These results underscore the importance of the RNA pseudoknot in stalling XRN1 and also demonstrate that it is the sole viral requirement for sfRNA production.

INTRODUCTION

The *Flavivirus* genus contains nearly 80 viruses distributed worldwide and includes important human pathogens such as dengue virus (DENV), yellow fever virus (YFV), Japanese encephalitis virus (JEV), West Nile virus (WNV), and tick-borne encephalitis virus (TBEV). Phylogenetic analysis clustered flaviviruses into the following three major groups, based on the vector of transmission: (i) mosquito-borne viruses, (ii) tick-borne viruses, and (iii) viruses with no known vector (NKV) [1,2].

Flaviviruses are small enveloped viruses containing a positive-sense single-stranded RNA genome of approximately 11 kb in length, with a 5′-cap structure and a 3′ non-polyadenylated terminus. The genomic RNA is flanked by 5′ and 3′ untranslated regions (UTRs) and encodes a single polyprotein that is co- and posttranslationally processed by viral and cellular proteases into three structural proteins (C, prM and E) and seven nonstructural proteins (NSs) (reviewed in reference [3]). Apart from the viral genome RNA and the replication-related replicative-form and intermediate RNAs [4,5], an additional small flavivirus RNA (sfRNA) has been detected in mice and both mammalian and insect cells infected with flaviviruses belonging to the JEV serogroup [6-8]. Recently, it was shown that production of sfRNA is not unique to JEV and closely related viruses but that all arthropod-borne flaviviruses generate an sfRNA upon infection of mammalian cells [9,10]. The lengths of these sfRNAs vary from 0.3 kb to 0.5 kb and are related to the length of the viral 3′ UTR. Surprisingly, these sfRNAs are not direct products of the viral transcription mechanism but result from incomplete degradation of the viral genomic RNA by the host 5′-3′ exonuclease XRN1, as shown for Kunjin virus (KUNV) by *in vitro* assays and RNA interference (RNAi) experiments. Although the exact role of the sfRNA in the viral life cycle is still elusive, production of sfRNA was shown to be essential for KUNV cytopathogenicity in cell culture and for viral pathogenicity in infected mice [10].

XRN1 is well conserved among eukaryotes and is the main cytoplasmic RNase associated with 5′-3′ mRNA decay that takes place in cytoplasmic processing bodies (P bodies), where the mRNA is decapped by the enzymes DCP1 and -2 and subsequently degraded 5′ to 3′ by XRN1 (reviewed in references [11-14]). XRN1 acts in a processive manner by hydrolyzing RNA with 5′-monophosphate end groups to 5′-mononucleotides [15,16]. Based on fluorescence *in situ* hybridization (FISH) analysis in KUNV-infected cells, the sfRNA was reported to colocalize with XRN1 in P bodies [10]. Interestingly, the role of XRN1 in a viral life cycle is not limited to flaviviruses. XRN1 has also been shown to have an antiviral activity by virtue of its exonuclease activity [17] and to act as a potent suppressor of viral RNA recombination in viruses such as tomato bushy stunt virus [18]. Studies have shown that XRN1 can be blocked to some degree by elements such as a poly(G) tract sequence or large, stable RNA stem-loop structures [19-21]. The 3′ UTR of the mosquito-borne flaviviruses is predicted to fold into a highly complex structure involving well-conserved

RNA sequences as well as strong secondary structures, such as the long 3′ stem-loop (3′ SL) and one or two dumbbell-like RNA structures [22,23]; for a review, see reference [24]. In addition, several RNA pseudoknots are predicted within the mosquito-borne flavivirus 3′ UTR [23,25]. Sequence alignments and computer-aided folding of several flavivirus 3′ UTRs indicated that the production of sfRNA most likely results from the stalling of XRN1 at a conserved RNA stem-loop structure designated SL-II. Deletions in the KUNV 3′ UTR that include SL-II abolished the production of KUNV sfRNA [10].

To determine the involvement of the flavivirus 3′ UTR SL-II structure in sfRNA generation, we performed a detailed analysis of sfRNA production in YFV-infected cells and mapped the XRN1 stalling site by using site-directed mutagenesis of a YFV-17D infectious clone and RNA structure probing. From our data, we concluded that YFV infection of mammalian cells results in the XRN1-mediated production of two sfRNAs that form a 5′ nested set. More importantly, we demonstrated that an RNA pseudoknot involving the YFV equivalent of SL-II is required for stalling XRN1 and therefore crucial for the generation of the flavivirus sfRNAs.

MATERIAL AND METHODS

Cell culture

The origin and culture conditions of the BHK-21J, Vero E6, and SW13 cells used in this study have been described before [26,27]. C6/36 cells [28] were obtained from the ATCC and grown in Eagle's minimal essential medium (EMEM) supplemented with 8% fetal calf serum (Bodinco, Netherlands) and 5% nonessential amino acids.

Recombinant DNA techniques and plasmid construction

Unless explained in more detail, standard nucleic acid methodologies were used [29,30]. Chemically competent *Escherichia coli* DH5α cells [31] were used for cloning. YFV-17D nucleotide numbering was according to Rice *et al.* [32] (GenBank accession no. X03700).

Plasmid pBlsrcptSK-YFV$_{9845\text{-}10861}$ (R. Molenkamp et al., unpublished data), which contains the complete 3′ UTR of YFV-17D, was used as a template for site-directed mutagenesis using the QuikChange strategy (Stratagene) to introduce mutations into the sequences of stem-loop E (SL-E) or in the nucleotides predicted to be involved in the formation of RNA pseudoknot 3 (PSK3) [23]. After sequencing to verify the presence of the introduced mutations and to exclude unintended nucleotide changes, the mutant pBlsrcptSK-

$YFV_{9845\text{-}10861}$ derivatives were digested with *Sfi I* and *Xba I*, and the 492-bp DNA fragments containing the mutated SL-E or PSK3 sequences were cloned into pACNR-FLYF17Da [33].

A 194-bp *Hind III – Xba I* fragment encompassing the region between nucleotides (nt) 10,520 and 10,714 in the 3'UTR of YFV was isolated from pHYF5'3'IVΔRS [26] and cloned into pBluescript SK(-). The resulting plasmid, $pBlscrptSK\text{-}YFV_{10520\text{-}10714}$, was used as a template to generate a minus-strand YFV RNA fragment for use as a probe in RNase protection assays.

Plasmid $pSinrep5\text{-}YFV_{10531\text{-}10611}$ was constructed by inserting a linker encompassing nt 10,531 to 10,611 of YFV flanked by a *Mlu I* and *Sph I* adapter into *Mlu I* and *Sph I* digested pSinrep5eGFP. The $pSINrep5\text{-}YFV_{10521\text{-}10662}$ recombinant was constructed by PCR. Plasmid $pBlsrcptSK\text{-}YFV_{9845\text{-}10861}$ was used as a template with oligonucleotides that contained either an *Mlu I* site (forward primer) or an *Sph I* site (reverse primer). The resulting PCR product was cloned into *Mlu I* and *Sph I* digested pSinrep5-eGFP [34].

Plasmid DNA of the pACNR-FLYF17Da mutants and recombinant pSinrep5 were linearized with *Afl II* and *Xho I*, respectively, and used for *in vitro* RNA transcription [27].

XRN1 RNA silencing

Stocks of lentivirus particles, each expressing a short hairpin RNA (shRNA) (TRCN-049675, TRCN-049676, or TRCN-049677) against the human XRN1 gene (GenBank accession no. NM_ 019001.3), and a lentivirus expressing the scrambled shRNA SHC-002 were prepared from the MISSION®TRC-Hs1.0 library (Sigma) according to the manufacturer's recommendations. The particle titers of the lentivirus stocks were determined using a p24 enzyme-linked immunosorbent assay (ELISA) (Zeptometrix).

To analyze the effect of XRN1 silencing on YFV sfRNA production, 5×10^5 SW13 cells were transduced with a combination of two different lentiviruses expressing one of the shRNAs specified above at a multiplicity of infection (MOI) of 5 (each) or with the SHC-002-expressing virus at an MOI of 10. At 72 h posttransduction, the cells were infected with YFV-17D at an MOI of 5, and 30 h later, the cells were lysed [35] and analyzed for the expression of the host XRN1 protein and actin by Western blotting [36] after PAGE on 5 and 10% gels, respectively. Antibodies directed against human XRN1 (Bethyl Laboratories) and actin (Santa Cruz) were used at dilutions of 1:5000 and 1:2000, respectively. Viral sfRNA production was analyzed by Northern blotting as described below.

In vitro XRN1 assay

Plasmid pYF-R.luc2A-RP [37] was linearized with *Xho I* and used as template for *in vitro* transcription of YFV replicon RNA in the absence of cap analogue. To remove the 5′ triphosphate, 7 µg of transcript was incubated with tobacco acid pyrophosphatase (TAP; Epicentre) as specified by the manufacturer. The TAP-treated RNA was purified by phenol-chloroform extraction and ethanol precipitation. The pellet was dissolved in 16 µl H_2O and used for digestion with XRN1 (Terminator 5′-phosphate-dependent exonuclease; Epicentre) under the conditions described by the manufacturer. The units of XRN1 are indicated in the relevant figure.

RNA transfection and analysis of viral RNA synthesis

BHK-21J cells were transfected with 5 µg of full-length YFV-17D or Sinrep5 transcripts as described previously [27]. In general, 2.5 ml (approximately 1.5 x 10^6 cells) of the transfected BHK-21J cell suspension was seeded in a 35-mm plate. Total RNA was isolated from the transfected cells at 8 h postelectroporation (p.e.) for the recombinant Sinrep5-transfected cells and at 24 h p.e. for the YFV-transfected cells. Analysis of RNA synthesis by [^{3}H]uridine labeling was performed as described previously [26]. Trizol (Invitrogen) was used for cell lysis and subsequent RNA purification. [^{3}H]uridine-labeled RNAs were denatured with glyoxal and analyzed in 0.8% agarose gels [30].

For Northern blotting, samples containing 10 µg of total RNA from electroporated cells or *in vitro*-transcribed XRN1-treated RNA mixed with 5 µg total RNA from BHK cells were denatured using formaldehyde, separated in a formaldehyde-containing 1.5% agarose gel, and blotted onto a Hybond-N$^+$ membrane (GE-Healthcare) [30]. The blots were hybridized with ^{32}P-labeled oligonucleotides as described previously [38,39], except for the analysis of sfRNA production in the XRN1 silencing experiments, in which case randomly primed, [α-^{32}P]dATP-labeled cDNA fragments directed against the YFV 3′ UTR (nt 10,555 to 10,862) and the human glyceraldehydes-3-phosphate dehydrogenase (GAPDH) gene (GenBank accession no. NM_00246; nt 321 to 724) were used as probes in 5 x SSC and 50% formamide at 42°C [38].

Virus stocks, infections and plaque assays

Medium was harvested from transfected cells to obtain virus stocks when a complete cytopathic effect (CPE) was observed. YFV-17D infections and plaque assays were per-

formed essentially as described before [27], except for the agarose in the overlay, which was replaced by 1.2% Avicel [40].

Primer extension assay

Primer extension analysis was performed as described by Sambrook *et al.* [30], with minor modifications. Total RNA (5 to 7 µg) from YFV or recombinant Sinrep5 RNA-transfected cells or XRN1-treated pYF-R.luc2A-RP transcript was annealed to ^{32}P-labeled oligonucleotide 1632 or 1648. Oligonucleotide 1632 is complementary to YFV nt 10,690 to 10,708, and oligonucleotide 1648 is complementary to YFV nt 10,580 to 10,598. After hybridization and subsequent ethanol precipitations, primer extension reactions were performed using 200 U of RevertAid H Minus Moloney murine leukaemia virus reverse transcriptase (Fermentas) as described by the manufacturer. After 1 h of incubation at 42°C, the samples were treated with RNase A (10 µg/µl; Qiagen) and purified by phenol-chloroform extraction and ethanol precipitation. The primer extension products were analyzed in a denaturing 5% polyacrylamide–8M urea sequencing gel. A ^{33}P-labeled Cycle Reader (Fermentas) sequence reaction mix using oligonucleotide 1632- or 1648-primed pBlsrcptSK-YFV$_{9845\text{-}10861}$ as a template was run on the same gel as the primer extension products and served as a size and sequence marker.

RNase protection assay

pBlscrptSK-YFV$_{10520\text{-}10714}$ was linearized using *Hind III,* purified by phenol-chloroform extraction, and used as a template for T7 RNA polymerase-mediated *in vitro* transcription to yield a 229-nt probe that contains 34 nt derived from the vector and 195 nt that are complementary to the YFV 10,520 to 10,714. The RNA probe was purified from a 6% polyacrylamide gel and hybridized overnight at 42°C to 10 µg of total RNA from mock- or YFV-infected BHK cells, using the solutions and protocols supplied with an RPAIII RNase protection assay kit (Ambion). As a positive control, 10 ng of a positive-strand T7 RNA polymerase transcript containing the 3' 1 kb of the YFV genome was mixed with 10 µg of total BHK cell RNA and treated similarly to the samples containing the YFV- or mock-infected total BHK cell RNA. After hybridization, the samples were treated with RNase A/T1, ethanol precipitated, and subsequently analyzed in a 6% denaturing sequencing gel. A sequence reaction mix that served as a size marker was run in parallel with the samples in the RNase protection assay.

RNA structure determination by selective 2′-hydroxyl acylation and primer extension (SHAPE) probing

A DNA template containing a T7 RNA polymerase promoter fused to the 3′ terminal 342 nt of YFV was generated by PCR, using *Pfu* DNA polymerase (Fermentas) as described by the manufacturer. The template was digested with *Eco RI* and *Hind III* and cloned in pUC9. After linearization with *Xba I*, the plasmid was used for *in vitro* RNA transcription (T7 MEGAscript kit; Ambion) to obtain a YFV RNA fragment of 188 nt (YFV nt 10,520 to 10,708) for probing of the RNA structure in the region encompassing the predicted RNA pseudoknot [23]. RNA was purified as described by the manufacturer (Ambion), and the yield was determined by spectrophotometry.

Probing was performed essentially as previously described [41]. Briefly, RNA (20 pmol) in 6 µl H_2O was heated at 95°C for 3 min, cooled down on ice, and subsequently incubated with 3 µl folding buffer at 37°C. After 25 min of incubation, 1 µl of 65 mM N-methylisatoic anhydride (NMIA; Sigma) in anhydrous dimethyl sulfoxide (DMSO) was added to the RNA and allowed to react for 60 min at 35°C. The control reaction mixture contained 1 µl of anhydrous DMSO and no NMIA. To determine the NMIA-induced modifications in the 188-bp YFV transcript, 4 pmol (2 µl) of the NMIA-treated RNA was annealed to 10 pM ^{32}P-labeled oligonucleotide 1632 and used for primer extension without any further purification.

RESULTS

Small virus-specific RNAs can be detected in YFV-infected mammalian and insect cells

Recent work [10] demonstrated that cells infected with arthropod-borne flaviviruses produced a small, virus-specific RNA derived from the 3′ UTR. To analyze whether the YFV sfRNA is also produced in infected cells other than BHK-21 cells [10], the mammalian cell lines Vero-E6 and SW13 and the mosquito cell line C6/36 were infected with YFV-17D. At 30 h postinfection (p.i.) (mammalian cell lines) or 36 h p.i. (insect cell line), total RNAs were isolated and analyzed by Northern blotting for YFV sfRNA production, using ^{32}P-labeled oligonucleotide 1632, complementary to nt 10,690 to 10,708 in the YFV 3′ UTR, as a probe. Apart from the YFV genomic RNA, three smaller YFV-specific RNAs were detected in the infected mammalian cells (fig. 1). Using *in vitro* RNA transcripts of various lengths as size markers (data not shown), these RNAs were estimated to be 630 nt (RNA A), 330 nt (RNA B), and 235 nt (RNA C) in length. Based on their estimated sizes and assuming that the RNAs were colinear with the 3′ end of the genome, the 5′ end of RNA A

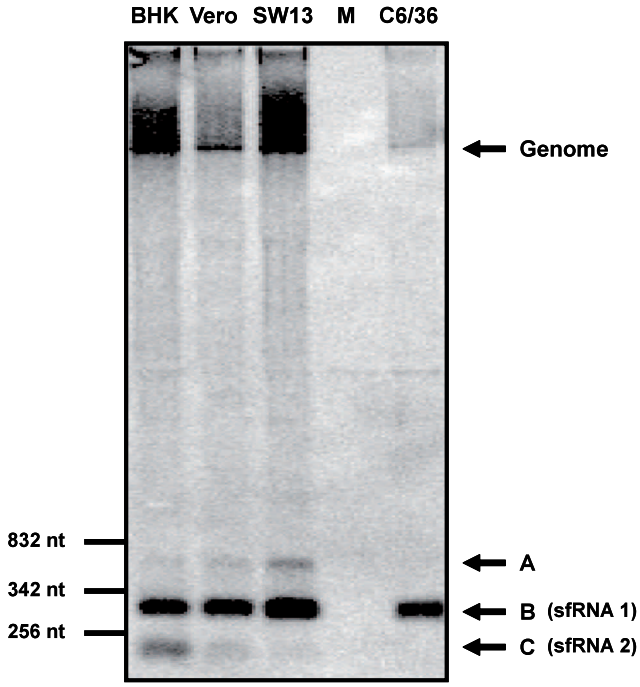


Fig. 1. YFV-17D sfRNA production in mammalian and insect cell lines.
The mammalian cell lines BHK-21J, Vero E6, and SW13 and the mosquito cell line C6/36 were infected with YFV-17D at an MOI of 10. At 30 h (mammalian cell lines) or 36 h (mosquito cell line) p.i., total RNA was isolated and analyzed by Northern blotting for the production of YFV-17D sfRNA. Lane M, total RNA isolated from mock-infected BHK-21J cells. Oligonucleotide 1632, complementary to nt 10,690 to 10,708 of the YFV 3′ UTR, was used as a probe. Size markers are indicated on the left. Bands corresponding to the YFV-17D genome and to three small viral RNAs (A, B, and C) are indicated by arrows. RNA B and RNA C are referred to as sfRNA1 and sfRNA2, respectively.

would be located within the carboxy-terminal coding region of NS5, whereas the 5′ ends of RNAs B and C would be in the 3′ UTR. RNA B was the most abundant of these RNAs, and the concentrations of RNA A and RNA C varied depending on the cell line studied. Hardly any RNA A was detected in the infected BHK cells, in which RNA C was relatively abundant, whereas the reverse was true for SW13 cells (fig. 1, lanes 1 and 3). Surprisingly, RNA B was the only small YFV-related RNA detected in the mosquito (C6/36 cells) (fig. 1). Of the three RNAs detected in the analyzed mammalian cell lines, RNA B appeared to be similar to the recently described YFV sfRNA [10]. To keep in line with the nomenclature used in previously published work, RNA B was named sfRNA1. The slightly smaller RNA C was named sfRNA2, which does not imply that this sfRNA is similar to WNV sfRNA2, which can be detected only when the XRN1 stalling site for WNV sfRNA1 is deleted [10].

YFV sfRNA1 and sfRNA2 have identical 5′ ends

Oligonucleotide 1632 (fig. 2.A), which was expected to bind to both sfRNA1 and sfRNA2, was used for primer extension analysis of total RNA isolated from YFV-infected or mock-infected BHK cells. As an additional control, a reaction was performed with a sample containing total RNA of uninfected cells mixed with 1 µg of full-length YFV genome transcript to detect products that resulted from strong stops of the reverse transcriptase on the YFV genome. As shown in fig. 2.B, lane 1, primer extension of oligonucleotide 1632 on total RNA of YFV-infected cells yielded a product that was not present in the RNA of uninfected cells (fig. 2.B, lane 2) or in the sample containing the full-length YFV transcript (fig. 2.B, lane 3). Two primer extension products were observed that differed in size by only one nucleotide. The 5′ ends of these primer extension products were mapped to the A residues at positions 10,532 and 10,533 by use of an oligonucleotide 1632-primed sequence reaction with pBluescript-$YFV_{9845\text{-}10861}$ as a size marker. Based on these results, the length of the YFV sfRNA is 329 or 330 nt, which is in agreement with the size estimate for sfRNA1 based on the Northern blots. The result of the primer extension analysis was verified using an RNase protection assay with a gel-purified, ^{32}P-labeled probe that is complementary to 34 nt of the vector and to YFV nt 10,520 to 10,714. This probe has extra YFV nucleotides at the 3′ end relative to the predicted 5′ end of the sfRNA to discriminate between YFV genome and sfRNA protected fragments. As shown in fig. 2.C, some large YFV genome-derived bands were detected in both total RNA of YFV-infected cells and the control sample (fig. 2.C, lanes 2 and 4, respectively). In addition, the probe protected a unique 181 or 182 nt fragment in the total RNA of YFV-infected BHK cells (fig. 2.C, lane 2). The size of the protected RNA fragment agreed with that predicted based on the results of the primer extension analysis and confirmed that the 5′ end of sfRNA1 is at position 10,532 or 10,533.

The most surprising result of these assays was that only these primer extension products and these RNase-resistant RNA fragments were detected, suggesting that sfRNA1 and sfRNA2 have a common 5′ end and that the observed difference in size between the two RNAs was due to a truncation at the 3′ end of sfRNA2. To verify this hypothesis, total RNAs were isolated from YFV-infected BHK-21, Vero E6, SW13 and C6/36 cells and analyzed by Northern blotting, using oligonucleotides 1648 and 1296 as probes. Oligonucleotide 1648 was complementary to nt 10,580 to 10,597 of the YFV genome (fig. 3.A), near the predicted 5′ end of the sfRNAs. Oligonucleotide 1296 should bind to the 3′ terminal 32 nt of the YFV genome (fig. 3.A). As shown in fig 3.B, oligonucleotide 1648 hybridized to both sfRNA1 and sfRNA2, whereas oligonucleotide 1296 only detected sfRNA1 (fig. 3.C.). These results demonstrated that the two YFV sfRNAs have a common 5′ end and that sfRNA2 is indeed truncated at the 3′ end.

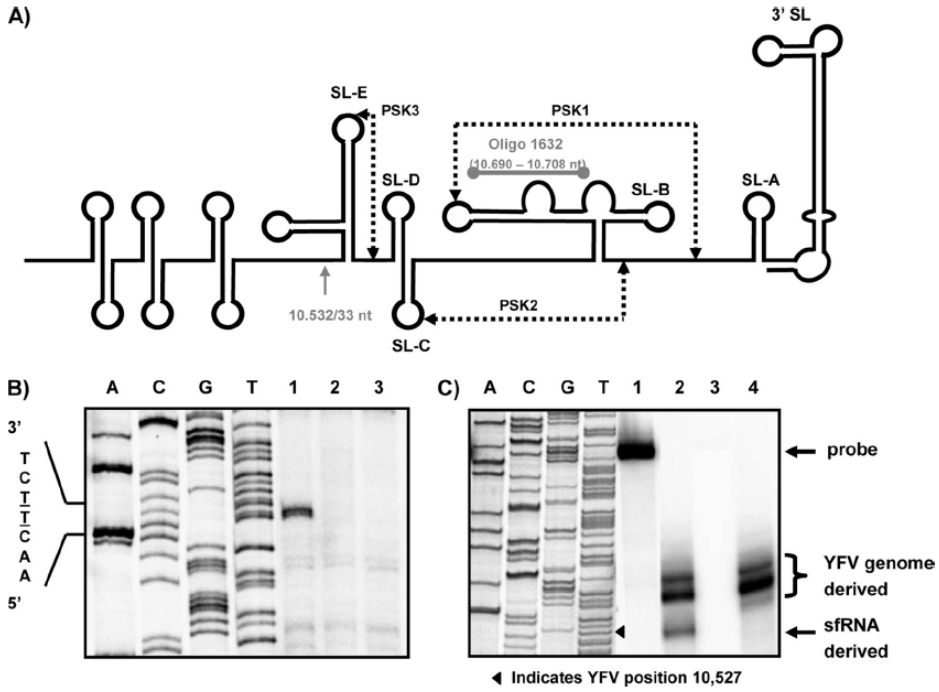


Fig. 2. Determining the 5′ end of the YFV 3′ UTR-specific sfRNAs produced in BHK-21J cells.
A) Schematic diagram of the predicted secondary structure of the YFV 3′ UTR [23]. **B)** Primer extension analysis using oligonucleotide 1632, which is complementary to YFV nt 10,690 to 10,708 (see panel A). pBluescript-YFV$_{9,845\text{-}10,861}$, containing the COOH-terminal part of the YFV NS5 gene and the complete 3′ UTR, was sequenced with oligonucleotide 1632 to obtain a sequencing ladder for determination of the 5′ end of the sfRNAs. Lanes 1 and 2, primer extension on total RNAs isolated from YFV-infected and uninfected BHK-21J cells, respectively; lane 3, primer extension on a full-length *in vitro* YFV transcript mixed with total RNA from uninfected BHK-21J cells. The underlined T residues of the depicted sequence correspond to the 5′ ends of the primer extension products and map the 5′ ends of the YFV sfRNAs to the A residues at positions 10,532 and 10,533 of the YFV genome. **C)** RNase protection assay using a 229-nt antisense RNA probe encompassing nt 10,520 to 10,714 of the YFV-17D 3′ UTR. Lane 1, ^{32}P-labeled RNA transcript used as the probe; lane 2, RNA fragments that were protected from RNase digestion after hybridization of the probe to total RNA isolated from YFV-17D-infected cells; lane 3, RNase protection assay on total RNA of mock-infected cells; lane 4, protected RNA fragments obtained when pBluescript-YFV$_{9,845\text{-}10,861}$ transcripts mixed with total RNA from mock-infected BHK-21J cells were analyzed. Bands corresponding to the protected fragments derived from either the YFV-17D genome or the sfRNA are indicated. pBluescript-YFV$_{9,845\text{-}10,861}$ was sequenced with oligonucleotide 1632 to obtain a marker to determine the sizes of the protected RNA fragments.

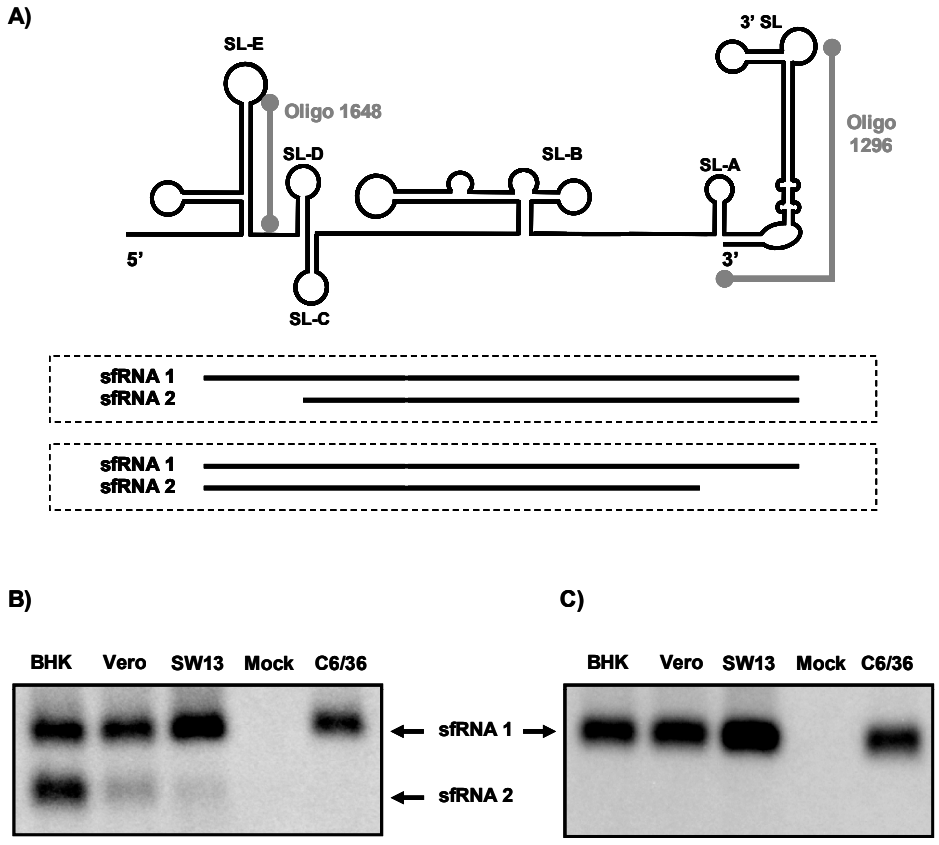


Fig. 3. YFV sfRNA2 is truncated at the 3' end.
A) Scheme of the YFV 3' UTR. The two possible orientations of the sfRNAs relative to the YFV 3' UTR are depicted. The positions of oligonucleotides 1648 and 1296, used to determine the orientation of the sfRNAs, are indicated. Total RNAs were isolated from the indicated YFV-infected cell lines or from mock-infected BHK-21J cells and subsequently analyzed by Northern blotting. **B)** Northern blot analysis using oligonucleotide 1648, complementary to YFV-17D nt 10,580 to 10,598. **C)** Northern blot analysis using oligonucleotide 1296, complementary to YFV-17D nt 10,830 to 10,862.

YFV sfRNA is generated by the host enzyme XRN1

As recently shown, the host exoribonuclease XRN1 hydrolyzes the WNV genome in a 5'-3' direction until it is stalled by a currently unknown signal yielding the WNV sfRNA [10]. To determine whether XRN1 was also involved in the production of the YFV sfRNAs, the effects of silencing of XRN1 expression on sfRNA production in YFV-infected cells and the *in vitro* production of sfRNA by incubation of YFV transcripts with purified XRN1 were studied. The human-derived SW13 cell line was used in the XRN1 silencing experiments because the nucleotide sequence of the human XRN1 gene is the only one known among the cell lines that were analyzed for YFV sfRNA expression (fig. 1). SW13 cells

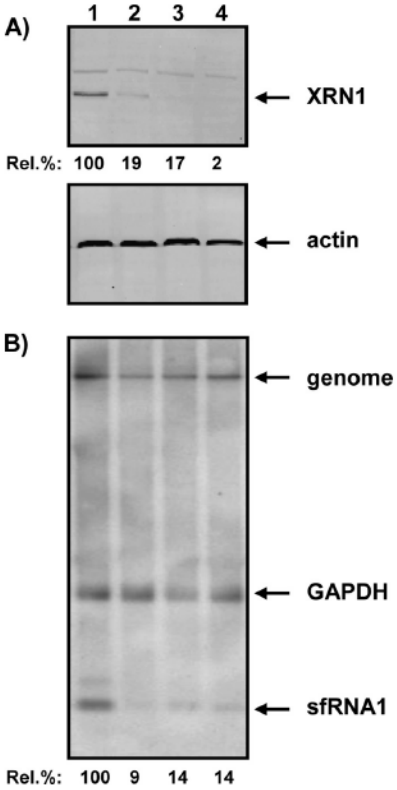


Fig. 4. shRNA-mediated XRN1 silencing decreases YFV sfRNA production.
SW13 cells were transduced with an shRNA-expressing lentivirus from the MISSION®TRC-Hs1.0 library (Sigma) and then infected with YFV-17D as described in Material and Methods. Lentiviruses expressed the following shRNAs: lane 1, scrambled shRNA (SHC-002); lane 2, shRNAs TRCN-049675 and TRCN-049676; lane 3, shRNAs TRCN-049675 and TRCN-049677; and lane 4, shRNAs TRCN-049676 and TRCN-049677. **A)** XRN1 and actin expression in Western blots. Rel. %, percentage of XRN1 expression in cells transduced with shRNAs against XRN1 compared to that in cells expressing the scrambled shRNA. **B)** YFV sfRNA production and GAPDH mRNA expression by Northern analysis. Rel. %, expression of YFV sfRNA in cells transduced with shRNAs against XRN1 compared to that in cells expressing the scrambled shRNA.

were transduced with lentiviruses expressing shRNA direct against the XRN1 transcript or a control shRNA, infected with YFV, and analyzed for XRN1 expression and YFV sfRNA production by Western and Northern blotting. Expression of each combination of two of the three shRNAs, TRCN-049675, TRCN-049676, and TRCN-049677, in SW13 cells resulted in significantly less production of XRN1 (fig. 4.A, lanes 2, 3, and 4) than that in cells transduced with a similar lentiviral vector expressing a scrambled shRNA (fig. 4.A, lane 1). The silencing of XRN1 expression by the indicated shRNAs was likely specific, since no effect on the expression of actin was observed in these cell lysates (fig. 4.A, bottom panel). As shown by the Northern blot analysis, silencing of XRN1 decreased the production of the YFV sfRNA by approximately 90% for all three combinations of shRNAs tested when compared to that in cells transduced with a lentivirus expressing a scrambled shRNA (fig. 4.B). The signals for the host cell GAPDH gene were similar in all these lysates, indicating that the observed differences in sfRNA production were not due to significant experimental error. Visual inspection of the cells expressing the shRNAs directed against XRN1 indicated that silencing affected the homeostasis of the cells. This might explain

why the hybridization signal for the YFV genome was also somewhat weaker than that in cells transfected with the scrambled shRNA (fig. 4.B).

To determine whether purified XRN1 was able to produce YFV sfRNA *in vitro*, linearized pYF-R.luc2A-RP [37] was used as a template for the production of an uncapped YFV transcript. After treatment with TAP to remove the XRN1-blocking 5′-triphosphate, this RNA was incubated with 0.1 or 1 unit of XRN1 and analyzed by Northern blot analysis for sfRNA production. Incubation of YF-R.luc2A-RP RNA with XRN1 resulted in the production of a small YFV-specific RNA that comigrated with the sfRNA1 produced in YFV-infected BHK-21J cells (fig. 5.A). Primer extension analysis was performed to provide additional evidence that this small RNA produced *in vitro* truly reflected the YFV sfRNA. As shown in fig. 5.B, primer extension using oligonucleotide 1632 resulted in similarly sized products for both XRN1-treated YF-R.luc2A-RP RNA transcript and total RNA from YFV-infected BHK cells. The fact that the smaller sfRNA2 was readily detectable in YFV-infected BHK-21J cells and not in the *in vitro* XRN1 assays is in line with our previous findings that sfRNA1 and sfRNA2 have a common 5′ end and suggests that additional RNA processing is required to yield sfRNA2.

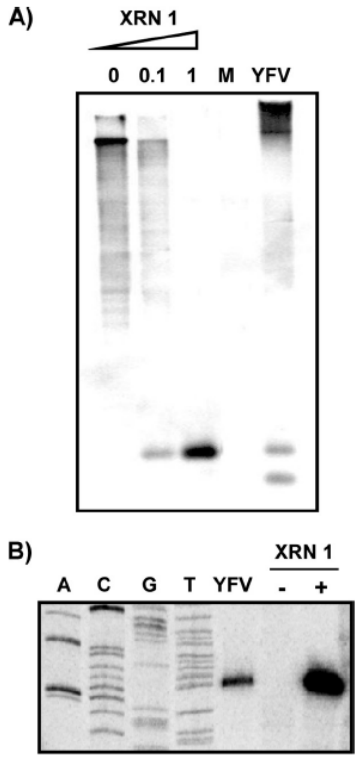


Fig. 5. *In vitro* production of YFV-17D sfRNAs by the host exoribonuclease XRN1.
TAP-treated *in vitro* RNA transcripts of YF-R.luc2A-RP were incubated with the indicated units of XRN1 and analyzed by hybridization after denaturing gel electrophoresis (**A**) and primer extension (**B**), using oligonucleotide 1632, complementary to YFV-17D nt 10,690 to 10,708, as a probe and primer, respectively. RNAs isolated from mock (M) and/or YFV-infected BHK-21J cells were used as controls in both experiments.

From the results obtained with the RNA silencing experiments and the *in vitro* YFV sfRNA production by purified XRN1, it was concluded that similar to the case for WNV, the host RNase XRN1 is responsible for the production of YFV sfRNAs.

An RNA pseudoknot in the YFV 3′ UTR is required for production of sfRNAs

Compared to the predicted RNA structure of the YFV 3′ UTR [23], the 5′ end of sfRNA1 and sfRNA2 is located just a few nucleotides upstream of SL-E (fig. 2.A), which is part of an RNA pseudoknot. To determine whether the predicted RNA structures that form this pseudoknot (PSK3) are required for YFV sfRNA production, site-directed mutagenesis was used to construct pACNR-FLYF-17DΔSL-E, which lacked nt 10,537 to 10,596 (which form SL-E) and was therefore unable to form PSK3. *In vitro*-transcribed RNA of *AflII*-linearized pACNR-FLYF-17DΔSL-E was electroporated into BHK-21J cells that were subsequently analyzed for viral RNA replication by [^{3}H]uridine labeling and for sfRNA synthesis by Northern blotting. Despite the relatively minor effects on viral replication (fig. 6.A), no YFV sfRNA could be detected in the cells transfected with YFV-ΔSL-E RNA (fig. 6.B). To demonstrate that the lack of sfRNA production in the cells transfected with YFV-ΔSL-E RNA was due to disruption of the XRN1 stalling site, transcripts of this mutant were incubated *in vitro* with purified XRN1. As shown in fig. 6.C, XRN1 was not stalled when SL-E was deleted and therefore no sfRNA was produced. These *in vitro* and *in vivo* results together demonstrate that SL-E encompasses RNA structures and/or sequences that are essential for stalling XRN1 and therefore for YFV sfRNA production. The YFV-ΔSL-E mutant virus was delayed in inducing CPE in BHK cells. It also showed a slight delay in initial virus production but reached a similar maximum titer to that of the parental virus (fig. 6.D).

Two sets of additional mutants were constructed to dissect the role of SL-E in YFV sfRNA synthesis in more detail. The first set of mutants focused on the role of the top stem structure of SL-E (referred to as e2) [23]. In mutant YFV-e2AA, the right arm of the stem (5′-GCAGU-3′; YFV nt 10,583 to 10,587) was replaced by the sequence of the left part of the stem (5′-ACUGC-3′) (fig. 7.A). Due to these mutations, the formation of SL-E e2 was expected to be disrupted. YFV-e2BB was the opposite of the YFV-e2AA mutant. In this mutant, the left part of the e2 stem (5′-ACUGC-3′; YFV nt 10,567 to 10,571) was replaced by the nucleotide sequence of the right arm of SL-E e2 (5′-GCAGU-3′) (fig. 7.A). In the YFV-e2BA mutant, the possibility to form stem e2 was restored by switching the parental YFV-17D original sequences of the left and right arms of the stem. BHK-21J cells were electroporated with *in vitro*-transcribed RNAs of the parental YFV-17D strain and the YFV-e2 stem mutants. Cells were analyzed for genome RNA synthesis by [^{3}H]uridine labeling and for sfRNA production by Northern blotting with oligonucleotide 1632. As shown in fig. 7.B, all three YFV-e2 mutants were able to synthesize viral genomic RNA

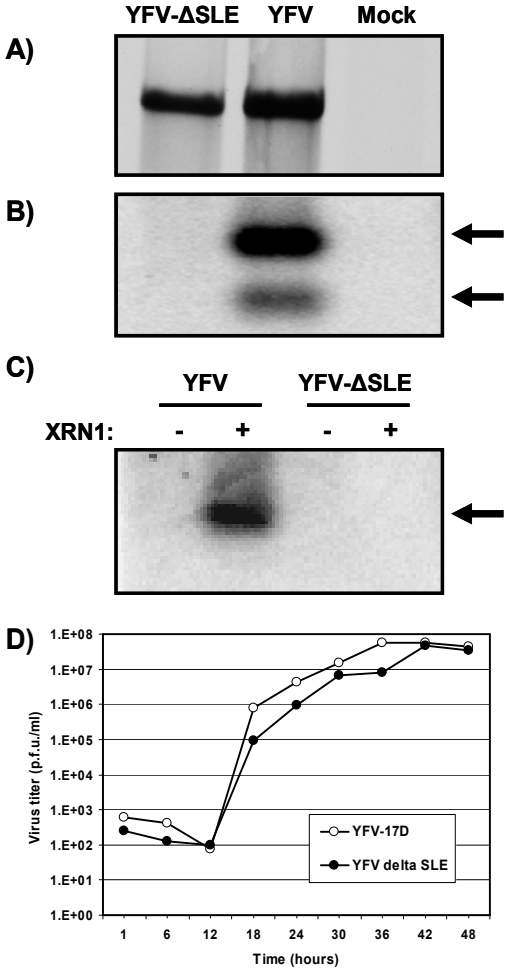


Fig. 6. Stem-loop structure SL-E in the YFV 3' UTR is required for the production of sfRNA.

A) Viral RNA synthesis in BHK-21J cells transfected with *in vitro*-transcribed genome-length RNAs of YFV-17D and YFV-ΔSL-E. Transfected cells were labelled with [^{3}H]uridine from 18 to 24 h posttransfection. Total RNAs were isolated and analyzed after denaturation by agarose gel electrophoresis, as described in Materials and Methods. **B)** sfRNA production in BHK-21J cells transfected with YFV-17D and YFV-ΔSL-E transcripts. Total RNAs were isolated at 24 h posttransfection and analyzed by Northern blotting and hybridization with oligonucleotide 1632 (complementary to YFV-17D nt 10,690 to 10,708). The sfRNAs are indicated by arrows. **C)** *In vitro* RNA transcripts of YFV-17D and YFV-ΔSL-E were incubated in the presence or absence of 1 unit of XRN1 and analyzed for the production of sfRNA by Northern blotting, using oligonucleotide 1632 as a probe. The sfRNA is indicated by an arrow. **D)** Viral growth kinetics of YFV-17D and the YFV-ΔSL-E mutant. BHK-21J cells were infected at an MOI of 5, and the medium of the infected cells was sampled at the indicated times postinfection. Titers were determined by plaque assays on BHK-21J cells.

efficiently. Strikingly, neither YFV-e2AA nor YFV-e2BB was able to produce detectable amounts of sfRNA (fig. 7.C). Restoring the possibility to form stem e2, as in YFV-e2BA, also restored the production of YFV sfRNA. Although in comparison to YFV-17D YFV-e2BA produced less sfRNA, both viruses induced CPE in the cells about 48 h after electroporation, whereas YFV-e2AA and YFV-e2BB were clearly delayed in CPE induction compared to the parental virus. The kinetics of virus production in BHK-21 cells were similar to that shown for YFV-ΔSL-E (e.g., YFV-e2AA) or closer to that for the parental virus (data not shown). The fact that the YFV sfRNAs were produced in cells infected with YFV-e2BA indicated that the SL-E structure was more important than the primary sequence for producing the sfRNA.

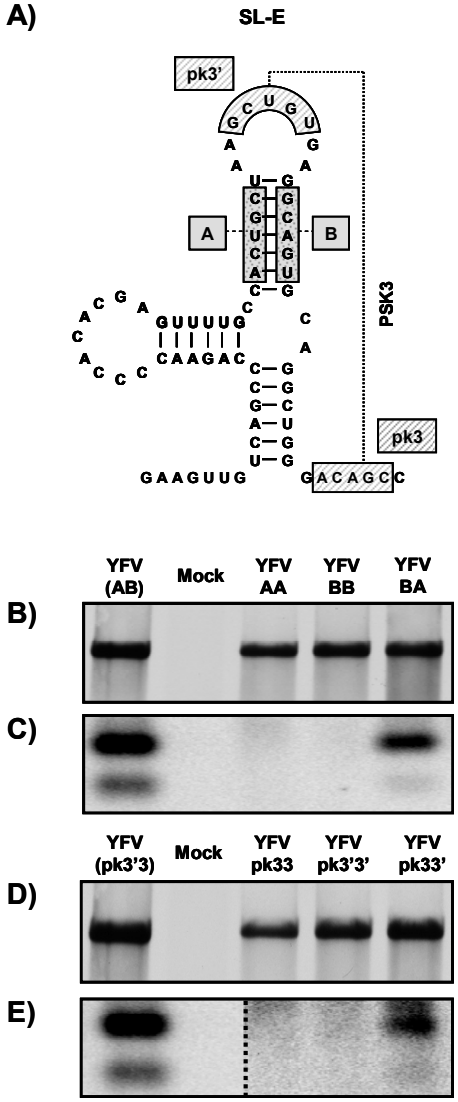


Fig. 7. YFV-17D RNA pseudoknot 3 is required for sfRNA production.

A) Schematic diagram of the YFV 3′ UTR SL-E structure. The primary sequences involved in stem-loop e2 (sequences A and B) and in the pseudoknot interaction (sequences pk3′ and pk3) are depicted. **B** and **D)** Viral RNA synthesis in BHK-21J cells transfected with *in vitro*-transcribed genomic RNAs of YFV-17D and YFV mutant viruses (at SL-E-e2 [**B**] and at PSK3 [**D**]). Transfected cells were labeled with [^{3}H]uridine from 18 to 24 h posttransfection. Total RNAs were isolated and analyzed after denaturation by agarose gel electrophoresis, as described in Materials and Methods. YFV(AB) and YFV(pk3′3) indicate wild-type YFV-17D for these different groups of mutants. **C** and **E)** Northern blot analysis of RNAs isolated from BHK-21J cells 30 h posttransfection with YFV-17D and YFV mutant viruses (at SL-E e2 [**C**] and at PSK3 [**E**]). Oligonucleotide 1632 (complementary to YFV-17D nt 10,690 to 10,708) was used as a probe. In panel E, the dotted line separates the wild-type YFV-17D and mock lanes from the same Northern blot at a higher exposure.

SL-E is predicted to be part of a more complex RNA structure in which the sequence 5′-GCUGU-3′ (pk3′; YFV nt 10,575 to 10,579) in the top loop of e2 is predicted to base pair with the sequence 3′-CGACA-5′ (pk3; YFV nt 10,598 to 10,602) immediately downstream of SL-E, forming an RNA pseudoknot (PSK3) (fig. 7.A) [23]. A second set of YFV mutants was constructed to determine whether the pk3′-pk3 interaction is important for YFV sfRNA production. In YFV-pk3′pk3′ and YFV-pk3pk3, the downstream pk3 sequence or the upstream pk3′ sequence was replaced by the complementary sequence. For both of these mutants, PSK3 formation was expected to be disrupted. In YFV-pk3pk3′, the possi-

bility to form PSK3 was restored, albeit with the positions of the pk3′ and pk3 sequences reversed in comparison to those in YFV-17D. [^{3}H]uridine labeling of BHK cells transfected with *in vitro*-transcribed full-length RNAs of these mutants demonstrated that they replicated quite efficiently (fig. 7.D). More importantly, no sfRNA could be detected in cells transfected with YFV-pk3′pk3′ or YFV-pk3pk3, whereas YFV sfRNA could be detected in cells electroporated with the YFV-pk3pk3′ mutant, in which the ability to form PSK3 was restored (fig. 7.E). However, the amount of sfRNA produced by YFV-pk3pk3′ was far less than that observed for the parental YFV-17D strain and required contrast enhancement of the phosphorimager data, which changed only the view of the data, not the data themselves. All three pk3-pk3′ mutant viruses showed a significant delay in the onset of CPE in BHK cells compared to YFV-17D. The kinetics of virus production in BHK-21J cells were similar to that for YFV-17D, except for YFV-pk3′pk3′, whose kinetics was like that of YFV-ΔSL-E (data not shown).

The combined results of the mutagenesis of the SL-E e2 stem and the pk3′-pk3 sequences strongly suggest that the predicted H-type RNA pseudoknot PSK3 is required to stall XRN1, resulting in the production of YFV sfRNA.

Insertion of the PSK3 sequence into a heterologous RNA results in production of an sfRNA-like RNA

The fact that the base pairings in SL-E stem e2 and between pk3′ and pk3 are important determinants of YFV sfRNA synthesis did not exclude the possibility that this region is part of a more complex RNA structure, including additional elements downstream of PSK3. To address this possibility, a synthetic *Mlu I – Sph I* DNA mimicking YFV nt 10,521 to 10,662 and YFV nt 10,531 to 10,611 was inserted into a Sinrep5 vector [34] that contained an enhanced green fluorescent protein (GFP) gene just upstream of the insertion site (fig. 8.A). This resulted in two constructs: pSinrep5-eGFP-YFV$_{10,521\text{-}10,662}$, encoding the SL-E to SL-C fragment of the YFV 3′ UTR, and pSinrep5-eGFP-YFV$_{10,531\text{-}10,611}$, encoding the YFV SL-E/PSK3 region. The GFP gene allowed for easy determination of the transfection efficiency and for discrimination between the Sindbis virus subgenomic RNA and a potentially produced sfRNA-like RNA. BHK-21J cells were transfected with *in vitro*-transcribed RNAs of the Sinrep5-eGFP and Sinrep5-eGFP-YFV mutants. RNAs were isolated at 8 h p.e. and analyzed for sfRNA-like RNA production by Northern blotting and hybridization, with two oligonucleotides as probes. Probe 1674 was complementary to nt 7,601 to 7,625 of Sindbis virus and was expected to hybridize to both the Sinrep5 genomic and subgenomic mRNAs, whereas probe 1648, which was complementary to YFV SL-E (nt 10,580 to 10,597), was expected to hybridize not only to Sinrep5 mRNAs but also to any sfRNA-like RNA produced in the transfected cells. As shown in fig. 8.B, hybridization of total RNA

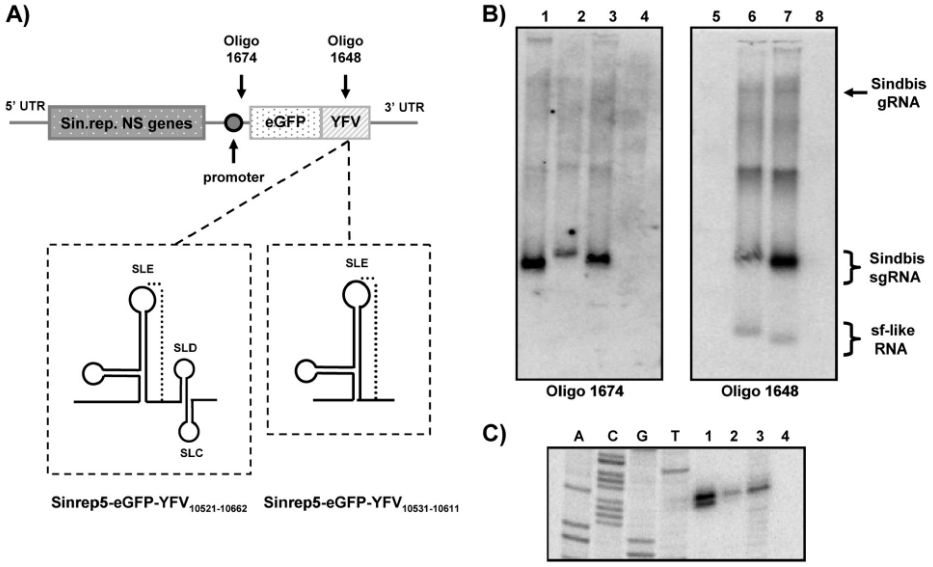


Fig. 8. Insertion of sequences required for the formation of PSK3 is sufficient to produce an sfRNA-like RNA in the context of a Sindbis virus replicon.
A) Scheme of the characteristics of the pSinrep5 vector and the predicted structures of the YFV-17D regions that were inserted into the vector. The YFV 3' UTR nucleotides cloned into pSinRep5-eGFP are indicated in the name of each construct. The promoter for the enhanced GFP (eGFP)-expressing subgenomic Sinrep mRNA and the binding sites for oligonucleotides 1674 and 1648 are indicated. **B)** Northern blot analysis of RNAs isolated from BHK-21J cells 8 h p.e. with Sinrep5-eGFP (lanes 1 and 5), Sinrep5-eGFPYFV$_{10,521-10,662}$ (lanes 2 and 6), and Sinrep5-eGFPYFV$_{10,531-10,611}$ (lanes 3 and 7) RNAs; lanes 4 and 8 correspond to uninfected BHK-21J cells. Oligonucleotide 1674 and oligonucleotide 1648 were used as probes, as specified in the figure. The Sindbis virus genomic (gRNA) and subgenomic (sgRNA) RNAs and the sfRNA-like RNAs are indicated. **C)** Primer extension analysis with oligonucleotide 1648 to determine the 5' end of the sfRNA-like RNAs produced with the Sinrep5 mutants. pBluescript-YFV$_{9,845-10,861}$ was sequenced with oligonucleotide 1648 to obtain a sequencing ladder. RNAs isolated from BHK-21J cells transfected with YFV-17D (lane 1), Sinrep5-eGFPYFV$_{10,521-10,662}$ (lane 2), and Sinrep5-eGFPYFV$_{10,531-10,611}$ (lane 3) were analyzed; lane 4 corresponds to mock-transfected cells.

of the transfected cells with oligonucleotide 1674 demonstrated that Sinrep5-eGFP and both Sinrep5-eGFP-YFV$_{10,521-10,622}$ and pSinrep5-eGFP-YFV$_{10,531-10,611}$ replicated efficiently and produced a subgenomic mRNA of the expected size in the transfected BHK cells. In addition to the recombinant Sinrep5 genomic and subgenomic mRNAs, hybridization with oligonucleotide 1648 revealed the production of an additional small RNA that was unique for the cells transfected with Sinrep5-eGFP-YFV$_{10,521-10,662}$ and pSinrep5-eGFP-YFV$_{10,531-10,611}$ (fig. 8.B, lanes 6 and 7). The size of this RNA was in agreement with what was expected for the production of an sfRNA-like RNA in pSinrep5-eGFP upon insertion of the YFV sequences. Primer extension of total RNAs from cells transfected with Sinrep5-eGFP-YFV$_{10,521-10,662}$ and pSinrep5-eGFP-YFV$_{10.531-10.611}$ (fig. 8.C) showed that the 5' end of

the sfRNA-like RNA produced in these transfected cells was similar to that of the sfRNA produced in YFV-infected cells. Taken together, these results clearly demonstrate that the YFV nt 10,531 to 10,611, containing the region that allows the formation of PSK3, are sufficient to stall XRN1. Furthermore, these results also indicate that RNA structures downstream of PSK3 are not required for the production of sfRNA.

Chemical RNA probing provides evidence of the formation of PSK3

To obtain additional evidence to support the formation of PSK3, *in vitro* structure probing by SHAPE with NMIA [41] was performed on an *in vitro*-synthesized RNA template containing YFV 3′ UTR nt 10,520 to 10,708, encompassing PSK3. As shown in fig. 9.B, chemical probing of this region of the wild-type YFV-17D genome essentially confirmed the previously predicted RNA structure model [23]. Apart from the nucleotides predicted to form the top loop of the side stem SL-E e3 (fig. 9.A), most of the other nucleotides did not react with NMIA, indicating that they were in a double-stranded conformation. This observation also includes the nucleotides of the pk3′ and pk3 sequences (fig. 9.B). The lack of reactivity to NMIA of these particular sequences was in agreement with

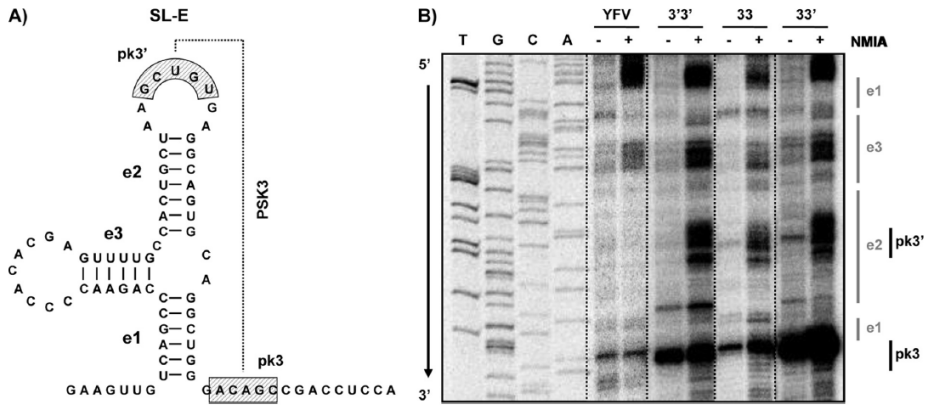


Fig. 9. The predicted pseudoknot PSK3 at the wild-type YFV 3′ UTR is genuine.
A) Schematic diagram of the minimal YFV 3′ UTR region (nt 10,531 to 10,611) required for stalling of XRN1. Sequences involved in the pseudoknot interaction are depicted. **B)** *In vitro* structure probing by SHAPE with wild-type YFV and the PSK3 mutants YFV-pk3′pk3′, YFV-pk3pk3, and YFV-pk3pk3′. After treatment with NMIA, which preferentially reacts with bases in a single-stranded conformation, i*n* vitro-synthesized RNA templates (YFV nt 10,520 to 10,708) were analyzed by primer extension with oligonucleotide 1632 (complementary to YFV nt 10,690 to 10,708). Samples were treated with either 65 mM NMIA in DMSO (+) or, as a control, DMSO only (-). The different substructures of SL-E and the pk3 and pk3′ sequences are indicated on the right relative to their positions in the NMIA reactivity pattern. pBluescript-YFV$_{9,845\text{-}10,861}$ was sequenced with oligonucleotide 1632 to obtain a sequencing ladder.

the prediction that these complementary sequences could base pair to form the PSK3 pseudoknot.

The mutants that should have disrupted (YFV-pk3'pk3' and YFV-pk3pk3) or restored (YFV-pk3pk3') PSK3 formation were also subjected to RNA structure probing. The NMIA reactivity patterns of these mutant RNAs showed that the overall SL-E structure was maintained but that especially the pk3' sequence, and to a somewhat lesser extent the pk3 sequence, were now susceptible to modification by NMIA (fig. 9.B), indicating that these sequences were no longer base pairing to form PSK3. Surprisingly, a similar result was also obtained for the putatively restored YFV-pk3pk3' mutant RNA, indicating that despite the nucleotide complementarity between the pk3 and pk3' sequences, the formation of PSK3 was not restored in this mutant.

Nonetheless, the structural probing data obtained with the RNA fragment encompassing YFV-17D nt 10,520 to 10,708 provided strong support for the previously proposed structural model [23].

DISCUSSION

The generally accepted concept that the viral genome, which also serves as the viral mRNA, and the genome-length, minus-strand RNA are the only RNA species that can be detected in cells infected with flaviviruses was recently challenged. Several studies have now shown that in addition to these genome-length viral RNAs, a small, positive-strand viral RNA (sfRNA) is produced in cells infected with arthropod-borne flaviviruses [6-10]. Recently, the sfRNA was shown to be a product of incomplete 5' to 3' degradation of the viral genome by the host exoribonuclease XRN1 and to serve as an important determinant of viral pathogenicity [10]. This study describes the production of sfRNA by XRN1 in YFV-infected cells and, more importantly, defines an RNA pseudoknot in the viral 3' UTR as a prerequisite for stalling of XRN1, resulting in the production of the sfRNA.

Not one but two sfRNAs are detected in YFV-infected cells. Another, larger YFV-specific RNA was also detected in infected mammalian cell lines, especially in SW13 cells. This RNA is likely an instable intermediate of 5' to 3' XRN1-mediated decay of the viral genome. Silencing of XRN1 in YFV-infected SW13 cells resulted in a significant decrease of YFV sfRNA1 as well as the larger YFV-specific RNA A. Digestion of YFV genome transcripts with purified XRN1 resulted in the *in vitro* production of sfRNA1, providing the ultimate proof for the role of XRN1 in YFV sfRNA production.

Mapping of the positions of sfRNA1 and sfRNA2 in the 3' UTR of the YFV genome revealed that sfRNA1 was colinear with the distal part of the 3' UTR. Both sfRNAs had the same 5' end, and the smaller size of sfRNA2 was due to a truncation at the 3' end. sfRNA2 was never detected upon *in vitro* digestion of YFV genome transcripts with puri-

fied XRN1. One hypothesis that could explain this truncation is that sfRNA2 is derived by XRN1 digestion of a YFV genomic RNA template that is already truncated at the 3′ end. Alternatively, sfRNA2 could be the result of additional processing by an exo- or an endoribonuclease at the 3′ end of sfRNA1. Currently, we have no evidence to favour one of these hypotheses. Interestingly, it has been shown that in insect cells the 5′ to 3′ decay pathway is the dominant route for overall RNA turnover [42], while in mammalian cells the relative contributions of 5′ to 3′ and 3′ to 5′ RNA degradation pathways to RNA turnover are still a subject of debate (reviewed in reference [13]). The fact that sfRNA2 was not detected in infected mosquito cells thus supports the hypothesis that sfRNA2 is the result of 3′ to 5′ processing.

Both the primer extension and RNase protection assays mapped the 5′ end of the YFV sfRNAs to nucleotide 10,532 or 10,533, which is just upstream of a predicted H-type RNA pseudoknot (PSK3) [23] that includes SL-E. The inability of XRN1 to produce an sfRNA both *in vivo* and *in vitro* when SL-E was disrupted strongly indicates that this region contains essential elements for the stalling of this host RNase. XRN1 can be blocked by elements such as strong secondary structures or G-track sequences [17,19-21]. The YFV PSK3 region is not particularly G rich but adopts a complex RNA structure. So far, no experimental evidence supporting the formation of PSK3 has been presented, nor has this structure been implicated in any biologically relevant function. Our RNA structure probing of the PSK3 region provided unequivocal evidence of the presence of this RNA pseudoknot in YFV 3′ UTR. Disruption of the SL-E structure resulted in viruses that replicated with a similar efficiency to that of YFV-17D but did not produce detectable amounts of sfRNA. Reconstituting the SL-E structure in a YFV mutant, albeit with different base pairs, resulted in sfRNA production in cells infected with this mutant. On the other hand, disrupting the predicted PSK3 resulted in mutant viruses that replicated efficiently but were no longer able to produce sfRNA. Surprisingly, when the base pairing possibility was restored (YFV-pk3pk3′), sfRNA production was hardly detectable. These data are supported by the probing results showing that although the overall SL-E structure was maintained in all PSK3 mutants, none of them showed any significant base pairing of the nucleotides involved in the equivalent of the parental pk3′ – pk3 interaction. Although restoration of the pseudoknot was expected in the YFV-pk3pk3′ mutant, this seemed to occur in only a minor fraction of the RNA molecules, which explains why they were not detected in the probing experiments and why enhancement of the contrast in the picture of the Northern blot was required. The experiments using the Sinrep5 expression system demonstrated that YFV nt 10,531 to 10,611, predicted to form PSK3, contain all of the RNA sequences and structures that are needed to stall XRN1. In an attempt to restore PSK3 interaction and subsequent sfRNA production, an additional YFV mutant was constructed in which the pk3′ and pk3 sequences were replaced by 5′ CCCGC 3′ and 5′ GCGGG 3′ sequences, respectively, to enhance the thermodynamic

stability of this interaction. Although this mutant was viable, it did not produce any sfRNA (data not shown), and probing provided no evidence of a pk3'-pk3 interaction (data not shown). Taken together, our mutational analysis and probing results support the actual formation of PSK3 in the 3' UTR of YFV but also indicate that the formation of this pseudoknot is sequence dependent. For many viral pseudoknots, the primary sequence is unimportant for function, as long as the conformation and overall stability of the structure are maintained (reviewed in reference [43]). Nonetheless, there are other pseudoknots in which subtle nucleotide changes interfere with pseudoknot thermodynamic stability and are deleterious for pseudoknot function [44,45]. Examples of viruses in which the primary sequence proved to be important for RNA pseudoknot stability and/or function include tobacco mosaic virus [46], beet western yellow virus [47], severe acute respiratory syndrome (SARS) coronavirus [48,49], and Visna-Maedi virus [50]. A similar situation can also be envisioned for YFV, in which specific constraints to sustain the overall stability/conformation of the higher-order RNA structure of PSK3 can be fulfilled only by the wild-type sequence. The fact that the pk3' and pk3 sequences are well conserved among the different YFV strains [51,52] supports this idea. The observation that disruption of PSK3 allows XRN1 to degrade the viral genome completely implies that the YFV genome has one unique stalling site for XRN1 within the 3' UTR. This observation differs from the case for KUNV, in which more than one stalling site was found [10].

It was shown that KUNV viruses deficient in sfRNA production are significantly more attenuated in cell culture and in mice than viruses that produce the sfRNA, suggesting an important role for the sfRNA in viral pathogenicity [10]. A similar observation was made with our YFV mutants: viruses that were unable to produce sfRNA were also unable to form plaques on SW13 cells, despite the fact that they were able to replicate efficiently in these cells (data not shown). Interestingly, it has previously been hypothesized that there is a correlation between the structure of the YFV region predicted to form PSK3 and the degree of virulence exhibited by the virus [53]. Although these observations for KUNV and YFV may hint at a function of the sfRNA, it is still unclear how sfRNA synthesis relates to viral pathogenesis. It has been suggested that the sfRNA could modulate the host antiviral responses by antagonizing or inactivating certain cellular RNA sensors (e.g., TLR3, RIG-I, and MDA5) or that it could act as a decoy for cellular microRNAs [54]. Since KUNV sfRNA was reported to be generated by XRN1 in P bodies [10], and since these P bodies contain, besides 5'-3' degradation components, proteins involved in translational repression, mRNA surveillance and RNA-mediated gene silencing (reviewed in references [11-13,55]), one could speculate that the sfRNA could indeed be involved in a pathway linked to P bodies, such as RNA-mediated gene silencing. However, a progressive decrease in the number of P bodies during the time course of DENV-2 and WNV infections of BHK cells has been reported [56]. Alternatively, either free in the cytoplasm or associated with

cellular structures, the sfRNA could also act as a decoy to sequester a host protein(s) that would otherwise bind to the genomic 3' UTR and exert a negative effect on the virus.

Additional studies are needed to determine the kinetics of sfRNA production, its subcellular localization, and potential sfRNA-interacting host factors. Previous studies with flaviviruses containing mutations or deletions in the genomic 3' UTR region should be re-evaluated in light of the potential effect of such mutations on the production and function of the sfRNA. For instance, deletions involving the SL-II region of DENV (which is comparable to the YFV SL-E) that resulted in restricted growth in cell culture, different plaque phenotypes and viral attenuation [57,58] could be explained by the lack of sfRNA production, illustrating the importance of the sfRNA in viral pathogenesis.

From this study and those of others, it is obvious that a simple deletion that disrupts the stalling site for XRN1 may be used to construct attenuated flaviviruses that could be considered vaccine candidates. However, it remains to be established whether the reduced pathogenicity of sfRNA-deficient viruses interferes with the broad immune response that is required to induce protection against wild-type virus infection or, even worse, results in persistent infections after administration of the vaccine.

ACKNOWLEDGEMENTS

We thank M. Rabelink and R.C. Hoeben for technical advice and support in preparing the lentiviruses for the XRN1 RNA silencing experiments and X. Jiang for providing the GAPDH cDNA fragment that was used to generate a GAPDH probe in the analysis of the XRN1 RNA silencing experiments.

REFERENCE LIST

1. **Cook, S. and E. C. Holmes**. 2006. A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: Flaviviridae) and the evolution of vector transmission. Arch. Virol. **151**:309-325.
2. **Kuno, G., G. J. Chang, K. R. Tsuchiya, N. Karabatsos, and C. B. Cropp**. 1998. Phylogeny of the genus Flavivirus. J. Virol. **72**:73-83.
3. **Lindenbach, B. D., H. J. Thiel, and C. M. Rice**. 2007. Flaviviridae: The Viruses and Their Replication *In* D. M. Knipe and P. M. Howley (eds.), Fields Virology. Lippincott Williams & Wilkins, Philadelphia.
4. **Chu, P. W. and E. G. Westaway**. 1985. Replication strategy of Kunjin virus: evidence for recycling role of replicative form RNA as template in semiconservative and asymmetric replication. Virology **140**:68-79.
5. **Cleaves, G. R., T. E. Ryan, and R. W. Schlesinger**. 1981. Identification and characterization of type 2 dengue virus replicative intermediate and replicative form RNAs. Virology **111**:73-83.
6. **Lin, K. C., H. L. Chang, and R. Y. Chang**. 2004. Accumulation of a 3′-terminal genome fragment in Japanese encephalitis virus-infected mammalian and mosquito cells. J. Virol. **78**:5133-5138.
7. **Scherbik, S. V., J. M. Paranjape, B. M. Stockman, R. H. Silverman, and M. A. Brinton**. 2006. RNase L plays a role in the antiviral response to West Nile virus. J. Virol. **80**:2987-2999.
8. **Urosevic, N., M. M. van, J. P. Mansfield, J. S. Mackenzie, and G. R. Shellam**. 1997. Molecular characterization of virus-specific RNA produced in the brains of flavivirus-susceptible and -resistant mice after challenge with Murray Valley encephalitis virus. J. Gen. Virol. **78 (Pt 1)**:23-29.
9. **Liu, R., L. Yue, X. Li, X. Yu, H. Zhao, Z. Jiang, E. Qin, and C. Qin**. 2010. Identification and characterization of small sub-genomic RNAs in dengue 1-4 virus-infected cell cultures and tissues. Biochem. Biophys. Res. Commun. **391**:1099-1103.
10. **Pijlman, G. P., A. Funk, N. Kondratieva, J. Leung, S. Torres, L. van der Aa, W. J. Liu, A. C. Palmenberg, P. Y. Shi, R. A. Hall, and A. A. Khromykh**. 2008. A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity. Cell Host. Microbe **4**:579-591.
11. **Anderson, P. and N. Kedersha**. 2006. RNA granules. J. Cell Biol. **172**:803-808.
12. **Eulalio, A., I. Behm-Ansmant, and E. Izaurralde**. 2007. P bodies: at the crossroads of post-transcriptional pathways. Nat. Rev. Mol. Cell Biol. **8**:9-22.
13. **Garneau, N. L., J. Wilusz, and C. J. Wilusz**. 2007. The highways and byways of mRNA decay. Nat. Rev. Mol. Cell Biol. **8**:113-126.
14. **Sheth, U. and R. Parker**. 2003. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science **300**:805-808.
15. **Stevens, A.** 1978. An exoribonuclease from Saccharomyces cerevisiae: effect of modifications of 5′ end groups on the hydrolysis of substrates to 5′ mononucleotides. Biochem. Biophys. Res. Commun. **81**:656-661.
16. **Stevens, A.** 1980. Purification and characterization of a Saccharomyces cerevisiae exoribonuclease which yields 5′-mononucleotides by a 5′ leads to 3′ mode of hydrolysis. J. Biol. Chem. **255**:3080-3085.
17. **Esteban, R., L. Vega, and T. Fujimura**. 2008. 20S RNA narnavirus defies the antiviral activity of SKI1/XRN1 in Saccharomyces cerevisiae. J. Biol. Chem. **283**:25812-25820.
18. **Cheng, C. P., E. Serviene, and P. D. Nagy**. 2006. Suppression of viral RNA recombination by a host exoribonuclease. J. Virol. **80**:2631-2640.

19. **Decker, C. J. and R. Parker**. 1993. A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation. Genes Dev. **7**:1632-1643.
20. **Muhlrad, D., C. J. Decker, and R. Parker**. 1994. Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5'-->3' digestion of the transcript. Genes Dev. **8**:855-866.
21. **Vreken, P. and H. A. Raue**. 1992. The rate-limiting step in yeast PGK1 mRNA degradation is an endonucleolytic cleavage in the 3'-terminal part of the coding region. Mol. Cell Biol. **12**:2986-2996.
22. **Hahn, C. S., Y. S. Hahn, C. M. Rice, E. Lee, L. Dalgarno, E. G. Strauss, and J. H. Strauss**. 1987. Conserved elements in the 3' untranslated region of flavivirus RNAs and potential cyclization sequences. J. Mol. Biol. **198**:33-41.
23. **Olsthoorn, R. C. and J. F. Bol**. 2001. Sequence comparison and secondary structure analysis of the 3' noncoding region of flavivirus genomes reveals multiple pseudoknots. RNA. **7**:1370-1377.
24. **Markoff, L.** 2003. 5'- and 3'-noncoding regions in flavivirus RNA. Adv. Virus Res. **59**:177-228.
25. **Shi, P. Y., M. A. Brinton, J. M. Veal, Y. Y. Zhong, and W. D. Wilson**. 1996. Evidence for the existence of a pseudoknot structure at the 3' terminus of the flavivirus genomic RNA. Biochemistry **35**:4222-4230.
26. **Bredenbeek, P. J., E. A. Kooi, B. Lindenbach, N. Huijkman, C. M. Rice, and W. J. Spaan**. 2003. A stable full-length yellow fever virus cDNA clone and the role of conserved RNA elements in flavivirus replication. J. Gen. Virol. **84**:1261-1268.
27. **Silva, P. A., R. Molenkamp, T. J. Dalebout, N. Charlier, J. H. Neyts, W. J. Spaan, and P. J. Bredenbeek**. 2007. Conservation of the pentanucleotide motif at the top of the yellow fever virus 17D 3' stem-loop structure is not required for replication. J. Gen. Virol. **88**:1738-1747.
28. **Igarashi, A.** 1978. Isolation of a Singh's Aedes albopictus cell clone sensitive to Dengue and Chikungunya viruses. J. Gen. Virol. **40**:531-544.
29. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl**. 2000. Current Protocols in Molecular Biology. Wiley Interscience, New York.
30. **Sambrook, J., T. Fritsch, and T. Maniatis**. 1989. Molecular Cloning: a Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
31. **Inoue, H., H. Nojima, and H. Okayama**. 1990. High efficiency transformation of Escherichia coli with plasmids. Gene **96**:23-28.
32. **Rice, C. M., E. M. Lenches, S. R. Eddy, S. J. Shin, R. L. Sheets, and J. H. Strauss**. 1985. Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. Science **229**:726-733.
33. **Molenkamp, R., E. A. Kooi, M. A. Lucassen, S. Greve, J. C. Thijssen, W. J. Spaan, and P. J. Bredenbeek**. 2003. Yellow fever virus replicons as an expression system for hepatitis C virus structural proteins. J. Virol. **77**:1644-1648.
34. **Bredenbeek, P. J., I. Frolov, C. M. Rice, and S. Schlesinger**. 1993. Sindbis virus expression vectors: packaging of RNA replicons by using defective helper RNAs. J. Virol. **67**:6439-6446.
35. **de Vries, A. A., E. D. Chirnside, M. C. Horzinek, and P. J. Rottier**. 1992. Structural proteins of equine arteritis virus. J. Virol. **66**:6294-6303.
36. **Liefhebber, J. M., B. W. Brandt, R. Broer, W. J. Spaan, and H. C. van Leeuwen**. 2009. Hepatitis C virus NS4B carboxy terminal domain is a membrane binding domain. Virol. J. **6**:62.
37. **Jones, C. T., C. G. Patkar, and R. J. Kuhn**. 2005. Construction and applications of yellow fever virus replicons. Virology **331**:247-259.

38. **Meinkoth, J. and G. Wahl**. 1984. Hybridization of nucleic acids immobilized on solid supports. Anal. Biochem. **138**:267-284.
39. **van der Most, R. G., P. J. Bredenbeek, and W. J. Spaan**. 1991. A domain at the 3′ end of the polymerase gene is essential for encapsidation of coronavirus defective interfering RNAs. J. Virol. **65**:3219-3226.
40. **Matrosovich, M., T. Matrosovich, W. Garten, and H. D. Klenk**. 2006. New low-viscosity overlay medium for viral plaque assays. Virol. J. **3**:63.
41. **Merino, E. J., K. A. Wilkinson, J. L. Coughlan, and K. M. Weeks**. 2005. RNA structure analysis at single nucleotide resolution by selective 2′-hydroxyl acylation and primer extension (SHAPE). J. Am. Chem. Soc. **127**:4223-4231.
42. **Bonisch, C., C. Temme, B. Moritz, and E. Wahle**. 2007. Degradation of hsp70 and other mRNAs in Drosophila via the 5′ 3′ pathway and its regulation by heat shock. J. Biol. Chem. **282**:21818-21828.
43. **Brierley, I., S. Pennell, and R. J. Gilbert**. 2007. Viral RNA pseudoknots: versatile motifs in gene expression and replication. Nat. Rev. Microbiol. **5**:598-610.
44. **Cornish, P. V., S. N. Stammler, and D. P. Giedroc**. 2006. The global structures of a wild-type and poorly functional plant luteoviral mRNA pseudoknot are essentially identical. RNA. **12**:1959-1969.
45. **Nixon, P. L., P. V. Cornish, S. V. Suram, and D. P. Giedroc**. 2002. Thermodynamic analysis of conserved loop-stem interactions in P1-P2 frameshifting RNA pseudoknots from plant Luteoviridae. Biochemistry **41**:10665-10674.
46. **Leathers, V., R. Tanguay, M. Kobayashi, and D. R. Gallie**. 1993. A phylogenetically conserved sequence within viral 3′ untranslated RNA pseudoknots regulates translation. Mol. Cell Biol. **13**:5331-5347.
47. **Su, L., L. Chen, M. Egli, J. M. Berger, and A. Rich**. 1999. Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. Nat. Struct. Biol. **6**:285-292.
48. **Baranov, P. V., C. M. Henderson, C. B. Anderson, R. F. Gesteland, J. F. Atkins, and M. T. Howard**. 2005. Programmed ribosomal frameshifting in decoding the SARS-CoV genome. Virology **332**:498-510.
49. **Plant, E. P., R. Rakauskaite, D. R. Taylor, and J. D. Dinman**. 2010. Achieving a golden mean: mechanisms by which coronaviruses ensure synthesis of the correct stoichiometric ratios of viral proteins. J. Virol. **84**:4330-4340.
50. **Pennell, S., E. Manktelow, A. Flatt, G. Kelly, S. J. Smerdon, and I. Brierley**. 2008. The stimulatory RNA of the Visna-Maedi retrovirus ribosomal frameshifting signal is an unusual pseudoknot with an interstem element. RNA. **14**:1366-1377.
51. **Bryant, J. E., P. F. Vasconcelos, R. C. Rijnbrand, J. P. Mutebi, S. Higgs, and A. D. Barrett**. 2005. Size heterogeneity in the 3′ noncoding region of South American isolates of yellow fever virus. J. Virol. **79**:3807-3821.
52. **Mutebi, J. P., R. C. Rijnbrand, H. Wang, K. D. Ryman, E. Wang, L. D. Fulop, R. Titball, and A. D. Barrett**. 2004. Genetic relationships and evolution of genotypes of yellow fever virus and other members of the yellow fever virus group within the Flavivirus genus based on the 3′ noncoding region. J. Virol. **78**:9652-9665.
53. **Proutski, V., M. W. Gaunt, E. A. Gould, and E. C. Holmes**. 1997. Secondary structure of the 3′-untranslated region of yellow fever virus: implications for virulence, attenuation and vaccine development. J. Gen. Virol. **78 (Pt 7)**:1543-1549.
54. **Fernandez-Garcia, M. D., M. Mazzon, M. Jacobs, and A. Amara**. 2009. Pathogenesis of flavivirus infections: using and abusing the host cell. Cell Host. Microbe **5**:318-328.

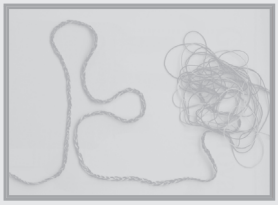
55. **Parker, R. and U. Sheth**. 2007. P bodies and the control of mRNA translation and degradation. Mol. Cell **25**:635-646.
56. **Emara, M. M. and M. A. Brinton**. 2007. Interaction of TIA-1/TIAR with West Nile and dengue virus products in infected cells interferes with stress granule formation and processing body assembly. Proc. Natl. Acad. Sci. U. S. A **104**:9041-9046.
57. **Alvarez, D. E., A. L. De Lella Ezcurra, S. Fucito, and A. V. Gamarnik**. 2005. Role of RNA structures present at the 3'UTR of dengue virus on translation, RNA synthesis, and viral replication. Virology **339**:200-212.
58. **Men, R., M. Bray, D. Clark, R. M. Chanock, and C. J. Lai**. 1996. Dengue type 4 virus mutants containing deletions in the 3' noncoding region of the RNA genome: analysis of growth restriction in cell culture and altered viremia pattern and immunogenicity in rhesus monkeys. J. Virol. **70**:3930-3937.

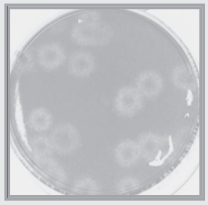
Chapter 4

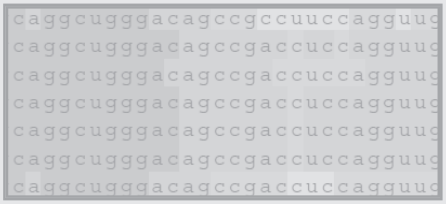
Characterization of the sfRNAs that are produced in cells infected with flaviviruses with no known vector and cell fusing agent

Patrícia A. G. C. Silva
Tim. J. Dalebout
Alexander P. Gultyaev
Rene C. Olsthoorn
Peter J. Bredenbeek

This manuscript has been submitted for publication in "RNA"







ABSTRACT

A virus-specific, non-coding RNA of 0.3 – 0.5 kb, co-linear with the genomic 3' UTR can be detected in cells and mice infected with arthropod-borne flaviviruses. This small flavivirus RNA (sfRNA) results from incomplete degradation of the viral genome by the host 5' – 3' exonuclease XRN1 and was shown to be important for viral pathogenicity. To determine whether sfRNA production is a unique feature of the Flavivirus genus or only restricted to vector-borne flaviviruses, flaviviruses with no known vector (NKV) and the insect flavivirus cell fusing agent virus (CFAV) have been analyzed for the production of the sfRNA. The data presented in this study clearly demonstrate that the XRN1-mediated production of sfRNA is not limited to the vector-borne flaviviruses and most likely is an unique common feature of all flaviviruses, implying that it could be considered an additional determinant to assign viruses to this genus. Computer-aided RNA structure predictions combined with *in vitro* XRN1 assays and cell culture experiments defined an RNA pseudoknot as the XRN1 stalling site for the production of these sfRNAs in NKV flaviviruses and CFAV. These data imply that the sfRNA is likely to be important for the flavivirus life cycle in both the mammalian host and the arthropod vector, as NKV flaviviruses restricted to mammalian hosts and the mosquito-restricted CFAV produce an sfRNA.

INTRODUCTION

The genus Flavivirus of the *Flaviviridae* family contains nearly 80 viruses, including many important human pathogens such as dengue virus (DENV), yellow fever virus (YFV), and West Nile virus (WNV). Based on phylogenetic analysis, flaviviruses were grouped into three major clusters that correlate with the type of vector used for their transmission: (i) mosquito-borne, (ii) tick-borne, and (iii) no known vector (NKV) flaviviruses [1,2]. No arthropod vector has yet been implicated in the transmission of NKV viruses. NKV flaviviruses have been isolated exclusively from rodents or bats and are divided into three groups: i) the Entebbe bat virus group which includes viruses like the Entebbe bat and Yokose virus (YOKV), ii) the Modoc virus group that comprises Modoc virus (MODV) and Apoi virus (APOIV) and iii) the Rio Bravo virus group which includes viruses like Rio Bravo virus (RBV) and Montana myotis leukoencephalitis virus (MMLV) [3]. In contrast to the MODV and RBV groups, whose members are unable to replicate in the mosquito C6/36 cell line, viruses belonging to the Entebbe bat group can replicate in these cells albeit to low titers [4].

Apart from the viruses that are assigned to one of the clusters within the Flavivirus genus, there are viruses like cell fusing agent virus (CFAV) that are considered tentative flaviviruses [5]. CFAV was isolated from a cell line derived from laboratory-reared *Aedes aegypti* mosquitoes [6] and has been classified as a tentative insect flavivirus with genome organization and gene expression strategy similar to that of the flaviviruses. However, CFAV can only be propagated in mosquito cells and not in cell lines of vertebrate origin [6,7]. Although CFAV has never been found in nature, CFAV-related viruses like Kamiti River virus (KRV) have been isolated from field-collected mosquitoes [8-10].

All flaviviruses have a positive single-stranded RNA genome of approximately 11 kb, with a 5′ cap structure and a 3′ non-polyadenylated end. The genome encodes one large open reading frame that is flanked by 5′ and 3′ untranslated regions (UTRs) that contain several conserved RNA sequences and structures that are involved in the regulation of translation and viral genome amplification. Translation of the viral genome results in a polyprotein that is co- and post-translationally processed by viral and cellular proteases into the individual viral proteins [11]. Northern blot analysis of viral RNA isolated from mammalian and insect cell lines or mice infected with arthropod-borne flaviviruses has revealed the production of a small, positive-stranded, non-coding flavivirus RNA (sfRNA) in addition to the viral genome [12-17]. This sfRNA is 0.3 – 0.5 kb long, co-linear with the 3′ end of the viral genome and originates from incomplete degradation of the viral genomic RNA by the host 5′-3′ exonuclease XRN1, due to stalling of this nuclease upstream an RNA pseudoknot located in the viral 3′ UTR [15,17,18].

Although the precise role of the sfRNA in the viral life cycle still needs to be elucidated, current data suggest that it is important for viral pathogenicity in the mammalian host [15,18], (Silva, Pereira, Dalebout, and Bredenbeek, unpublished results). Despite the fact that

sfRNA production has also been described in mosquito cells infected with mosquito-borne flaviviruses [13,15,17], nothing is known about the potential role of the sfRNA in the arthropod host. If production of the sfRNA is only required for efficient completion of the viral life cycle in either the mammalian or the arthropod host, it is not unlikely that the ability to produce an sfRNA might be lacking in either the NKV flaviviruses or CFAV. To address this hypothesis, the production of sfRNA in mammalian cells infected with several NKV flaviviruses and of mosquito cells infected with CFAV was analyzed. Surprisingly, all the flaviviruses that were included in this study produced at least one sfRNA that was co-linear with the 3' end of the viral genome. As has been shown for arthropod-borne flaviviruses, production of sfRNA by these NKV viruses and CFAV is also mediated by the host 5' – 3' exoribonuclease XRN1, which is well conserved in eukaryotes. In addition, the minimal sequence within the viral 3' UTR required for the stalling of XRN1 on the genome of MODV, MMLV and also CFAV was determined and found to form an RNA pseudoknot.

MATERIAL AND METHODS

Cell culture

The origin and culture conditions of the BHK-21J cells have been described before [19,20]. C6/36 cells [21] were obtained from the ATCC and grown in EMEM supplemented with 8% fetal calf serum (Bodinco, The Netherlands) and 5% none-essential amino acids.

Recombinant DNA techniques and plasmid constructions

Unless described in more detail, standard nucleic acid methodologies were used [22,23]. Chemically competent *E. coli* DH5α cells [24] were used for cloning. The nucleotide numbering was according to the sequence files for which the accession numbers can be found in table 1.

Infections were performed essentially as described before [20]. Total RNA was isolated with Trizol (Invitrogen) at 30 hr p.i. from BHK-21J cells infected with MODV, APOIV, MMLV or RBV or at 36 hr p.i. from CFAV infected C6/36 cells. RNA was dissolved in 30 μl H_2O and 5 μg was used for first strand cDNA synthesis using M-MuLV reverse transcriptase (Fermentas). The PCR was performed with GoTaq Flexi DNA Polymerase (Promega) as described by the manufacturer. Oligonucleotides used in the PCR contained either a *Mlu I* site (forward primer) or a *Sph I* site (reversed primer). The RT-PCR products were cloned using the TOPO TA Cloning system (Invitrogen). Inserts with the correct sequence were isolated after digestion of the plasmids with *Mlu I – Sph I* and cloned into Sinrep5eGFP

Table 1.	**Oligonucleotides that were used to identify and characterize the sfRNAs of NKV flaviviruses and CFAV.** The oligonucleotides that were used in this study, the virus to which they were directed, the NCBI accession number used to obtain the sequence and the actual nucleotide sequence are indicated. All oligonucleotides are complementary to the viral genome. Abbreviations in the column "Purpose" refer to Northern blotting and hybridizations (Hyb.), primer extension (Prim.Ex.) and DNA sequencing (Seq.).			
Oligo	**Virus**	**NCBI Number**	**Sequence (5' to 3')**	**Purpose**
NKV2	MMLV	AJ299445	CCGCTCAATCTCGAGAGGAGCGA	Hyb/Prim.Ex.
NKV3	APOIV	AF452050	CTCAGGCGCTAAAGGATGCCGCTA	Hyb.
NKV4	MODV	AJ242984	GGGTCTCCACTAACCTCTAGTCCT	Hyb.
NKV6	APOIV	AF452050	CGCTCAAAGAGAGAAGGGTCGC	Hyb.
NKV19	RBV	AF452049	ACTCGGTCAGTTGGGATCATCCCAC	Hyb.
NKV20	MODV	AJ242984	CCCTAACCTATTTACAATGACTGGC	Hyb./Prim.Ex.
NKV21	CFAV	NC_008604	AGATGGGCCGCCACCACCATCTTAG	Hyb./Prim.Ex.
NKV24	YOKV	NC_005039	TCCATGCGTAGGAGAGGGTCTCC	Hyb.
NKV31	RBV	AF452049	CACCCTATCAGGGTTGACTGGCTCA	Prim.Ex.
NKV33	APOIV	AF452050	CCCACTGGAATGCAATGCTGGCC	Prim.Ex.
YFV1632	YFV-17D	X03700	ACCCCGTCTTTCTACCACC	Seq.
YFV1676	SinV	NC_001547	GTACCAGCCTGATGCATTATGCACATC	Hyb.

[25,17]. Plasmid DNAs of these pSinrep5eGFP recombinants containing either a MODV, MMLV or CFAV insert were linearized with *Not I* and used for *in vitro* RNA transcription [20].

In vitro XRN1 assay

Plasmid DNA of the pSinrep5eGFP recombinants was prepared for *in vitro* RNA transcription without the addition of a cap analog as described above. The RNA transcripts were pre-treated with tobacco acid pyrophosphatase (TAP, Epicentre) to create a 5' mono-phosphate and incubated with 1 unit of XRN1 (available as Terminator 5'-phosphate-dependent exonuclease, Epicentre) as described before [17].

RNA transfection and analysis of viral RNA

BHK-21J cells were transfected with 5 or 20 µg of Sinrep5eGFP and recombinant RNAs as described before [20]. In general, 2.5 ml (approximately 1.5 x 10^6 cells) of the transfected BHK-21J cell suspension was seeded in a 35 mm plate. Total RNA was isolated from the transfected cells at 8 hr post electroporation (p.e.). Trizol (Invitrogen) was used for cell lysis and subsequent RNA purification.

For Northern blotting, samples containing 7.5 – 10 µg of total RNA isolated from either infected or electroporated cells, or obtained from *in vitro* XRN1 assays, were denatured using formaldehyde and separated on a formaldehyde containing 1.5% agarose gel and blotted to Hybond-N$^+$ (GE-Healthcare) [23]. The blots were hybridized with random hexameer primed cDNA probes or ^{32}P-labelled oligonucleotides [26,23,27] that were targeted at the 3' UTR of the virus under study.

Primer extension assay

Primer extension analysis was performed as reported by Sambrook *et al.* [23] with minor modifications [17]. Briefly, 5 – 7 µg total RNA from virus infected cells or XRN1-treated recombinant SINrep5 transcripts, were annealed to a ^{32}P-labeled oligonucleotide that was specific for the studied viral RNA. The primer extension products were analyzed on a denaturing 5% polyacrylamide/8M urea sequence gel. A ^{33}P-labeled Cycle Reader sequence reaction (Fermentas) using oligonucleotide 1632 primed pBlsrcptSK-YFV$_{9845\text{-}10861}$ as a template served as a sequence marker.

RNA structure prediction

RNA structure was predicted as described by Olsthoorn and Bol [28]. The viruses included in this analysis were: MODV, MMLV, RBV, APOIV and CFAV. The NCBI accession numbers for the sequences can be found in table 1.

RESULTS

Production of a small 3' subgenomic RNA is a unique feature of all Flaviviruses

Recently it was shown that many, if not all, of the arthropod-borne flaviviruses produce an sfRNA that is collinear with the distal part of the viral 3' UTR [13,15-17]. These sfRNAs were generated in infected mammalian as well as in insect cells. BHK-21J cells were infected with the NKV flaviviruses MODV, APOIV, RBV, MMLV and YOKV to determine whether such sfRNAs were also produced by the flaviviruses that lack an arthropod vector. Total RNA was isolated from the infected cells at 30 hr p.i. and analyzed for sfRNA production by Northern blot analysis using ^{32}P-labelled oligonucleotides directed against the distal part of the 3' UTR of the NKV viruses as a probe. In addition to the viral genomic RNA,

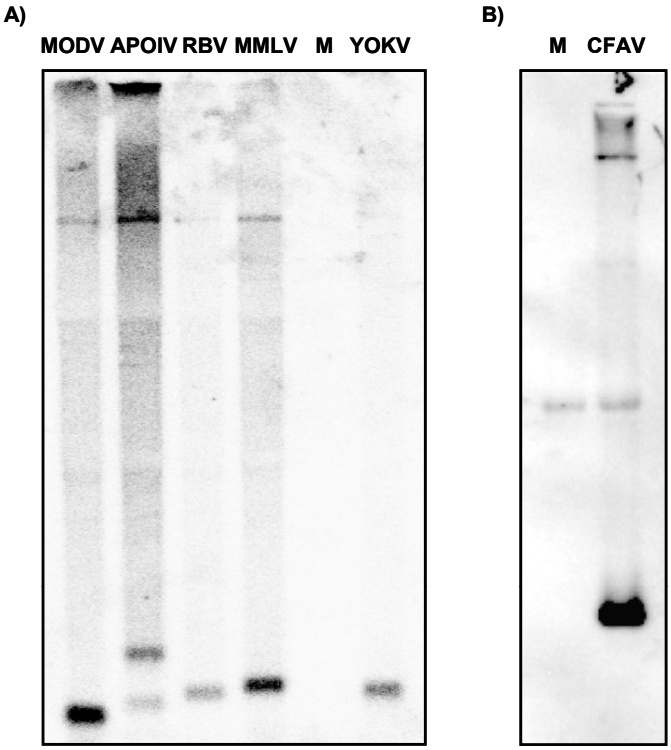


Fig. 1. sfRNA production in mammalian and insect cells infected with NKV flaviviruses and CFAV. A) Northern blot analysis of viral RNAs isolated at 30 hr p.i. from BHK-21J cells infected with MODV, APOIV, RBV, MMLV and YOKV respectively. Kinased oligonucleotides complementary to the distal part of the respective virus 3' UTR were mixed and used as probes. **B)** Northern blot analysis of CFAV RNAs isolated from infected C6/36 cells at 36 hr p.i. A ^{32}P-labelled oligonucleotide complementary to the distal part of the CFAV 3' UTR was used as a probe. Lane M corresponds to total RNA isolated from mock-infected cells.

a small virus-specific RNA was detected for the NKV flaviviruses MODV, RBV, MMLV and YOKV (fig. 1.A; lanes 1, 3, 4 and 6), whereas two small virus-specific RNAs were detected in cells infected with APOIV (fig. 1.A; lane 2).

Apart from these NKV flaviviruses, the tentative insect flavivirus CFAV was also tested for the production of a small virus specific RNA originating from the viral 3' UTR. Mosquito C6/36 cells were infected with CFAV and at 36 hr p.i. total RNA was isolated and analyzed by Northern blotting and hybridization. As shown in fig. 1.B, a small RNA was readily detected in CFAV-infected C6/36 cells. From these results it was concluded that, similar to the arthropod-borne flaviviruses, MODV, APOIV, RBV, MMLV and YOKV, representing the three different groups of NKV flaviviruses, produced at least one sfRNA. These data demonstrated that sfRNA production is a distinguishing feature for all Flaviviruses. Even CFAV, a virus tentatively assigned to the Flavivirus genus, was shown to produce an sfRNA.

Determining the 5′ end of the sfRNAs from flaviviruses with no known vector and CFAV

Primer extension analysis on total RNA isolated from infected cells was used to determine the 5′ end of the sfRNA of MODV, APOIV, MMLV, RBV and CFAV. As shown in fig. 2.A, primer extension on total RNA isolated from MODV-infected BHK-21J cells (lane 1) resulted in the production of two unique cDNA products that were only one nucleotide apart in length and not present in total RNA isolated from uninfected cells (lane 2). Using the sequence ladder that was run in parallel as a marker, the 5′ end of the MODV sfRNA was mapped to nt position 10.262 or 10.263. Based on these results, the MODV sfRNAs

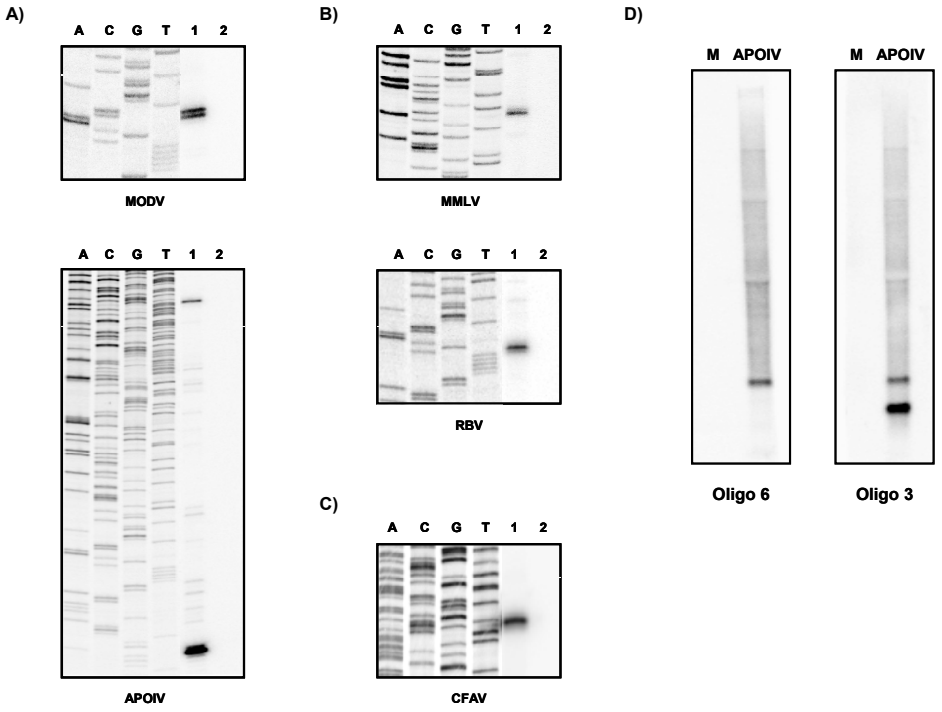


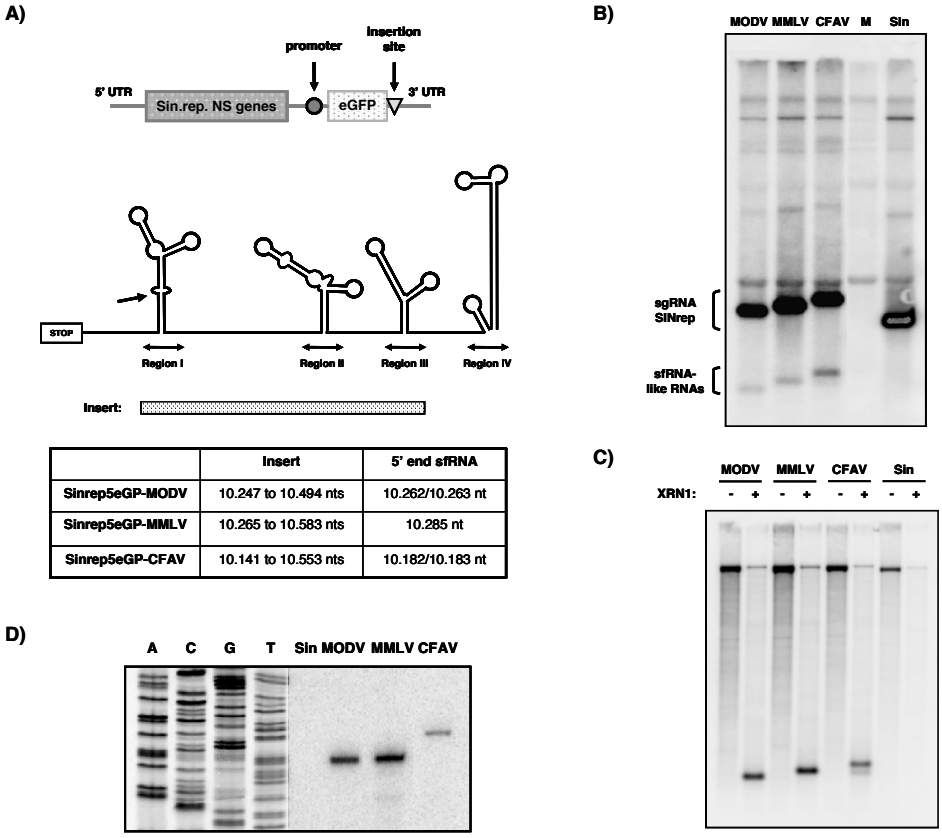
Fig. 2. Mapping the relative position of the NKV and CFAV sfRNAs to the viral genome. Primer extension analysis was performed to determine the 5′ end of the sfRNAs produced by the rodent NKV viruses MODV and APOIV (panel **A**), the bat NKV viruses MMLV and RBV (panel **B**), and the tentative insect flavivirus CFAV (panel **C**). RNA was isolated from infected BHK-21J cells at 30 hr p.i. for the NKV flaviviruses and at 36 hr p.i. from CFAV-infected C6/36 cells. For panels A to C, lanes 1 and 2 correspond to primer extension on total RNA isolated from infected and uninfected cells, respectively. Information on the oligonucleotides that were used as probes for these viruses is presented in table 1. A sequence reaction using oligonucleotide YFV1632 on pBluescript-$YFV_{9,845\text{-}10,861}$ [17] was used as a DNA size marker. **D)** Northern blot analysis of viral RNA isolated from APOIV-infected BHK-21J cells to determine the relative orientation on the viral genome of the two APOIV sfRNAs using oligonucleotides NKV3 and NKV6 (see table 1) as a probe. Lane M corresponds to total RNA isolated from mock-infected cells.

were calculated to be 337 to 338 nts in length. Primer extension on APOIV RNA also resulted in two cDNA products; however, in contrast to MODV, these products showed a significant size difference. This was actually expected given the results of the hybridization presented in fig. 1.A. Based on the length of the primer extension products, the longest sfRNA was calculated to be approximately 566 nts, whereas the smaller sfRNA was predicted to have a length of approximately 371 nts. Primer extension analysis on MMLV, RBV and CFAV RNA resulted in unique products (fig. 2, panels B and C). The 5′ ends were mapped to positions 10.285 in the MMLV genome, 61 nts into the 3′ UTR of RBV, and positions 10.182 – 10.183 in the genome of CFAV. Based on these primer extension results, the sfRNAs of MMLV, RBV and CFAV were calculated to be 405 nt, 425 nt, and 512 – 513 nt, respectively.

The combined results of the Northern blot (fig. 1.A, lane 2) and primer extension (fig. 2.A), suggested that the two detected APOIV sfRNAs would form a 3′ nested set. To determine whether this hypothesis was correct, the position of the APOI sfRNAs relative to the viral genome was analyzed by Northern blotting using oligonucleotides NKV3 and NKV6 as probes. Oligonucleotide NKV3 is complementary to the 3′ end of the viral genome and will recognize both APOIV sfRNAs if they form a 3′ nested set. Oligonucleotide NKV6 hybridizes to a position upstream of the determined 5′ end of the smaller APOIV sfRNA and is predicted to detect only the larger sfRNA if the hypothesis is correct. The results presented in fig. 2.D clearly demonstrated that the APOIV sfRNAs form a 3′ nested set. Both sfRNAs hybridized to the 3′ end-specific oligonucleotide NKV3, whereas only the largest APOIV sfRNA hybridized with oligonucleotide NKV6.

The host exoribonuclease XRN1 is required for sfRNA production of NKV flaviviruses and CFAV

It has now been firmly established that the sfRNAs of WNV and YFV are produced by incomplete degradation of the viral genome by the host 5′ – 3′ exoribonuclease XRN1 [15,17]. To analyze whether XRN1 was also required for the production of the NKV flaviviruses sfRNAs, MODV and MMLV were selected to represent two different groups of NKV flaviviruses that are associated with rodents and bats, respectively. In addition, whenever possible, the tentative flavivirus species CFAV was included in these studies. Unfortunately, no full-length cDNA clone for the transcription of infectious RNA is available for any of these three viruses. To circumvent this handicap, cDNA fragments encompassing the XRN1 stalling site were cloned into the Sindbis virus derived RNA driven expression vector Sinrep5eGFP [25]. This strategy allowed the use of the same RNA templates for both *in vivo* and *in vitro* studies. A schematic representation and relevant details of these constructs is shown in fig. 3.A. BHK-21J cells were transfected with *in*



	Insert	5' end sfRNA
Sinrep5eGP-MODV	10.247 to 10.494 nts	10.262/10.263 nt
Sinrep5eGP-MMLV	10.265 to 10.583 nts	10.285 nt
Sinrep5eGP-CFAV	10.141 to 10.553 nts	10.182/10.183 nt

Fig. 3. Insertion of the MODV, MMLV and CFAV 3' UTR sequence directs the *in vivo* and *in vitro* production of an sfRNA-like RNA in the context of a Sindbis replicon RNA. A) Schematic representation of the pSinrep5eGFP vector characteristics and the predicted general RNA folding of the NKV 3' UTR according to Charlier and colleagues [29]. The 5' end of the sfRNAs relative to this RNA structure is indicated by an arrow. The box below the structure indicates the region of the MODV, MMLV or CFAV 3' UTR cloned into the Sinrep5eGFP vector. The names of the constructs and the exact nucleotide numbers of the cDNA fragments cloned into Sinrep5eGFP are indicated in the table in panel A. **B)** Northern blot analysis of total RNA isolated from BHK-21J cells transfected with Sinrep5eGFP-MODV (lane 1), Sinrep5eGFP-MMLV (lane 2), Sinrep5eGFP-CFAV (lane 3), mock transfected cells (lane 4) and Sinrep5eGFP (lane 5). Oligonucleotide 1676 complementary to Sinrep5eGFP nucleotides downstream of the insertion site of flavivirus 3' UTR was used as probe. **C)** *In vitro* production of the MODV, MMLV and CFAV sfRNA-like RNA by incubation with XRN1. The "–" symbol refers to incubation without the enzyme, while the "+" signal corresponds to the addition of XRN1. **D)** Primer extension using oligonucleotide 1676 to determine the 5' end of the sfRNA-like RNAs produced by *in vitro* incubation of RNA transcribed from the Sinrep5eGFP constructs with XRN1. "Sin" represents the control reaction on pSinrep5eGFP.

vitro transcribed RNA of these Sinrep constructs to demonstrate that an sfRNA-like RNA could be produced in cell culture. As shown in fig. 3.B, Northern blotting of total RNA isolated from these transfected cells using oligonucleotide 1676 as a probe revealed that Sinrep5eGFP-MODV, Sinrep5eGFP-MMLV and Sinrep5eGFP-CFAV produced an sfRNA-like RNA. This additional RNA was not detected in cells that were transfected with Sinrep5eGFP that did not contain any flavivirus sequences (lane Sin) or mock transfected cells (lane M). RNA transcripts of Sinrep5eGFP, Sinrep5eGFP-MODV, Sinrep5eGFP-MMLV and Sinrep5eGFP-CFAV, were incubated *in vitro* with commercially available purified XRN1 and analysed by Northern blotting to demonstrate that this enzyme was also responsible for the production of the NKV and CFAV sfRNAs. The results showed that an sfRNA-like RNA was generated upon incubation with XRN1 of the Sinrep5eGFP transcripts that contain either the MODV, MMLV or CFAV insert encompassing the predicted XRN1 stalling site (fig. 3.C). Primer extension analysis on these sfRNA-like RNAs produced *in vitro* demonstrated that the 5' ends that were produced using this heterologous expression system were identical to the 5' ends of the sfRNAs as detected in BHK-21J cells infected with either MODV, MMLV or CFAV (fig. 3.D). The combined data of the *in vivo* and *in vitro* experiments using the Sinrep5eGFP templates provided strong evidence that XRN1 is also responsible for the production of the sfRNA in cells infected with NKV flaviviruses or CFAV.

An RNA pseudoknot in the 3' UTR of NKV flaviviruses and CFAV is required for the production of the sfRNAs

Current data for YFV and WNV strongly suggest that the stalling of XRN1 that is required for the production of sfRNA in arthropod-borne flaviviruses is directed by an RNA pseudoknot within the viral 3' UTR [17,18]. Therefore, the nucleotide sequences immediately downstream of the predicted 5' end of the NKV flaviviruses and CFAV sfRNAs were analyzed for their potential to fold into an RNA pseudoknot structure. An alignment of the primary sequence in this region of the NKV flaviviruses genomes showed significant sequence similarity interspaced by insertions or deletions of a few nucleotides (fig. 4.A). A similar result was obtained when CFAV was compared to KRV and Aedes virus. The latter two viruses were recently isolated from mosquitoes and shown to be closely related to CFAV. Subsequent extensive RNA structure modelling predicted the formation of RNA pseudoknots for all the NKV flaviviruses as well as for CFAV and the related KRV and Aedes virus. The predicted RNA pseudoknot structures for MODV, MMLV and CFAV that could potentially serve as stalling sites for XRN1 are depicted in fig. 4.B. The structure for MODV and MMLV was predicted to be very similar. Compared to the structure for YFV [17], the sequence indicated by pk that was predicted to base pair with the pk' sequence

A)

```
MODV:   5' ugaaagagugu-u-gaggg-caaccaguggggc-uagccac--------augggu-----augacgcaccc-acccucu-gca-uucuuguaaa-uacuuug-gccagucauugu----- 3'
MMLV:   5' aaaaaaugcgagu-gaggg-caac-ucugggu-uagcuca--------augggu-----gugacg-acccuacccuuccgca-uuu--guaaa-uaa-uugagccagucauuuc----- 3'
RBV:    5' -----uucggugc-u-gaggggcaac-ucuaggacuagccua-------augggu-----gugacg-acccuaccccaaugcacuu---guaaa-uaa-uugagccagucaacc------- 3'
APOIV:  5' ---aauaagcgccuggagag-cgac-cuu-ggacuguggggguccaaagggcuugcgacccuucucucuuugagcgcuuugggacaacuauuug-gccag-cauugcauuc 3'

CFAV:   5' ccacaggagcagggcaugaaaaugucgggcaugacgaacccgcucccccgaguc-cccug-gcaa-caggg- 3'  (1st putative stalling signal)
        5' ---guucgagcagggcacauuggugucgggcgugacgcacccgcucccuc-aguc-cccuguguaa-caggg- 3'  (2nd putative stalling signal)
KRV:    5' --------agagcagggcacgaaagugucgggcaugacgcacccgcucccccgaguc-cccug-aaaa-uaggg- 3'
Aedes:  5' --------agagcagggcacaauagugucgggccugacgacccgcucccccgagucgcccaacggaguuuggc 3'
```

B)

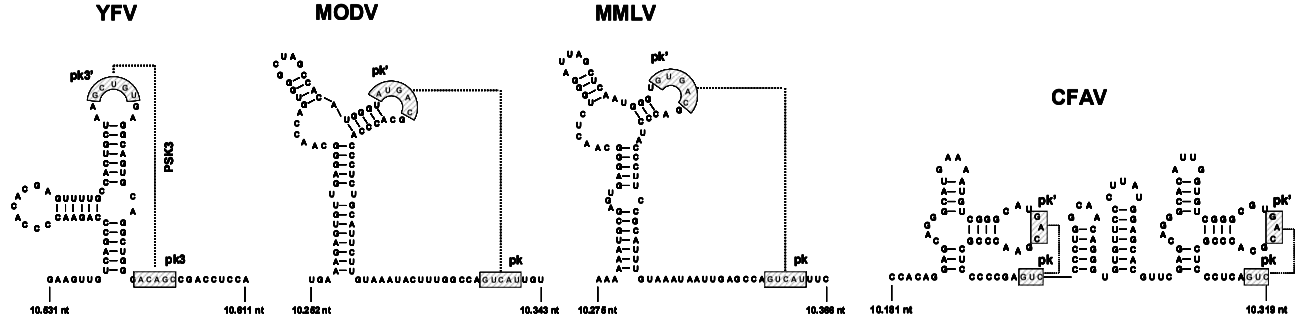
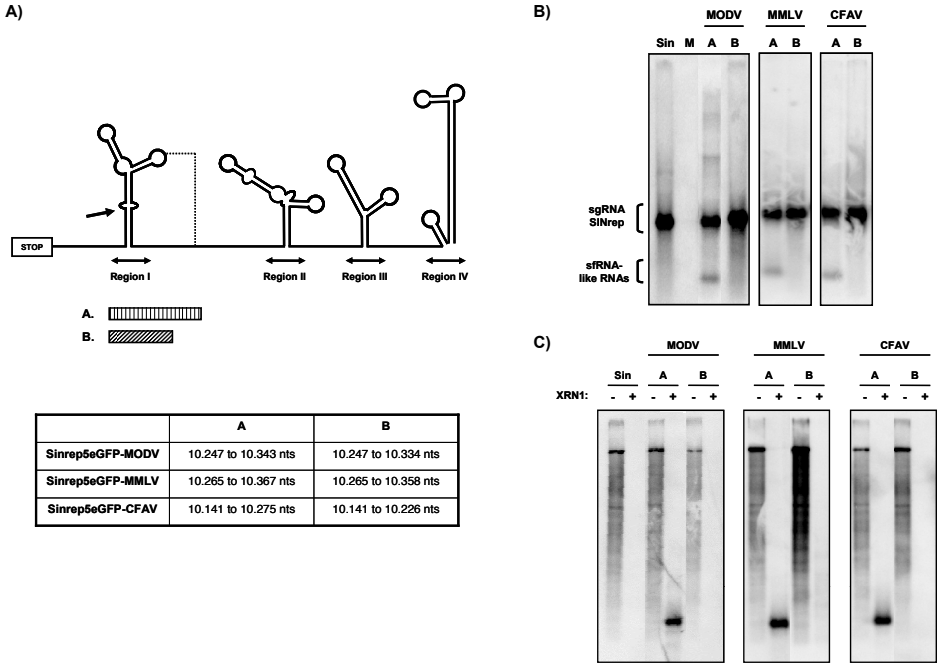


Fig. 4. Sequence alignment and predicted RNA structures of the region that is involved in stalling XRN1 in MODV, MMLV and CFAV. A) Alignment of the sequences downstream the 5' end of the sfRNAs of the NKV flaviviruses MODV, MMLV, RBV and APOIV (top panel) and the tentative flaviviruses CFAV, KRV and Aedes virus (bottom panel). The second predicted stalling signal for XRN1 in CFAV is also included. Base paired nucleotides are underlined. Conserved nucleotides are indicated in bold. Putative pk and pk' sequences are in italics and indicated by boxes. Gaps were introduced to maximize the nucleotide identity. The NCBI accession numbers for KRV and Aedes virus are NC_012671 and NC_012932, respectively. Accession numbers for the other viruses can be found in table 1. **B)** RNA structures were predicted using a combinatorial approach involving Mfold, phylogeny and manual sequence-structure analysis. Potential RNA pseudoknot interactions between the pk and pk' sequences [28] are indicated by grey boxes. For comparison, the YFV-17D RNA pseudoknot that has been shown to stall XRN1 for the production of the YFV sfRNA is depicted [17].

(nomenclature according to Olsthoorn and Bol [28]), is located further downstream of the stem. For both MODV and MMLV, the presumed pk' – pk interaction comprised five nucleotides, which is similar to the YFV pseudoknot that was proven to act as a stalling signal for XRN1 [17]. The initial impression of the folding for the comparable region of the CFAV RNA suggested that it was rather different from the proposed RNA pseudoknots for YFV, MODV and MMLV. However, closer inspection revealed that the overall folding was actually relatively similar to the proposed structures for NKV flaviviruses and YFV. The main differences of the CFAV structure versus that of the other viruses were in the length of the depicted stem structures and the pk' – pk interaction, which was proposed to involve only three nucleotides. Interestingly, the formation of a second, very similar RNA pseudoknot structure was predicted for the nucleotide sequence just downstream of the first predicted pseudoknot (fig. 4.B).

To provide support for the actual formation of these predicted RNA pseudoknots, two cDNA fragments of MODV, MMLV and CFAV were cloned into Sinrep5eGFP (fig. 5.A). The cDNA fragment A of MODV, MMLV and CFAV contained all the nucleotides that were predicted to be involved in the formation of the RNA pseudoknot required to stall XRN1. Fragment B contained a 3' deletion compared to the sequences contained in fragment A. This 3' truncation was expected to disrupt the pseudoknot formation by deletion of the pk sequence. The ability of the cDNA fragments A and B of MODV, MMLV and CFAV to direct the production of an sfRNA-like RNA was analyzed *in vivo* as well as *in vitro* using the Sinrep5eGFP based expression system. RNA transcripts of the NKV virus and CFAV Sinrep5eGFP constructs containing cDNA fragment A or B were electroporated into BHK-21J cells. At 8 hrs. p.e., total RNA was isolated and analyzed by Northern blotting using the Sindbis virus 3' UTR specific oligonucleotide 1676 as a probe. As shown in fig. 5.B, all the cells transfected with Sinrep5eGFP RNA containing cDNA fragment A of MODV, MMLV or CFAV were able to produce an sfRNA-like RNA in the transfected cells. As expected, this sfRNA-like RNA was not detected in BHK cells that were either mock transfected (lane M) or electroporated with Sinrep5eGFP RNA lacking any flavivirus insert (lane Sin). More importantly, this sfRNA-like RNA was not detected in cells that were transfected with RNA of the Sinrep5eGFP recombinants containing cDNA fragment B of MODV, MMLV, and CFAV.

In vitro experiments were performed with purified XRN1 and RNA transcribed from the Sinrep5eGFP recombinants containing the MODV, MMLV or CFAV cDNA fragments A and B to provide additional evidence for the role of the predicted RNA pseudoknot in the stalling of XRN1. Sinrep5eGFP RNA containing the NKV flavivirus or CFAV cDNA fragment A, did produce an sfRNA-like RNA after incubation with XRN1, whereas Sinrep5eGFP constructs encompassing cDNA fragment B were completely degraded and did not yield this additional subgenomic RNA (fig. 5.C). The outcome of the *in vivo* and in *vitro* experiments presented in fig. 5 supports the predicted model in which an RNA



	A	B
Sinrep5eGFP-MODV	10.247 to 10.343 nts	10.247 to 10.334 nts
Sinrep5eGFP-MMLV	10.265 to 10.367 nts	10.265 to 10.358 nts
Sinrep5eGFP-CFAV	10.141 to 10.275 nts	10.141 to 10.226 nts

Fig. 5. RNA sequences that are predicted to be involved in the formation of an RNA pseudoknot in the 3′ UTR of MODV, MMLV and CFAV are required for the production of an sfRNA. A) Generalized overall RNA folding of the NKV 3′ UTR based on Charlier and colleagues [29]. The relative position of the RNA pseudoknot that is predicted to be involved in sfRNA production is indicated by a dashed line. The 5′ end of the sfRNAs relative to this RNA structure is indicated with an arrow. The boxes below the structure indicate the regions of the flavivirus 3′ UTR that were cloned into the Sinrep5eGFP vector from MODV, MMLV and CFAV. The exact nucleotide numbers of the cDNA fragments cloned into Sinrep5eGFP are indicated in the table. **B)** Northern blot analysis of RNA isolated from BHK-21J cells transfected with Sinrep5eGFP (Sin) and Sinrep5eGFP with the 3′ UTR A and B inserts of MODV, MMLV and CFAV, respectively. M corresponds to uninfected BHK-21J cells. Oligonucleotide 1676 was used as probe. **C)** Northern blot analysis of RNA transcripts of Sinrep5eGFP and derivatives containing fragment A or B of MODV, MMLV or CFAV incubated with XRN1. The "–" symbol refers to incubation without the enzyme, while the "+" signal corresponds to the addition of XRN1. Oligonucleotide 1676 was used as a probe.

pseudoknot is required to stall XRN1 for the production of an sfRNA-like RNA. In contrast to the Sinrep5eGFP recombinants containing MODV, MMLV or CFAV cDNA fragment A, formation of the RNA pseudoknot that stalls XRN1 was no longer possible in fragment B due to deletion of the pk sequence, explaining why an sfRNA-like RNA was not produced.

DISCUSSION

The family *Flaviviridae* comprises three genera: Flaviviruses, Pestiviruses, and Hepaciviruses [5]. All the members of this virus family are enveloped, positive-stranded RNA viruses with a similar genome organization and expression strategy. Despite their evolutionary relatedness, these viruses show significant differences in host range, tropism, pathogenicity, and various aspects of their molecular biology. Over the last few years it has been shown that arthropod-borne Flaviviruses produce a small flavivirus RNA (sfRNA) in addition to their genome-length negative- and positive-stranded RNAs [12-17]. This RNA results from partial 5′ to 3′ degradation of the viral genome by the host exoribonuclease XRN1 [15,17]. No such small viral subgenomic RNA has been detected in cells infected with Pestiviruses and Hepaciviruses [15], suggesting that the production of the sfRNA might be a new feature specific for Flaviviruses. However, flaviviruses are divided into three groups and the previous studies only analyzed sfRNA production in cells and mice infected with either mosquito- or tick-borne viruses. In this study, viruses that belong to the third, poorly studied cluster of Flaviviruses with no known vector, were analyzed for their ability to direct sfRNA synthesis. The analysis included MODV and APOIV from the rodent-associated MODV-related flaviviruses and MMLV and RBV, along with YOKV, representing the bat-associated RBV and Entebbe bat virus subgroup of the NKV flaviviruses, respectively [3]. In addition to these NKV flaviviruses, the mosquito cell-infecting virus CFAV, which has been tentatively assigned to the Flavivirus genus, was also included in this study. The results of this analysis showed that all NKV flaviviruses as well as CFAV are able to produce at least one sfRNA in infected cells. APOIV even generates two sfRNAs that form a 3′ nested set. Primer extension analysis on total RNA mapped the 5′ end of the NKV sfRNAs to an internal bulge in the Y-structure that is indicated as region I in the 3′ UTR of NKV flaviviruses [29]. Due to recombination, deletions and sequence duplications, the length of the Flavivirus 3′ UTR is rather heterogeneous (reviewed in [30]). This size heterogeneity is also reflected in the length of the sfRNAs. In general, the sfRNAs are co-linear with the 3′ distal part of the viral 3′ UTR and contain all the conserved RNA elements that are present in this region of the genome. These results demonstrate that production of a subgenomic sfRNA is a new feature of viruses that belong to the genus Flavivirus and that it can be considered as an additional characteristic to the established ICTV criteria for classification of newly discovered viruses into the genus Flavivirus.

Using mutagenesis of the available infectious cDNA clones, *in vitro* assays and RNA silencing experiments, previous studies on WNV and YFV have demonstrated that the sfRNAs of the arthropod-borne flaviviruses are produced by incomplete degradation of the viral genomic RNA by the host 5′-3′ exoribonuclease XRN1 [15,17]. No infectious cDNA clone is currently available for any of the NKV flaviviruses or CFAV. Therefore, a

Sindbis virus-based expression system [25] was used to determine whether XRN1 was also required for the production of sfRNA in cells infected with NKV flaviviruses or CFAV. The cDNA fragments that contained the nearly complete distal part of the MODV, MMLV or CFAV 3' UTR were cloned into the Sinrep5-eGFP and analyzed for their ability to direct the synthesis of an sfRNA-like RNA *in vitro* and *in vivo*. Transfection of BHK-21J cells with RNA transcribed from these constructs did indeed result in the production of an sfRNA-like RNA. An identical RNA was detected upon incubation of the Sinrep5eGFP recombinants containing the MODV, MMLV or CFAV insert with purified XRN1. Furthermore, the *in vitro* produced sfRNA-like RNAs have an identical 5' end as the sfRNAs that are produced during infection of cell cultures with MODV, MMLV or CFAV. These results strongly suggest that XRN1 is also required for the production of the sfRNA in NKV flaviviruses and CFAV. These results are actually not very surprising since XRN1 is well conserved among eukaryotes as it plays a vital role in 5' to 3' mRNA decay (reviewed in [31-34]), and therefore in the homeostasis of the host. Several experiments have been performed to provide additional proof for the role of XRN1 in the sfRNA production of MODV and MMLV by using lentiviruses expressing small-hairpin RNAs to silence human XRN1 expression. However, silencing of XRN1 has a significant impact on the condition of the cells and this, together with the relatively poor replication of MODV and MMLV in the tested SW13 and Huh7 cells, unfortunately resulted in inconclusive data (Silva, Dalebout and Bredenbeek, unpublished results).

It has previously been shown that XRN1 is stalled by an RNA pseudoknot in the 3' UTR of YFV and WNV to produce the sfRNA [17,18]. However, no RNA pseudoknot structure has been suggested in the 3' UTR region predicted to stall XRN1 in either the NKV flaviviruses or CFAV [29,7]. Sequence alignment combined with RNA structure modelling have been used in this study to predict an RNA pseudoknot just downstream of the mapped 5' end of the sfRNA for every NKV virus, similar to what has been encountered in WNV and YFV. Unfortunately, the limited amount of sequence data available for these NKV flaviviruses does not allow for co-variance analysis to obtain additional support for the predicted structures. An RNA pseudoknot that can serve as a stalling site for XRN1 has also been predicted for CFAV. On first sight, this RNA structure may seem rather unlikely, due to the relatively short stem-loop structures and a pseudoknot interaction that only involves three nucleotides. A nearly perfect duplication of this pseudoknot was predicted downstream of the first pseudoknot (fig. 4, panels A and B). Careful inspection of the data presented in fig. 3, where the stalling of XRN1 was analyzed in the background of the Sindbis expression system, reveals a minor, slightly smaller sfRNA-like RNA that can be explained by the XRN1 stalling on the second, more downstream located pseudoknot. This minor band is no longer detected in the data presented in fig. 5, because the sequence involved in the formation of this second pseudoknot is not present in the CFAV inserts of those particular Sindbis constructs. So these experiments actually

provide support for the proposed CFAV pseudoknot. The sequence of the CFAV, KRV and Aedes virus 3' UTR in the region in which the pseudoknot structure is predicted is nearly identical (fig. 4.A); therefore any co-variance in the primary sequence to support this structure is very limited.

The sequence that encompasses the 5' end of the sfRNA and the predicted RNA pseudoknots of MODV, MMLV and CFAV was cloned into the Sinrep5eGFP expression vector and shown to be capable of producing an sfRNA-like RNA. These results demonstrate that the ability for stalling XRN1 is contained within a relatively small sequence of MODV, MMLV, and CFAV and that no other sequences are required to produce the sfRNA. Deletion of the pk sequence, which is predicted to interact with the pk' sequence to form the RNA pseudoknot, from the flavivirus insert in the Sinrep5eGFP constructs abolished the production of the sfRNA-like RNA. This result supports the model in which, similar to what was found in arthropod-borne flaviviruses, RNA pseudoknots are required to stall XRN1. In addition, N-methylisatoic anhydride (NMIA) based chemical probing of the RNA structure required for stalling XRN1 showed that the pk' and pk sequences were in a double-stranded conformation, which is in agreement with the predicted RNA pseudoknots. The RNA probing results of the sequences that encompass the XRN1 stalling sites of MODV, MMLV, and CFAV, however, did not fully support the stem-loop part of the predicted pseudoknots, suggesting that a more complex structure might actually be present (Silva, Dalebout and Bredenbeek, unpublished results). Unfortunately, NMIA-based probing only discriminates between single- and double-stranded nucleotides and does not allow identification of individual nucleotides within a structure.

Although the precise function of the sfRNA in the virus life cycle remains to be elucidated, current evidence for WNV and YFV have shown that it is an important determinant for the pathogenicity of these viruses in their mammalian hosts [15,18] (Silva, Pereira, Dalebout, and Bredenbeek, unpublished results). Nothing is known about the requirement of the sfRNA of these viruses in the arthropod host. The fact that the ability to produce the sfRNA is maintained in flaviviruses that have a more limited host range, like the NKV flaviviruses and CFAV, strongly suggest that the sfRNA is essential for successful survival of these viruses in the mammalian or the insect host. However, this does not necessarily imply that the function of the NKV virus or CFAV sfRNA is similar to that of the sfRNA of the arthropod-borne viruses in either their vertebrate or arthropod host. Determining the function of these Flavivirus sfRNAs in the various virus-host systems and unravelling the link between the sfRNA and viral pathogenicity will be an interesting challenge for further research.

REFERENCE LIST

1. **Kuno, G., G. J. Chang, K. R. Tsuchiya, N. Karabatsos, and C. B. Cropp**. 1998. Phylogeny of the genus Flavivirus. J. Virol. **72**:73-83.
2. **Cook, S. and E. C. Holmes**. 2006. A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: Flaviviridae) and the evolution of vector transmission. Arch. Virol. **151**:309-325.
3. **ICTV**. 2005. Virus Taxonomy - Eight Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego.
4. **Varelas-Wesley, I. and C. H. Calisher**. 1982. Antigenic relationships of flaviviruses with undetermined arthropod-borne status. Am. J. Trop. Med. Hyg. **31**:1273-1284.
5. **Thiel, H. J., M. S. Collet, E. A. Gould, F. X. Heinz, G. Meyers, R. H. Purcell, C. M. Rice, and M. Houghton**. 2005. Flaviridae, p. 981-998. *In* L. A. Ball (ed.), Virus Taxonomy - Eight Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego.
6. **Stollar, V. and V. L. Thomas**. 1975. An agent in the Aedes aegypti cell line (Peleg) which causes fusion of Aedes albopictus cells. Virology **64**:367-377.
7. **Cammisa-Parks, H., L. A. Cisar, A. Kane, and V. Stollar**. 1992. The complete nucleotide sequence of cell fusing agent (CFA): homology between the nonstructural proteins encoded by CFA and the nonstructural proteins encoded by arthropod-borne flaviviruses. Virology **189**:511-524.
8. **Cook, S., S. N. Bennett, E. C. Holmes, C. R. de, G. Moureau, and L. de, X**. 2006. Isolation of a new strain of the flavivirus cell fusing agent virus in a natural mosquito population from Puerto Rico. J. Gen. Virol. **87**:735-748.
9. **Sang, R. C., A. Gichogo, J. Gachoya, M. D. Dunster, V. Ofula, A. R. Hunt, M. B. Crabtree, B. R. Miller, and L. M. Dunster**. 2003. Isolation of a new flavivirus related to cell fusing agent virus (CFAV) from field-collected flood-water Aedes mosquitoes sampled from a dambo in central Kenya. Arch. Virol. **148**:1085-1093.
10. **Crabtree, M. B., R. C. Sang, V. Stollar, L. M. Dunster, and B. R. Miller**. 2003. Genetic and phenotypic characterization of the newly described insect flavivirus, Kamiti River virus. Arch. Virol. **148**:1095-1118.
11. **Lindenbach, B. D., H. J. Thiel, and C. M. Rice**. 2007. Flaviviridae: The Viruses and Their Replication *In* D. M. Knipe and P. M. Howley (eds.), Fields Virology. Lippincott Williams & Wilkins, Philadelphia.
12. **Urosevic, N., M. M. van, J. P. Mansfield, J. S. Mackenzie, and G. R. Shellam**. 1997. Molecular characterization of virus-specific RNA produced in the brains of flavivirus-susceptible and -resistant mice after challenge with Murray Valley encephalitis virus. J. Gen. Virol. **78 (Pt 1)**:23-29.
13. **Lin, K. C., H. L. Chang, and R. Y. Chang**. 2004. Accumulation of a 3'-terminal genome fragment in Japanese encephalitis virus-infected mammalian and mosquito cells. J. Virol. **78**:5133-5138.
14. **Scherbik, S. V., J. M. Paranjape, B. M. Stockman, R. H. Silverman, and M. A. Brinton**. 2006. RNase L plays a role in the antiviral response to West Nile virus. J. Virol. **80**:2987-2999.
15. **Pijlman, G. P., A. Funk, N. Kondratieva, J. Leung, S. Torres, L. van der Aa, W. J. Liu, A. C. Palmenberg, P. Y. Shi, R. A. Hall, and A. A. Khromykh**. 2008. A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity. Cell Host. Microbe **4**:579-591.
16. **Liu, R., L. Yue, X. Li, X. Yu, H. Zhao, Z. Jiang, E. Qin, and C. Qin**. 2010. Identification and characterization of small sub-genomic RNAs in dengue 1-4 virus-infected cell cultures and tissues. Biochem. Biophys. Res. Commun. **391**:1099-1103.

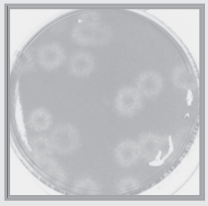
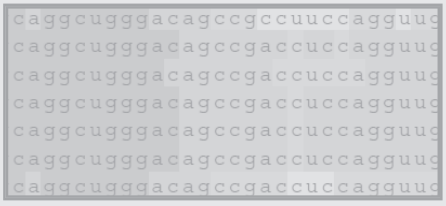
17. **Silva, P. A., C. F. Pereira, T. J. Dalebout, W. J. Spaan, and P. J. Bredenbeek**. 2010. An RNA Pseudoknot Is Required for Production of Yellow Fever Virus Subgenomic RNA by the Host Nuclease XRN1. J. Virol. **84**:11395-11406.
18. **Funk, A., K. Truong, T. Nagasaki, S. Torres, N. Floden, M. E. Balmori, J. Edmonds, H. Dong, P. Y. Shi, and A. A. Khromykh**. 2010. RNA structures required for production of subgenomic flavivirus RNA. J. Virol. **84**:11407-11417.
19. **Bredenbeek, P. J., E. A. Kooi, B. Lindenbach, N. Huijkman, C. M. Rice, and W. J. Spaan**. 2003. A stable full-length yellow fever virus cDNA clone and the role of conserved RNA elements in flavivirus replication. J. Gen. Virol. **84**:1261-1268.
20. **Silva, P. A., R. Molenkamp, T. J. Dalebout, N. Charlier, J. H. Neyts, W. J. Spaan, and P. J. Bredenbeek**. 2007. Conservation of the pentanucleotide motif at the top of the yellow fever virus 17D 3′ stem-loop structure is not required for replication. J. Gen. Virol. **88**:1738-1747.
21. **Igarashi, A.** 1978. Isolation of a Singh's Aedes albopictus cell clone sensitive to Dengue and Chikungunya viruses. J. Gen. Virol. **40**:531-544.
22. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl**. 2000. Current Protocols in Molecular Biology. Wiley Interscience, New York.
23. **Sambrook, J., T. Fritsch, and T. Maniatis**. 1989. Molecular Cloning: a Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
24. **Inoue, H., H. Nojima, and H. Okayama**. 1990. High efficiency transformation of Escherichia coli with plasmids. Gene **96**:23-28.
25. **Bredenbeek, P. J., I. Frolov, C. M. Rice, and S. Schlesinger**. 1993. Sindbis virus expression vectors: packaging of RNA replicons by using defective helper RNAs. J. Virol. **67**:6439-6446.
26. **Meinkoth, J. and G. Wahl**. 1984. Hybridization of nucleic acids immobilized on solid supports. Anal. Biochem. **138**:267-284.
27. **van der Most, R. G., P. J. Bredenbeek, and W. J. Spaan**. 1991. A domain at the 3′ end of the polymerase gene is essential for encapsidation of coronavirus defective interfering RNAs. J. Virol. **65**:3219-3226.
28. **Olsthoorn, R. C. and J. F. Bol**. 2001. Sequence comparison and secondary structure analysis of the 3′ noncoding region of flavivirus genomes reveals multiple pseudoknots. RNA. **7**:1370-1377.
29. **Charlier, N., P. Leyssen, C. W. Pleij, P. Lemey, F. Billoir, L. K. Van, A. M. Vandamme, C. E. De, L. de, X, and J. Neyts**. 2002. Complete genome sequence of Montana Myotis leukoencephalitis virus, phylogenetic analysis and comparative study of the 3′ untranslated region of flaviviruses with no known vector. J. Gen. Virol. **83**:1875-1885.
30. **Gritsun, T. S. and E. A. Gould**. 2007. Origin and evolution of 3′UTR of flaviviruses: long direct repeats as a basis for the formation of secondary structures and their significance for virus transmission. Adv. Virus Res. **69**:203-248.
31. **Sheth, U. and R. Parker**. 2003. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science **300**:805-808.
32. **Anderson, P. and N. Kedersha**. 2006. RNA granules. J. Cell Biol. **172**:803-808.
33. **Eulalio, A., I. Behm-Ansmant, and E. Izaurralde**. 2007. P bodies: at the crossroads of post-transcriptional pathways. Nat. Rev. Mol. Cell Biol. **8**:9-22.
34. **Garneau, N. L., J. Wilusz, and C. J. Wilusz**. 2007. The highways and byways of mRNA decay. Nat. Rev. Mol. Cell Biol. **8**:113-126.

Chapter 5

Characterization of a stable full-length cDNA clone for the transcription of infectious RNA of a Flavivirus with no known vector

Patrícia A. G. C. Silva[#]
Xiaohong Jiang[#]
Tim. J. Dalebout
Charles M. Rice
Peter J. Bredenbeek
[#]These authors contributed equally

The work described in this chapter is part of a manuscript entitled "An infectious Modoc Virus cDNA as a tool to study RNA replication signals in flaviviruses with no known vector"

ABSTRACT

The *Flavivirus* genus can be divided into three different groups depending on the vector of transmission: i) mosquito-borne, ii) tick-borne, and iii) no known vector (NKV) flaviviruses. The third group is less-well studied, which in part is due to the lack of full-length cDNA clones that can be used for the transcription of infectious RNA. In this chapter the construction and characterization of a stable infectious full-length cDNA clone of the NKV Modoc virus (MODV) is described. The full-length MODV cDNA was constructed in the low copy number plasmid pACNR1180. An initial screening of plasmids containing full-length MODV cDNAs for the production of infectious RNA transcripts resulted in four plasmids whose transcripts yielded immunofluorescent positive cells. Subsequent analysis revealed that only one of these clones could be used for the production of transcripts with a high enough specific infectivity to allow semi-first cycle analysis of virus replication. This clone (pACNR-FLMODV6.1) was shown to be genetically stable in *E. coli*. The viruses derived from this clone showed similar plaque morphology and growth kinetics as the parental MODV virus. This first infectious cDNA clone for a NKV flavivirus will be a valuable tool to increase our understanding of general and virus-specific characteristics in flavivirus biology.

INTRODUCTION

Phylogenetic analysis of members of the Flavivirus genus revealed three clusters of related viruses that largely coincided with the route of transmission. Apart from the clusters of mosquito- and tick-borne flaviviruses, which contain important human pathogens, the third cluster comprises a less-well studied group of viruses that have been exclusively isolated from bats and rodents and for which no arthropod vector has been implicated in transmission [1,2]. An increasing number of these no known vector (NKV) flaviviruses has been sequenced [3-5], revealing a similar organization of the approximately 10.5 kb long NKV genome as that of the arthropod-borne flaviviruses. The viral genome encompasses one large open reading frame that encodes the following proteins: (5') C, prM, E, NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5 (3'). Similar to other flaviviruses, protease and NTPase/helicase/RNA triphosphatase domains were found in the NS3 N-terminal and C-terminal regions, respectively, of Modoc virus (MODV) [5]. Furthermore, the N-terminal region of the MODV NS5 protein was shown to encode motifs important for methyltransferase activity, while the C-terminal region encoded highly conserved RNA-dependent RNA polymerase (RdRp) domains [5]. The crystal structure of the methyltransferase domain of MODV NS5 was recently determined [6].

The genomic RNA is not polyadenylated; instead it terminates with a large 3' stem-loop structure (SL) that is characteristic for all flaviviruses. Apart from this 3' SL structure, the 5' and 3' UTRs of NKV viruses contain RNA elements that have also been identified and characterized in the UTRs of arthropod-borne flaviviruses, like the conserved sequence (CS) 2 and complementary sequences predicted to be involved in circularization of the viral genome [4,5]. Recently, it has been shown that cells infected with NKV flaviviruses also produce a small subgenomic flavivirus RNA (sfRNA) [7] as was initially reported for arthropod-borne flavivirus infections [8-13]. Synthesis of this sfRNA results from incomplete degradation of the genomic RNA by the host ribonuclease XRN1 [10,12,7].

NKV flaviviruses are divided into three groups: i) the Entebbe bat virus group, which includes viruses isolated from bats like the Entebbe bat virus, Sokuluk virus and Yokose virus (YOKV), ii) the Modoc virus group that comprises viruses isolated from rodents such as Modoc virus (MODV) and Apoi virus (APOIV) and iii) the Rio Bravo virus group encompassing viruses like Rio Bravo virus (RBV) and Montana myotis leukoencephalitis virus (MMLV) isolated from bats [14]. Little is know about how NKV flaviviruses spread among their hosts. It has been postulated that NKV viruses are transmitted by nasal or oral contact between infected and uninfected animals [15-17]. Initially it was thought that these viruses were unable to infect arthropods due to a block at the level of entry. However, recently it has been shown that at least for MODV, the inability to infect arthropods is not at the level of entry, but at a later stage of the viral life cycle [18].

MODV was initially isolated from white-footed deer mouse (*Peromyscus maniculatis)* in Modoc County in California [19] and was shown to cause a persistent infection in rodents [20]. MODV has not been implicated in human disease, although there has been an indirect reference to an apparently fatal infection of a boy. A serological study provided evidence for the occurrence of natural infection without disease among human inhabitants of Alberta [17]. MODV is neuroinvasive and causes lethal encephalitis in SCID mice and hamsters similar to flaviviral encephalitis in humans, making MODV a potential model virus to study flaviviruses infections [21]. The viral prM and/or E proteins have been shown to be important for the neuroinvasive characteristics of MODV in SCID mice [22]. Sequence analysis and comparison of the MODV 3' UTR with the 3' UTR of APOIV, RBV and MMLV revealed four regions with similar secondary RNA structures [4].

The construction of full-length flavivirus genome cDNAs that can be used for the *in vitro* transcription of infectious RNA can be rather challenging, as the plasmid backbone used and/or the bacterial host strain selected can greatly influence the outcome [23]. For some of these viruses like Kunjin virus (KUNV) [24], West Nile virus (WNV) [25], some dengue (DENV) strains [26,27], and tick-borne encephalitis virus (TBEV) [28], this approach has been straight forward enabling the construction of stable, full-length genome cDNA copies in *Escherichia coli*. For other flaviviruses, like YFV and JEV, the *in vitro* production of infectious RNA transcripts required the use of labor-intensive *in vitro* ligation procedures [29-31] or the use of low copy number plasmid vectors [32,33] to circumvent the genetic instability of the full-length cDNA insert in the *E. coli* host. This report describes the construction and characterization of the first stable full-length NKV cDNA clone. RNA transcribed from this MODV clone yielded infectious virus upon transfection of BHK cells. Furthermore, the virus derived from the clone showed similar growth kinetics when compared with the wild-type virus.

MATERIAL AND METHODS

Cell culture and virus

The origin and culture conditions of the BHK-21J cells that were used throughout this study were described before [32]. The Modoc virus strain M455 was obtained from Prof. J. Neyts (Leuven, Belgium) during collaborative research [22] and was originally purchased from the American Tissue Culture Collection (ATCC, Manassas, USA). Stocks of MODV M455 were produced by infecting BHK-21J cells at a multiplicity of infection (MOI) of 0.1 in PBS containing 2% fetal calf serum (FCS) for 1 hr and subsequent incubation at 37°C and 5% CO_2 in DMEM/2%FCS. After 3 to 4 days, depending on the severity of the cytophatic effect (CPE), the medium was harvested and centrifuged at 3000 x g for

5 min to remove cell debris. The supernatant was used as a virus stock. Stocks of the cDNA-derived viruses were obtained by electroporating BHK-21J cells with full-length RNA transcripts [34]. For analysis of the viral growth kinetics, BHK-21J cells were infected at an MOI of 1; medium was subsequently collected and replaced by the same volume of fresh medium at the indicated times. MODV titers were determined by plaque assays on BHK-21J cells as described previously [34], except for the agarose in the overlay, which was replaced by Avicel [35].

Recombinant DNA techniques and plasmid constructions

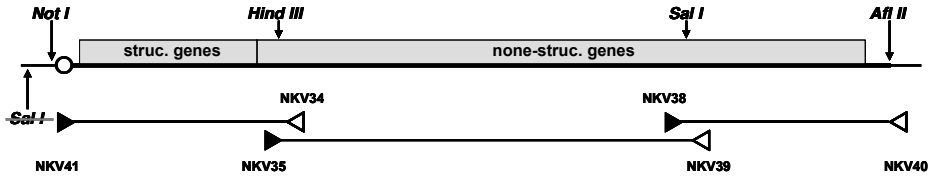
General standard nucleic acid methodologies were used throughout this study [36] unless described in more detail. Chemically competent *E. coli* DH5α cells were used for cloning [37]. Nucleotide numbering of the various constructs containing MODV-derived inserts and the resulting full-length clones was based on the MODV sequence deposited in GenBank (AJ242984; [5]).

MODV cDNA was prepared using a one-step RT-PCR system containing a modified M-MLV reversed transcriptase for the cDNA reaction and a mixture of Taq polymerase and *Pyrococcus* GB-D polymerase (Invitrogen, Carlsbad, USA) for the PCR. The RT-PCR reactions contained 1 μg of total RNA that was isolated from MODV-infected cells at 36 hr post infection (p.i.), using reaction conditions as suggested by the supplier. Oligonucleotides were designed based on the published MODV sequence. The most 5' oligonucleotide (NKV41, fig.1) contained the T7 Φ2.5 promoter [38], so that T7 RNA polymerase driven transcription would start on the "A" residue that is the first nucleotide of the MODV genome. In the oligonucleotide that hybridized to the extreme 3' end of MODV (NKV40, fig. 1) the complement of the two last viral nucleotides were fused to 3'TAAG 5' to yield an unique *Afl II* restriction enzyme site. Relative positions of the oligonucleotides used for the cDNA reconstruction of MODV are depicted in fig. 1.

RNA transcription

Plasmid DNA for *in vitro* run-off RNA transcription was purified using the Nucleobond AX DNA isolation kit. pACNR-FLMODV plasmids were linearized with *Afl II* and purified by proteinase K treatment and phenol/chloroform extraction. Approximately 2 μg of linearized DNA was used as a template for *in vitro* transcription using the Ampliscribe T7 high yield transcription kit (Epicentre, Madison, USA). For the production of 5'-capped full-length MODV transcripts, UTP, GTP and CTP concentration was 7.5 mM, whereas the ATP concentration was 2 mM. G(5')ppp(5')A (NEB, Ipswich, USA) was added as RNA cap

A)



B)

Name	Sequence	Position	Remarks
NKV34	CTGCCAGGAAAGACCATTGCGGCCAG	2719 – 2694	sense
NKV35	GGTATGGGGGTACATGTACTACCCAG	2454 – 2480	anti-sense
NKV38	GTTCCTCTGGAGACCGCATGGCCGGTGT	7716 – 7689	sense
NKV39	GGGCATTTGTTGGAGTGCCCCCACACT	7536 – 7563	anti-sense
NKV40	AGCGCTTAAGCGGAGGTCATATTCATGACCACACAG	10600 – 10573	Contains *Afl II* run-off site sense
NKV41	GACGCGGCCGCAGTAATACGACTCACTATTAGTTGATCCTGCCAGCGGTGGGTCGCTAC	1 – 29	Contains *Not I* site and T7 RNA pol. ϕ 2.5 promoter. anti-sense

Fig. 1. Schematic representation of the construction of pACNR-MODV.
A) The large boxes represent the viral ORF encoding the structural and non-structural proteins. The oligonucleotides that were used to generate the cDNA fragments for constructing the clone are indicated by triangles. The open circle upstream the MODV insert represents the T7 Φ2.5 promoter. The restriction sites that were used to assemble the clone as well as the deleted *Sal I* site are indicted. **B)** Table with detailed information over the sequence, position and orientation of the oligonucleotides that were used in the construction of this clone.

analog to a final concentration of 6 mM. After 2 hr of incubation at 42°C, DNase I was added and the incubation was continued for another 15 min at 37°C. The RNA transcripts were subsequently purified by LiCl precipitation and the concentration was determined by spectrophotometry.

RNA transfection and analysis of viral RNA synthesis

BHK-21J cells were transfected with 5 µg of full-length MODV as described previously [34]. For RNA analysis, 2.5 ml (approximately 1.5 x 10^6 cells) of the transfected BHK-21J cell suspension was seeded in a 10 cm^2 plate. Total RNA was isolated from the transfected cells at 30 hr post electroporation (p.e). Analysis of RNA synthesis by [^{3}H]-uridine labelling was performed as described before [32]. Trizol (Invitrogen, Carlsbad, USA) was used for

cell lysis and subsequent RNA isolation. [^{3}H]-Uridine labelled RNAs were denatured with glyoxal and analyzed on 0.8% agarose gels [36].

Immunofluorescence

At times indicated in the legend of the figures, control and infected cells were washed once with PBS and prepared for immunofluorescence as described previously [32]. Commercially available immune ascitic fluid obtained from mice infected with MODV (ATCC, Manassas, USA) was used in a 1:1000 dilution as primary antibody.

RESULTS AND DISCUSSION

Construction of a full-length MODV cDNA for the *in vitro* transcription of infectious RNA

It is notoriously difficult to construct full-length cDNAs for the *in vitro* transcription of infectious RNA for many flaviviruses due to genetic instability of these clones in *E. coli* [29-31,33]. Therefore, the pACNR1180-derived vector backbone, previously used for the successful construction of a stable full-length YFV-17D clone [32], was selected as the vector for the MODV cDNA inserts. pACNR-FLYF17D$_a$, in which the standard *Xho I* run-off site is replaced by *Afl II*, was cut with *BamH I* and religated. Subsequently the *Sal I* site upstream the Sp6 promoter in this plasmid was inactivated by site-directed mutagenesis. The resulting plasmid pACNR-FLYF17D$_a$ Δ2576-9294 contained unique *Not I*, *Hind III, Sal I* and *Afl II* restriction enzyme sites that were used to assemble the MODV full-length cDNA.

Figure 1 is a schematic representation of the MODV genome and the position of the oligonucleotides used to prime RT-cDNA reactions on MODV RNA isolated from infected cells; the three PCR products that were used to assemble the full-length MODV cDNA are also depicted. Oligonucleotide NKV41 contained the ϕ2.5 promoter for T7 RNA polymerase fused to the most 5′ 19 nucleotides of MODV, whereas in oligonucleotide NKV40, the 28 3′ nucleotides of MODV were fused to a unique *Afl II* recognition sequence, serving as the transcription run-off site. Construction of the full-length MODV clone in pANCR-FLYF17D$_a$ Δ2576-9294 was performed in three steps in the 5′ to 3′ direction during which the remaining YFV-17D sequences were replaced by MODV cDNAs fragments. First, the 2597 bp fragment obtained after *Not I* and *Hind III* digestion of the NKV41 – NKV39 PCR product was cloned into *Not I – Hind III* digested pANCR-FLYF17D$_a$ Δ2576-9294. Subsequently, the 5063 bp *Hind III – Sal I* fragment and the 3′ *2952 bp Sal I – Afl II* fragment were cloned stepwise, to complete the construction of the full-length MODV cDNA. Many

plasmids harbouring what appeared to be a full-length MODV genome cDNA insert based on various restriction enzyme digests were obtained using this strategy.

Identification and characterization of an infectious MODV cDNA clone

Plasmids that appeared to contain a full-length MODV genome cDNA insert were used to produce full-length genomic MODV RNAs by *in vitro* RNA transcription. The RNA transcripts were electroporated into BHK-21J cells; immunofluorescence assay (IFA) were performed 48 hr p.e. for MODV antigen expression. Four (clone numbers 2.4, 4.2, 6.1 and 8.7) out of the ten original clones yielded a clearly positive signal in IFA (data not shown). ^{3}H-uridine labelling of cells electroporated with *in vitro* transcribed RNA of these four plasmids revealed that only clone 6.1 allowed effective first cycle analysis of RNA synthesis (fig. 2). No MODV genome-sized ^{3}H-labeled RNA could be detected in cells electroporated with RNA transcripts from clones 2.4, 4.2 or 8.7. From these results it was concluded that pACNR-FLMODV6.1 contained the most robust template for the *in vitro* production of infectious MODV RNA. This clone and its infectious RNA transcripts were characterized in more detail and compared to the parental MODV virus.

The genetic stability of the full-length MODV6.1 clone in *E.coli* DH5α was analyzed by repeated passaging, i.e., growing the bacteria for more than 12 hr in 2 ml of LB medium plus 50 μg/ml ampicilline, followed by a streak on selective medium to obtain a single colony for the next cycle. After the 10th streak, a bacterial colony was picked and used to prepare plasmid DNA (pACNR-FLMODV6.1-p10). The purified plasmid DNA was digested with various restriction enzymes as an indicator for its genetic stability. No difference in the restriction pattern was observed from the plasmid at passage 10 compared to the original pACNR-FLMODV6.1 DNA (data not shown). Full-length MODV RNA transcripts were then prepared from both pACNR-FLMODV6.1 and pACNR-FLMODV6.1-p10 and

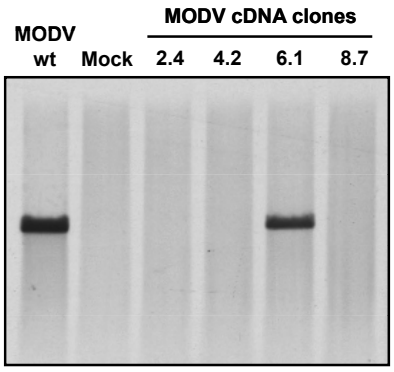


Fig. 2. Analysis of viral RNA synthesis in BHK-21J cells transfected with transcripts of pACNR-MODV clones 2.4, 4.2, 6.1 and 8.7.
BHK-21J cells that were electroporated with full-length MODV transcripts were labelled with ^{3}H-uridine from 24 – 30 hr post transfection. Total RNA was isolated and analyzed after denaturation by agarose gel electrophoresis as described in the Materials and Methods section. BHK-21J cells infected with the parental MODV virus (wt) were used as a control.

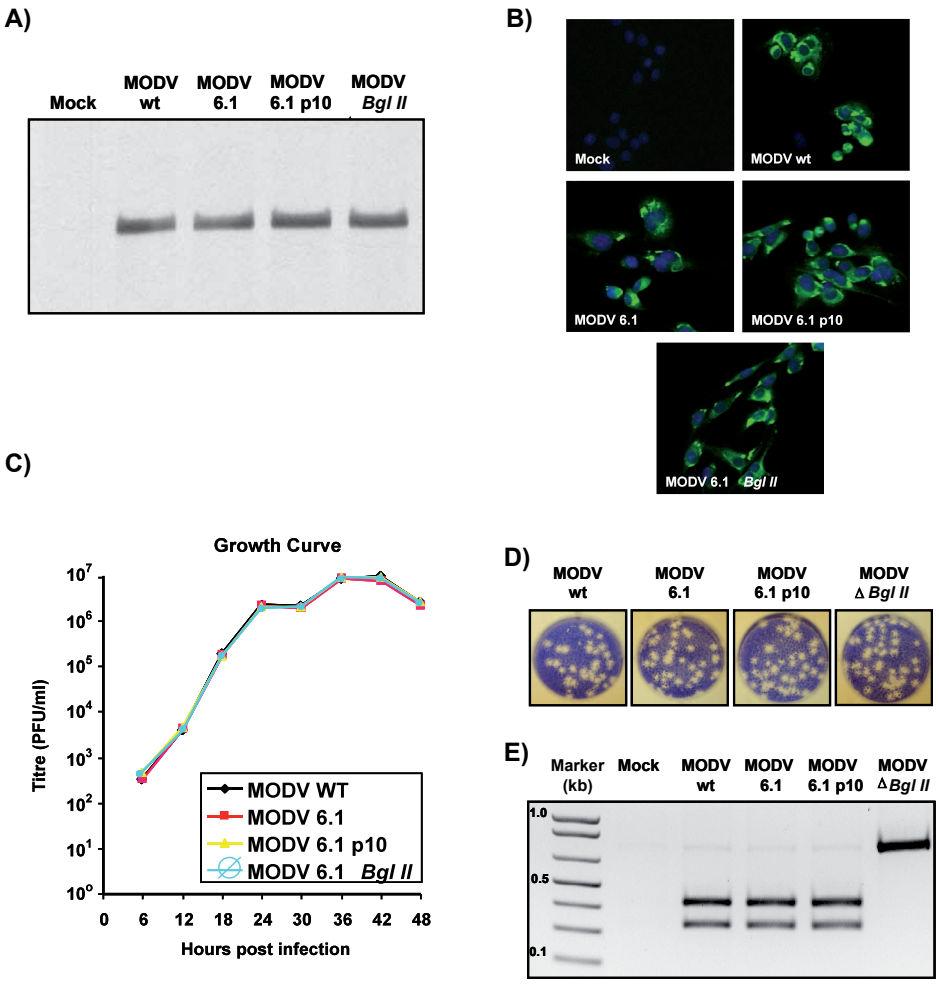


Fig. 3. Characterization of the pACNR-MODV6.1 infectious cDNA clone.
A) Viral RNA synthesis in BHK-21J cells transfected with *in vitro* transcribed RNA of the pACNR-MODV 6.1, 6.1-p10 and 6.1Δ*Bgl II*. Transfected cells were labelled with ^{3}H-uridine from 24 - 30 hr post transfection. Total RNA was isolated and analyzed after denaturation by agarose gel electrophoresis as described in the Materials and Methods section. BHK-21J cells were infected with the parental MODV virus as a control. **B)** Immunofluorescence staining of BHK-21J cells that were infected with wt-MODV or transfected with full-length MODV RNA derived from pACNR-MODV6.1 and related plasmids. Cells were fixed at 30 hr p.e. and stained with the MODV hyperimmune serum. **C)** Viral growth kinetics; BHK cells were infected with a MOI of 1; the medium of the infected cells was harvested at the indicated times p.i. and the viral titer was determined by plaque assays. **D)** Plaque morphology of the wild-type MODV, pACNR-MODV6.1, pACNR-MODV6.1 p10 and pACNR-MODV Δ*Bgl II* in BHK-21J cells. **E)** *Bgl II* restriction analysis of the RT-PCR products generated from RNA isolated from BHK-21J cells infected with the indicated viruses. The sizes of the DNA marker (lane 1) are given on the left side.

electroporated into BHK-21J cells. As a control, BHK cells were infected in parallel with the parental MODV virus. Electrophoresis of ^{3}H-uridine-labelled total intracellular RNA isolated from MODV-6.1-p10 transfected cells, showed the synthesis of MODV genome RNA that co-migrated with the ^{3}H-uridine labelled genome RNA isolated from MODV-infected or MODV6.1 RNA transfected cells (fig. 3.A; lanes 2, 3 and 4). IFA on the electroporated cells with the MODV hyperimmune serum as the primary antibody indicated similar transfection efficiency with pACNR-FLMODV6.1 and pACNR-FLMODV6.1-p10 (fig. 3.B). These results confirmed high specific infectivity as well as genetically stability of pACNR-FLMODV clone 6.1. Furthermore, viruses derived from both pACNR-FLMODV6.1 and pACNR-FLMODV6.1-p10 showed similar growth kinetics when compared to the wt MODV (fig. 3.C). No differences in plaque morphology were observed (fig. 3.D).

In order to rule out the possibility of accidental cross-contamination of BHK-21J cells with the parental MODV virus, thus the obtained results, a mutation was introduced at position 8756 of pACNR-FLMODV6.1, abolishing the *Bgl II* as a genetic marker. The mutation was silent and did not impair viral RNA synthesis (fig. 3.A, lane 5), infectivity of the RNA transcript (fig. 3.B) or viral growth kinetics (fig. 3.C). The wt MODV, MODV6.1, MODV6.1-p10 and MODV6.1Δ*Bgl II* virus stocks harvested for the growth curve experiment (fig. 3.C) at 48 hr p.i. were used to infect BHK-21J cells at an M.O.I. 5. At 30 hr p.i., total intracellular RNA was isolated; a fragment spanning MODV nts. 8435 to 9176 was amplified by RT-PCR, followed by *Bgl II* digestion. As shown in fig. 3.E, the MODV6.1Δ*Bgl II* still lacks this restriction site, demonstrating that the infectious clone was stable and the viruses generated by transfection of permissive cells with transcripts of pACNR-FLMODV6.1 were indeed derived from the cDNA.

The nucleotide sequence of clone 6.1 was determined. Compared to the MODV sequence AJ242984 available in GenBank, 22 nucleotide differences were identified in clone 6.1. As summarized in table 1, these differences were present throughout the coding sequence and five of them corresponded to silent mutations. Although the E protein is the most variable between flaviviruses, only one amino acid difference was observed when comparing the MODV GenBank sequence with clone 6.1. Apart from a high variation in NS5, the largest protein, a relatively high number of amino acid changes were detected in the rather small and generally well conserved NS4B protein (around 254 amino acids). The function of NS4B in the flavivirus life cycle remains to be established, although studies have implicated the involvement of NS4B in modulating the host interferon response [39,40]. None of the mutations found in NS3 or NS5 were part of the functional domains proposed for these viral proteins [5]. Taken together, clone pACNR-MODV6.1 generated in this study can yield viable and infectious MODV particles, and is comparable to wild-type MODV regarding infectivity, plaque morphology and growth kinetics.

Table 1. Summary of the nucleotide differences between the MODV AJ242984 and the MODV clone 6.1 sequences. Changes are grouped by encoded viral proteins. Positions, actual nucleotide change as well as the amino acid substitutions are shown. Nucleotide or amino acid to the right indicates MODV NCBI AJ242984 followed by the nucleotide or amino acid encountered in the MODV6.1 genome.

Gene	Position	Nucleotide	Amino Acid
prM	610	U → C	silent
Env	1543	U → G	Phe → Leu
NS1	2776	C → U	silent
	3089	C → A	Gln → Arg
NS2A	3529	A → G	Ile → Met
NS2B	4410	A → G	Glu → Gly
NS3	4861	A → G	silent
	6312	U → A	Leu → Gln
NS4B	6837	U → C	Ile → Thr
	7098	G → A	Ser → Asn
	7444	C → G	silent
	7445	C → G	Leu → Val
	7503	U → A	Leu → His
NS5	7756	G → A	silent
	7767	G → C	Ser → Thr
	7938	G → C	Arg → Thr
	8141	C → G	Arg → Ala
	8142	G → C	Arg → Ala
	8612	C → G	Gln → Glu
	8920	U → A	Ser → Arg
	9120	A → G	Lys → Arg
	9990	A → G	Asp → Gly

The strategy to clone the MODV inserts directly into the low copy number plasmid pACNR1180, which is known to accept viral sequences that are not well tolerated by *E. coli* [41,32] probably contributed significantly to the successful construction of clone 6.1. pACNR-FLMODV6.1 proved to be stable upon repeated passages in *E.coli*. Attempts to clone the full-length MODV insert from pACNR-FLMODV6.1 into high copy number plas-

mids like pBluescribe or pUC has met no success thus far and only resulted in plasmids that contained recombined MODV inserts harbouring deletions (unpublished results). The pACNR-FLMODV6.1 clone will be a useful tool to increase our understanding of the molecular determinants important for flaviviral replication and tropism. The biological relevance of predicted conserved RNA structures in NKV viruses can now be studied using a reverse genetic approach. Another useful application for this clone is the construction of chimeric viruses between NKVs and arthropod-borne flaviviruses that can provide valuable insights into host-range restrictions and pathogenicity, as illustrated by studies using a yellow fever virus in which prM and E were replaced by the analogous proteins of MODV [22,18]. In addition, this clone and the viruses derived from it are likely to be useful in the screening of antiviral compounds. A very convenient model to monitor MODV infection in the Syrian hamster has been developed to test small molecules with anti-flavivirus activity in a relevant small animal model [21]. The availability of this MODV infectious clone will be a valuable addition to these studies by allowing a more comprehensive analysis of viral resistance.

REFERENCE LIST

1. **Kuno, G., G. J. Chang, K. R. Tsuchiya, N. Karabatsos, and C. B. Cropp**. 1998. Phylogeny of the genus Flavivirus. J. Virol. **72**:73-83.
2. **Cook, S. and E. C. Holmes**. 2006. A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: Flaviviridae) and the evolution of vector transmission. Arch. Virol. **151**:309-325.
3. **Billoir, F., C. R. de, H. Tolou, M. P. de, E. A. Gould, and L. de, X**. 2000. Phylogeny of the genus flavivirus using complete coding sequences of arthropod-borne viruses and viruses with no known vector. J. Gen. Virol. **81 Pt 9**:2339.
4. **Charlier, N., P. Leyssen, C. W. Pleij, P. Lemey, F. Billoir, L. K. Van, A. M. Vandamme, C. E. De, L. de, X, and J. Neyts**. 2002. Complete genome sequence of Montana Myotis leukoencephalitis virus, phylogenetic analysis and comparative study of the 3' untranslated region of flaviviruses with no known vector. J. Gen. Virol. **83**:1875-1885.
5. **Leyssen, P., N. Charlier, P. Lemey, F. Billoir, A. M. Vandamme, C. E. De, L. de, X, and J. Neyts**. 2002. Complete genome sequence, taxonomic assignment, and comparative analysis of the untranslated regions of the Modoc virus, a flavivirus with no known vector. Virology **293**:125-140.
6. **Jansson, A. M., E. Jakobsson, P. Johansson, V. Lantez, B. Coutard, L. de, X, T. Unge, and T. A. Jones**. 2009. Structure of the methyltransferase domain from the Modoc virus, a flavivirus with no known vector. Acta Crystallogr. D. Biol. Crystallogr. **65**:796-803.
7. **Silva, P. A. G. C., T. J. Dalebout, and P. J. Bredenbeek**. 2010. Characterization of the sfRNAs that are produced in cells infected with flaviviruses with no known vector and cell fusing agent. -.
8. **Lin, K. C., H. L. Chang, and R. Y. Chang**. 2004. Accumulation of a 3'-terminal genome fragment in Japanese encephalitis virus-infected mammalian and mosquito cells. J. Virol. **78**:5133-5138.
9. **Liu, R., L. Yue, X. Li, X. Yu, H. Zhao, Z. Jiang, E. Qin, and C. Qin**. 2010. Identification and characterization of small sub-genomic RNAs in dengue 1-4 virus-infected cell cultures and tissues. Biochem. Biophys. Res. Commun. **391**:1099-1103.
10. **Pijlman, G. P., A. Funk, N. Kondratieva, J. Leung, S. Torres, L. van der Aa, W. J. Liu, A. C. Palmenberg, P. Y. Shi, R. A. Hall, and A. A. Khromykh**. 2008. A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity. Cell Host. Microbe **4**:579-591.
11. **Scherbik, S. V., J. M. Paranjape, B. M. Stockman, R. H. Silverman, and M. A. Brinton**. 2006. RNase L plays a role in the antiviral response to West Nile virus. J. Virol. **80**:2987-2999.
12. **Silva, P. A., C. F. Pereira, T. J. Dalebout, W. J. Spaan, and P. J. Bredenbeek**. 2010. An RNA Pseudoknot Is Required for Production of Yellow Fever Virus Subgenomic RNA by the Host Nuclease XRN1. J. Virol. **84**:11395-11406.
13. **Urosevic, N., M. M. van, J. P. Mansfield, J. S. Mackenzie, and G. R. Shellam**. 1997. Molecular characterization of virus-specific RNA produced in the brains of flavivirus-susceptible and -resistant mice after challenge with Murray Valley encephalitis virus. J. Gen. Virol. **78 (Pt 1)**:23-29.
14. **ICTV**. 2005. Virus Taxonomy - Eight Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego.
15. **BELL, J. F. and L. A. THOMAS**. 1964. A NEW VIRUS, "MML", ENZOOTIC IN BATS (MYOTIS LUCIFUGUS) OF MONTANA. Am. J. Trop. Med. Hyg. **13**:607-612.
16. **CONSTANTINE, D. G. and D. F. WOODALL**. 1964. LATENT INFECTION OF RIO BRAVO VIRUS IN SALIVARY GLANDS OF BATS. Public Health Rep. **79**:1033-1039.

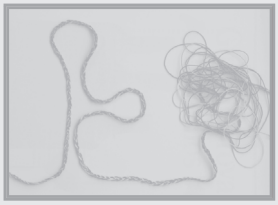
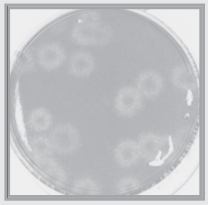
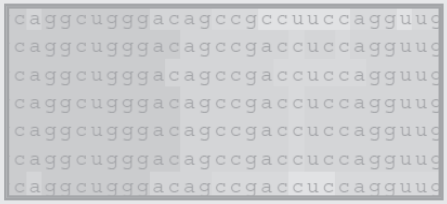
17. **Zarnke, R. L. and T. M. Yuill**. 1985. Modoc-like virus isolated from wild deer mice (Peromyscus maniculatus) in Alberta. J. Wildl. Dis. **21**:94-99.
18. **Charlier, N., A. Davidson, K. Dallmeier, R. Molenkamp, C. E. De, and J. Neyts**. 2010. Replication of not-known-vector flaviviruses in mosquito cells is restricted by intracellular host factors rather than by the viral envelope proteins. J. Gen. Virol. **91**:1693-1697.
19. **Johnson, H. N.** 1967. Ecological implications of antigenically related mammalian viruses for which arthropod vectors are unknown and avian associated soft tick viruses. Jpn. J. Med. Sci. Biol. **20 Suppl**:160-166.
20. **Davis, J. W. and J. L. Hardy**. 1974. Characterization of persistent Modoc viral infections in Syrian hamsters. Infect. Immun. **10**:328-334.
21. **Leyssen, P., L. A. Van, C. Drosten, H. Schmitz, C. E. De, and J. Neyts**. 2001. A novel model for the study of the therapy of flavivirus infections using the Modoc virus. Virology **279**:27-37.
22. **Charlier, N., R. Molenkamp, P. Leyssen, J. Paeshuyse, C. Drosten, M. Panning, C. E. De, P. J. Bredenbeek, and J. Neyts**. 2004. Exchanging the yellow fever virus envelope proteins with Modoc virus prM and E proteins results in a chimeric virus that is neuroinvasive in SCID mice. J. Virol. **78**:7418-7426.
23. **Hurrelbrink, R. J., A. Nestorowicz, and P. C. McMinn**. 1999. Characterization of infectious Murray Valley encephalitis virus derived from a stably cloned genome-length cDNA. J. Gen. Virol. **80 (Pt 12)**:3115-3125.
24. **Khromykh, A. A. and E. G. Westaway**. 1994. Completion of Kunjin virus RNA sequence and recovery of an infectious RNA transcribed from stably cloned full-length cDNA. J. Virol. **68**:4580-4588.
25. **Shi, P. Y., M. Tilgner, and M. K. Lo**. 2002. Construction and characterization of subgenomic replicons of New York strain of West Nile virus. Virology **296**:219-233.
26. **Kinney, R. M., S. Butrapet, G. J. Chang, K. R. Tsuchiya, J. T. Roehrig, N. Bhamarapravati, and D. J. Gubler**. 1997. Construction of infectious cDNA clones for dengue 2 virus: strain 16681 and its attenuated vaccine derivative, strain PDK-53. Virology **230**:300-308.
27. **Lai, C. J., B. T. Zhao, H. Hori, and M. Bray**. 1991. Infectious RNA transcribed from stably cloned full-length cDNA of dengue type 4 virus. Proc. Natl. Acad. Sci. U. S. A **88**:5139-5143.
28. **Mandl, C. W., M. Ecker, H. Holzmann, C. Kunz, and F. X. Heinz**. 1997. Infectious cDNA clones of tick-borne encephalitis virus European subtype prototypic strain Neudoerfl and high virulence strain Hypr. J. Gen. Virol. **78 (Pt 5)**:1049-1057.
29. **Kapoor, M., L. Zhang, P. M. Mohan, and R. Padmanabhan**. 1995. Synthesis and characterization of an infectious dengue virus type-2 RNA genome (New Guinea C strain). Gene **162**:175-180.
30. **Rice, C. M., A. Grakoui, R. Galler, and T. J. Chambers**. 1989. Transcription of infectious yellow fever RNA from full-length cDNA templates produced by in vitro ligation. New Biol. **1**:285-296.
31. **Sumiyoshi, H., C. H. Hoke, and D. W. Trent**. 1992. Infectious Japanese encephalitis virus RNA can be synthesized from in vitro-ligated cDNA templates. J. Virol. **66**:5425-5431.
32. **Bredenbeek, P. J., E. A. Kooi, B. Lindenbach, N. Huijkman, C. M. Rice, and W. J. Spaan**. 2003. A stable full-length yellow fever virus cDNA clone and the role of conserved RNA elements in flavivirus replication. J. Gen. Virol. **84**:1261-1268.
33. **Gualano, R. C., M. J. Pryor, M. R. Cauchi, P. J. Wright, and A. D. Davidson**. 1998. Identification of a major determinant of mouse neurovirulence of dengue virus type 2 using stably cloned genomic-length cDNA. J. Gen. Virol. **79 (Pt 3)**:437-446.

34. **Silva, P. A., R. Molenkamp, T. J. Dalebout, N. Charlier, J. H. Neyts, W. J. Spaan, and P. J. Bredenbeek**. 2007. Conservation of the pentanucleotide motif at the top of the yellow fever virus 17D 3' stem-loop structure is not required for replication. J. Gen. Virol. **88**:1738-1747.
35. **Matrosovich, M., T. Matrosovich, W. Garten, and H. D. Klenk**. 2006. New low-viscosity overlay medium for viral plaque assays. Virol. J. **3**:63.
36. **Sambrook, J., T. Fritsch, and T. Maniatis**. 1989. Molecular Cloning: a Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
37. **Inoue, H., H. Nojima, and H. Okayama**. 1990. High efficiency transformation of Escherichia coli with plasmids. Gene **96**:23-28.
38. **Coleman, T. M., G. Wang, and F. Huang**. 2004. Superior 5' homogeneity of RNA from ATP-initiated transcription under the T7 phi 2.5 promoter. Nucleic Acids Res. **32**:e14.
39. **Munoz-Jordan, J. L., G. G. Sanchez-Burgos, M. Laurent-Rolle, and A. Garcia-Sastre**. 2003. Inhibition of interferon signaling by dengue virus. Proc. Natl. Acad. Sci. U. S. A **100**:14333-14338.
40. **Munoz-Jordan, J. L., M. Laurent-Rolle, J. Ashour, L. Martinez-Sobrido, M. Ashok, W. I. Lipkin, and A. Garcia-Sastre**. 2005. Inhibition of alpha/beta interferon signaling by the NS4B protein of flaviviruses. J. Virol. **79**:8004-8013.
41. **Ruggli, N., J. D. Tratschin, C. Mittelholzer, and M. A. Hofmann**. 1996. Nucleotide sequence of classical swine fever virus strain Alfort/187 and transcription of infectious RNA from stably cloned full-length cDNA. J. Virol. **70**:3478-3487.

Chapter 6

Functions and requirements of conserved RNA structures in the 3' untranslated region of Flaviviruses

Patrícia A. G. C. Silva
Peter J. Bredenbeek

ABSTRACT

Flaviviruses are small enveloped viruses with a positive, single-stranded RNA genome of approximately 11 kb in length, with a 5′ cap structure and a 3′ non-polyadenylated end. The Flavivirus genus has been divided into three different clusters that correlate with the vector that is used for their transmission: i) mosquito-borne, ii) tick-borne, and iii) no known vector flaviviruses. The 3′ untranslated region (UTR) of flaviviruses can be roughly divided into a proximal part, which exhibits extensive heterogeneity in both length and sequence and is present immediately downstream of the stop codon of the open reading frame, and a more conserved distal part that has been defined as the core element of the 3′ UTR as it contains the majority of the elements involved in viral translation, replication, and assembly. A number of small but well conserved RNA sequence elements as well as secondary and tertiary RNA structures have been identified in the flaviviruses 3′ UTR. Some of these have been recognized in all flaviviruses studied thus far, whereas others are characteristic for a particular cluster of the genus. This review describes the characteristics and function of conserved RNA sequences and structures in the 3′ UTR of all three Flavivirus clusters. For this purpose, the flavivirus 3′ UTR was divided into five domains in a 3′-to-5′ direction according to the position in the genome, sequence and structural similarity. Special emphasis has been given to the RNA structures that were reported to play a role in viral replication and pathogenicity.

INTRODUCTION

Positive-strand RNA viruses are unique in the viral world as their genome serves a dual role as both mRNA and as a template for minus-strand RNA synthesis. In general, these viral genomes are characterized by a 5′ untranslated region (UTR), one or more open reading frames (ORF) and a 3′ UTR. Apart from the coding information for the viral proteins present within the ORFs, the viral genome contains a substantial amount of information in the form of RNA sequences and structures that is required for RNA synthesis, translation and encapsidation. Identifying and characterizing the function of these RNA elements is important for understanding the regulation of the several, often mutually exclusive, activities in which the viral genome is involved. Research has shown that these RNA signals can in principal be anywhere in the viral genome, but that the 5′ and especially the 3′ UTR harbor the majority of them. This review focuses on the conserved RNA sequences and structures in the 3′ UTR of viruses belonging to the genus Flavivirus.

The Flavivirus genus belongs to the *Flaviviridae* family, which also includes the Pestivirus and Hepacivirus genera [1]. Flaviviruses are small enveloped viruses containing a positive, single-stranded RNA genome of approximately 11 kb in length, with a 5′ cap structure and a 3′ non-polyadenylated end. Nearly 80 viruses belong to the *Flavivirus* genus and many of them are considered important human pathogens, namely dengue virus (DENV), yellow fever virus (YFV), Japanese encephalitis virus (JEV), West Nile virus (WNV), and tick-borne encephalitis virus (TBEV). Phylogenetic analysis based on the complete coding sequence divided the flaviviruses into three clusters that correlate with the vector of transmission: mosquito-borne, tick-borne and no known vector (NKV) flaviviruses [2,3]. In general, the 3′ UTR of viruses that have a similar mode of transmission show a higher similarity in terms of conserved sequences and RNA structures [4-8]. Despite these differences, certain RNA elements are, however, characteristic for the 3′ UTR of every flavivirus. This review describes the characteristics and function of conserved RNA sequences and structures in the 3′ UTR of all three Flavivirus clusters. It should be emphasized however that this review may seem rather biased towards the mosquito-borne flaviviruses; this is hard to avoid since the 3′ UTR has been studied more extensively for these viruses in comparison to the NKV flaviviruses, and to a lesser extent, the tick-borne flaviviruses.

The length of the viral 3′ UTR varies from approximately 350 to 800 nts depending on the virus, and it can differ even between strains of the same virus. This heterogeneity in length originates primarily from the proximal part of the 3′ UTR, immediately following the stop codon of the viral ORF, where deletions, insertions, sequence repeats and even internal poly(A) tracts have been observed; in contrast, the distal part of the 3′ UTR exhibits a more similar RNA topology and regions with significant sequence similarity [9-15].

This distal region has been defined as the core element of the flavivirus 3′ UTR in which important elements for viral translation and replication are located [5].

In this review, the flavivirus 3′ UTR has been divided into five domains in a 3′-to-5′ direction according to the position in the genome, sequence and structural similarity (fig. 1). The identified domains are usually separated from each other by U-A rich sequences predicted to be single-stranded [16].

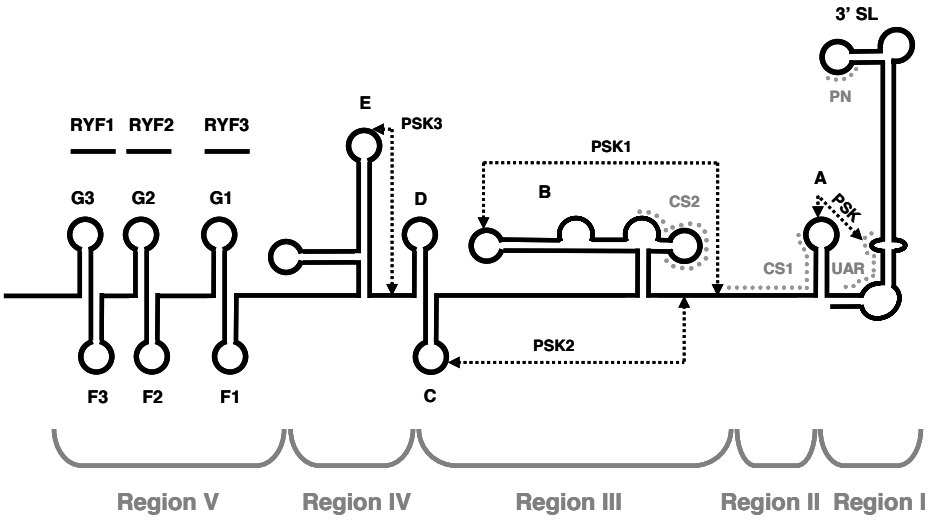


Fig. 1. Schematic model of the predicted RNA folding of the 3′ UTR of the prototype flavivirus yellow fever virus (YFV). The 3′ UTR was divided into five different regions. Nomenclature for the individual stem-loop (except 3′ SL) and pseudoknot structures was adopted from Olsthoorn and Bol [6]. The pseudoknot (PSK) predicted by Shi and colleagues [22] is also depicted. The pentanucleotide motif (PN), the conserved sequence 2 (CS2) and the 3′ cyclization motifs CS1 and UAR (upstream AUG region) are indicated. The third 3′ cyclization motif DAR (downstream AUG region) overlaps with CS1 in the stem-loop A region. The predicted pseudoknot structures (PSK) and the repeated sequences of the yellow fever virus (RYF) are also depicted.

3′ UTR region I: the 3′ stem-loop structures

Region I comprises the last 100 to 120 nucleotides of the viral genome. RNA structure analysis of this region revealed that it can fold into two RNA stem-loop structures. The more upstream structure is relatively small, encompassing only 14 to 20 nts, whereas the last 80 to 100 nts are predicted to form a long stem-loop structure (3′ SL; fig. 1) [4,9,14,17-22]. Even though the general structure of 3′ SL is well conserved among all flaviviruses, the sequence similarity is restricted to the pentanucleotide (PN) motif 5′-CACAG-3′, that is located in the bulge at the top of 3′ SL (fig. 1 and 2), and the terminal 3′ dinucleotide

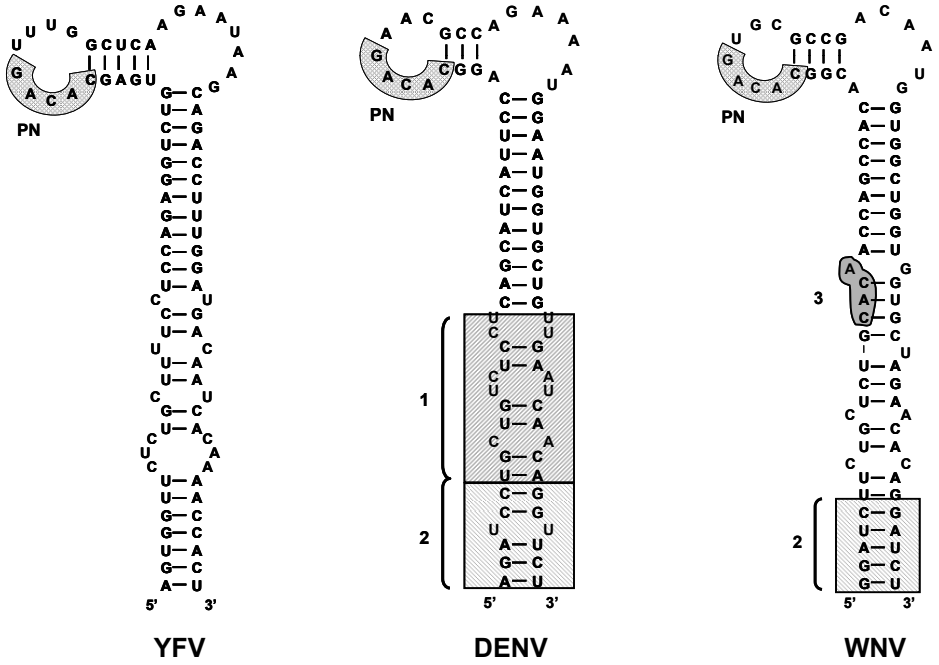


Fig. 2. Structural model of the terminal 3′ SL structure of YFV, DENV and WNV. The pentanucleotide motif (PN) is indicated. The shaded box 1 represents a region important for virus viability while box 2 represents the region that is required for replication in mosquito cells. The region indicated by number 3 corresponds to the eEF-1α-binding sequence.

5′-CU-3′, which is complementary to the dinucleotide 5′-AG-3′ present at the very 5′ end of every flavivirus genome [19-21]. Limited sequence variation has been observed in the PN motif of the NKV flaviviruses, which either contain a "C" or a "U" residue at the second position. APOI virus contains an additional "C" to "U" change at the third position [7].

Because of the absence of a poly-A tail at the 3′ end of the flavivirus genome it was suggested that the 3′ SL structure was functionally replacing the poly-A tail by signalling the integrity of the viral genome and therefore protecting it from degradation. Experiments using either reporters expressing WNV replicon RNAs or synthetic mRNAs containing flavivirus 5′ and 3′ UTRs, yielded contradicting results varying from no effect to a modest stimulation on translation by the 3′ SL [23-26], to actually inhibiting translation [27,28].

Deletion of the 3′ SL showed that this conserved RNA structure is absolutely required for flavivirus RNA synthesis [29-34]. Detailed mutagenesis of the 3′ SL revealed that the terminal dinucleotide 3′ CU_{OH} is essential for efficient viral replication and that it probably functions as a recognition site for the replication complex to initiate RNA synthesis [33,35]. Deletion of the conserved PN motif does not influence translation efficiency of the viral genome, but is lethal for viral RNA synthesis. Mutational analysis of the 5′ CACAG 3′ sequence revealed that the "G" residue at the 5th position and base paring of the nucleo-

tide at the 1st position with a complementary nucleotide four positions downstream of the PN motif, are the most critical elements of this conserved RNA sequence. Nucleotide changes at either the 2nd or 4th positions of the WNV or YFV PN motif were shown to be well tolerated. Mutations at the 3rd position were reported to be detrimental for WNV replication, but showed only a relatively minor to no effect in YFV-17D [25,33,36,37]. The exact role of the PN motif in flavivirus RNA replication is currently unknown. It is interesting to note that although some mutations within the PN motif seem to be well tolerated, competition experiments revealed that a virus with the wild-type PN sequence has a clear advantage over viruses with a mutated PN motif [37]. Phosphorodiamidate morpholino oligomers (PMOs) targeted at regions including the PN motif inhibited the replication of DENV and WNV replicons [38,39]. Interestingly, part of the loop that contains the PN motif was show to be involved in binding eukaryotic translation elongation factor-1α (eEF-1α) [40]. Recently, the RNA structure of the top of 3′ SL was resolved by NMR for representatives of the three flavivirus groups. Surprisingly, the results implied that the structure was not well conserved and indicated clear differences in the stacking pattern of the nucleotides that form the top of the 3′ SL, including the PN motif, among the different groups of flaviviruses [41].

Immediately upstream the 3′ SL is a smaller stem-loop structure (SL-A according to the nomenclature of Olsthoorn and Bol [6]) that comprises 14 to 18 nucleotides and that is present in every flavivirus. Biochemical and biophysical probing and *in silico* RNA modelling showed that the four nucleotides in the SL-A loop of mosquito-borne flaviviruses were involved in the formation of an RNA pseudoknot by base pairing with nucleotides in the lower part of the 3′ stem of 3′ SL [22] (fig. 1). This pseudoknot appears well conserved and can also be predicted for TBEV and the NKV flaviviruses MMLV, MODV, RBV and APOIV. SL-A and the RNA pseudoknot were shown to be required for efficient viral RNA synthesis in *in vitro* DENV RdRp assays [30]. Furthermore, a PMO targeting the pseudoknot interaction moderately inhibited the replication of a WNV reporter replicon [38]. Exchanging SL-A of DENV2 by the WNV SL-A sequence was relatively well tolerated. This substitution, which is predicted to maintain the pseudoknot formation, illustrates that the formation of the structure is more important than the primary sequences [29]. Deletion of SL-A was shown to decrease the translation efficiency of a DENV reporter construct [26]. SL-A was suggested to be a minor binding site for the host eEF-1α protein in WNV [40].

Previous studies demonstrated that the complete DENV and WNV 3′ SL nucleotide sequences were not interchangeable [29,34]. Specific nucleotide sequence elements within the 3′ SL were found to be important for viability as illustrated by the RNA elements indicated in fig. 2 as number 1 in the DENV 3′ SL [29], and number 3 in the WNV 3′ SL (5′-CACA-3′, the eEF-1α-binding sequence) [34,40]. Others were found to be important in a host cell-specific manner; for example the RNA structure represented by number 2 (fig. 2). Substitution of this sequence in DENV by the comparable region of WNV resulted in

a mutant that grew well in monkey kidney cells but was severely restricted in mosquito cells [29]. The U-U bulge in this region of the DENV 3′ SL was shown to be important for efficient replication of these WNV chimeras in C6/36 cells, suggesting that it acts as an enhancer for both DENV and WNV replication in mosquito cells, while it is dispensable for replication in mammalian cells [34]. The results demonstrate that specific bulges and regions in the flavivirus 3′ SL are critical determinants for viral RNA replication. These bulges were suggested to likely represent important binding sites for viral and cellular proteins to assemble the flavivirus replication complex [29,34]. It should be noted that the bottom part of the right-hand side of 3′ SL also contains other elements like the 3′ UAR (upstream AUG region) and the 3′ DAR (downstream AUG region) sequences which have been predicted to be required for cyclization of the viral genome (see a more detailed description on region II).

3′ UTR region II: the cyclization sequences

Immediately upstream of the 3′ SL, and partially overlapping with SL-A, is a 10 to 18 nucleotides region that is predicted to be relatively unstructured but forms the key RNA element for the initiation of viral minus-strand RNA synthesis. The importance of this region was initially suggested by Hahn et al. [4], whom reported the presence of a conserved sequence shared by all the mosquito-borne flaviviruses. This conserved sequence was originally named CS1 (conserved sequence 1) (fig. 1). The most conserved part of CS1 is located upstream of SL-A but the functional part of this sequence is likely to involve the nucleotides of SL-A as well. The key observation has been that CS1 is complementary to a conserved sequence (5′ CS) within the N-terminal coding region of the capsid protein [4]. This complementarity suggested a long-range RNA interaction that would promote circularization of the flavivirus genome resulting in the formation of a panhandle-like RNA structure required for viral RNA synthesis [4]. The critical role of CS1 in this 5′ – 3′ RNA interaction for RNA synthesis in mosquito-borne flaviviruses has now been firmly established by using *in vitro* RdRp assays [30,42-44] and mutational analysis of infectious flavivirus cDNAs [32,45-50]. Physical evidence for the 5′CS-3′CS1 interaction in DENV was obtained by atomic force microscopy [50]. Although formally a requirement for RNA circularization during flavivirus (+) strand RNA synthesis cannot be ruled out, current data demonstrate that it is crucial for (-) strand RNA synthesis [44]. Genome cyclization is dispensable for viral translation [26,46,49].

Apart from the 5′CS-3′CS1 interaction, a second long-range RNA interaction was shown to be required for genome cyclization. In mosquito-borne flaviviruses this interaction is mediated by a sequence at the 5′ end, located immediately upstream of the start codon of the ORF, and a sequence that is part of the bottom of the 3′ SL [50-52]. These complementary sequences were named 5′-3′UAR (upstream AUG region) and were shown to

be important for viral replication [50]. Similar to the 5'CS-3'CS1 interaction, 5'-3'UAR base pairing was also demonstrated by atomic force microscopy [50]. Current evidence suggests that the base pairing involving 5'CS and 3'CS1 initiates the circularization of the genome and promotes the 5'-3' UARs interaction to increase the stability of this long-range RNA interaction [53]. Interestingly, additional RNA base pairing contributing to flavivirus genome circularization were recently identified between nucleotides downstream of the AUG region (5' DAR) and nucleotides downstream CS1 in the SL-A stem (3' DAR) [54,55]. The 5' and 3' DAR motifs were shown to be important for genome circularization and RNA replication in DENV and WNV [54,56]. In WNV, the 5'-3' DAR interaction actually consists of two stretches of complementary sequences [55]. In YFV, the DAR motifs [56] were originally included in the 18 nt found to be part of the 5' CS and 3' CS1 elements [47] (see fig. 3). A general model was proposed for flaviviruses in which the 5'-3' DAR interaction extends the initial circularization between 5' and 3' CS, and together with the 5'-3' UAR interaction, unwinds the 3' SL [54,56]. The various 5' and 3' RNA sequences that were shown to be required for RNA circularization in mosquito-borne flaviviruses are shown in figure 3.

RNA cyclization sequences have also been identified in the other two flavivirus clusters. In tick-borne flaviviruses these complementary sequences are named 5'-CS-A and 3'-CS-A. 5'-CS-A is located upstream of the translation initiation codon at approximately 100 nt from the 5' end, whereas 3'-CS-A is present in the 3' SL, approximately 80 nts from the 3' end [21]. There is no sequence similarity between the 5'- and 3'-CS-A sequences and 5'CS and CS1 of the mosquito-borne flaviviruses [21,51] and the position of 5'- and 3'-CS-A actually resembles the location of the sequences involved in the 5'-3' UAR interaction of mosquito-borne flaviviruses. Mutagenesis of the 5'- and 3'-CSA revealed that complementarity between these RNA sequences was required for viral RNA synthesis [21,51,57]. In addition to the 5'–3' CS-A interaction, another pair of complementary sequences (CS-B) has been identified at positions in the genome of tick-borne viruses (fig. 3). These sequences are reminiscent of the 5'-CS and CS1 in the mosquito-borne flaviviruses [45]. Although the 5'- and 3'-CS-B sequences are well conserved in TBEV strains and related viruses like Powassan, Vasilchenko and louping ill virus [10,13,21], their interaction is not essential for viral replication [57]. Sequence complementarity between 5' and 3' ends in the NKV flaviviruses MODV and MMLV has also been reported. Complementary RNA sequences that could promote circularization of the viral genome were identified at the 5' end of the genome, encompassing the AUG codon and the nucleotides encoding the N-terminus of the capsid protein; the 3' end counterpart was located immediately upstream the SL-A, resembling the location of CS1 in mosquito-borne flaviviruses [7,58].

A particular stem-loop structure (SLA) in the flavivirus 5' UTR was found to be important for RNA replication [30,59-61]. In fact, it was shown that this SLA structure was the responsible for promoting viral RNA synthesis and not the cyclization sequences per se [59]. This was supported by the fact that the SLA was found to be specifically recognized by the viral

YFV	**5' CS:** 5' – C**CCUGG**GCGUCAAUAUGGU – 3'	**3' CS1:** 5' – ACCAUAUUGACG**CCAGG**G – 3'
	5' UAR: 5' – AGCAGAGAACUG – 3'	**3' UAR:** 5' – UGGUUCUCUGCU – 3'
	5' DAR: 5' – CCUGG – 3'	**3' DAR:** 5' – CCAGG – 3'
DENV	**5' CS:** 5' – UCAAUAUGCUG – 3'	**3' CS1:** 5' – CAGCAUAUUGA – 3'
	5' UAR: 5' – AGAGAGCAGAUCUCUG – 3'	**3' UAR:** 5' – CAGAGAUCCUGCUGUCU – 3'
	5' DAR: 5' – CCAACG – 3'	**3' DAR:** 5' – CGCUGG – 3'
WNV	**5' CS:** 5' – UGUCAAUAUGCU – 3'	**3' CS1:** 5' – AGCAUAUUGACA – 3'
	5' UAR: 5' – AGCACGAAGAUCUC – 3'	**3' UAR:** 5' – GAGAUCUUCUGCU – 3'
	5' DAR: 5' – GUCUA.....CCAGG – 3'	**3' DAR:** 5' – CCUGG..UAGAC – 3'
JEV	**5' CS:** 5' – UCAAUAUGUG – 3'	**3' CS1:** 5' – CACAUAUUGA – 3'
	5' UAR: 5' – UAGAACGGAAGAUAACCAUG – 3'	**3' UAR:** 5' – UGGGAGAUCUUCUGCUCUA – 3'
	5' DAR: 5' – CCAGG – 3'	**3' DAR:** 5' – CCUGG – 3'
TBEV	**5' CS-A:** 5' – GGAGAACAAGAGCUG – 3'	**3' CS-A:** 5' – CGGUUCUUGUUCUCC – 3'
	5' CS-B: 5' – GGGGCGGUCCC – 3'	**3' CS-B:** 5' – GGGAGGCCCC – 3'
MODV	**5' CS:** 5' – AAUGUGCGAAAAUAACAGGA	**3' CS:** 5' – UCCUGUUAUUUUCCAAAUU – 3'

Fig. 3. Potential 5′ and 3′ cyclization motifs in the mosquito-borne flaviviruses YFV, DENV, WNV and JEV, the tick-borne flavivirus TBEV and the no known vector MODV. The YFV nucleotides presented in bold in the 5′ CS and 3′ CS1 sequences [47] correspond to nucleotides that were afterwards described as the DAR elements [56]. The underlined nucleotides indicate unpaired nucleotides. The dots in the DAR elements of WNV represent the nucleotides between the two stretches of nucleotides characteristic of the DAR interaction of WNV.

RdRp NS5 [59,60,62]. These data have led to a model for the initiation of viral (-) strand RNA synthesis in which the viral RdRp would bind to a conserved stem-loop structure at the 5′ end of the genome. This binding would initiate genome circularization by the 5′CS-3′CS1 interaction, which is subsequently extended and therefore stabilized by base pairing of the additional RNA UAR and DAR motifs. Because the 3′UAR is located at the bottom of the 3′ SL, the 5′UAR-3′UAR interaction might be responsible for destabilizing the 3′ SL structure, which enforces the 3′ end of the genome in a single-strand conformation, thereby making it accessible for the RdRp to use it as a template for initiation of viral (-) strand RNA synthesis [59,60,63]. This model implies that the interaction between the 5′- and 3′ ends of the viral genome functions as a riboswitch that can be induced by the binding of NS5 to the 5′ end of the genome. This ensures refolding of the RNA from a "linear" conformation that is used for translation and packaging, into a circular conformation that is required for the initiation of viral RNA synthesis [26,50,64,65]. Genome circularization might be beneficial for viral replication for several reasons: (i) as a control mechanism to amplify only full-length templates, (ii) to regulate transcription versus translation in space and time, (iii) to bring the replication complex in close vicinity of the transcription initiation site at the 3′ end of the viral genome, and (iv) to control the level of (-) strand RNA synthesis [4,59,65].

3' UTR region III: RNA structures required for enhancement of RNA synthesis

This region essentially involves the nucleotides downstream of the XRN1 stalling signal (see region IV) and upstream of the CS1 or the circularization sequences in the mosquito-borne and NKV flaviviruses or 3' CS-B in the TBE-like viruses. Even flaviviruses that belong to the same cluster show a significant variation in the length of this region due to duplication of RNA sequences. Region III varies in the mosquito-borne flaviviruses from approximately 100 nts for YFV to 170 nts in DENV and in the NKV viruses from ± 150 nts for MODV to 225 nts for RBV. In terms of RNA structure, this region is characterized by the predicted formation of one or two so-called dumbbell-like or Y-shaped RNA stem-loop structures [6,7,13]. The dumbbell structure of the mosquito-borne flaviviruses contains a conserved sequence that has originally been named CS2 and has ~24 nucleotides in length (fig. 1). Mosquito-borne flaviviruses like JEV, WNV and DENV, which show a duplication of this dumbbell-like structure, contain two CS2 elements (CS2 and RCS2 [4]). In NKV flaviviruses an RNA sequence with significant sequence homology to CS2 can be found in a similar position in the predicted Y-shaped RNA structures as to CS2 of the mosquito-borne flaviviruses [7,58]. So far a CS2-like sequence has not been identified within region III of TBEV or related viruses. Deletion of CS2 has a relatively minor effect on viral RNA synthesis and virus production, but seems to decrease viral pathogenicity [32,46,48,49,66].

In addition to CS2, the region III of flaviviruses is characterized by two predicted RNA pseudoknots (fig. 1) [6] (Jiang, Silva, Dalebout and Bredenbeek unpublished results). These RNA pseudoknots, named PSK1 and PSK2, involve four to five nts of the loop on the left-hand side of the dumbbell structures. These nucleotides are predicted to base pair with a complementary sequence downstream of the dumbbell structures. Experimental evidence for the formation of both these pseudoknots in DENV, YFV and WNV has been obtained by RNA structure probing and mutational analysis [67,68] (Molenkamp, Dalebout and Bredenbeek, unpublished results). Disruption of either PSK1 or PSK2 significantly impairs viral RNA synthesis and virus production. This effect is enhanced when neither of the two pseudoknot structures can be formed. Formation of PSK1 and PSK2 is not required for efficient viral translation (Molenkamp, Dalebout and Bredenbeek, unpublished results).

Many other deletion mutants involving sequences of region III have been described for several flaviviruses [46,48,49,66]. Unfortunately, most of these deletions were not guided by the predicted RNA structures and/or did not take into account the potential redundancy of the effect of the introduced deletions due to duplication of conserved RNA elements. Therefore, and also due to the differences in experimental design, it is rather difficult to truly compare the outcome of the various studies. In general, the results support a

role for CS2 and the RNA pseudoknots in viral RNA synthesis. Deletion of the (nearly) complete dumbbell structure that contains CS2 has a more dramatic effect on viral RNA synthesis than deleting only CS2. This can be explained by the fact that such deletions exhibit the cumulative effect of disrupting one of the RNA pseudoknots and the loss of the CS2 sequence. Removing both dumbbell structures of region II of DENV results in very crippled to none-viable viruses when analyzed in mammalian cells [49,66]. Surprisingly, one of these DENV deletion mutants was able to replicate in C6/36 mosquito cells. Another mutant, in which the left part of the dumbbell structures was deleted but that maintained CS2 and RCS2 sequences, exhibited the reverse phenotype as it was able to replicate efficiently in Vero cells but not in C6/36 cells [66]. The molecular basis of these apparently host-specific effects of deletions in region III of DENV is unknown. YFV mutants that are either unable to form PSK1 or PSK2 or lack CS2 yield a similar phenotype in the mammalian BHK and SW13 cells and in the C6/36 mosquito cells (Molenkamp, Dalebout and Bredenbeek; unpublished results).

Of particular interest is a 30-nucleotide deletion (nucleotides 10.478-10.507 in the dumbbell structure 1; involving PSK1) in the 3' UTR of DENV4 genome. This rDENV4Δ30 mutant was shown to be attenuated in rhesus monkeys [48] and well tolerated and highly immunogenic in human volunteers [69,70]. In addition, rDENV4Δ30 also exhibited a limited ability to infect the midgut of mosquitoes [71]. Introduction of the Δ30 mutation into the homologous region of DENV1 yielded similar results as for rDENV4Δ30 [72,73]. Unfortunately, introduction of the Δ30 mutation in DENV2 and DENV3 did not result in significant attenuation when tested in rhesus monkeys [74,75]. The molecular basis of the different phenotypes caused by this deletion in either DENV1 and 4 versus DENV2 and 3 is currently unknown, but it should be the subject of further studies in order to produce a safe and effective tetravalent DENV vaccine based on this Δ30 deletion.

3' UTR region IV: the XRN1-stalling region

Region IV of the flavivirus 3' UTR varies in length due to sequence duplication especially in members of the JEV subgroup that have a third conserved sequence (CS3), which is also repeated (RCS3) [4,76]. These sequences were shown to be important for efficient RNA replication in KUNV and WNV [46,77]. DENV mutants with deletions in region III also exhibited reduced growth properties, suggesting that this region is essential for an efficient viral replication [78].

Recently it has become evident that the major characteristic of this region is an RNA pseudoknot structure that can be formed in all flaviviruses (fig. 4). Viruses that contain a sequence duplication within region IV are predicted to form two pseudoknots. This RNA pseudoknot structure is often followed by a small stem-loop structure (see fig. 1) with

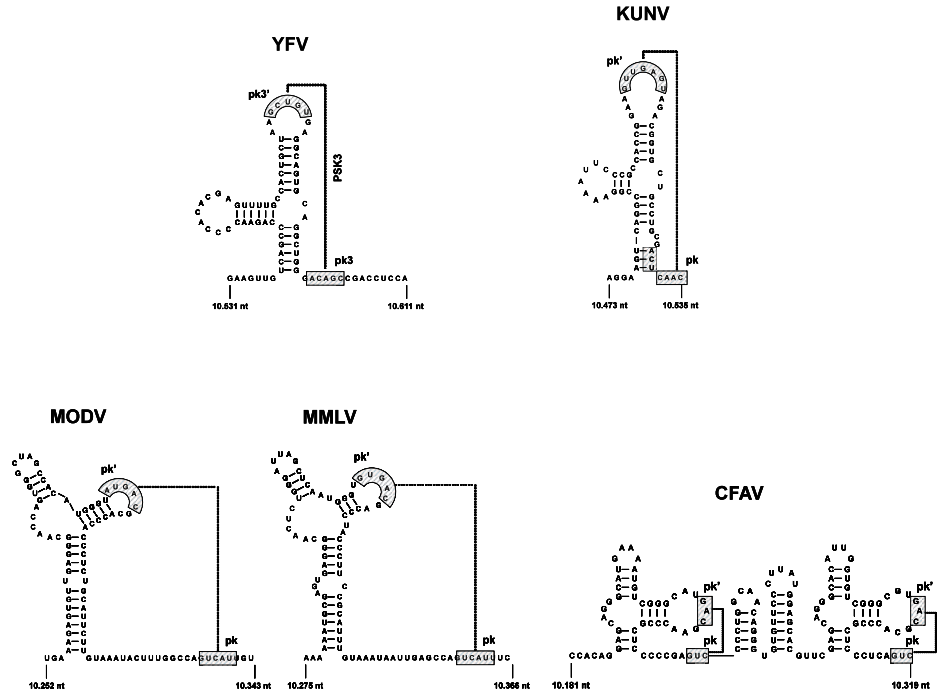


Fig. 4. RNA sequence and predicted RNA structures for the part of region III that is required for stalling XRN1. RNA structures were predicted using a combinatorial approach involving Mfold, phylogeny and manual sequence-structure analysis. The predicted RNA pseudoknot formation by interaction of the pk and pk' sequences are indicated by grey boxes and dotted lines. The NCBI accession numbers are X03700 for YFV-17D, D00246.1 for KUNV, AJ242984 for MODV, AJ299445 for MMLV, and NC_008604 for CFAV.

unknown function. The RNA pseudoknot has been identified as an important determinant for the synthesis of a small flavivirus subgenomic RNA (sfRNA) that is colinear with the 3' end of the viral genome [68,79,80]. This sfRNA is not produced by viral transcription, but by 5' to 3' degradation of the viral genome by the host exoribonuclease XRN1 [79,81]. XRN1 is the main host RNase associated with cellular 5' to 3' mRNA decay (reviewed in [82,83]). The RNA pseudoknot has been shown to serve as a stalling site for XRN1-mediated RNA decay, resulting in the production of the sfRNA [68,79,81]. Interestingly, this sfRNA was shown to be an important determinant for virus pathogenicity as recombinant flaviviruses that are unable to produce the sfRNA show an attenuated phenotype. Disruption of the RNA pseudoknot structure has only a moderate effect on flavivirus genome RNA synthesis [68,79,81].

These results provide a prime example of how an RNA structure can (in)directly be involved in viral pathogenesis. The mechanism by which the sfRNA influences the viral pathogenicity is not known. It has been suggested that it may act as decoy for host miRNAs or proteins that are directed against the viral genome to decrease replication

[81,84]. Another intriguing hypothesis is that the sfRNA itself may serve as a precursor for the production of a virus-encoded miRNA [79]. It is interesting to note that XRN1-mediated RNA decay occurs within the subcellular P bodies that among others also contain Dicer and Argonaute proteins, which are required for the production of functional miRNAs (reviewed in [82,83]). It is important to realize that the sfRNA contains all the RNA sequences and most likely the RNA structures that are characteristic for the distal part of the flavivirus 3' UTR. Therefore, the effect of mutations within this region is not necessarily linked to the viral genome, but can be the result of their effect on the function of the sfRNA.

3' UTR region V: the variable region

The RNA sequence that starts immediately downstream of the stop codon of the viral ORF up to the XRN1-stalling site is known as the variable region (VR). Natural isolates of arthropod-borne flaviviruses often show a significant sequence and size variability due to relatively large nucleotide insertions and deletions in this region of the 3' UTR [9-12,15,85]. The heterogeneity of the VR in natural isolates is well documented for YFV. The originally reported YFV-17D VR corresponds to a unique set of three closely spaced repeated sequences (RYF1, RYF2 and RYF3), each of approximately 40 nucleotides in length [17] (fig. 1) that were shown to be characteristic for West-African strains of YFV [86]. It was subsequently demonstrated that Central and East-African strains had only two repeats (RYF1 and RYF3), whereas South America genotypes had one single copy (RYF3) [12,87]. Interestingly, passaging of an YFV strain harboring all three repeats in mice and cell culture resulted in the deletion of both RYF1 and RYF2 [87]. Similar results were reported for TBEV isolates, as they were also shown to accumulate deletions in the VR region during propagation in either cell lines or mice [13,88]. These results suggest that the majority of the 3'VR is not essential for efficient replication of natural virus isolates. Deletions in the background of infectious cDNA clones of YFV-17D, KUNV, and TBEV demonstrated that actually the complete 3' VR could be deleted without significant impact on replication or virulence for these mutant viruses in either cell culture or mice [32,77,88-90]. Mutagenesis studies on the 3'VR of DENV resulted in a slightly more complicated picture. Relatively small deletions up to 19 nts did not in general have any significant effect on virus replication in either insect or mammalian cell lines [91]. However, larger deletions in the VR resulted in a significant decrease in RNA synthesis and virus production especially in mammalian cells [49,78]. The significance of these often subtle differences is currently unknown and may in part result from the use of different cell lines and virus isolates as well as the rationale of the introduced deletions.

No characteristic conserved RNA structure has been predicted in the VR. It has been suggested that the VR acts as a spacer element to allow proper folding of the RNA struc-

tures in the distal part of the 3′ UTR [16]. However, in the case of NKV flaviviruses, the VR sequences are in general rather short for serving as a spacer element (e.g. 20 and 41 nts for MODV and MMLV, respectively), although they are relatively A-U rich and therefore likely to be unstructured [7].

In order to truly understand the importance of the VR in the 3′ UTR of flaviviruses, recombinant deletion mutants should be assessed *in vivo* in mosquitoes/ticks and animal models. The fact that this region is present in almost every flavivirus known indicates that it serves a purpose in the viral life cycle in nature. Furthermore, the fact that natural isolates tend to have a longer VR suggests a selective advantage that is either associated with replication or pathogenicity in the natural situation.

Flavivirus 3′ UTR-binding proteins

Many host proteins have been found to play an important role in flavivirus infection (for a review see [92]). Several of these proteins were shown to interact with the viral 5′ and 3′ UTRs. Unfortunately, for only a few of these host factors the function of the interaction has been identified. This section briefly describes the host and viral proteins that were reported to interact specifically with the flavivirus 3′ UTR and summarizes our limited knowledge about their role in the flavivirus life cycle.

Host proteins interacting with the 3′ UTR of the viral genome

Apart from the usual suspects like polypyrimidine tract-binding protein (PTB), La and to some extend Poly(A)-binding protein (PABP), only a few host proteins that specifically interact with the 3′ UTR of the flavivirus genome have been identified. These include the eukaryotic translation elongation factor-1α (EF-1α), Y box binding protein-1 (YB-1) and the protein Mov34.

The eukaryotic translation elongation factor-1α (EF-1α) was shown to bind specifically to the right-hand side of the 3′ SL region of WNV [40] and DENV4 [93]. The binding of EF-1α to the 3′ SL is primarily determined by an only four nucleotides long 5′-CACA-3′ motif, although additional sequences that are located in the top 3′ SL loop and in the smaller adjacent stem-loop SL-A are also involved [40]. Mutational analysis of a WNV infectious cDNA clone revealed that the interaction between eEF-1α and the WNV 3′ SL was required for viral (-) strand RNA synthesis [94]. In addition, in WNV and DENV4 infected cells, eEF-1α colocalized with dsRNA and the viral proteins NS3 and NS5, suggesting that this host protein was required for specific recognition of the 3′ SL by the RdRp to promote (-) strand RNA synthesis [94]. It is currently unknown whether, apart from DENV4, eEF-1α also binds to the 3′ UTR of other flaviviruses.

Y box binding protein-1 (YB-1) is a member of the highly conserved Y box proteins and functions as pleiotropic transcription factor that is mainly involved in the regulation of the expression of stress induced genes. YB-1 was shown to bind specifically to the DENV 3′ SL and to repress replication. This antiviral effect can in part be explained by YB-1 mediated inhibition of DENV RNA translation [95]. In addition, it is speculated that the YB-1 serves as a transcription factor to promote activation of innate immune response genes such as ISG54 and ISG56, which can down-regulate translation. It is currently unknown whether YB-1 binds to the 3′ UTR of other flaviviruses, or whether its role as virus repressor is limited to DENV4 virus.

Mouse Mov34 protein is another cellular protein that was reported to bind to the 3′ SL RNA of JEV [96]. It belongs to a family of proteins that share a so-called MPN-like domain. Mov34 serves as regulatory subunit of the 26 proteasome and it is therefore not clear why this protein would interact with the flavivirus 3′ SL.

In addition to the above proteins, various nuclear ribonucleoproteins like hnRNP A1, hnRNPA2/B1 and hnRNP Q were also shown to bind to the DENV 3′ SL by RNA affinity chromatography [95]. However, no *in vivo* data that demonstrate the relevance of these hnRNP- 3′ SL interactions is available.

The promiscuous RNA-binding proteins polypyrimidine tract-binding (PTB) and La protein have also been shown to interact with the 3′ UTR of several flaviviruses like DENV and JEV [93,97]. These proteins are normally restricted to the nucleus, but were shown to translocate to the cytoplasm upon infection [98,99]. Binding for both La and PTB is mapped in the CS1, SL-A and 3′ SL RNA structures in which putative binding sites have been suggested [93]. However, the precise binding sites for these proteins in the viral 3′ UTR remain to be determined. Several studies suggest that PTB is part of the viral replication complex [93,99-101]. However, the experimental data on the role of PTB in the viral life cycle is often conflicting as illustrated by experiments in which inhibition of the expression of PTB by siRNA did not have a significant effect on YFV replication, whereas it severely inhibited DENV production [100]. There appears to be consensus that La is required for efficient virus replication, although the suggested functions are rather diverse [93,97,98,102,103]. Several studies suggest that La serves as an RNA chaperone or an indirect factor that by interacting with NS3 and NS5 helps in the transition from the "linear" translation competent RNA structure to the circular RNA structure that is required for transcription/replication [93,98,102]. Other studies suggest that the La-associated helicase activity is required in the flavivirus replication complex [97,103]. This would be rather surprising since the viral NS3 protein, which is present in the viral replication complex, also contains a helicase activity.

Poly(A)-binding protein (PABP) is a unique translation initiation factor that stimulates translation by promoting mRNA circularization through simultaneous interactions with eIF4G and the 3′ poly(A) tail (reviewed in [104]). Despite the fact that the genome RNA of

flaviviruses is not polyadenylated, PABP was shown to bind to the DENV 3′ UTR [105]. This binding most likely involves the relatively A-rich sequences flanking the RNA dumbbell structures in region III of the viral 3′ UTR. PABP interaction with the 3′ UTR appears to be required for efficient translation initiation of the viral genome [105].

Viral proteins interacting with the 3′ UTR of the viral genome

Purified recombinant NS5 protein that contains the viral RdRp activity was shown to interact *in vitro* with the 3′ UTR, especially the 3′ SL, of JEV and DENV [106,107]. In addition to NS5, purified recombinant NS3 that contained the viral RNA helicase activity, was also reported to interact with the 3′ UTR of DENV [108]. The NS3 and NS5 interactions with the 3′ SL were also demonstrated by UV cross-linking of proteins from JEV-infected cell lysates with RNA probes mimicking the viral 3′ UTR [109]. The studies cited above suggested that the interaction of NS3 and NS5 with RNA sequences or structures of region I or II reflect the (partial) formation of the viral replication complex on the 3′ UTR to initiate viral (-) strand RNA synthesis. However, these data are contradicted by recent data on the formation of DENV replication complex, which shows that DENV NS5 has a strong affinity for stem-loop A at the 5′ UTR of the viral genome [59]. The binding of NS5 to and/or the formation of a replication complex at the 5′ UTR is postulated to serve as a trigger to initiate the circularization of the viral genome that is required for flavivirus RNA synthesis. This resulted in a very attractive model that explains many of the experimental data on the initiation of flavivirus (-) strand RNA replication (reviewed in [65,110]). Taking into account their function as viral RNA helicase and RdRp it is not surprising that both NS3 and NS5 demonstrated affinity for RNA, but the biological relevance of these interactions with RNA elements in the 3′ UTR is in our opinion questionable. Specific binding of KUNV recombinant NS2A with the 3′ UTR of this virus has also been reported [111].

A FINAL NOTE

Understanding the biology of flaviviruses at the molecular level is crucial for the development of new vaccines and the rational design of novel antiviral strategies. Although the E protein is recognized as a major determinant for the tropism and virulence of flaviviruses (reviewed in [112]), it is evident that RNA sequences and structures located in the 3′ UTRs can also influence their virulence [48,88,113-118]. Several of these RNA elements within the viral 3′ UTR represent interesting targets for an antiviral strategy based on antisense oligonucleotides that either inhibit translation or replication, or serve as siRNA [38,39,119-123].

The overall structural integrity of the flavivirus 3′ UTR is clearly important for efficient replication of the virus [16]. Deletion or mutation of RNA elements that do not appear to result in a significant decrease in viral replication when analyzed in cell culture, may

actually show a crippled replication or pathogenicity of the virus when analyzed in a more relevant model system or in competition with the wt-virus [37,68,79,81].

The biological significance and molecular function of many of the RNA structures within the viral 3′ UTR, as well as the interactions between host and/or viral proteins with the viral 3′ UTR, are still unknown. Future research should verify the predicted RNA structures and determine their role in the virus life cycle, as well as the role of certain protein interactions with the flavivirus RNA under *in vivo* conditions. Furthermore, the composition and assembly of the viral replication complex and the mechanism by which this complex initiates the transcription of viral RNA should also be studied in more detail. The results from such studies will increase our understanding over the role of RNA structures in flavivirus replication, and help us in identifying and validating new targets for antiviral strategies and/or for the development of attenuated viruses that can be used as vaccines.

REFERENCE LIST

1. **ICTV**. 2005. Virus Taxonomy - Eight Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego.
2. **Kuno, G., G. J. Chang, K. R. Tsuchiya, N. Karabatsos, and C. B. Cropp**. 1998. Phylogeny of the genus Flavivirus. J. Virol. **72**:73-83.
3. **Cook, S. and E. C. Holmes**. 2006. A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: Flaviviridae) and the evolution of vector transmission. Arch. Virol. **151**:309-325.
4. **Hahn, C. S., Y. S. Hahn, C. M. Rice, E. Lee, L. Dalgarno, E. G. Strauss, and J. H. Strauss**. 1987. Conserved elements in the 3′ untranslated region of flavivirus RNAs and potential cyclization sequences. J. Mol. Biol. **198**:33-41.
5. **Proutski, V., E. A. Gould, and E. C. Holmes**. 1997. Secondary structure of the 3′ untranslated region of flaviviruses: similarities and differences. Nucleic Acids Res. **25**:1194-1202.
6. **Olsthoorn, R. C. and J. F. Bol**. 2001. Sequence comparison and secondary structure analysis of the 3′ noncoding region of flavivirus genomes reveals multiple pseudoknots. RNA. **7**:1370-1377.
7. **Charlier, N., P. Leyssen, C. W. Pleij, P. Lemey, F. Billoir, L. K. Van, A. M. Vandamme, C. E. De, L. de, X, and J. Neyts**. 2002. Complete genome sequence of Montana Myotis leukoencephalitis virus, phylogenetic analysis and comparative study of the 3′ untranslated region of flaviviruses with no known vector. J. Gen. Virol. **83**:1875-1885.
8. **Gritsun, T. S. and E. A. Gould**. 2007. Origin and evolution of 3′UTR of flaviviruses: long direct repeats as a basis for the formation of secondary structures and their significance for virus transmission. Adv. Virus Res. **69**:203-248.
9. **Mandl, C. W., C. Kunz, and F. X. Heinz**. 1991. Presence of poly(A) in a flavivirus: significant differences between the 3′ noncoding regions of the genomic RNAs of tick-borne encephalitis virus strains. J. Virol. **65**:4070-4077.
10. **Wallner, G., C. W. Mandl, C. Kunz, and F. X. Heinz**. 1995. The flavivirus 3′-noncoding region: extensive size heterogeneity independent of evolutionary relationships among strains of tick-borne encephalitis virus. Virology **213**:169-178.
11. **Poidinger, M., R. A. Hall, and J. S. Mackenzie**. 1996. Molecular characterization of the Japanese encephalitis serocomplex of the flavivirus genus. Virology **218**:417-421.
12. **Wang, E., S. C. Weaver, R. E. Shope, R. B. Tesh, D. M. Watts, and A. D. Barrett**. 1996. Genetic variation in yellow fever virus: duplication in the 3′ noncoding region of strains from Africa. Virology **225**:274-281.
13. **Gritsun, T. S., K. Venugopal, P. M. Zanotto, M. V. Mikhailov, A. A. Sall, E. C. Holmes, I. Polkinghorne, T. V. Frolova, V. V. Pogodina, V. A. Lashkevich, and E. A. Gould**. 1997. Complete sequence of two tick-borne flaviviruses isolated from Siberia and the UK: analysis and significance of the 5′ and 3′-UTRs. Virus Res. **49**:27-39.
14. **Rauscher, S., C. Flamm, C. W. Mandl, F. X. Heinz, and P. F. Stadler**. 1997. Secondary structure of the 3′-noncoding region of flavivirus genomes: comparative analysis of base pairing probabilities. RNA. **3**:779-791.
15. **Shurtleff, A. C., D. W. Beasley, J. J. Chen, H. Ni, M. T. Suderman, H. Wang, R. Xu, E. Wang, S. C. Weaver, D. M. Watts, K. L. Russell, and A. D. Barrett**. 2001. Genetic variation in the 3′ noncoding region of dengue viruses. Virology **281**:75-87.

16. **Proutski, V., T. S. Gritsun, E. A. Gould, and E. C. Holmes**. 1999. Biological consequences of deletions within the 3′-untranslated region of flaviviruses may be due to rearrangements of RNA secondary structure. Virus Res. **64**:107-123.
17. **Rice, C. M., E. M. Lenches, S. R. Eddy, S. J. Shin, R. L. Sheets, and J. H. Strauss**. 1985. Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. Science **229**:726-733.
18. **Grange, T., M. Bouloy, and M. Girard**. 1985. Stable secondary structures at the 3′-end of the genome of yellow fever virus (17 D vaccine strain). FEBS Lett. **188**:159-163.
19. **Wengler, G. and E. Castle**. 1986. Analysis of structural properties which possibly are characteristic for the 3′-terminal sequence of the genome RNA of flaviviruses. J. Gen. Virol. **67 (Pt 6)**:1183-1188.
20. **Brinton, M. A., A. V. Fernandez, and J. H. Dispoto**. 1986. The 3′-nucleotides of flavivirus genomic RNA form a conserved secondary structure. Virology **153**:113-121.
21. **Mandl, C. W., H. Holzmann, C. Kunz, and F. X. Heinz**. 1993. Complete genomic sequence of Powassan virus: evaluation of genetic elements in tick-borne versus mosquito-borne flaviviruses. Virology **194**:173-184.
22. **Shi, P. Y., M. A. Brinton, J. M. Veal, Y. Y. Zhong, and W. D. Wilson**. 1996. Evidence for the existence of a pseudoknot structure at the 3′ terminus of the flavivirus genomic RNA. Biochemistry **35**:4222-4230.
23. **Tilgner, M. and P. Y. Shi**. 2004. Structure and function of the 3′ terminal six nucleotides of the west nile virus genome in viral replication. J. Virol. **78**:8159-8171.
24. **Holden, K. L. and E. Harris**. 2004. Enhancement of dengue virus translation: role of the 3′ untranslated region and the terminal 3′ stem-loop domain. Virology **329**:119-133.
25. **Tilgner, M., T. S. Deas, and P. Y. Shi**. 2005. The flavivirus-conserved penta-nucleotide in the 3′ stem-loop of the West Nile virus genome requires a specific sequence and structure for RNA synthesis, but not for viral translation. Virology **331**:375-386.
26. **Chiu, W. W., R. M. Kinney, and T. W. Dreher**. 2005. Control of translation by the 5′- and 3′-terminal regions of the dengue virus genome. J. Virol. **79**:8303-8315.
27. **Li, W. and M. A. Brinton**. 2001. The 3′ stem loop of the West Nile virus genomic RNA can suppress translation of chimeric mRNAs. Virology **287**:49-61.
28. **Wei, Y., C. Qin, T. Jiang, X. Li, H. Zhao, Z. Liu, Y. Deng, R. Liu, S. Chen, M. Yu, and E. Qin**. 2009. Translational regulation by the 3′ untranslated region of the dengue type 2 virus genome. Am. J. Trop. Med. Hyg. **81**:817-824.
29. **Zeng, L., B. Falgout, and L. Markoff**. 1998. Identification of specific nucleotide sequences within the conserved 3′-SL in the dengue type 2 virus genome required for replication. J. Virol. **72**:7510-7522.
30. **You, S., B. Falgout, L. Markoff, and R. Padmanabhan**. 2001. In vitro RNA synthesis from exogenous dengue viral RNA templates requires long range interactions between 5′- and 3′-terminal regions that influence RNA structure. J. Biol. Chem. **276**:15581-15591.
31. **Shi, P. Y., M. Tilgner, and M. K. Lo**. 2002. Construction and characterization of subgenomic replicons of New York strain of West Nile virus. Virology **296**:219-233.
32. **Bredenbeek, P. J., E. A. Kooi, B. Lindenbach, N. Huijkman, C. M. Rice, and W. J. Spaan**. 2003. A stable full-length yellow fever virus cDNA clone and the role of conserved RNA elements in flavivirus replication. J. Gen. Virol. **84**:1261-1268.
33. **Khromykh, A. A., N. Kondratieva, J. Y. Sgro, A. Palmenberg, and E. G. Westaway**. 2003. Significance in replication of the terminal nucleotides of the flavivirus genome. J. Virol. **77**:10623-10629.

34. **Yu, L. and L. Markoff**. 2005. The topology of bulges in the long stem of the flavivirus 3′ stem-loop is a major determinant of RNA replication competence. J. Virol. **79**:2309-2324.
35. **Nomaguchi, M., M. Ackermann, C. Yon, S. You, and R. Padmanabhan**. 2003. De novo synthesis of negative-strand RNA by Dengue virus RNA-dependent RNA polymerase in vitro: nucleotide, primer, and template parameters. J. Virol. **77**:8831-8842.
36. **Elghonemy, S., W. G. Davis, and M. A. Brinton**. 2005. The majority of the nucleotides in the top loop of the genomic 3′ terminal stem loop structure are cis-acting in a West Nile virus infectious clone. Virology **331**:238-246.
37. **Silva, P. A., R. Molenkamp, T. J. Dalebout, N. Charlier, J. H. Neyts, W. J. Spaan, and P. J. Bredenbeek**. 2007. Conservation of the pentanucleotide motif at the top of the yellow fever virus 17D 3′ stem-loop structure is not required for replication. J. Gen. Virol. **88**:1738-1747.
38. **Deas, T. S., I. Binduga-Gajewska, M. Tilgner, P. Ren, D. A. Stein, H. M. Moulton, P. L. Iversen, E. B. Kauffman, L. D. Kramer, and P. Y. Shi**. 2005. Inhibition of flavivirus infections by antisense oligomers specifically suppressing viral translation and RNA replication. J. Virol. **79**:4599-4609.
39. **Holden, K. L., D. A. Stein, T. C. Pierson, A. A. Ahmed, K. Clyde, P. L. Iversen, and E. Harris**. 2006. Inhibition of dengue virus translation and RNA synthesis by a morpholino oligomer targeted to the top of the terminal 3′ stem-loop structure. Virology **344**:439-452.
40. **Blackwell, J. L. and M. A. Brinton**. 1997. Translation elongation factor-1 alpha interacts with the 3′ stem-loop region of West Nile virus genomic RNA. J. Virol. **71**:6433-6444.
41. **Lescrinier, E., N. Dyubankova, K. Nauwelaerts, R. Jones, and P. Herdewijn**. 2010. Structure determination of the top-loop of the conserved 3′-terminal secondary structure in the genome of flaviviruses. Chembiochem. **11**:1404-1412.
42. **You, S. and R. Padmanabhan**. 1999. A novel in vitro replication system for Dengue virus. Initiation of RNA synthesis at the 3′-end of exogenous viral RNA templates requires 5′- and 3′-terminal complementary sequence motifs of the viral RNA. J. Biol. Chem. **274**:33714-33722.
43. **Ackermann, M. and R. Padmanabhan**. 2001. De novo synthesis of RNA by the dengue virus RNA-dependent RNA polymerase exhibits temperature dependence at the initiation but not elongation phase. J. Biol. Chem. **276**:39926-39937.
44. **Nomaguchi, M., T. Teramoto, L. Yu, L. Markoff, and R. Padmanabhan**. 2004. Requirements for West Nile virus (-)- and (+)-strand subgenomic RNA synthesis in vitro by the viral RNA-dependent RNA polymerase expressed in Escherichia coli. J. Biol. Chem. **279**:12141-12151.
45. **Khromykh, A. A., H. Meka, K. J. Guyatt, and E. G. Westaway**. 2001. Essential role of cyclization sequences in flavivirus RNA replication. J. Virol. **75**:6719-6728.
46. **Lo, M. K., M. Tilgner, K. A. Bernard, and P. Y. Shi**. 2003. Functional analysis of mosquito-borne flavivirus conserved sequence elements within 3′ untranslated region of West Nile virus by use of a reporting replicon that differentiates between viral translation and RNA replication. J. Virol. **77**:10004-10014.
47. **Corver, J., E. Lenches, K. Smith, R. A. Robison, T. Sando, E. G. Strauss, and J. H. Strauss**. 2003. Fine mapping of a cis-acting sequence element in yellow fever virus RNA that is required for RNA replication and cyclization. J. Virol. **77**:2265-2270.
48. **Men, R., M. Bray, D. Clark, R. M. Chanock, and C. J. Lai**. 1996. Dengue type 4 virus mutants containing deletions in the 3′ noncoding region of the RNA genome: analysis of growth restriction in cell culture and altered viremia pattern and immunogenicity in rhesus monkeys. J. Virol. **70**:3930-3937.

49. **Alvarez, D. E., A. L. De Lella Ezcurra, S. Fucito, and A. V. Gamarnik**. 2005. Role of RNA structures present at the 3'UTR of dengue virus on translation, RNA synthesis, and viral replication. Virology **339**:200-212.
50. **Alvarez, D. E., M. F. Lodeiro, S. J. Luduena, L. I. Pietrasanta, and A. V. Gamarnik**. 2005. Long-range RNA-RNA interactions circularize the dengue virus genome. J. Virol. **79**:6631-6643.
51. **Thurner, C., C. Witwer, I. L. Hofacker, and P. F. Stadler**. 2004. Conserved RNA secondary structures in Flaviviridae genomes. J. Gen. Virol. **85**:1113-1124.
52. **Song, B. H., S. I. Yun, Y. J. Choi, J. M. Kim, C. H. Lee, and Y. M. Lee**. 2008. A complex RNA motif defined by three discontinuous 5-nucleotide-long strands is essential for Flavivirus RNA replication. RNA. **14**:1791-1813.
53. **Polacek, C., J. E. Foley, and E. Harris**. 2009. Conformational changes in the solution structure of the dengue virus 5' end in the presence and absence of the 3' untranslated region. J. Virol. **83**:1161-1166.
54. **Friebe, P. and E. Harris**. 2010. Interplay of RNA elements in the dengue virus 5' and 3' ends required for viral RNA replication. J. Virol. **84**:6103-6118.
55. **Zhang, B., H. Dong, H. Ye, M. Tilgner, and P. Y. Shi**. 2010. Genetic analysis of West Nile virus containing a complete 3'CSI RNA deletion. Virology **408**:138-145.
56. **Friebe, P., P. Y. Shi, and E. Harris**. 2010. The 5' and 3' Downstream of AUG region (DAR) elements are required for mosquito-borne flavivirus RNA replication. J. Virol.
57. **Kofler, R. M., V. M. Hoenninger, C. Thurner, and C. W. Mandl**. 2006. Functional analysis of the tick-borne encephalitis virus cyclization elements indicates major differences between mosquito-borne and tick-borne flaviviruses. J. Virol. **80**:4099-4113.
58. **Leyssen, P., N. Charlier, P. Lemey, F. Billoir, A. M. Vandamme, C. E. De, L. de, X, and J. Neyts**. 2002. Complete genome sequence, taxonomic assignment, and comparative analysis of the untranslated regions of the Modoc virus, a flavivirus with no known vector. Virology **293**:125-140.
59. **Filomatori, C. V., M. F. Lodeiro, D. E. Alvarez, M. M. Samsa, L. Pietrasanta, and A. V. Gamarnik**. 2006. A 5' RNA element promotes dengue virus RNA synthesis on a circular genome. Genes Dev. **20**:2238-2249.
60. **Dong, H., B. Zhang, and P. Y. Shi**. 2008. Terminal structures of West Nile virus genomic RNA and their interactions with viral NS5 protein. Virology **381**:123-135.
61. **Li, X. F., T. Jiang, X. D. Yu, Y. Q. Deng, H. Zhao, Q. Y. Zhu, E. D. Qin, and C. F. Qin**. 2010. RNA elements within the 5' untranslated region of the West Nile virus genome are critical for RNA synthesis and virus replication. J. Gen. Virol. **91**:1218-1223.
62. **Lodeiro, M. F., C. V. Filomatori, and A. V. Gamarnik**. 2009. Structural and functional studies of the promoter element for dengue virus RNA replication. J. Virol. **83**:993-1008.
63. **Zhang, B., H. Dong, D. A. Stein, P. L. Iversen, and P. Y. Shi**. 2008. West Nile virus genome cyclization and RNA replication require two pairs of long-distance RNA interactions. Virology **373**:1-13.
64. **Edgil, D. and E. Harris**. 2006. End-to-end communication in the modulation of translation by mammalian RNA viruses. Virus Res. **119**:43-51.
65. **Villordo, S. M. and A. V. Gamarnik**. 2009. Genome cyclization as strategy for flavivirus RNA replication. Virus Res. **139**:230-239.
66. **Blaney, J. E., Jr., N. S. Sathe, L. Goddard, C. T. Hanson, T. A. Romero, K. A. Hanley, B. R. Murphy, and S. S. Whitehead**. 2008. Dengue virus type 3 vaccine candidates generated by introduction of deletions in the 3' untranslated region (3'-UTR) or by exchange of the DENV-3 3'-UTR with that of DENV-4. Vaccine **26**:817-828.

67. **Romero, T. A., E. Tumban, J. Jun, W. B. Lott, and K. A. Hanley**. 2006. Secondary structure of dengue virus type 4 3′ untranslated region: impact of deletion and substitution mutations. J. Gen. Virol. **87**:3291-3296.
68. **Funk, A., K. Truong, T. Nagasaki, S. Torres, N. Floden, M. E. Balmori, J. Edmonds, H. Dong, P. Y. Shi, and A. A. Khromykh**. 2010. RNA structures required for production of subgenomic flavivirus RNA. J. Virol. **84**:11407-11417.
69. **Durbin, A. P., R. A. Karron, W. Sun, D. W. Vaughn, M. J. Reynolds, J. R. Perreault, B. Thumar, R. Men, C. J. Lai, W. R. Elkins, R. M. Chanock, B. R. Murphy, and S. S. Whitehead**. 2001. Attenuation and immunogenicity in humans of a live dengue virus type-4 vaccine candidate with a 30 nucleotide deletion in its 3′-untranslated region. Am. J. Trop. Med. Hyg. **65**:405-413.
70. **Durbin, A. P., S. S. Whitehead, J. McArthur, J. R. Perreault, J. E. Blaney, Jr., B. Thumar, B. R. Murphy, and R. A. Karron**. 2005. rDEN4delta30, a live attenuated dengue virus type 4 vaccine candidate, is safe, immunogenic, and highly infectious in healthy adult volunteers. J. Infect. Dis. **191**:710-718.
71. **Troyer, J. M., K. A. Hanley, S. S. Whitehead, D. Strickman, R. A. Karron, A. P. Durbin, and B. R. Murphy**. 2001. A live attenuated recombinant dengue-4 virus vaccine candidate with restricted capacity for dissemination in mosquitoes and lack of transmission from vaccinees to mosquitoes. Am. J. Trop. Med. Hyg. **65**:414-419.
72. **Whitehead, S. S., B. Falgout, K. A. Hanley, J. E. J. Blaney Jr, L. Markoff, and B. R. Murphy**. 2003. A live, attenuated dengue virus type 1 vaccine candidate with a 30-nucleotide deletion in the 3′ untranslated region is highly attenuated and immunogenic in monkeys. J. Virol. **77**:1653-1657.
73. **Durbin, A. P., J. McArthur, J. A. Marron, J. E. Blaney, Jr., B. Thumar, K. Wanionek, B. R. Murphy, and S. S. Whitehead**. 2006. The live attenuated dengue serotype 1 vaccine rDEN1Delta30 is safe and highly immunogenic in healthy adult volunteers. Hum. Vaccin. **2**:167-173.
74. **Blaney, J. E., Jr., C. T. Hanson, K. A. Hanley, B. R. Murphy, and S. S. Whitehead**. 2004. Vaccine candidates derived from a novel infectious cDNA clone of an American genotype dengue virus type 2. BMC. Infect. Dis. **4**:39.
75. **Blaney, J. E., Jr., C. T. Hanson, C. Y. Firestone, K. A. Hanley, B. R. Murphy, and S. S. Whitehead**. 2004. Genetically modified, live attenuated dengue virus type 3 vaccine candidates. Am. J. Trop. Med. Hyg. **71**:811-821.
76. **Khromykh, A. A. and E. G. Westaway**. 1994. Completion of Kunjin virus RNA sequence and recovery of an infectious RNA transcribed from stably cloned full-length cDNA. J. Virol. **68**:4580-4588.
77. **Khromykh, A. A. and E. G. Westaway**. 1997. Subgenomic replicons of the flavivirus Kunjin: construction and applications. J. Virol. **71**:1497-1505.
78. **Tajima, S., Y. Nukui, T. Takasaki, and I. Kurane**. 2007. Characterization of the variable region in the 3′ non-translated region of dengue type 1 virus. J. Gen. Virol. **88**:2214-2222.
79. **Silva, P. A., C. F. Pereira, T. J. Dalebout, W. J. Spaan, and P. J. Bredenbeek**. 2010. An RNA Pseudoknot Is Required for Production of Yellow Fever Virus Subgenomic RNA by the Host Nuclease XRN1. J. Virol. **84**:11395-11406.
80. **Silva, P. A. G. C., T. J. Dalebout, and P. J. Bredenbeek**. 2010. Characterization of the sfRNAs that are produced in cells infected with flaviviruses with no known vector and cell fusing agent. -.
81. **Pijlman, G. P., A. Funk, N. Kondratieva, J. Leung, S. Torres, L. van der Aa, W. J. Liu, A. C. Palmenberg, P. Y. Shi, R. A. Hall, and A. A. Khromykh**. 2008. A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity. Cell Host. Microbe **4**:579-591.

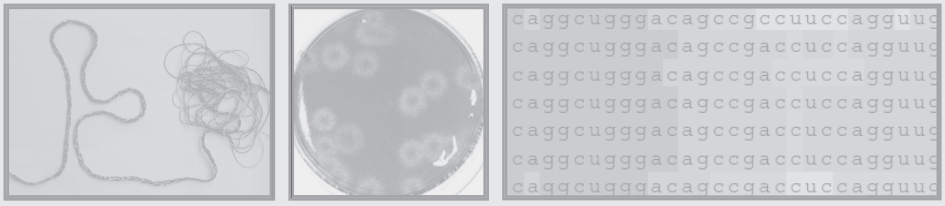
82. **Anderson, P. and N. Kedersha**. 2006. RNA granules. J. Cell Biol. **172**:803-808.
83. **Garneau, N. L., J. Wilusz, and C. J. Wilusz**. 2007. The highways and byways of mRNA decay. Nat. Rev. Mol. Cell Biol. **8**:113-126.
84. **Fernandez-Garcia, M. D., M. Mazzon, M. Jacobs, and A. Amara**. 2009. Pathogenesis of flavivirus infections: using and abusing the host cell. Cell Host. Microbe **5**:318-328.
85. **Jan, L. R., K. L. Chen, C. F. Lu, Y. C. Wu, and C. B. Horng**. 1996. Complete nucleotide sequence of the genome of Japanese encephalitis virus ling strain: the presence of a 25-nucleotide deletion in the 3'-nontranslated region. Am. J. Trop. Med. Hyg. **55**:603-609.
86. **Hahn, C. S., J. M. Dalrymple, J. H. Strauss, and C. M. Rice**. 1987. Comparison of the virulent Asibi strain of yellow fever virus with the 17D vaccine strain derived from it. Proc. Natl. Acad. Sci. U. S. A **84**:2019-2023.
87. **Mutebi, J. P., R. C. Rijnbrand, H. Wang, K. D. Ryman, E. Wang, L. D. Fulop, R. Titball, and A. D. Barrett**. 2004. Genetic relationships and evolution of genotypes of yellow fever virus and other members of the yellow fever virus group within the Flavivirus genus based on the 3' noncoding region. J. Virol. **78**:9652-9665.
88. **Mandl, C. W., H. Holzmann, T. Meixner, S. Rauscher, P. F. Stadler, S. L. Allison, and F. X. Heinz**. 1998. Spontaneous and engineered deletions in the 3' noncoding region of tick-borne encephalitis virus: construction of highly attenuated mutants of a flavivirus. J. Virol. **72**:2132-2140.
89. **Bryant, J. E., P. F. Vasconcelos, R. C. Rijnbrand, J. P. Mutebi, S. Higgs, and A. D. Barrett**. 2005. Size heterogeneity in the 3' noncoding region of South American isolates of yellow fever virus. J. Virol. **79**:3807-3821.
90. **Hoenninger, V. M., H. Rouha, K. K. Orlinger, L. Miorin, A. Marcello, R. M. Kofler, and C. W. Mandl**. 2008. Analysis of the effects of alterations in the tick-borne encephalitis virus 3'-noncoding region on translation and RNA replication using reporter replicons. Virology **377**:419-430.
91. **Tajima, S., Y. Nukui, M. Ito, T. Takasaki, and I. Kurane**. 2006. Nineteen nucleotides in the variable region of 3' non-translated region are dispensable for the replication of dengue type 1 virus in vitro. Virus Res. **116**:38-44.
92. **Pastorino, B., A. Nougairede, N. Wurtz, E. Gould, and L. de, X**. 2010. Role of host cell factors in flavivirus infection: Implications for pathogenesis and development of antiviral drugs. Antiviral Res. **87**:281-294.
93. **De Nova-Ocampo, M., N. Villegas-Sepulveda, and R. M. del Angel**. 2002. Translation elongation factor-1alpha, La, and PTB interact with the 3' untranslated region of dengue 4 virus RNA. Virology **295**:337-347.
94. **Davis, W. G., J. L. Blackwell, P. Y. Shi, and M. A. Brinton**. 2007. Interaction between the cellular protein eEF1A and the 3'-terminal stem-loop of West Nile virus genomic RNA facilitates viral minus-strand RNA synthesis. J. Virol. **81**:10172-10187.
95. **Paranjape, S. M. and E. Harris**. 2007. Y box-binding protein-1 binds to the dengue virus 3'-untranslated region and mediates antiviral effects. J. Biol. Chem. **282**:30497-30508.
96. **Ta, M. and S. Vrati**. 2000. Mov34 protein from mouse brain interacts with the 3' noncoding region of Japanese encephalitis virus. J. Virol. **74**:5108-5115.
97. **Vashist, S., M. Anantpadma, H. Sharma, and S. Vrati**. 2009. La protein binds the predicted loop structures in the 3' non-coding region of Japanese encephalitis virus genome: role in virus replication. J. Gen. Virol. **90**:1343-1352.
98. **Yocupicio-Monroy, M., R. Padmanabhan, F. Medina, and R. M. del Angel**. 2007. Mosquito La protein binds to the 3' untranslated region of the positive and negative polarity dengue virus RNAs and relocates to the cytoplasm of infected cells. Virology **357**:29-40.

99. **Agis-Juarez, R. A., I. Galvan, F. Medina, T. Daikoku, R. Padmanabhan, J. E. Ludert, and R. M. del Angel**. 2009. Polypyrimidine tract-binding protein is relocated to the cytoplasm and is required during dengue virus infection in Vero cells. J. Gen. Virol. **90**:2893-2901.
100. **Anwar, A., K. M. Leong, M. L. Ng, J. J. Chu, and M. A. Garcia-Blanco**. 2009. The polypyrimidine tract-binding protein is required for efficient dengue virus propagation and associates with the viral replication machinery. J. Biol. Chem. **284**:17021-17029.
101. **Jiang, L., H. Yao, X. Duan, X. Lu, and Y. Liu**. 2009. Polypyrimidine tract-binding protein influences negative strand RNA synthesis of dengue virus. Biochem. Biophys. Res. Commun. **385**:187-192.
102. **Garcia-Montalvo, B. M., F. Medina, and R. M. del Angel**. 2004. La protein binds to NS5 and NS3 and to the 5′ and 3′ ends of Dengue 4 virus RNA. Virus Res. **102**:141-150.
103. **Yocupicio-Monroy, R. M., F. Medina, V. J. Reyes-del, and R. M. del Angel**. 2003. Cellular proteins from human monocytes bind to dengue 4 virus minus-strand 3′ untranslated region RNA. J. Virol. **77**:3067-3076.
104. **Derry, M. C., A. Yanagiya, Y. Martineau, and N. Sonenberg**. 2006. Regulation of poly(A)-binding protein through PABP-interacting proteins. Cold Spring Harb. Symp. Quant. Biol. **71**:537-543.
105. **Polacek, C., P. Friebe, and E. Harris**. 2009. Poly(A)-binding protein binds to the non-polyadenylated 3′ untranslated region of dengue virus and modulates translation efficiency. J. Gen. Virol. **90**:687-692.
106. **Tan, B. H., J. Fu, R. J. Sugrue, E. H. Yap, Y. C. Chan, and Y. H. Tan**. 1996. Recombinant dengue type 1 virus NS5 protein expressed in Escherichia coli exhibits RNA-dependent RNA polymerase activity. Virology **216**:317-325.
107. **Kim, Y. G., J. S. Yoo, J. H. Kim, C. M. Kim, and J. W. Oh**. 2007. Biochemical characterization of a recombinant Japanese encephalitis virus RNA-dependent RNA polymerase. BMC. Mol. Biol. **8**:59.
108. **Cui, T., R. J. Sugrue, Q. Xu, A. K. Lee, Y. C. Chan, and J. Fu**. 1998. Recombinant dengue virus type 1 NS3 protein exhibits specific viral RNA binding and NTPase activity regulated by the NS5 protein. Virology **246**:409-417.
109. **Chen, C. J., M. D. Kuo, L. J. Chien, S. L. Hsu, Y. M. Wang, and J. H. Lin**. 1997. RNA-protein interactions: involvement of NS3, NS5, and 3′ noncoding regions of Japanese encephalitis virus genomic RNA. J. Virol. **71**:3466-3473.
110. **Simon, A. E. and L. Gehrke**. 2009. RNA conformational changes in the life cycles of RNA viruses, viroids, and virus-associated RNAs. Biochim. Biophys. Acta **1789**:571-583.
111. **Mackenzie, J. M., A. A. Khromykh, M. K. Jones, and E. G. Westaway**. 1998. Subcellular localization and some biochemical properties of the flavivirus Kunjin nonstructural proteins NS2A and NS4A. Virology **245**:203-215.
112. **McMinn, P. C.** 1997. The molecular basis of virulence of the encephalitogenic flaviviruses. J. Gen. Virol. **78 (Pt 11)**:2711-2722.
113. **Leitmeyer, K. C., D. W. Vaughn, D. M. Watts, R. Salas, I. Villalobos, C. de, C. Ramos, and R. Rico-Hesse**. 1999. Dengue virus structural differences that correlate with pathogenesis. J. Virol. **73**:4738-4747.
114. **Pletnev, A. G.** 2001. Infectious cDNA clone of attenuated Langat tick-borne flavivirus (strain E5) and a 3′ deletion mutant constructed from it exhibit decreased neuroinvasiveness in immunodeficient mice. Virology **282**:288-300.
115. **Gritsun, T. S., A. Desai, and E. A. Gould**. 2001. The degree of attenuation of tick-borne encephalitis virus depends on the cumulative effects of point mutations. J. Gen. Virol. **82**:1667-1675.

116. **Chiou, S. S. and W. J. Chen**. 2001. Mutations in the NS3 gene and 3'-NCR of Japanese encephalitis virus isolated from an unconventional ecosystem and implications for natural attenuation of the virus. Virology **289**:129-136.
117. **Blaney, J. E., Jr., D. H. Johnson, G. G. Manipon, C. Y. Firestone, C. T. Hanson, B. R. Murphy, and S. S. Whitehead**. 2002. Genetic basis of attenuation of dengue virus type 4 small plaque mutants with restricted replication in suckling mice and in SCID mice transplanted with human liver cells. Virology **300**:125-139.
118. **Edgil, D., M. S. Diamond, K. L. Holden, S. M. Paranjape, and E. Harris**. 2003. Translation efficiency determines differences in cellular infection among dengue virus type 2 strains. Virology **317**:275-290.
119. **Kinney, R. M., C. Y. Huang, B. C. Rose, A. D. Kroeker, T. W. Dreher, P. L. Iversen, and D. A. Stein**. 2005. Inhibition of dengue virus serotypes 1 to 4 in vero cell cultures with morpholino oligomers. J. Virol. **79**:5116-5128.
120. **Deas, T. S., C. J. Bennett, S. A. Jones, M. Tilgner, P. Ren, M. J. Behr, D. A. Stein, P. L. Iversen, L. D. Kramer, K. A. Bernard, and P. Y. Shi**. 2007. In vitro resistance selection and in vivo efficacy of morpholino oligomers against West Nile virus. Antimicrob. Agents Chemother. **51**:2470-2482.
121. **Zhang, B., H. Dong, D. A. Stein, and P. Y. Shi**. 2008. Co-selection of West Nile virus nucleotides that confer resistance to an antisense oligomer while maintaining long-distance RNA/RNA base pairings. Virology **382**:98-106.
122. **Stein, D. A., C. Y. Huang, S. Silengo, A. Amantana, S. Crumley, R. E. Blouch, P. L. Iversen, and R. M. Kinney**. 2008. Treatment of AG129 mice with antisense morpholino oligomers increases survival time following challenge with dengue 2 virus. J. Antimicrob. Chemother. **62**:555-565.
123. **Yoo, J. S., C. M. Kim, J. H. Kim, J. Y. Kim, and J. W. Oh**. 2009. Inhibition of Japanese encephalitis virus replication by peptide nucleic acids targeting cis-acting elements on the plus- and minus-strands of viral RNA. Antiviral Res. **82**:122-133.

Chapter 7

Epilogue



Positive-stranded RNA viruses represent the biggest fraction of all known viruses [1] and are responsible for many human and animal diseases. RNA viruses have the highest mutation rate among living species, which explains why they can adapt relatively easy to new environments (reviewed in [2]). As Stuart Nichol [3] wrote "There is a sense that RNA viruses enjoy life in the evolutionary fast lane; however we, slow-moving DNA-based life forms, may have some opportunities to outwit them yet". In order to achieve this, a thorough understanding of the virus life cycle in general and the molecular interactions between the virus and the host is required. Certain steps in the viral life cycle were found to correlate closely with conformational changes in the viral RNA [4]. RNA elements present in the 3' untranslated region (3' UTR) of the genome in particular, are known to undergo such dynamic conformational changes and to play a role in critical steps such as translation and replication. In this thesis, RNA structures and motifs present in the 3' UTR of flaviviruses were studied in more detail, including their role in the virus life cycle.

The 3' UTR of the flavivirus genome is predicted to fold into a complex structure that includes well conserved primary sequences, multiple RNA stem-loop structures as well as a number of RNA pseudoknots [5,6] (chapter 6 of this thesis). Some of these RNA structures have been studied extensively, while others have only been predicted based on phylogenetic studies and computer-aided RNA folding and still require functional characterization.

As shown in chapter 2 and discussed in chapter 6, the well conserved pentanucleotide (PN) sequence CACAG at the top bulge of the 3' terminal stem-loop (3' SL) is an essential element for viral RNA synthesis [7-9]. Unfortunately, the function of this PN motif is not known. Several studies have indicated that it is part of the binding site for certain host proteins such as the eukaryotic translation elongation factor-1α (eEF-1α) in West Nile virus (WNV) [10,11], Y box binding protein-1 (YB-1) in dengue virus (DENV) [12] and La protein in Japanese encephalitis virus (JEV) [13]. Although none of these studies claim that the PN motif is the major interface for this protein-RNA interaction, it is rather surprising that such a conserved sequence appears to be part of the binding site for these different and functionally unrelated proteins. An alternative explanation is that the PN motif is actually part of an RNA kissing or pseudoknot interaction; however, the fact that the yellow fever virus (YFV) PN motif tolerates point mutations at almost every position would also argue against this hypothesis. Recently, the structure of the top of the 3' SL including the PN motif was solved for mosquito-borne, tick-borne and no known vector (NKV) flaviviruses by nuclear magnetic resonance (NMR) spectroscopy [14]. Different from what was expected, the NMR analysis did not yield a conserved RNA structure for the studied viruses. This lack of conservation can be explained by the current technical limitations associated to NMR. Only the very top part of the 3' SL was analyzed, therefore i) the folding might have been incorrect in comparison with the natural situation, ii) relevant long-range RNA-RNA interactions involving the top of the 3' SL structure could not be taken into account. Alternatively, the absence of particular host or viral proteins may

have prevented the RNA from adopting the natural conformation. Our study and those of others have clearly demonstrated that the PN motif is crucial for flaviviruses RNA synthesis and showed that the wild-type CACAG sequence confers a selective advantage. The mutational analysis of this region mainly involved point mutations. Therefore, it will be interesting to perform an *in vivo* RNA SELEX experiment to reveal the actual sequence requirements of the PN motif as well as of the nucleotides immediately downstream. Viable virus mutants could be selected by their ability to form plaques. Continued passaging of poorly replicating mutants isolated from such a screen can potentially identify efficiently replicating "second-site revertant" viruses that may yield valuable information about the function and interaction partners of the PN motif.

The RNA pseudoknot that serves as a stalling site for XRN1, resulting in the generation of a small flavivirus (sf) RNA, is another RNA element in the flavivirus 3' UTR that was described in detail in this thesis [15]. An RNA pseudoknot is essentially an RNA structure that is minimally composed of two helical segments, which are connected by single stranded regions [16]. These tertiary RNA structures were initially recognized in the 3' UTR of turnip yellow mosaic virus [17] and subsequently found to be one of the most widespread structural RNA domains. Currently, pseudoknots have been shown to be essential in the catalytic center of the hepatitis delta virus and related ribozymes, telomerases, the regulation of gene expression by ribosomal frameshifting, IRES-driven translation initiation, and to be involved in the transcription and replication of many RNA viruses (reviewed in [16,18]). The stalling of the host ribonuclease XRN1 by an RNA pseudoknot in the flavivirus 3' UTR (chapters 3 and 4 of this thesis and by others [19,20]), points yet to another function of these RNA structures. Although the role of XRN1 in 5' to 3' RNA decay is well documented (reviewed in [21-24]), these particular RNA pseudoknots are the first tertiary RNA structures that are capable of stalling XRN1. In addition, these results provide an interesting link between RNA structure and pathogenicity, as the sfRNA that results from the incomplete degradation of the viral genome has been demonstrated to play an important role in the pathogenicity of mosquito-borne flaviviruses in cell culture and in mice [15,19,20]. The precise role of the sfRNA in the viral life cycle is currently unknown. It has been speculated that the sfRNA is involved in evading the host innate immune response, for instance by antagonizing cellular RNA sensors like RIG-I and MDA5 [25]. Alternatively, it has been suggested that the sfRNA could serve as a decoy to protect the genomic RNA from cellular proteins or miRNAs that would otherwise bind to the 3' UTR of the viral genome and inhibit viral RNA synthesis or translation [25]. There is currently no evidence that supports any of the above hypotheses. It is, however, interesting to note in this context that the cellular YB-1 protein, which binds to the DENV 3' SL structure, has been suggested to mediate antiviral activity [12].

An intriguing alternative hypothesis is that the sfRNA itself serves as a precursor for a virus-encoded miRNA. The ribonuclease XRN1, that is required for the production of the

sfRNA, is enriched in cytoplasmic P bodies that also harbor proteins like Dicer, GW182 and Argonaute, which are involved in RNA interference [21-23,26]. Fluorescence *in situ* hybridization (FISH) analysis of Kunjin-infected cells suggested accumulation of the sfRNA in these P bodies [19]. Recent experiments have shown that recombinant flaviviruses containing the sequence of a cellular miRNA can produce functional miRNAs upon infection [27,28]. These results demonstrate that flavivirus RNA is somehow able to enter the miRNA biogenesis pathway. However, in apparent contradiction with the above hypothesis, is a progressive decrease in the number of P bodies that has been observed in cells during the time course of DENV-2 and WNV infections [29]. To increase our understanding regarding sfRNA production, a more detailed analysis of the kinetics and the subcellular sites of sfRNA synthesis is required. Important questions that need to be address are: i) How do flavivirus genomes end up in P bodies and are these important for sfRNA generation ii) Are the sfRNAs the final product or are they intermediates that undergo further processing iii) Do sfRNAs or their derivatives leave the P bodies and is this required for their function? Various experiments can be envisioned to deal with these questions. "Pulse-chase"-like experiments using temperature-sensitive viruses impaired in viral RNA synthesis or small molecule inhibitors of flavivirus replication are required to determine the kinetics of sfRNA production and turn-over. Reagents that either induce or disrupt P bodies and the associated stress granules should be tested as they can be useful in determining the role of these subcellular structures in sfRNA synthesis. In addition, expanding the RNA silencing experiments that so far have been limited to XRN1, by including other targets like stress granules markers [e.g. Ras-GTPase-activating protein SH3-domain-binding protein (G3BP), T-cell intracellular antigen 1 (TIA-1) or TIA-1-related protein (TIAR)], P body markers (e.g. RNA decapping enzymes DCP1 and 2) and various proteins involved miRNA production (Dicer as an obvious candidate).

One of the important findings of the research described in this thesis is that sfRNA production is not limited to arthropod-borne flaviviruses, but instead is a rather unique feature of all flaviviruses irrespective of the nature of their transmission cycle. This does not only imply that sfRNA production can be used as an unique additional criterion to define currently unassigned and newly discovered RNA viruses to the Flavivirus genus, but also indicates that the sfRNA production is required in the natural life cycle of all these viruses. It remains to be established whether the sfRNA of the different flaviviruses serves a similar function in the different hosts or whether it has evolved to meet specific virus requirements related to (a) particular host(s). The fact that CFAV produces an sfRNA hints at the possibility that the sfRNA of the arthropod-borne flaviviruses may not only be required in the vertebrate host but also in mosquitoes or ticks. One way to address these questions is by expression of sfRNAs in *trans* in cells that are infected with mutant viruses that no longer produce the sfRNA [19]. Such a system can be used not only to determine whether a heterologous sfRNA is able to complement the defect in pathogenicity

of an sfRNA-minus flavivirus mutant, but will also allow to determine which part of the sfRNA is required for its biological function.

It will be interesting to discover whether any of the conserved sequences or RNA structures in the flavivirus 3′ UTR is actually required for sfRNA function. One possible candidate is the conserved sequence CS2 which currently has no clearly defined function. Deletion of CS2 in mosquito-borne flaviviruses was shown to yield viable mutants, albeit with a slightly delayed replication [30-33]. It is interesting to note that in YFV, deletion of CS2 results in turbid plaques on SW13 cells when compared to the clear plaques observed with the wild-type virus [30]. This turbidity in the plaques can be interpreted as a sign of decreased pathogenicity, similar to what is observed with mutants deficient in sfRNA production. However, it should also be noted that CS2 is absent from TBEV and CFAV.

Flaviviruses are not the only RNA viruses that apparently use a host ribonuclease for the production of viral subgenomic RNA. Plants and protoplasts infected with the positive-strand RNA Red clover necrotic mosaic virus (RCNMV) produce a small viral noncoding RNA, designated SR1f, by incomplete degradation of the genomic RNA 1 possibly by a host enzyme with 5′-3′ exoribonuclease activity [34]. SR1f has no effect on the pathogenicity of the virus but it inhibits translation of the viral proteins resulting in a decrease in negative-strain RNA synthesis [34]. A 58-nucleotide sequence at the 5′ end of this RCNMV is sufficient to stall the host ribonuclease and yield SR1f. There is currently no evidence that XRN1 is involved in SR1f production, but it is interesting to note that the 5′ end of SR1f can fold into a pseudoknot structure that could act as a stalling site for the involved ribonuclease (P.A.G.C. Silva and P.J. Bredenbeek; data not shown).

The involvement of XRN1 in the life cycle of viruses is not without precedent. XRN1 has also been shown to act as a suppressor of viral RNA recombination by rapidly degrading 5′ truncated RNAs that can serve as substrates for viral recombination in tomato bushy stunt virus, a positive-stranded plant RNA virus belonging to Tombusviridae [35].

The last experimental chapter of this thesis describes the construction and characterization of a full-length MODV cDNA that can be used to transcribe infectious MODV RNA. The development of these tools is important in order to truly understand the molecular biology of the flaviviruses and to identify the key factors that determine host range and tropism.

Many unanswered questions still remain. In the end, a full understanding of the 3′ UTR structures and specific conformation adopted during the virus life cycle and the function they perform will require the determination of their three-dimensional structure. In this respect, X-ray crystallography and NMR spectroscopy are major challenges to be conquered in future studies. Such knowledge will provide us with further insight into the regulatory mechanisms behind the viral life cycle and can ultimately reveal targets and offer us, "DNA-based life forms", tools to outwit infection by the rapidly evolving RNA viruses.

REFERENCE LIST

1. **ICTV**. 2005. Virus Taxonomy - Eight Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego.
2. **Domingo, E. and J. J. Holland**. 1997. RNA virus mutations and fitness for survival. Annu. Rev. Microbiol. **51**:151-178.
3. **Nichol, S.** 1996. RNA viruses. Life on the edge of catastrophe. Nature **384**:218-219.
4. **Simon, A. E. and L. Gehrke**. 2009. RNA conformational changes in the life cycles of RNA viruses, viroids, and virus-associated RNAs. Biochim. Biophys. Acta **1789**:571-583.
5. **Markoff, L.** 2003. 5′ - and 3′ - noncoding regions of flavivirus RNA, p. 177-228. *In* T. J. Chambers and T. P. Monath (eds.), The Flaviviruses; Structure, Replication, and Evolution. Elsevier Academic Press.
6. **Markoff, L.** 2003. 5′- and 3′-noncoding regions in flavivirus RNA. Adv. Virus Res. **59**:177-228.
7. **Tilgner, M., T. S. Deas, and P. Y. Shi**. 2005. The flavivirus-conserved penta-nucleotide in the 3′ stem-loop of the West Nile virus genome requires a specific sequence and structure for RNA synthesis, but not for viral translation. Virology **331**:375-386.
8. **Elghonemy, S., W. G. Davis, and M. A. Brinton**. 2005. The majority of the nucleotides in the top loop of the genomic 3′ terminal stem loop structure are cis-acting in a West Nile virus infectious clone. Virology **331**:238-246.
9. **Silva, P. A., R. Molenkamp, T. J. Dalebout, N. Charlier, J. H. Neyts, W. J. Spaan, and P. J. Bredenbeek**. 2007. Conservation of the pentanucleotide motif at the top of the yellow fever virus 17D 3′ stem-loop structure is not required for replication. J. Gen. Virol. **88**:1738-1747.
10. **Blackwell, J. L. and M. A. Brinton**. 1997. Translation elongation factor-1 alpha interacts with the 3′ stem-loop region of West Nile virus genomic RNA. J. Virol. **71**:6433-6444.
11. **Davis, W. G., J. L. Blackwell, P. Y. Shi, and M. A. Brinton**. 2007. Interaction between the cellular protein eEF1A and the 3′-terminal stem-loop of West Nile virus genomic RNA facilitates viral minus-strand RNA synthesis. J. Virol. **81**:10172-10187.
12. **Paranjape, S. M. and E. Harris**. 2007. Y box-binding protein-1 binds to the dengue virus 3′-untranslated region and mediates antiviral effects. J. Biol. Chem. **282**:30497-30508.
13. **Vashist, S., M. Anantpadma, H. Sharma, and S. Vrati**. 2009. La protein binds the predicted loop structures in the 3′ non-coding region of Japanese encephalitis virus genome: role in virus replication. J. Gen. Virol. **90**:1343-1352.
14. **Lescrinier, E., N. Dyubankova, K. Nauwelaerts, R. Jones, and P. Herdewijn**. 2010. Structure determination of the top-loop of the conserved 3′-terminal secondary structure in the genome of flaviviruses. Chembiochem. **11**:1404-1412.
15. **Silva, P. A., C. F. Pereira, T. J. Dalebout, W. J. Spaan, and P. J. Bredenbeek**. 2010. An RNA Pseudoknot Is Required for Production of Yellow Fever Virus Subgenomic RNA by the Host Nuclease XRN1. J. Virol. **84**:11395-11406.
16. **Staple, D. W. and S. E. Butcher**. 2005. Pseudoknots: RNA structures with diverse functions. PLoS. Biol. **3**:e213.
17. **Rietveld, K., R. Van Poelgeest, C. W. Pleij, J. H. Van Boom, and L. Bosch**. 1982. The tRNA-like structure at the 3′ terminus of turnip yellow mosaic virus RNA. Differences and similarities with canonical tRNA. Nucleic Acids Res. **10**:1929-1946.
18. **Brierley, I., S. Pennell, and R. J. Gilbert**. 2007. Viral RNA pseudoknots: versatile motifs in gene expression and replication. Nat. Rev. Microbiol. **5**:598-610.

19. **Pijlman, G. P., A. Funk, N. Kondratieva, J. Leung, S. Torres, L. van der Aa, W. J. Liu, A. C. Palmenberg, P. Y. Shi, R. A. Hall, and A. A. Khromykh**. 2008. A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity. Cell Host. Microbe **4**:579-591.
20. **Funk, A., K. Truong, T. Nagasaki, S. Torres, N. Floden, M. E. Balmori, J. Edmonds, H. Dong, P. Y. Shi, and A. A. Khromykh**. 2010. RNA structures required for production of subgenomic flavivirus RNA. J. Virol. **84**:11407-11417.
21. **Anderson, P. and N. Kedersha**. 2006. RNA granules. J. Cell Biol. **172**:803-808.
22. **Eulalio, A., I. Behm-Ansmant, and E. Izaurralde**. 2007. P bodies: at the crossroads of post-transcriptional pathways. Nat. Rev. Mol. Cell Biol. **8**:9-22.
23. **Garneau, N. L., J. Wilusz, and C. J. Wilusz**. 2007. The highways and byways of mRNA decay. Nat. Rev. Mol. Cell Biol. **8**:113-126.
24. **Sheth, U. and R. Parker**. 2003. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science **300**:805-808.
25. **Fernandez-Garcia, M. D., M. Mazzon, M. Jacobs, and A. Amara**. 2009. Pathogenesis of flavivirus infections: using and abusing the host cell. Cell Host. Microbe **5**:318-328.
26. **Parker, R. and U. Sheth**. 2007. P bodies and the control of mRNA translation and degradation. Mol. Cell **25**:635-646.
27. **Heiss, B. L., O. A. Maximova, and A. G. Pletnev**. 2011. Insertion of microRNA targets into the flavivirus genome alters its highly neurovirulent phenotype. J. Virol. **85**:1464-1472.
28. **Rouha, H., C. Thurner, and C. W. Mandl**. 2010. Functional microRNA generated from a cytoplasmic RNA virus. Nucleic Acids Res. **38**:8328-8337.
29. **Emara, M. M. and M. A. Brinton**. 2007. Interaction of TIA-1/TIAR with West Nile and dengue virus products in infected cells interferes with stress granule formation and processing body assembly. Proc. Natl. Acad. Sci. U. S. A **104**:9041-9046.
30. **Bredenbeek, P. J., E. A. Kooi, B. Lindenbach, N. Huijkman, C. M. Rice, and W. J. Spaan**. 2003. A stable full-length yellow fever virus cDNA clone and the role of conserved RNA elements in flavivirus replication. J. Gen. Virol. **84**:1261-1268.
31. **Men, R., M. Bray, D. Clark, R. M. Chanock, and C. J. Lai**. 1996. Dengue type 4 virus mutants containing deletions in the 3′ noncoding region of the RNA genome: analysis of growth restriction in cell culture and altered viremia pattern and immunogenicity in rhesus monkeys. J. Virol. **70**:3930-3937.
32. **Alvarez, D. E., A. L. De Lella Ezcurra, S. Fucito, and A. V. Gamarnik**. 2005. Role of RNA structures present at the 3′UTR of dengue virus on translation, RNA synthesis, and viral replication. Virology **339**:200-212.
33. **Blaney, J. E., Jr., N. S. Sathe, L. Goddard, C. T. Hanson, T. A. Romero, K. A. Hanley, B. R. Murphy, and S. S. Whitehead**. 2008. Dengue virus type 3 vaccine candidates generated by introduction of deletions in the 3′ untranslated region (3′-UTR) or by exchange of the DENV-3 3′-UTR with that of DENV-4. Vaccine **26**:817-828.
34. **Iwakawa, H. O., H. Mizumoto, H. Nagano, Y. Imoto, K. Takigawa, S. Sarawaneeyaruk, M. Kaido, K. Mise, and T. Okuno**. 2008. A viral noncoding RNA generated by cis-element-mediated protection against 5′->3′ RNA decay represses both cap-independent and cap-dependent translation. J. Virol. **82**:10162-10174.
35. **Cheng, C. P., E. Serviene, and P. D. Nagy**. 2006. Suppression of viral RNA recombination by a host exoribonuclease. J. Virol. **80**:2631-2640.

SUMMARY

Due to the highly mobile and interconnected societies of today's world there are countless opportunities for the spread of infectious diseases. According to the World Health Organization (WHO), infectious diseases are now spreading much faster than at any time in history. RNA viruses in particular, are the causative agents of many of the emerging and re-emerging diseases of the past few decades. Among the RNA viruses, a dramatic increase in frequency and magnitude of flavivirus infections has been observed. This is most likely potentiated by factors like the increase in human population density, urbanization, transportation of goods, animals and agricultural products, global warming and the wider dispersal of competent vectors. Flaviviruses are responsible for important human and animal diseases that are usually characterized by hemorrhagic fever or encephalitis. A brief description of the flaviviruses biology and the burden that these viruses represent is presented in chapter 1. Flaviviruses of major global concern include yellow fever virus (YFV), dengue virus (DENV), Japanese encephalitis virus (JEV), West Nile virus (WNV), and tick-borne encephalitis virus (TBEV). Currently, vaccination to protect humans from disease caused by flaviviruses is limited to YFV, JEV, and TBEV. To reduce and prevent the impact of flavivirus infection on society, vaccines against other flaviviruses (especially DENV) and effective therapies are required. However, this can only be achieved by increasing our knowledge regarding fundamental aspects of the molecular biology of flaviviruses and a better understanding of the interactions between the virus, the host and the vector.

The 3′ UTR of RNA viruses is known to be important for several steps of the viral life cycle, namely in translation, replication and assembly. The flavivirus 3′ UTR contains well conserved RNA sequences and is predicted to fold into a highly complex structure involving several stem-loop structures and RNA pseudoknots. Some of these motifs and structures have been studied in detail and attributed a biological function. The aim of this thesis was to characterize and determine the biological function of some of the RNA elements in the flavivirus 3′ UTR. The 3′ terminal 80 – 90 nucleotides of every flavivirus are predicted to form a conserved stem-loop structure (3′ SL). The 3′ SL is not conserved in the nucleotide sequence, except for the pentanucleotide CACAG in a bulge at the top of the SL, and the dinucleotide "CU" at the 3′ end of the genome. Studies by others using WNV indicated that except for the nucleotides at the 2nd and 4th position, all the other nucleotides of the pentanucleotide motif are required for viral RNA replication. Surprisingly, we discovered that the sequence requirements for the YFV pentanucleotide motif were less strict than for WNV. In chapter 2 we showed that point mutations at either the 2nd, 3rd, or 4th position were generally well tolerated. Only the "G" residue at the 5th position and base pairing of the nucleotide at the 1st position were absolutely required for efficient replication. Although these mutations at the 2nd, 3rd, and 4th position did not

seem to have a significant effect on viral RNA synthesis and virus production, the wild-type pentanucleotide sequence CACAG offers an advantage for YFV-17D in cell culture as the mutant viruses were generally outcompeted by the parental virus upon repeated passaging in competition experiments.

In addition to the positive- and negative-stranded genome length RNAs, the production of a positive-stranded, small flavivirus (sf) RNA in both mammalian and insect cells as well as in mice infected with arthropod-borne flaviviruses is now well documented. The length of these sfRNAs varies from 0.3 kb to 0.5 kb and they are collinear with the distal part of the viral 3′ UTR. It was shown recently that sfRNA production results from incomplete degradation of the viral genome by the host 5′-3′ exoribonuclease XRN1 and that the sfRNA is an important determinant for viral pathogenicity. In chapter 3 we determined the molecular signal in the 3′ UTR that is required for the production of the sfRNA by stalling XRN1. A detailed analysis of sfRNA production in YFV-infected cells revealed that, different from other arthropod-borne flaviviruses, YFV generates not one but two sfRNAs that unexpectedly form a 5′ nested set. The precise 5′ end of the YFV sfRNAs was mapped and found to be just upstream of a previously predicted RNA pseudoknot (PSK3). RNA structure probing and mutagenesis studies supported the actual formation of this RNA pseudoknot and demonstrated that it functioned as the molecular signal to stall XRN1. An important consequence of this emerging picture on sfRNA production and function is that previous reports describing the effects of mutations in the distal part of the flavivirus 3′ UTR solely in the context of the viral genome, have to be re-evaluated in light of the potential effect of these same mutations on either the sfRNA production or function. Furthermore, we propose that abolishing sfRNA generation by simple disruption of this pseudoknot should be carefully analyzed as an additional target to develop flaviviruses vaccines based on attenuated viruses.

Production of sfRNA was previously shown only for the mosquito- and tick-borne members of the flaviviruses. However, the genus Flavivirus also comprises a 3rd group of viruses that do not appear to require an arthropod vector for their transmission. To determine if sfRNA production is restricted to the arthropod-borne flaviviruses or whether it is a hallmark of every flavivirus, we have analyzed sfRNA production in cells infected with no known vector (NKV) flaviviruses as well as with the insect virus cell fusing agent virus (CFAV) that has tentatively been assigned to the Flavivirus genus (chapter 4). From these experiments we concluded that sfRNA production is a hallmark of flaviviruses since all the analyzed NKV viruses, as well as CFAV, produced an sfRNA. Detailed analysis of the molecular determinants of sfRNA production in cells infected with NKV viruses or CFAV is hampered by the lack of an infectious clone for any of these viruses. However, by using *in vitro* assays we were able to show that, like for the arthropod-borne flaviviruses, the host protein XRN1 is likely required for sfRNA production. In addition, we used a Sindbis virus-based expression to determine the sequence requirements for sfRNA production

in these viruses. The presented data predict that also in the 3' UTR of NKV viruses and CFAV an RNA pseudoknot serves as a stalling site for XRN1. These results are not only important in identifying sfRNA production as a new, additional hallmark to assign viruses with a similar genomic organization to the Flavivirus genus; but more importantly, they indicate that the sfRNA may serve an essential and perhaps different function during the life cycle in the mammalian as well as in the insect host of these viruses.

Flavivirus full-length cDNA clones that can be used for the production of infectious RNA are often notoriously difficult to construct due to genetic instability in prokaryotic hosts. In chapter 5 we describe the construction and characterization of the first infectious cDNA clone for a NKV flavivirus. Using the low-copy number vector pACNR1180, that was previously used to construct a stable full-length YFV clone, we have been able to construct and propagate a stable infectious cDNA of Modoc virus. As shown in chapter 5, RNA transcribed from this full-length cDNA clone is highly infectious upon transfection of suitable host cells and produced virus with similar characteristics in cell culture as the parental MODV virus that was used to generate the full-length cDNA copy. This infectious cDNA clone can serve as a valuable tool for a detailed comparison of the life cycle of the apparently mammals restricted NKV flaviviruses versus the less host-restricted arthropod-borne flaviviruses.

Chapter 6 provides an extensive literature review of the published data on structural and functional characterization of the various RNA elements that have been identified within the 3' UTR of mosquito- and tick-borne flaviviruses, as well as of NKV flaviviruses. Special emphasis has been given to the RNA structures that were reported to play a role in viral replication and pathogenicity.

The final chapter of this thesis is a short epilogue in which the results that are presented in the chapters 2 to 5 are discussed against the background of recently published findings and in which future research directions are indicated to address the new hypotheses and questions that emerged from the results described in this thesis.

SAMENVATTING

Door de grote toename van onze mobiliteit en de toenemende verwevenheid van de verschillende samenlevingen zijn er tegenwoordig meer mogelijkheden voor de verspreiding van besmettelijke ziekten. Volgens de Wereldgezondheidsorganisatie (World Health Organization, WHO) is de snelheid waarmee infectieziekten zich momenteel verspreiden groter dan ooit. Met name RNA virussen zijn een belangrijke oorzaak van deze nieuwe, dan wel hernieuwde, en zich steeds verder verspreidende virusuitbraken. Flavivirussen vormen hierop geen uitzondering. Met name de verspreiding van dengue virus (DENV) en West Nile virus (WNV) is in het laatste decennium sterk toegenomen en daarmee ook het aantal geïnfecteerde individuen. Deze toenemende overlast door flavivirusinfecties is waarschijnlijk het gevolg van factoren als de groei van de wereldbevolking, toenemende verstedelijking, dieren- en goederentransport en het broeikaseffect met als gevolg een grotere verspreiding van geschikte vectoren. Flavivirussen zijn verantwoordelijk voor een aantal belangrijke ziekten, welke gepaard kunnen gaan met ernstige hemorrhagische koorts of encephalitis. Hoofdstuk 1 van dit proefschrift geeft een samenvatting van de biologie van flavivirussen. De moleculaire biologie en genexpressie-strategie, alsmede de onderlinge verwantschap van deze virussen en de transmissieroute van de verschillende flavivirussen wordt in dit hoofdstuk beschreven. De voor de mens belangrijkste flavivirussen zijn: gele koorts virus (YFV), knokkelkoorts (ook wel dengue virus genoemd; DENV), Japanse encefalitis virus (JEV), West Nile virus (WNV) en tickborne encefalitis virus (TBEV). Momenteel is bescherming tegen ziekten door vaccinatie alleen mogelijk voor YFV, JEV en TBEV. Om de invloed van flavivirusinfecties op de samenleving te voorkomen dan wel te beperken zijn vaccins en effectieve therapieën tegen andere flavivirussen, met name DENV, noodzakelijk. Om dit te bereiken is een uitgebreide kennis omtrent de fundamentele aspecten van de biologie van flavivirussen, inclusief de interacties tussen het virus, de gastheer en de vector, essentieel.

De sequentie en structuur van de 3' untranslated region (3' UTR) van het genoom van RNA virussen is van groot belang voor verschillende stappen in de virale levenscyclus en speelt een belangrijke rol bij de regulatie van de RNA synthese, translatie en de assemblage van nieuwe virusdeeltjes. De flavivirus 3' UTR kenmerkt zich door een aantal geconserveerde RNA sequenties en een complexe secundaire en tertiaire RNA structuur, die verschillende stem-loop structuren en een aantal RNA pseudoknots omvat. Aan sommige van deze sequentiemotieven en RNA structuren is op basis van experimenteel onderzoek een biologische functie toegeschreven, van andere RNA elementen is de functie nog onbekend. Het doel van het in dit proefschrift beschreven onderzoek was gericht op het vergroten van onze kennis van de functie van de verschillende RNA structuren in de flavivirus 3' UTR. De laatste 80 tot 90 nucleotiden van het 3' uiteinde van

het genoom van ieder flavivirus vormen een geconserveerde haarspeld structuur (3' SL). Met uitzondering van de pentanucleotide sequentie (PN) "CACAG" in een uitstulping aan de top van de 3' SL structuur en het dinucleotide "CU" aan het uiteinde van het virale genoom, is de primaire sequentie van de 3' SL structuur niet geconserveerd. Uit onderzoek aan de PN sequentie van WNV is gebleken dat, met uitzondering van de nucleotiden op de 2e en 4e positie, alle overige nucleotiden van het PN motief noodzakelijk zijn voor efficiënte replicatie van het virale RNA. Hoofdstuk 2 van dit proefschrift beschrijft een gedetailleerde analyse van de PN sequentie in de 3' SL structuur van YFV-17D. Uit onze resultaten blijkt dat conservering van het PN motief in YFV minder noodzakelijk lijkt dan voor WNV. Niet alleen puntmutaties van de nucleotiden op de 2e en 4e positie, maar ook op de 3e positie van het PN motief blijken toegestaan. Alleen het "G" residu op positie 5 en baseparing van de nucleotide op de 1ste positie zijn strikt noodzakelijk voor efficiënte replicatie. Ondanks dat de mutaties op de 2e, 3e en 4e positie van de PN sequentie in celkweek geen significant effect lijken te hebben op virale RNA synthese en virusproductie, blijkt uit competitie-experimenten dat het wildtype (wt) virus toch beter repliceert dan de virussen met een mutatie in het PN motief. Bij herhaald passeren worden de mutante virussen door het niet gemuteerde gele koorts virus verdrongen uit de viruspopulatie.

In de afgelopen jaren is uit verschillende studies met door muggen en teken overgebrachte flavivirussen gebleken dat, zowel in celkweek als in geïnfecteerde muizen, naast het positief- en negatief-strengige genoom RNA ook een klein positief-strengig subgenoom RNA maken. Dit "small flavivirus" (sf) RNA heeft een lengte van 0,3 tot 0,5 kb en is co-lineair met het distale gedeelte van de virale 3' UTR. Onlangs is aangetoond dat dit sfRNA ontstaat als gevolg van onvolledige afbraak van het virale genoom RNA door het 5'-3' exoribonuclease XRN1 van de gastheer. Productie van het sfRNA blijkt een belangrijke determinant voor de pathogeniciteit van het virus. In hoofdstuk 3 worden de karakteristieken van de YFV sfRNAs in detail beschreven. In tegenstelling tot de andere arthropod-borne flavivirussen worden in YFV-17D geïnfecteerde cellen niet één, maar twee sfRNAs geproduceerd, die hetzelfde 5' uiteinde hebben. De positie van dit 5' uiteinde op het virale genoom werd bepaald en bleek juist stroomopwaarts van een eerder voorspelde RNA pseudoknot (PSK3) te liggen. Uit de resultaten van biochemische experimenten en mutagenese studies kan worden geconcludeerd dat deze voorspelde RNA pseudoknot daadwerkelijk wordt gevormd en functioneert als een signaal om XRN1 te blokkeren. Een belangrijk gevolg van deze en andere resultaten met betrekking tot sfRNA productie en –functie is dat eerder gepubliceerde artikelen, waarin de effecten van mutaties in het distale gedeelte van de 3' UTR binnen de context van het virale genoom worden beschreven, kritisch tegen het licht moeten worden gehouden. Het is nu immers onduidelijk of de gerapporteerde effecten gerelateerd zijn aan effecten op de functie van de bestudeerde RNA structuren op het virale genoom dan wel op de functie van het bijbehorende sfRNA. Ook noemenswaardig is de waarneming dat

flavivirussen die niet langer een sfRNA maken een geattenueerd phenotype hebben. Dit opent interessante mogelijkheden voor de constructie van nieuwe of verbeterde flavivirus vaccins.

In eerdere studies is de productie van een sfRNA aangetoond voor flavivirussen die door muggen en teken worden verspreid. Er is echter nog een derde groep van zgn. no known vector (NKV) flavivirussen waarbij voor virusoverdracht geen vector nodig lijkt te zijn. In hoofdstuk 4 wordt het onderzoek naar mogelijke sfRNA productie in cellen die geïnfecteerd zijn met NKV flavirussen beschreven. Deze studie omvat ook het cell fusing agent virus (CFAV), dat alleen in insectencellen groeit en vermoedelijk ook tot de flavivirussen behoort. Uit de verkregen resultaten kan worden geconcludeerd dat sfRNA productie een karakteristieke eigenschap is voor alle flavivirussen. De geanalyseerde NKV flavivirussen en ook CFAV produceren allen sfRNAs. Een analyse van de moleculaire mechanismen van de sfRNA productie van met NKV flavivirus of CFAV geïnfecteerde cellen wordt bemoeilijkt door het ontbreken van een infectieuze cDNA kloon voor deze virussen. Uit de analyse van de *in vitro* experimenten met gezuiverd XRN1 en *in vitro* vervaardigde RNA transcipten als substraat kan echter worden geconcludeerd dat XRN1 ook verantwoordelijk is voor de sfRNA productie van de NKV flavivirussen en CFAV. Door gebruik te maken van een op het Sindbis virus gebaseerd expressiesysteem is vervolgens de RNA structuur in het gebied net stroomafwaarts van de plaats waar XRN1 vastloopt bij de productie van de sfRNAs van de NKV virussen en CFAV onderzocht. Deze experimenten, gecombineerd met een voorspelling van de plaatselijke RNA structuur, suggereren dat ook bij deze virussen een RNA pseudoknot verantwoordelijk is voor het blokkeren van XRN1. Uit de verkregen resultaten kan geconcludeerd worden dat de produktie van een sfRNA een uniek kenmerk is voor alle flavivirussen en dat het als zodanig gebruikt kan worden om virussen met een gelijksoortige genoomorganisatie aan het genus Flavivirus toe te wijzen. Uit deze studie blijkt ook dat het mechanisme voor de productie van flavivirus sfRNAs geconserveerd is. Dit wil echter niet zeggen dat de geproduceerde sfRNAs een zelfde functie hebben in de verschillende gastheren.

Volledige cDNA klonen van flavivirussen worden vaak geplaagd door genetische instabiliteit in prokaryote gastheren. Al naar gelang het virus kan het daarom zeer lastig zijn om een dergelijke kloon te maken. In hoofdstuk 5 wordt de succesvolle constructie en karakterisatie van de eerste infectieuze cDNA kloon van een NKV virus beschreven. Door gebruik te maken van het low-copy number plasmide pACNR1180, dat al eerder is gebruikt om stabiele, volledige YFV cDNA klonen te ontwikkelen, is het gelukt om een stabiel, infectieus cDNA te maken van Modoc virus (MODV). *In vitro* vervaardigde RNA transcripten van deze MODV cDNA blijken na transfectie in een geschikte gastheercellijn infectieus en de getransfecteerde cellen produceren MODV met overeenkomstige karakteristieken als het oorspronkelijke virus. Deze infectieuze cDNA kloon is een belangrijk hulpmiddel voor een gedetailleerde vergelijking van de replicatie signalen en

virus-gastheer interacties tussen de tot zoogdieren beperkte NKV flavivirussen en de minder gastheer gelimiteerde arthropod-borne flavivirussen.

Hoofdstuk 6 bevat een uitgebreid literatuuroverzicht van gepubliceerde data over de structurele- en functionele karakterisatie van de verschillende RNA elementen die geïdentificeerd zijn in de 3' UTR van de arthopod-borne - en NKV flavivirussen. De nadruk in dit overzichtsartikel ligt op de RNA sequenties en structuren, waarvan bekend is dat ze een effect hebben op de replicatie en translatie van het virale RNA of op de pathogeniciteit van deze virussen.

Het laatste hoofdstuk van dit proefschrift is een epiloog, waarin de in hoofdstuk 2 t/m 5 beschreven resultaten worden bediscussieerd in de context van recent gepubliceerde data en waarin suggesties worden gedaan met betrekking tot toekomstig onderzoek aan de geconserveerde sequenties en structuren in de 3' UTR van flavivirussen.

CURRICULUM VITAE

Patrícia Agostinho Gonçalves Costa da Silva was born on April 27 1983 in Leiria, Portugal. In 2001, she finished high school in Natural Sciences at Colégio Dr. Luís Pereira da Costa, Monte Redondo. In the same year, she entered Minho University in Braga, where in 2005 she obtained the degree in Biologia Aplicada (Applied Biology). Within this period, she received an Erasmus fellowship and developed her final project at the Department of Medical Microbiology, Leiden University Medical Center, under the supervision of Dr. P. J. Bredenbeek and Dr. R. Molenkamp. In October 2005 she started at the same department as a PhD student under the supervision of Dr. P. J. Bredenbeek and Professor Dr. W.J. M. Spaan. Her research focused on the characterization and understanding of the biological function of conserved elements in the 3' UTR of flaviviruses. In 2001, Patrícia has obtained a post-doctoral research fellowship to join the Molecular Parasitology Unit in the Institute of Molecular Medicine in Lisbon, Portugal, with Dr. Gunnar Mair.