The handle http://hdl.handle.net/1887/22114 holds various files of this Leiden University dissertation

**Author:** Jiang, Xiaohong  
**Title:** Construction characterization and application of Flavivirus infectious clones  
**Issue Date:** 2013-11-06
Chapter 4

An infectious Modoc virus cDNA as a tool to study conserved 3’-UTR RNA elements in flaviviruses with no known vector

Xiaohong Jiang, Patrícia A. G. C. Silva, Tim J. Dalebout, Charles M. Rice and Peter J. Bredenbeek
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, partly due to the lack of full-length cDNA clones. In this paper, the construction and characterization of the first infectious cDNA clone of an NKV flavivirus, Modoc virus (MODV) is described. The full-length MODV clone pACNR-FLMODV6.1 was shown to be genetically stable in *E. coli*. The viruses derived from this clone exhibited similar plaque morphology and growth kinetics as the parental MODV. Chemical probing was performed on the 3’-UTR region of the MODV genome, to corroborate the highly-ordered RNA structures predicted to be formed in its 3’-UTR. Subsequently, deletions and substitutions were introduced into the full-length clone to study the functional significance of these conserved RNA motifs and structures in the 3’-UTR of the NKV flaviviruses. The Y-shaped structure (region III) and the 3’ stem-loop structure (region IV) were found to be indispensable for viral replication, whereas more upstream regions were not absolutely required for viral RNA synthesis, although they are probably involved in other aspects of virus life cycle, such as cytopathogenicity and/or virus dissemination.
Introduction

Based on phylogenetic analysis, members of the Flavivirus genus can be grouped into three clusters of related viruses that largely reflect their route of transmission. The clusters of mosquito-borne and tick-borne flaviviruses include important human pathogens such as dengue virus (DENV), yellow fever virus (YFV), and tick-borne encephalitis virus (TBEV). The third cluster, on the other hand, comprises less-well studied flaviviruses that have been exclusively isolated from bats (Montana myotis leukoencephalitis virus [MMLV] and Rio Bravo virus [RBV] [1]) and rodents (Modoc virus [MODV] [2] and Apoi virus [APOIV] [3]), and for which no arthropod vector has been implicated in transmission. An increasing number of these no known vector (NKV) flaviviruses have been isolated and sequenced [4-6]. Like arthropod-borne flaviviruses, NKV flaviviruses encode a single large polyprotein that is co- and post-translationally cleaved into three structural and seven non-structural proteins: C, prM/M, E, NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5 [4-6].

Modoc virus (MODV) was initially isolated from white-footed deer mouse (Peromyscus maniculatus) in Modoc County, California [7], and was shown to cause a persistent infection in rodents [8]. The virus has not been implicated in human disease, although a serological survey provided evidence for the occurrence of natural human infection without disease among inhabitants of Alberta, Canada [2]. MODV is neuroinvasive and causes lethal encephalitis in SCID mice and hamsters, rendering it a potential model for the study of flavivirus-induced encephalitis in humans [9]. The viral prM and/or E proteins were shown to be important for the neuroinvasive properties of MODV in SCID mice [10]. Protease and NTPase/helicase/RNA triphosphatase domains were identified in the NS3 of MODV [6], whereas conserved motifs associated with methyltransferase activity and RNA-dependent RNA polymerase (RdRp) domains were mapped to the MODV NS5 protein [6]. The crystal structure of the methyltransferase domain of MODV NS5 was recently solved [11].

The single ORF of the flavivirus genome is flanked by 5'- and 3'-untranslated regions (UTRs). Conserved RNA motifs and structures within the UTRs of mosquito-borne flaviviruses have been recognized as essential elements for genome replication, translation and/or pathogenesis, such as the well conserved stem-loop structure formed by the terminal ~100 nucleotides of the 3’-UTR (3’SL) and the pentanucleotide (PN) motif located in the top loop of the 3’SL, the conserved sequences CS1 and CS2, the RNA pseudoknot interactions PSK1 and PSK2, as well as the RNA structure involved in the production of subgenomic flavivirus RNA (sfRNA) [12-23].

Computer-aided sequence analysis and RNA folding of MODV, APOIV, RBV and MMLV genome revealed the presence of four regions within the distal part of the viral 3’-UTR that appeared to be conserved in RNA structures [5]. Region I was predicted to form a long hairpin...
with a branching stem–loop. The loop sequence within this region (nt 10,294 – 10,298) was proposed to be involved in a pseudoknot interaction with downstream nucleotides [24]. The formation of this particular type of pseudoknot was reported to be essential for the stalling of the host ribonuclease XRN-1 in the production of small flavivirus RNA (sfRNA) [21;22;24;25]. Region II was predicted to form a dumbbell like structure, the 3’ arm of which contains the CS2 sequence [5’-G(A/U)CUAGAGGUAGAGGAGACCC-3’] that is well conserved among NKV flaviviruses and mosquito-borne flaviviruses [5;12]. CS2 was shown to be nonessential for replication of mosquito-borne flaviviruses in cell culture, although YFV-17D and DENV viruses lacking CS2 appear to be less pathogenic [13;14]. Region III was proposed to fold into a Y-shaped structure that bears resemblance to the Y-shaped structure present in tick-borne flaviviruses [26], while mosquito-borne flaviviruses seem to lack similar structures. The two loops of the “Y” structure are formed by a conserved stretch of nucleotides. However, the sequences of the stems carrying these loops are not conserved, but rather show a large number of compensatory base changes [5]. The 3’-terminal nucleotides of the 3’-UTR of the NKV flaviviruses were predicted to form a long stem-loop (3’ SL), a feature conserved among all flaviviruses. At the 5’ side of the 3’ SL, a small stem–loop was predicted which, together with 3’SL, belongs to region IV. Despite the predicted similarity in the RNA folding of this region, sequence conservation is restricted to the pentanucleotide (PN) motif and terminal dinucleotide only. The predicted structures for the 3’-UTR of the NKV flaviviruses have not been verified by RNA structure probing and investigation into their function(s) was hindered by the lack of infectious cDNA clones for the NKV flaviviruses.

This chapter describes the construction and characterization of the first stable full-length NKV cDNA clone, which serves as a versatile tool to address basic questions concerning e.g. replication, tropism, and host-virus interactions of the NKV flaviviruses. In this study, the biological relevance of conserved RNA sequences and predicted RNA structures in NKV 3’-UTR was analyzed in detail. The predicted RNA topology of the MODV 3’-UTR was first corroborated by using chemical probing. Selected mutations were subsequently introduced into the full-length clone to study the functional significance of these conserved RNA elements in the MODV replicative cycle.
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 [13]. The Modoc virus strain M544 was obtained from Prof. J. Neyts (Leuven, Belgium) during collaborative research [10] and was originally purchased from the American Tissue Culture Collection (ATCC, Manassas, USA). Stocks of MODV M544 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 with 5% CO₂ in DMEM/2%FCS. After 3 to 4 days, when cytopathic effect (CPE) became visible, the medium was harvested and centrifuged at 3,000 x g for 5 min to remove cellular 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 [27].

Recombinant DNA techniques and plasmid constructions

General standard nucleic acid methodologies were used throughout this study [28] unless described in more detail. Chemically competent E. coli DH5α cells were used for cloning [29]. 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 [6]).

MODV cDNA was prepared using a one-step RT-PCR system containing a modified M-MLV reverse transcriptase for the cDNA reaction and a mixture of Taq polymerase and Pyrococcus GB-D polymerase (Invitrogen, Carlsbad, USA) for the PCR. RT-PCR reactions contained 1 µg of total RNA from MODV-infected cells. Reaction conditions were as suggested by the supplier. Oligonucleotides were designed based on the published MODV sequence. The most 5’ oligonucleotide (NKV41, fig.1) contained the T7 Ф2.5 promoter [30], 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 AflII restriction enzyme site. Relative positions of the oligonucleotides used for the cDNA reconstruction of MODV are depicted in Fig. 1.

3’-UTR mutants were constructed in the background of pACNR-FLMODV6.1. Quickchange mutagenesis was used for nucleotide deletion/substitution in the construction of pACNR-FLMODV-YFVpk3’, pACNR-FLMODV-YFVpk3, pACNR-FLMODV-YFVpk3’pk3, pACNR-FLMODV-Δ10370-10393, pACNR-FLMODV-ΔCS2, pACNR-FLMODV-AACC, pACNR-FLMODV-BB, pACNR-FLMODV-AA, pACNR-FLMODV-BA as well as pACNR-FLMODV-Δpn; while deletion of the predicted RNA secondary structures, i.e., region I, region II, region III,
region IV, was achieved by engineering unique restriction enzyme recognition sites at the ends of the PCR products that were joined with the use of these restriction sites.

**RNA transcription**

Plasmid DNA for *in vitro* run-off RNA transcription was purified using the Nucleobond AX DNA isolation kit (Macherey-Nagel, Düren, Germany). pACNR-FLMODV plasmids were linearized with *AflII*, followed by proteinase K treatment and phenol/chloroform extraction. Approximately 1 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, the UTP, GTP and CTP concentration was 7.5 mM, whereas the ATP concentration was adjusted to 2 mM. G(5’)ppp(5’).A (NEB, Ipswich, USA) was added as RNA cap analog to a final concentration of 6 mM. After a 2 hr incubation at 37°C, DNase I was added and the incubation was continued for another 15 min. 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 RNA as described previously [27]. For RNA analysis, 2.5 ml (approximately 1.5 x 10⁶ cells) of the transfected BHK-21J cell suspension was seeded in a 10 cm² plate. Total RNA was isolated from the transfected cells at 30 h post electroporation (p.e). Analysis of RNA synthesis by ³[H]-uridine labeling was performed as described before [13]. Tripure (Roche, Mannheim, Germany) 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 [28].

**Northern blotting and hybridization**

Samples containing 10 µg of total RNA from electroporated cells were denatured using formaldehyde, separated in a formaldehyde-containing 1.5% agarose gel, and blotted onto a Hybond-N+ membrane (GE Healthcare, Buckinghamshire, UK). The blots were hybridized with ³²P-labeled oligonucleotides as described previously [31;32].

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

A pBluescript plasmid containing MODV nucleotides 9651-10506 was used to generate a PCR fragment of 383 nts, with oligonucleotide NKV45 that contained a T7 promoter sequence fused to nts complementary to MODV 10244-10269 and M13 reverse oligonucleotide as
primers. The resulting PCR product was used as template for \textit{in vitro} RNA transcription (T7 MEGAscript kit; Ambion). Transcripts were treated with DNase I and purified by LiCl precipitation as recommended by Ambion, and used for SHAPE probing as described previously [22;33]. NMIA-induced modifications in the MODV transcript were identified by primer extension using $^{32}$P-labeled oligonucleotide T3 primer, NKV 20, NKV 53 and NKV 161, which binds to the T3 promoter, nt 10332-10356 (linker between region I and II), nt 10399-10420 (the CS2 motif), and nt 10433-10452 (linker between region II and III and the 5’ arm of the bottom stem of region III) respectively.

**Immunofluorescence microscopy**

At 30 h p.e., control and infected cells were washed once with PBS and prepared for immunofluorescence microscopy as described previously [13]. Commercially available immune ascitic fluid obtained from mice infected with MODV (ATCC, Manassas, USA) was used as primary antibody in a 1:1000 dilution. Alexa Fluor® 488 goat anti-mouse IgG (Invitrogen) was used in a 1:500 dilution as secondary antibody.

**Immunostaining of plaque assays**

After the regular 4 days of incubation, the Avicel overlay was removed and the cells were fixed with 3.5% formaldehyde for 20 min, followed by 3 washes with PBS-Glycine (10mM), and permeabilized with 0.5% Triton X-100 in PBS for 15 min. After 3 additional washes with PBS, the individual 10 cm$^2$-wells were incubated with 200 μl of a 1:500 dilution of the MODV hyperimmune serum in PBS/10% horse serum/0.05% Tween-80 for 1 h, followed by 3 x 5 min washes with PBS/0.05% Tween-80, and subsequent incubation with Goat-anti-Mouse-HRP (Dako, Glostrup, Denmark), diluted 1:1000 in PBS/10% horse serum/0.05% Tween-80 for 1 h. After 3 washes with PBS/0.05% Tween-80 and a short rinse with 0.1 M Sodium Acetate buffer (pH 5.0), cells were incubated with 3-amino-9-ethylcarbazole as substrate in the dark at RT for >30 min. The peroxidase reaction was stopped by washing with milli-Q water when foci of infected cells were clearly visible.

**Viral growth kinetics determination by qRT-PCR**

For analysis of the viral growth kinetics, BHK-21J cells were infected at an MOI of 1 and the medium was subsequently collected and replaced by the same volume of fresh medium at 6-h intervals. Plaque assays were used when comparing the growth kinetics of parental MODV and the MODV-6.1(wt) viruses, whereas qRT-PCR was used to determine the growth kinetics of the MODV 3’-UTR mutants. Briefly, Triton X-100 was added to the collected medium to a final concentration of 0.1% for the disruption of virion structure and thus releasing the viral genome. 5 μl of the treated samples were used without any additional
purification steps in a qRT-PCR reaction using a one-step RT-PCR kit (Qiagen) and primers as previously described [9].

Results

Construction and characterization of a full-length MODV cDNA clone

The construction of full-length cDNAs for the transcription of infectious flavivirus RNA was often hampered by genetic instability of these clones in *E. coli* [34-39]. Given the positive experience with a pACNR1180-derived vector [40] for the construction of a stable full-length YFV-17D clone [13], this low-copy number vector was selected as the vector for the construction of an infectious MODV cDNA.

The full-length MODV cDNA was assembled in a modified pACNR-FLYF17D$_a$ vector [13], in which most of the YFV sequences were deleted while simultaneously creating unique *NotI*, *HindIII*, and *Sall* restriction sites. Three RT-PCR products (Fig. 1A) were sequentially inserted in the 5’ to 3’ direction to create the full-length MODV cDNA. Primer NKV41 contained the T7 RNA polymerase φ2.5 promoter sequence fused to the first 19 nucleotides of MODV 5’-UTR, whereas in oligonucleotide NKV40, nucleotides complementary to the last 28 nt of the MODV genome were fused to an *AflIII* recognition sequence, which served as run-off site for *in vitro* transcription. Upon transfection into BHK-21J cells, *in vitro*-generated RNA transcripts of several plasmids harbouring the full-length MODV cDNA were analyzed for viral RNA synthesis and protein expression by $^{3}$H-uridine labeling and immunofluorescence assays, respectively. Among all the analyzed clones, the transcripts of pACNRFMODV 6.1 yielded higher virus titers compared to the other tested clones (data not shown) and virus derived from this clone was characterized and compared in more detail to the parental MODV.

The genetic stability of pACNR-FLMODV6.1 in *E. coli* strain DH5α was evaluated by repeated passaging. One passage was defined as growing the bacteria for more than 12 h in 2 ml of LB medium containing 50 μg/ml ampicillin, 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). No differences were observed in immunofluorescence and $^{3}$H-RNA labeling (fig.1B and 1C) of cells transfected with RNA transcripts of the original plasmid and the passage 10 plasmid. More importantly, viruses derived from both pACNR-FLMODV6.1 and pACNR-FLMODV6.1-p10 showed very similar growth kinetics and plaque morphology in BHK-21J cells when compared to the parental MODV (Fig. 1D and 1E).
Figure 1. Construction and characterization of the pACNR-MODV6.1 infectious cDNA clone

A) Schematic representation of the construction of the MODV full-length clone.

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 of the MODV insert represents the T7 Φ2.5 promoter. The restriction sites that were utilized for the assembly of the clone as well as the SalI site that was deleted are indicated.

B) Immunofluorescence staining of BHK-21J cells infected with wt-MODV or transfected with in vitro transcribed full-length MODV RNA derived from pACNR-MODV6.1 and the passage 10 plasmid respectively. Cells were fixed at 30 hrs p.i. or p.e. and stained with the MODV hyperimmune serum.

C) Viral RNA synthesis in BHK-21J cells transfected with in vitro transcribed RNA of the pACNR-MODV6.1 (lane 2), pACNR-MODV6.1-p10 (lane 3). RNA isolated from cells infected with parental MODV virus (lane 2) and that from mock infected cells (lane 1) were loaded as controls.

D) Viral growth kinetics. BHK-21J cells were infected with the parental MODV (diamond), virus derived from pACNR-MODV6.1 (square) or virus derived from pACNR-MODV6.1-p10 (triangle) at an MOI of 1; the medium of the infected cells was harvested at the indicated time points p.i. and the viral titer was determined by plaque assays.

The nucleotide sequence of pACNR-FLMODV6.1 was determined and compared to the only other full-length MODV sequence (accession number AJ242984) that was available in GenBank. As summarized in table 1, the 22 nucleotide differences detected when comparing pACNR-FLMODV6.1 and the Genbank entry were scattered throughout the coding sequence. Most amino acid substitutions (8) were present in NS5, the largest virus-encoded protein. Surprisingly, the rather small (254-amino acid) NS4B protein, which is generally well conserved among flaviviruses, contained 4 amino acid substitutions. None of the mutations found in NS3 or NS5 were located within the functional domains that were previously proposed for these viral proteins [6].

MODV 3′-UTR folds into four discrete stem-loop regions.

As schematically depicted in Fig. 2A, MODV 3′-UTR as well as the distal part of the 3′-UTR of other NKV viruses, was proposed to fold into four highly ordered RNA structures [5]. To verify the predicted RNA topology of the MODV 3′-UTR, nt 9,651-10,506 were subjected to selective 2′-hydroxyl acylation and primer extension (SHAPE) probing. This region covers the complete 3′-UTR except for the terminal ~100 nt, which are predicted to fold into the well conserved long stem-loop structure (3′SL) designated as region IV [5] (Fig. 2A). The susceptibility of the nucleotides to N-methylisatoic anhydride (NMIA) treatment is summarized in Fig. 2C, with colors indicating strong (red), weak (yellow) or noisy (blue) reactivity. Apart from nucleotides comprising the CS2 motif, the bases predicted to be involved in basepairing showed little or no reactivity to NMIA, supporting their participation in higher-order RNA structures. Nucleotides predicted to be present in loop structures also correlated rather well with the chemical probing data, although their reactivity varied. Strong reactivity was observed for most of the nucleotides positioned in between region I and II, confirming the presence of a non-structured linker sequence between the depicted stem-loop structures. Discrepancy between chemical probing result and the RNA structure prediction, however, was noticed for nucleotides separating region II and III, which showed weak or no reactivity, instead of high reactivity that was expected for the predicted single-stranded conformation.

To investigate the roles that these primary and secondary RNA structures play in viral life cycle, deletion and substitution were introduced into these regions. The constructed mutants are listed in Table 2, where detailed information is provided on the nucleotides that were deleted or mutated.
Table 1. Summary of the nucleotide differences between the MODV AJ242984 and the MODV clone 6.1 sequence
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.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Position</th>
<th>Nucleotide</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>prM</td>
<td>610</td>
<td>U → C</td>
<td>silent</td>
</tr>
<tr>
<td>Env</td>
<td>1543</td>
<td>U → G</td>
<td>Phe → Leu</td>
</tr>
<tr>
<td>NS1</td>
<td>2776</td>
<td>C → U</td>
<td>silent</td>
</tr>
<tr>
<td>NS1</td>
<td>3089</td>
<td>C → A</td>
<td>Gln → Arg</td>
</tr>
<tr>
<td>NS2A</td>
<td>3529</td>
<td>A → G</td>
<td>Ile → Met</td>
</tr>
<tr>
<td>NS2B</td>
<td>4410</td>
<td>A → G</td>
<td>Glu → Gly</td>
</tr>
<tr>
<td>NS3</td>
<td>4861</td>
<td>A → G</td>
<td>silent</td>
</tr>
<tr>
<td></td>
<td>6312</td>
<td>U → A</td>
<td>Leu → Gln</td>
</tr>
<tr>
<td>NS4B</td>
<td>6837</td>
<td>U → C</td>
<td>Ile → Thr</td>
</tr>
<tr>
<td></td>
<td>7098</td>
<td>G → A</td>
<td>Ser → Asn</td>
</tr>
<tr>
<td></td>
<td>7444</td>
<td>C → G</td>
<td>silent</td>
</tr>
<tr>
<td></td>
<td>7445</td>
<td>C → G</td>
<td>Leu → Val</td>
</tr>
<tr>
<td></td>
<td>7503</td>
<td>U → A</td>
<td>Leu → His</td>
</tr>
<tr>
<td>NS5</td>
<td>7756</td>
<td>G → A</td>
<td>silent</td>
</tr>
<tr>
<td></td>
<td>7767</td>
<td>G → C</td>
<td>Ser → Thr</td>
</tr>
<tr>
<td></td>
<td>7938</td>
<td>G → C</td>
<td>Arg → Thr</td>
</tr>
<tr>
<td></td>
<td>8141</td>
<td>C → G</td>
<td>Arg → Ala</td>
</tr>
<tr>
<td></td>
<td>8142</td>
<td>G → C</td>
<td>Arg → Ala</td>
</tr>
<tr>
<td></td>
<td>8612</td>
<td>C → G</td>
<td>Gln → Glu</td>
</tr>
<tr>
<td></td>
<td>8920</td>
<td>U → A</td>
<td>Ser → Arg</td>
</tr>
<tr>
<td></td>
<td>9120</td>
<td>A → G</td>
<td>Lys → Arg</td>
</tr>
<tr>
<td></td>
<td>9990</td>
<td>A → G</td>
<td>Asp → Gly</td>
</tr>
</tbody>
</table>
Figure 2. Prediction and verification of RNA secondary structure of MODV 3’-UTR

A) Schematic representation of predicted RNA structure of MODV 3’-UTR. Sequences conserved in the 3’-UTR of NKV flaviviruses are highlighted: CS2 (orange), the Pentanucleotide motif (PN) (red) and the UUGG loop in Region III (purple). The sequences (PK’ and PK) involved in the formation of the predicted RNA pseudoknot are highlighted in blue.

B) A representative radiograph of a SHAPE probing experiment. Template RNA was produced and treated with NMIA and analyzed as described in the Material and Method section. Primer extension was performed using primer NKV 161 that binds to nt10,433-10,452 of the MODV genome. Primer extension was performed on NMIA treated (+) or not treated (-) RNA.

C) Results of SHAPE probing for nt 10,252–10,492 within MODV 3’-UTR. NMIA reactivity is summarized with colors indicating differences in reactivity: red, strong reactivity; yellow, weak reactivity; blue, similar reactivity in mock and NMIA treated RNA; no color indicates that these nucleotides were not reactive to NMIA.
The 3'-terminal stem-loop structure is indispensable for MODV replication.

Deletion of the 3'-terminal SL structure (region IV according to the nomenclature of [5]) was lethal to the virus, as no viral RNA synthesis was detected when [3H]-RNA labeling was performed on BHK-21J cells transfected with the construct in which the whole region IV was deleted (Fig. 3A, lane 8), nor could any virus be detected by plaque assays with the medium collected from the transfected cells. In fact, the mere deletion of the MODV 5' CUCAG 3' pentanucleotide (PN) motif (nt 10552–10556) was sufficient to completely abolish viral RNA synthesis (Fig. 3A, lane 9).

The Y-shaped RNA structure upstream of the 3' SL is required for MODV replication.

The proposed RNA folding of region III into a Y-shaped structure correlated well with the results of the SHAPE probing (Fig. 2C). Nucleotides in the loop region were quite accessible

---

Table 2. Characteristics of MODV 3'-UTR mutants constructed and analyzed in this study
Column one lists all the mutants constructed. Details of the nucleotides substituted or deleted are shown in column two, and the positions of changed nucleotides within the full-length genome are indicated in column three.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Number of deleted (Δ) or substituted (x→y) nts.</th>
<th>Position of the mutated nts. in the MODV genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 pACNR-FLMODV-Δregion IV</td>
<td>Δ 93</td>
<td>10506 – 10599</td>
</tr>
<tr>
<td>2 pACNR-FLMODV-Δpn</td>
<td>Δ 5</td>
<td>10552 – 10556</td>
</tr>
<tr>
<td>3 pACNR-FLMODV-Δregion III</td>
<td>Δ 39 (+AGAT)</td>
<td>10447 – 10485</td>
</tr>
<tr>
<td>4 pACNR-FLMODV-AACC</td>
<td>4 (TTGG → AACC)</td>
<td>10460 – 10463</td>
</tr>
<tr>
<td>5 pACNR-FLMODV-BB</td>
<td>6 (TTGGTG → CGCCAA)</td>
<td>10451 – 10456</td>
</tr>
<tr>
<td>6 pACNR-FLMODV-AA</td>
<td>6 (CGCCAA → TTGGTG)</td>
<td>10468 – 10473</td>
</tr>
<tr>
<td>7 pACNR-FLMODV-BA</td>
<td>6 (TTGGTG → CGCCAA) 6 (CGCCAA → TTGGTG)</td>
<td>10451 – 10456 10468 – 10473</td>
</tr>
<tr>
<td>8 pACNR-FLMODV-Δregion II</td>
<td>Δ 64</td>
<td>10360 – 10423</td>
</tr>
<tr>
<td>9 pACNR-FLMODV-Δ10370-10393</td>
<td>Δ 24</td>
<td>10370 – 10393</td>
</tr>
<tr>
<td>10 pACNR-FLMODV-ΔCS2</td>
<td>Δ 20</td>
<td>10400 – 10419</td>
</tr>
<tr>
<td>11 pACNR-FLMODV-Δregion I</td>
<td>Δ 67 (+ AAGCTT)</td>
<td>10253 – 10319</td>
</tr>
<tr>
<td>12 pACNR-FLMODV-YFVpk3'</td>
<td>5 (ATGAC → GCTGT)</td>
<td>10294 – 10298</td>
</tr>
<tr>
<td>13 pACNR-FLMODV-YFVpk3</td>
<td>5 (GTCAT → ACAGC)</td>
<td>10336 – 10340</td>
</tr>
<tr>
<td>14 pACNR-FLMODV-YFVpk3'pk3</td>
<td>5 (pk3'; ATGAC → GCTGT) 5 (pk3; GTCAT → ACAGC)</td>
<td>10294 – 10298 10336 – 10340</td>
</tr>
</tbody>
</table>
to modification by NMIA, while nucleotides in the predicted stems were in general non-reactive, with discrepancies being observed for only a few nucleotides.

This structure was shown to be absolutely required for viral replication, since no RNA synthesis was detected in cells transfected with MODV-ΔregionIII RNA (Fig. 3A, lane 3). To dissect the functional elements within this region, additional mutations were introduced to either change the loop sequence 5’ UUGG 3’ that is conserved in all NKV flaviviruses,
to 5’ AACC 3’ (highlighted in purple, Fig. 2A), or to disrupt the base-pairing of the stem (highlighted in green, Fig. 2A). Substituting the sequence 5’ GUGGUU 3’ (nt. 10,451 – 10,456) with 5’ CGCCAA 3’, which is present at position 10,473-10,468, or vice versa, resulted in the construction of mutants MODV-BB and MODV-AA, respectively. Base pairing of the 5’ stem of region III was predicted to be disrupted in these two mutants, whereas the base pairing potential was restored in mutant MODV-BA, in which the 5’ and 3’ arm of the duplex were swapped. As shown in Fig. 3A, no viral RNA synthesis was detected in cells transfected with either MODV-BB (lane 5) or MODV-AA RNA (lane 6). In contrast, cells that were electroporated with the MODV-BA transcript showed efficient viral RNA synthesis (Fig. 3A, lane 7), underlining the importance of this stem-loop structure for MODV replication.

Accordingly, no virus was detected in the medium of the cells transfected with either the MODV-BB or the -AA transcript, whereas cells electroporated with the MODV-BA transcript produced similar amounts of virus with identical plaque and foci morphology as the wt MODV-infected cells (Fig. 3B). Although the conserved loop sequence 5’ UUGG 3’ was dispensable for viral replication, its mutagenesis dramatically influenced the efficiency of RNA synthesis (Fig. 3A, lane 4). Virus derived from MODV-AACC RNA transfection produced smaller plaques and foci on BHK-21J cells compared to wtMODV (Fig. 3B). From these results it can be concluded that this particular RNA stem-loop structure is important for optimal MODV replication.

The CS2-containing dumbbell-like RNA structure is nonessential for viral replication.

The SHAPE probing results (Fig. 2C) did not fully support the predicted RNA structure for region II. Although the lack of reactivity to NMIA of the nucleotides that formed the bottom and the longer 5’ stem structure of region II was in agreement with the predicted RNA structure, most nucleotides within the CS2 motif reacted strongly with NMIA, indicating that this part of region II was not involved in the formation of secondary or any higher-order RNA structures. The CS2 motif was therefore depicted as a single-stranded loop in Fig. 2C.

Three deletion mutants were constructed to analyze the role of region II in the viral life cycle. In MODV-ΔregionII, the complete sequence of region II (nt 10360-10423) was deleted, while in MODV-Δ10370-10393 and MODV-ΔCS2 the 5’ arm of the dumbbell-like structure and the CS2 motif were deleted respectively. Surprisingly, cells transfected with RNA from these three mutants all showed a similar level of viral RNA synthesis as cells electroporated with the wt MODV control transcript (Fig. 4A). Crystal violet staining of MODV-ΔregionII plaque assays did not show visible plaques on BHK-21J cells whereas plaque assays for MODV-Δ10370-10393 and MODV-ΔCS2 revealed the formation of rather small and not well-defined plaques on BHK-21J cells (Fig. 4B). However, these viruses were able to infect
Figure 4. Characteristics of the MODV Region II mutants

A) Analysis of MODV RNA synthesis by [3H]-RNA labeling. BHK-21J cells were infected (M.O.I. 1) with MODV-6.1(wt) (lane 1), MODV-Δregion II (lane 3), MODV-Δ10370-10393 (lane 4), MODV-ΔCS2 (lane 5) and mock infected (lane 2). RNA was labeled, isolated, and analyzed as described in the Materials and Methods section.

B) Analysis of plaque (upper panel) and infectious center formation (lower panel) on BHK-21J cells by MODV-6.1(wt) (column 1), MODV-Δregion II (column 2), MODV-Δ10370-10393 (column 3) and MODV-ΔCS2 (column 4). Plaque assays and immunostaining methods were performed as described in the Materials and Methods section.

C) Viral growth kinetics of the MODV 3-UTR region II mutants. BHK-21J cells were infected with MODV-6.1(wt), MODV-Δregion II, MODV-Δ10370-10393 and MODV-ΔCS2 at an MOI of 1; the medium of the infected cells was harvested at the indicated time points p.i. and the amount of viral genome copies was determined by qRT-PCR as described in the Materials and Methods section.
Region I is involved in sfRNA production.

The results of the SHAPE probing correlated rather well with the RNA structure predicted for region I, especially for the nucleotides that were predicted to form the stems of this particular RNA structure. Surprisingly little or even no reactivity to NMIA was observed for the nucleotides that were predicted to form the loop regions of the predicted structure. For instance the sequence 5’ CAACC 3’ (nt 10269-10273) that was predicted to form a single-stranded bulge in between two stems did not show reactivity to NMIA (Fig. 2C). Interestingly, the sequence 5’ AUGAC 3’ (nt 10,294-10,298) that was predicted to form an RNA pseudoknot [24] by baseparing with the downstream sequence 5’ GUCAU 3’ (nt 10,336-10,340) showed a weak but reproducible reactivity to NMIA. For the predicted pseudoknot partner 5’ GUCAU 3’, three (GUC) out of the five nucleotides did not react to NMIA, while the other two nucleotides reacted strongly. These results suggested that the predicted RNA pseudoknot was at least not the thermodynamically favored RNA structure under the conditions used for these probing experiments. Despite the ambiguous probing results for these particular nucleotides, these sequences were found to be essential for MODV sfRNA production. No MODV sfRNA could be detected when the pk’ sequence (Fig. 2, nomenclature according to [20]) AUGAC was mutated to GCUGU (sequence involved in pseudoknot formation and sfRNA production in YFV 3’-UTR [20;22]; Fig. 5B). Likewise, mutagenesis of the predicted interacting pk sequence GUCAU to ACAGC (interaction partner for GCUGU in the YFV 3’-UTR) also abolished sfRNA production (Fig. 5B). Surprisingly, MODV-YFVpk3’pk3 in which the above-mentioned mutations were combined to restore the predicted RNA pseudoknot formation, was still unable to produce detectable amounts of MODV sfRNA, as was MODV Δregion I mutant. Viral genomic RNA synthesis, on the other hand, appeared to be only slightly compromised for all of the above-mentioned mutants, as judged by the intensity of the MODV genomic RNA band in the Northern blot and the results of the intracellular [H]-uridine labeling of viral RNA (Fig. 5A and 5B). Interestingly, BHK-21J cells electroporated with in vitro transcribed RNA of MODV-Dregion I and the MODV
Figure 5. Characteristics of the MODV Region I mutants

A) Analysis of viral RNA synthesis was analyzed by $^3$H-RNA labeling. BHK-21J cells were infected (M.O.I. 1) with MODV-6.1(wt) (lane 1), MODV-Δregion I (lane 3), MODV-YFVpk3’ (lane 4), MODV-YFVpk3 (lane 5), pACNR-FLMODV-YFVpk3’pk3 (lane 6) or mock infected (lane 2). RNA was labeled and analyzed as described in the Materials and Methods section.

B) Analysis of MODV sRNA production by Northern blotting. BHK-21J cells were infected (M.O.I. 1) with MODV-6.1(wt) (lane 1), MODV-Δregion I (lane 2), MODV-YFVpk3’ (lane 3), MODV- YFVpk3 (lane 4), MODV-YFVpk3’pk3 (lane 5) and FLMODV-Δregion II (lane 6). Intracellular RNA was isolated at 30hrs p.i. and analyzed by Northern blotting as described in the Materials and Methods section.

C) Analysis of plaque (upper panel) and infectious center formation (lower panel) on BHK-21J cells by MODV-6.1(wt) (column 1) and MODV-Δregion I (column 2). Plaque assays and immunostaining were done as described in the Materials and Methods section.

D) Viral growth kinetics. BHK-21J cells were infected with MODV-6.1(wt), MODV-Δregion I, MODV-YFVpk3’, MODV- YFVpk3, or MODV- YFVpk3’pk3 at an MOI of 1; the medium of the infected cells was harvested at the indicated time points p.i. and the amount of viral genome copies was determined by qRT-PCR as described in the Materials and Methods section.
pk mutants did not show any visual CPE at time points when CPE was clearly detectable in wtMODV RNA transfected cells (data not shown). In addition, no plaques were observed by crystal violet staining on BHK-21J cells for any of these MODV mutants that were unable to produce sfRNA. Nevertheless, immunostaining of cells infected with these MODV region I mutants showed clearly detectable yet small foci (Fig. 5C). The viral growth kinetics of MODV region I mutants on BHK-21J cells revealed that these viruses produced virus titers 5- to 10-fold lower compared to those of wtMODV during the exponential phase of the growth curve (Fig. 5D).

Discussion

Unlike for the mosquito- and tick-borne flaviviruses, no arthropod vectors have been identified for NKV flaviviruses thus far. It is therefore interesting to study the similarities of this group of flaviviruses compared to their vector-borne counterparts, as well as to identify unique features that distinguish them from arbo-flaviviruses. Unfortunately, these basic studies have been hindered by the lack of infectious full-length cDNA clones for NKV flaviviruses, while those of vector-borne flaviviruses have significantly facilitated research on yellow fever virus (YFV) [13;36], Kunjin virus (KUNV) [41], West Nile virus (WNV) [42], dengue virus (DENV) [43;44] and tick-borne encephalitis virus (TBEV) [45]. Construction of stable full-length clones for flaviviruses, however, has not always been straightforward, as genetic instability in E. coli hosts was often encountered. For certain flaviviruses [35-37], labor-intensive in vitro ligation procedures were used initially for the production of full-length cDNA templates for infectious RNA transcripts, until the instability problem was circumvented by cloning the cDNA into low-copy number plasmid vectors [13;34].

The low-copy number plasmid pACNR1180, previously used for the cloning of viral sequences that are not well tolerated by E. coli [13;40], was therefore chosen as the vector for the construction of full-length MODV clone. Clone pACNR-FLMODV6.1 was proven to be genetically stable upon repeated passages in E. coli. Attempts to clone the full-length MODV insert from pACNR-FLMODV6.1 into high copy number plasmids like pBluescript or pUC met no success (unpublished results). Viruses derived from the full-length clone (pACNR-FLMODV6.1) possessed similar characteristics as the parental MODV, in terms of genome replication efficiency, growth kinetics, and plaque morphology, establishing a reverse genetics system for the study of MODV molecular biology. Besides the functional analysis of RNA structures present in the 3'-UTR as described in this chapter, the construction of chimeric flaviviruses by exchanging the prM and E genes of NKV and arthropod-borne flaviviruses could create valuable tools to elucidate the molecular basis for flavivirus host-
range restrictions and pathogenicity, as illustrated by studies using a yellow fever virus in which prM and E were replaced by their counterparts of MODV [10;46]. These studies concluded that the neuroinvasive character of MODV in mice was mediated by prM and E, while the inability of MODV to replicate in arthropod hosts was determined by a post-entry event [10;46]. The availability of an infectious MODV cDNA allows construction of similar MODV chimeric viruses in which the MODV prM and E are replaced by the comparable genes of YFV or other flaviviruses to corroborate the results obtained with the YFV-MODVprM-E chimera.

Flavivirus genomes contain a single open reading frame (ORF) flanked by 5'-UTR and 3'-UTR. The 3'-UTR varies in length from about 350 to 800 nucleotides. This heterogeneity in length is mainly due to the variation in the so-called variable region (VR), which is located proximal to the stop codon. It is characterized by extensive sequence duplications and deletions, even among strains of the same virus, and by an apparent lack of sequence conservation. In contrast, the distal part of the 3'-UTR exhibits more sequence and/or structure similarities among flaviviruses. Since this part of the 3'-UTR contains RNA elements essential for viral translation, replication, and possibly assembly [47-54], it is referred to as the ‘core element’. With a length of only 366 nucleotides, MODV has the shortest 3'-UTR among all flaviviruses sequenced thus far. In comparison to other flaviviruses, the MODV 3'-UTR comprises only the core element and probably represents the minimal requirements for the 3'-UTR of NKV flaviviruses. The region was predicted to fold into four well-conserved RNA structures [5]. In this study, SHAPE RNA probing was used to validate three out of the four predicted RNA structures for the MODV 3'-UTR. Except for the CS2 sequence in region II, probing results were generally in agreement with the predicted structures. These RNA structures were deleted and mutated accordingly, and the effect of these mutations on viral RNA synthesis and virus productions was analyzed to determine the role of these RNA elements in the viral life cycle.

The well-conserved 3'-SL and the PN motif (within region IV) were indispensable for MODV replication, as was shown for vector-borne flaviviruses [13;42;55-58]. Using a primer extension analysis and an alphavirus-based expression system [24], region I was previously predicted to serve as the stalling site for the host ribonuclease XRN-1, resulting in the production of the MODV sfRNA [21;22;24]. As demonstrated in this study, region I is non-essential for viral RNA synthesis or virus production but plays a critical role in stalling the host ribonuclease XRN-1, resulting in the production of sfRNA that is characteristic for every flavivirus [21;22;24]. Although mutation of the sequences proposed to be involved in pseudoknot formation [24] illustrated the importance of these nucleotides in sfRNA production, SHAPE probing results did not confirm their participation in pseudoknot formation. This discrepancy between the biochemical probing results and the model proposed could be due to the fact that the RNA structure probing involved only part of the
viral 3′-UTR and that folding of this particular region might be different in the context of the whole viral genome and the presence of viral and/or host factors interacting with the viral 3′-UTR. An alternative explanation which poses more challenges for future research is that the current model of the folding of region I or in more general terms the XRN-1 stalling site in the 3′-UTR of the flavivirus genome is too simplistic and might be incorrect. After all it is rather surprising that such a relatively simple RNA structure is able to stall the highly processive XRN-1. In vivo probing with a recently described method that allowed structural determination of RNA molecules in a physiologically more relevant setting would hopefully provide more insights into the RNA structure of this region in living cells [59], and it would be interesting to determine the 3D structure of the MODV or a related flavivirus XRN-1 stalling site by NMR.

In between region I (XRN-1 stalling site) and region IV (3′SL), two other conserved highly-ordered RNA structures were predicted for all NKV flaviviruses [5]. Interestingly, the structure of region II is predicted to be similar to the CS2-containing structure in mosquito-borne flaviviruses which appears to be absent in tick-borne flaviviruses, whereas a conserved Y-shaped RNA structure similar to the region III of NKV flaviviruses has been identified in tick-borne flaviviruses and has no counterparts in mosquito-borne flaviviruses. It appeared that the conserved structures of region II and III in NKV flaviviruses also bear functional similarities to their counterparts of the vector-borne flaviviruses. Region II encompassing the CS2 containing dumbbell-like RNA structure is not required for RNA synthesis and virus production. Given the effect of mutations in MODV region II on cytopathogenicity in cell culture, the RNA elements within this region are probably involved in virus-host interactions. Similar observations have been reported for mutations involving the CS2-containing region of DENV and YFV [13;14;60]. The Y-shaped structure (region III) that has a structurally and positionally similar counterpart in TBEV is, just like in TBEV, indispensable for viral replication [13;14;60;61].

Using the first infectious cDNA clone for a NKV flavivirus, this study provides valuable information on the functional significance of conserved RNA sequences/structures in the life cycle of these flaviviruses. It still remains to be elucidated at which stage of the viral life cycle the changes of these motifs and/or structures exert their influence and which viral or cellular proteins are involved in the interaction. Besides, it is still beyond our understanding how an element promoting cytopathology could benefit NKV flaviviruses that in general establish persistent infection in mammalian hosts. Moreover, the identification of insect-specific flaviviruses [62-64] has added a novel angle to the analysis of flavivirus host restriction. Hopefully the access to full-length clones representing each cluster of flaviviruses will help to solve these basic questions in flavivirus biology.
Acknowledgements

We thank Prof. J. Neyts (Leuven, Belgium) for providing the Modoc virus strain M544.
Reference List


102


Chapter 4


[56] Elghonemy S, Davis WG, Brinton MA. 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 2005 Jan 20;331(2):238-46.


