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

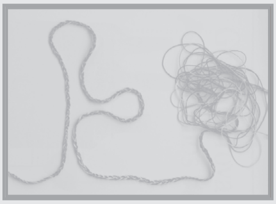
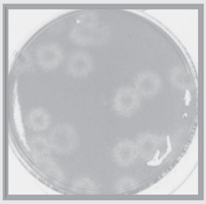
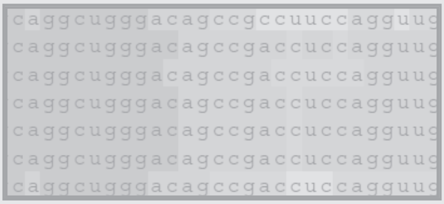
Agostinho Gonçalves Costa da Silva, P.

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

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].

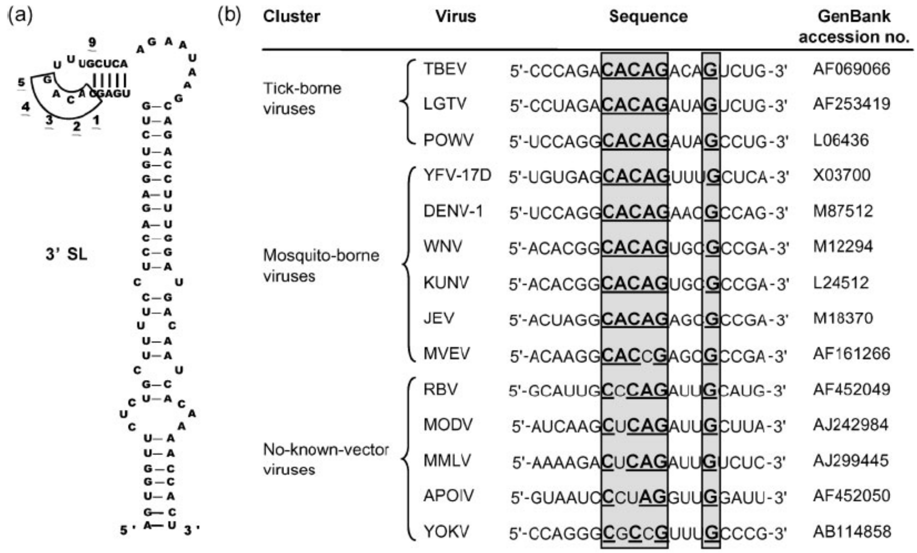


Fig. 1.

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.

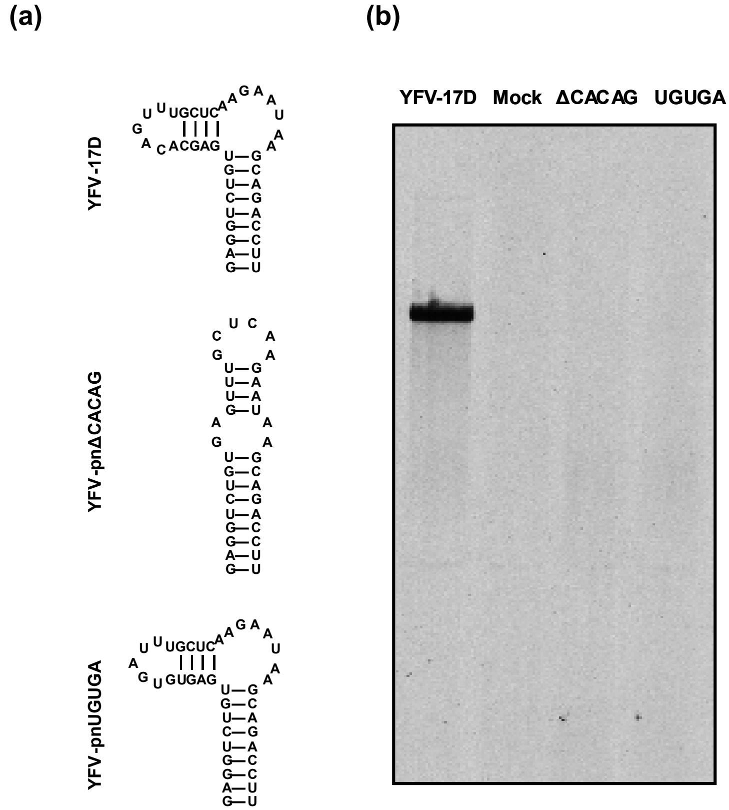


Fig. 2.
A) Secondary RNA structure model for the top of the 3′ SL structure of YFV-17D, YFV-ΔpnCACAG and YFV-pn**<u>UGUGA</u>** as predicted by MFOLD. **B)** Viral RNA synthesis in BHK-21J cells transfected with *in vitro*-transcribed RNA of YFV-17D, YFV-ΔpnCACAG and YFV-pn**<u>UGUGA</u>**.

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>U</u>**CAG) or RBV (C**<u>C</u>**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**<u>CU</u>**AG).

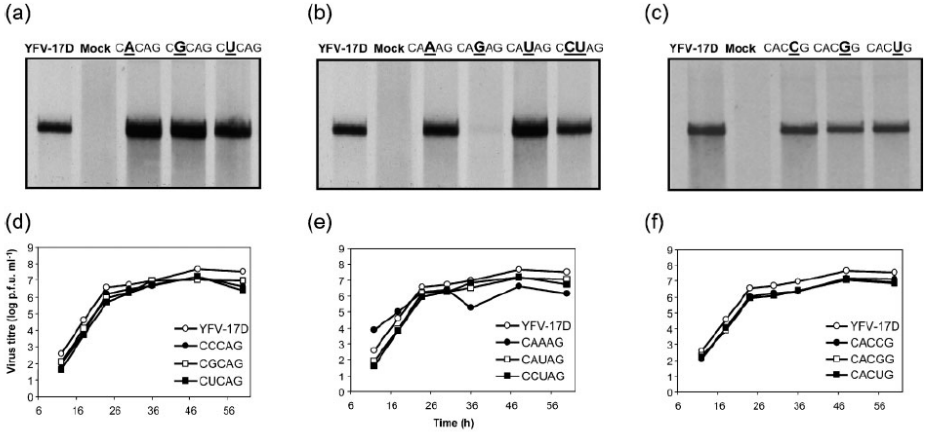


Fig. 3.
Effect of mutations at the 2nd, 3rd and 4thpositions of the PN motif on YFV replication. (**A** – **C**) Analysis of RNA synthesis in BHK-21J cells transfected with RNA of mutants at the 2nd, 2nd and 3rd, and 4th PN positions. The mutants tested are indicated above each lane. (**D** – **F**) 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 4^{th} 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 1^{st} position of the PN motif reveals the importance of base pairing

The C residue at the 1^{st} 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 9^{th} 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 1^{st} and 9^{th} positions is more important than the actual nucleotides. To verify this hypothesis, additional mutants were created by introducing mutations at the 9^{th} position in combination with the 1^{st} 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 9^{th} position was changed to the same nucleotide as in the 1^{st} 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,

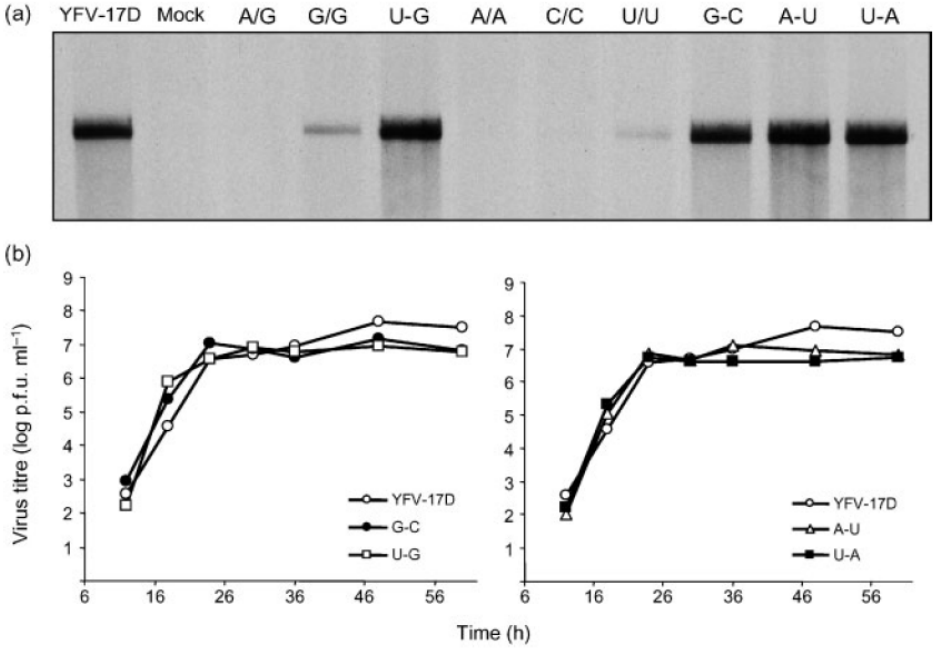


Fig. 4.
Characterization of YFV with mutations at the 1st PN position and the 9th position. **A)** RNA synthesis in BHK-21J cells transfected with RNA of mutants: YFV-pn**A**ACAG9G, YFV-pn**G**ACAG9G, YFV-pn**U**ACAG9G, YFV-pn**A**ACAG9**A**, YFV-pnCACAG9**C**, YFV-pn**U**ACAG9**U**, YFV-pn**G**ACAG9**C**, 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.

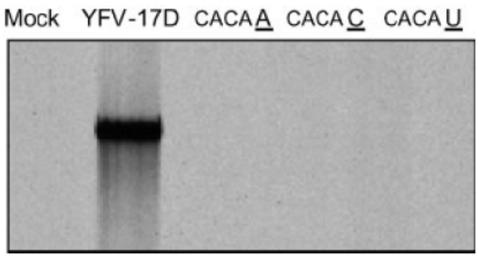


Fig. 5.
RNA synthesis in BHK-21J cells transfected with *in vitro*-transcribed YFV RNA containing mutations at the 5th position of the PN motif.

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**<u>C</u>**CAG	20	3	17	3:17	YFV
2	pnC**<u>U</u>**CAG	18	1	17	1:17	YFV
3	pnCA**<u>U</u>**AG	22	0	22	0:22	YFV
4	pnC**<u>CU</u>**AG	19	0	19	0:19	YFV
5	pnCAC**<u>C</u>**G	19	12	7	12:7	CAC**<u>C</u>**G
6	pn**<u>G</u>**ACAG9**<u>C</u>**	26	20	6	10:3	**<u>G</u>**ACAG9**<u>C</u>**
7	pn**<u>U</u>**ACAG9**<u>G</u>**	20	0	20	0:20	YFV
8	pn**<u>U</u>**ACAG9**<u>U</u>**	19	0	19	0:19	YFV
9	pn**<u>A</u>**ACAG9**<u>U</u>**	29	27	2	27:2	**<u>A</u>**ACAG9**<u>U</u>**
10	pn**<u>U</u>**ACAG9**<u>A</u>**	26	18	8	9:4	**<u>U</u>**ACAG9**<u>A</u>**

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>U</u>**ACAG9G was completely outcompeted by the parental virus (Table 1). The result obtained for YFV-pn**<u>U</u>**ACAG9**<u>A</u>** was essentially similar to that of the mutant virus YFV-pnCAC**<u>C</u>**G. At the 10th passage, the YFV-pn**<u>U</u>**ACAG9**<u>A</u>** 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**<u>G</u>**ACAG9**<u>C</u>** and YFV-pn**<u>A</u>**ACAG9**<u>U</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>U</u>**ACAG9**<u>U</u>**, in which the base pairing was disrupted, was easily outcompeted by YFV-17D.

Although the results obtained with YFV-pn**<u>G</u>**ACAG9**<u>C</u>** and YFV-pn**<u>A</u>**ACAG9**<u>U</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.

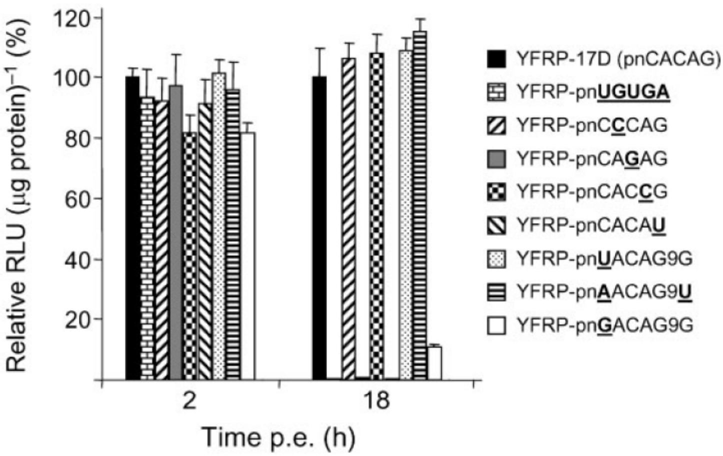


Fig. 6.
Luciferase expression of YFV replicons containing mutations in the PN motif at 2 and 18 h post-electroporation. The following replicons were analyzed: YFRP-17D (pnCACAG), YFRP-pn**UGUGA**, YFRP-pnC**C**CAG, YFRP-pnCA**G**AG, YFRP-pnCAC**C**G, YFRP-pnCACA**U**, YFRP-pn**U**ACAG9G, YFRP-pn**A**ACAG9**U** 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 nucleotide 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.