Musashi Binding Elements in Zika and Related Flavivirus 3'UTRs

Michael T. Wolfinger[1,2] and Adriano de Bernardi Schneider[3]

[1]Department of Theoretical Chemistry, University of Vienna, Währingerstraße 17, 1090 Vienna, Austria
[2]Bioinformatics and Computational Biology Research Group, University of Vienna, Währingerstraße 17, 1090 Vienna, Austria
[3]Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, Charlotte, North Carolina, USA

Contact: michael.wolfinger@univie.ac.at - http://www.tbi.univie.ac.at/~mtw

Background and outline

Emerging and re-emerging arthropod-borne viruses such as Japanese encephalitis virus (JEV), Dengue virus (DENV), Yellow fever virus (YFV), and Chikungunya virus (CHIKV) are a growing global health threat. **Zika virus** (ZIKV) is a neurotropic flavivirus (FV) that can cause **congenital infection**, which can result in **microcephaly** and **fetal demise**. Recently, the translational regulator protein **Musashi-1** (Msi1) has been attributed to **promoting ZIKV replication, neurotropism, and pathology** [1]. Msi1 predominantly binds single-stranded UAG motifs in the 3'UTR of RNA [2].

Here we systematically analyzed the thermodynamic properties of **Musashi binding elements** (MBEs) in the 3'UTR of 76 arbovirus genomes *in silico*. Our results indicate that MBEs in the ZIKV 3'UTR occur predominantly in unpaired, single-stranded structural context, thus **corroborating experimental observations** of Msi1 binding affinity with a thermodynamic model of RNA structure formation.

Opening energy and single-strandedness

We use the **ViennaRNA Package** [3] to model the thermodynamics of RNA secondary structure formation. The partition function Z allows for computation of the equilibrium probability of secondary structure s

$$Z = \sum_s e^{-E(s)/RT} \qquad P(s) = \frac{e^{-E(s)/RT}}{Z}$$

The **accessibility** (i.e. the probability that a region $i \ldots j$ along the RNA is single-stranded) can be derived from Z. The **opening energy** (i.e. the free energy required to force the region to be in a single stranded structural context) is computed as $\Delta G_{\text{open}} = -RT \ln P(\text{unpaired})$.

Low opening energy indicates single-strandedness. We compute local pairing probabilities of trinucleotides to assess the likelihood of MBE single-strandedness in a genomic context. Comparison to a large sample of randomized sequences allows to compute a z score for each trinucleotide.

$$z = \frac{\Delta G_{\text{open}}(x) - \mu}{\sigma}$$

Flavivirus 3'UTR mediates pathogenicity

Flaviviruses are small (+)ssRNA viruses of 10-12kB length with highly structured UTRs. Upon infection, accumulation of stable long non-coding viral RNA, subgenomic flaviviral RNA (**sfRNA**), is observed. sfRNAs can modulate cellular function and are linked to pathogenicity. They are stable decay intermediates produced by partial degradation of the viral genome by 5'-3' exoribonuclease Xrn1, which is efficiently stalled at **evolutionary conserved xrRNA** structures.

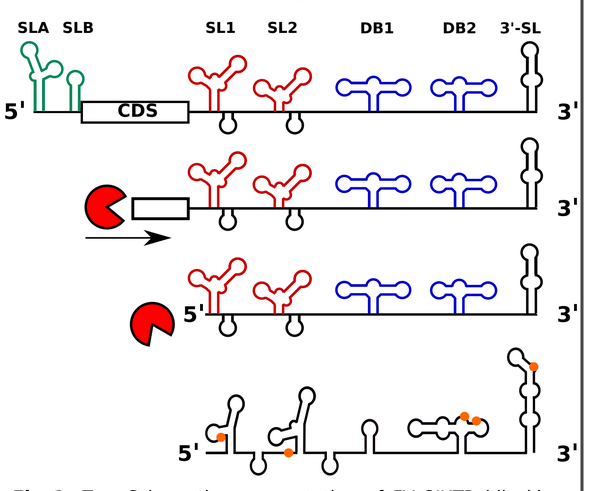

Fig 1. *Top:* Schematic representation of FV 3'UTR hijacking the host mRNA degradation pathway. Conserved xrRNA stem-loop (SL) and dumbbell (DB) elements are located in single or tandem copies within 3'UTRs and stall the host exonuclease Xrn1 (red pac-man). *Bottom:* ZIKV has two SL and one DB element. Location of UAG MBEs are highlighted in orange.

MBEs are highly accessible in ZIKV 3'UTR

We analyzed the accessibility of all trinucleotides in the coding region (CDS) and 3'UTR of ZIKV from Brazil and found a marked difference in the distribution of z scores, suggesting different sequence composition. Musashi-binding **UAG trinucleotides are maximally accesible** in the 3'UTR, which **corroborates experimental studies** [1,4].

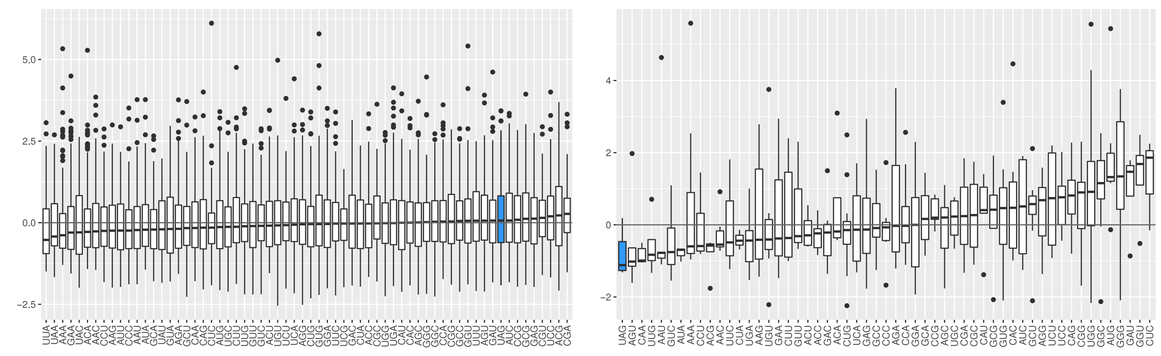

Fig 2. Distribution of z scores for all trinucleotides in the CDS (left) and 3'UTR (right) of ZIKV-BR, sorted by median z score. Interquartile ranges are homogeneous within the CDS region. The MBE motif UAG (blue) is maximally accessible in the 3'UTR.

MBE accessibility in related viruses

To address the broader question whether other viruses have a similar neurotropic potential to ZIKV in the developing fetus, we analyzed MBEs in the 3'UTR of related arbovirus genomes. Low overall z scores indicate unpaired UAG motifs, suggesting high Msi1 affinity. We therefore use it as an **estimator for teratogenicity**.

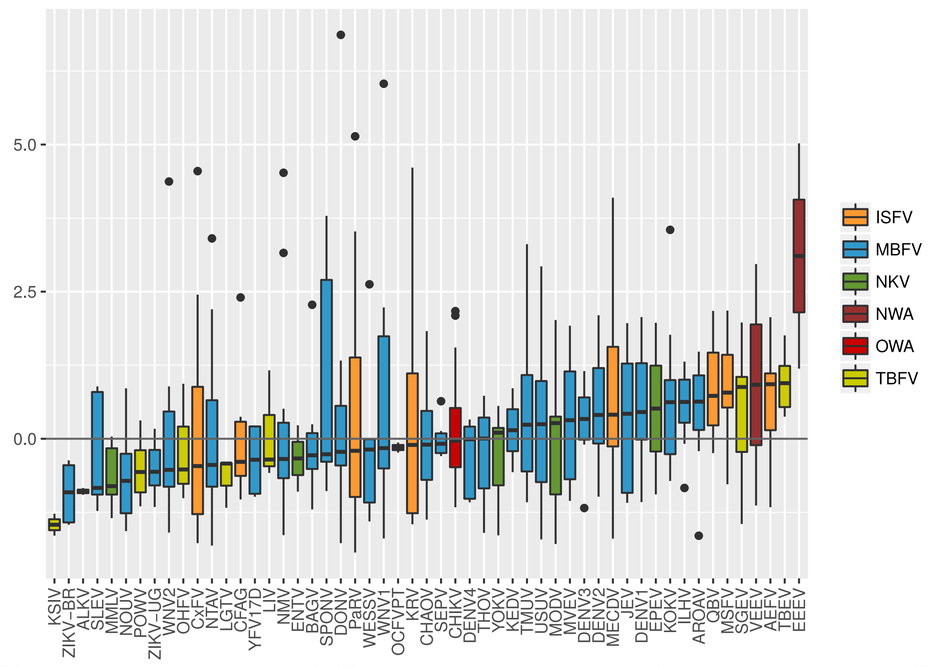

Fig 3. MBE opening energy z scores in the 3'UTR of flaviviruses and alphaviruses. Color indicates virus group (ISFV: Insect-specific flaviviruses; MBFV: Mosquito-borne flaviviruses; NKV: No known vector flaviviruses; NWA: New World alphaviruses; OWA: Old World Alphaviruses; TBFV: Tick-borne flaviviruses).

The Brazilian ZIKV isolate has the lowest median MBE opening energy z scores among mosquito-borne FV, followed by the neurotropic viruses SLEV, WNV, and the tick-borne POWV, which can cause transplacental infection, severe neuropathology and fetal demise [5].

Conserved xrRNAs contain MBEs

We localized MBEs in the 3'UTR of FV and found a **perfectly conserved UAG trinucleotide pair** in the dumbbell (DB) element. Both MBEs appear in an unpaired structural context, rendering them perfect targets for the two Musashi RNA recognition motif domains.

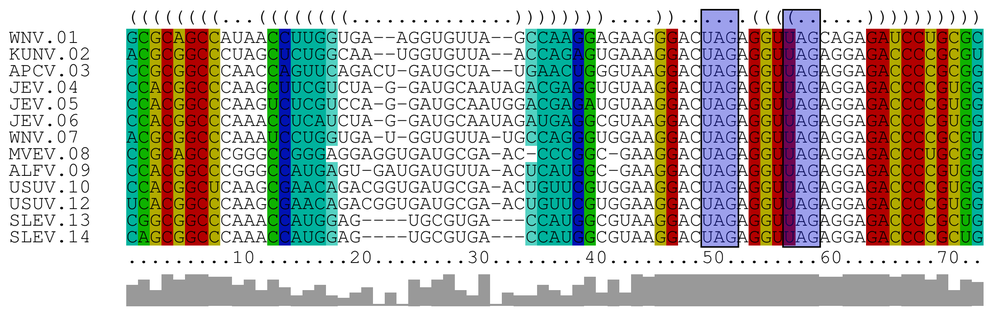

Fig 4. Structural alignment of DB elements in the 3'UTR of Japanese encephalitis group viruses. Nucleotide coloring indicates covariation. Conserved UAG trinucleotides are found in the cental multiloop and the distal hairpin loop (blue boxes). Bars below the alignment indicate+++ the level of sequence conservation.

Conclusion

We employed an established biophysical model of RNA structure formation to **analyze thermodynamic properties of MBEs** *in silico*. Our results underline experimental studies suggesting that ZIKV is not alone in its capacity to cause severe neuropathology [5]. While several tick- or mosquito-borne viral species like Karshi virus (KSIV), Alkhumra hemorrhagic fever virus (ALKV) or Nounané virus (NOUV) line up with ZIKV in our thereretical model, their tropism might have been overseen due to the lack of reported significant outbreaks. However, they appear to have a similar neurotropic potential.

Acknowledgements: This work was partly funded by the Austrian Science Fund FWF project F43.

[1] Chavali, P.L. et al. (2017), **Neurodevelopmental protein Musashi 1 interacts with the Zika genome and promotes viral replication**, *Science*, eaam92433.
[2] Uren, P.J. et al. (2015), **RNA-binding protein Musashi1 is a central regulator of adhesion pathways in glioblastoma**, *Mol Cell Biol*, 35.17, 2965-2978
[3] Lorenz, R. et al. (2011), **ViennaRNA Package 2.0**, *Algorithm Mol Biol* 6:1, 26
[4] Zearfoss, N.R. et al. (2014), **A conserved three-nucleotide core motif defines Musashi RNA binding specificity**, J Biol Chem 289(51):35530-35541
[5] Platt, D.J. et al. (2018), **Zika virus-related neurotropic flaviviruses infect human placental explants and cause fetal demise in mice**, *Sci Transl Med*, 10(426) eaao7090.