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

Agostinho Gonçalves Costa da Silva, P.

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.