Zika virus produces non-coding RNAs using a multi-pseudoknot structure that confounds a cellular exonuclease

Benjamin M. Akiyama1, Hannah M. Laurence1,2,3,*, Aaron R. Massey4,*, David A. Costantino1, Xuping Xie5, Yujiao Yang5, Pei-Yong Shi5, Jay C. Nix6, J. David Beckham4, and Jeffrey S. Kieft1,7,†

1Department of Biochemistry and Molecular Genetics, University of Colorado Denver School of Medicine, Aurora, Colorado, 80045, USA
2Howard Hughes Medical Institute
3University of California, Davis, School of Veterinary Medicine, Davis, California, 95616
4Department of Medicine, Division of Infectious Diseases, University of Colorado Denver School of Medicine, Aurora, Colorado, 80045, USA
5Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, Texas, 77555, USA
6Molecular Biology Consortium, Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, California, 94720, USA
7RNA BioScience Initiative, University of Colorado Denver School of Medicine, Aurora, Colorado, 80045, USA

Abstract

The outbreak of Zika virus (ZIKV) and associated fetal microcephaly mandates efforts to understand the molecular processes of infection. Related flaviviruses produce non-coding subgenomic flaviviral RNAs (sfRNAs) that are linked to pathogenicity in fetal mice. These viruses make sfRNAs by co-opting a cellular exoribonuclease using structured RNAs called xrRNAs. Here, we demonstrate that ZIKV infected monkey and human epithelial cells, mouse neurons, and mosquito cells produce sfRNAs. The RNA structure that is responsible for ZIKV sfRNA production forms a complex fold that is likely found in many pathogenic flaviviruses. Mutations that disrupt the structure affect exonuclease resistance in vitro and sfRNA formation during infection. The complete ZIKV xrRNA structure clarifies the mechanism of exonuclease resistance and identifies features that may modulate function in diverse flaviviruses.
Globalization, urbanization, and climate change contribute to the spread of pathogenic mosquito-borne viruses, typified by the outbreak of Zika virus (ZIKV). ZIKV infection can cause fetal microcephaly and Guillain-Barré syndrome, motivating efforts to understand the molecular drivers of pathology. ZIKV is a (+)-sense single-stranded RNA mosquito-borne flavivirus (MbFV) related to Dengue (DV), Yellow Fever (YFV), and West Nile (WNV) viruses. The structured 3′ UTRs of many MbFVs are the source of non-coding subgenomic flavivirus RNAs (sfRNAs) that accumulate during infection when RNA elements resist degradation by host 5′→3′ exonuclease Xrn1. These sfRNAs are directly linked to cytopathic and pathologic effects; they dysregulate RNA decay pathways and bind cellular proteins important for antiviral responses. Preventing sfRNA production could be a strategy for targeted therapeutics or for generating attenuated virus for vaccines.

sfRNA formation during ZIKV infection has not been reported, therefore we infected multiple cell lines with ZIKV strain PRVABC59, isolated in 2015 from an infected U.S. mainland-Puerto Rico traveler. Northern blot analysis of total RNA isolated from infected cells showed discrete bands containing parts of the ZIKV 3′ UTR, consistent with sfRNAs. Mouse primary neuron infection resulted in very little infectious virus and produced three weak sfRNA bands. Infection of C6/36 (Aedes albopictus mosquito) cells produced two predominant sfRNAs, whereas Vero (monkey) and A549 (human) epithelial cell infection produced additional bands. Different cell types produced different sfRNA patterns but the largest sfRNA was present in all. The significance of this cell type-dependent variation in the sfRNA patterns is unknown, although studies with DV suggest sfRNA production is modulated to enable host adaptation.

The production of ZIKV sfRNAs suggests the existence of Xrn1-resistant structures (xrRNAs) in the viral 3′ UTR. Two areas of the UTR match the sequence pattern and potential secondary structure of known MbFV xrRNAs (fig. 1B&C; fig S1B). The putative xrRNAs are in series near the 5′ end of the UTR, a location and pattern similar to other MbFVs. Xrn1 halting at putative ZIKV xrRNA1 and xrRNA2 would result in sfRNAs of sizes matching the two produced in all cell types tested (fig. 1A; fig S1C). To test if these putative elements are indeed Xrn1 resistant, we challenged in vitro transcribed full-length ZIKV 3′ UTR RNA with recombinant Xrn1. Although multiple sfRNAs are observed during ZIKV infection, in vitro the upstream xrRNA1 quantitatively halts the enzyme. However, a UTR lacking the upstream xrRNA (ΔxrRNA1) allows the enzyme to stop at the downstream xrRNA2. The size of the Xrn1-resistant RNAs matches the infection-produced sfRNAs (fig. S1D). Thus ZIKV contains two xrRNAs and cellular factors could modulate xrRNA function during infection to produce a multiple-sfRNA pattern.

Insight into the structural basis of sfRNA formation comes from a previously-solved structure from Murray Valley encephalitis virus (MVE), but our understanding is incomplete. Specifically, in the MVE xrRNA structure an important pseudoknot (L3–S4, grey line, fig 1B) was not formed, and thus the structure of a fully-folded xrRNA remains unsolved. Also, the MVE structure was of the downstream xrRNA of two in series (xrRNA2) and evidence now suggests there are differences between xrRNA1s and...
xrRNA2s (16, 18); xrRNA1s are predicted to have a more stable fold, and WNV infection models suggest loss of xrRNA1 has a greater effect on viral infection (4, 24). Fully understanding sfRNA formation therefore requires a detailed structural description of a fully-folded xrRNA1. ZIKV xrRNA1 is ideal for such studies as it shows robust Xrn1 resistance in vitro (fig. 1E), formed sfRNAs in all tested cell types (fig. 1A), its pseudoknot is predicted to be stabilized by four consecutive G-C base pairs, and ZIKV is of pressing medical importance.

We solved the structure of ZIKV xrRNA1 by x-ray crystallography (fig. 2A; table S1), using the MR766-NIID African isolate sequence (PRVABC59 xrRNA1 is 97% identical in this region) with sequence alterations to improve transcription and crystallization (fig. S2). This RNA maintained Xrn1 resistance and thus is correctly folded (fig. 2B). All 71 nucleotides were visible (fig. S3), and unlike the MVE xrRNA2 structure the RNA is fully folded; a pseudoknot is formed between L3 and S4 completely encircling and constraining the 5′ end of the RNA (fig. 2C). In both the MVE and ZIKV xrRNA structures the fold is organized around the P1-P2-P3 three-way junction; the P1 and P3 helices form a ring through which the 5′ end passes from one side of the structure to the other, positioned by nucleotides at the 5′ end base-pairing to nucleotides in the junction and a U4●A24-U42 base triple (fig. 2D). The presence of these interactions in both structures strongly suggests their functional importance; disrupting them may be a way to attenuate diverse MbFVs.

The ZIKV xrRNA1 structure reveals multiple previously unobserved interactions. A37 and U51 form a reverse Watson-Crick (or trans) long-range base pair that closes the ring structure to “lasso” the RNA that passes through (fig. 3A). Contrary to predictions, A37 is flipped out of the helix, a position created by an unexpected structure in the P3-L3 stem-loop that contains a G-U wobble pair between G38 and U28 and a Hoogsteen pair between A36 and U29 (fig. 3B). The A37-U51 base pair led us to hypothesize that this interaction may precisely define the size of the ring, which is 14 nucleotides in the ZIKV xrRNA1. Indeed, sequence alignment shows that 30 out of 33 confirmed or putative xrRNAs have Watson-Crick base-pairing partners 14 nucleotides apart in analogous positions (fig. S4A&B). Although these are apparent Watson-Crick partners, the structure suggests that they form non-canonical pairs. Indeed, YFV, Sepik virus, and Wesselsbron virus have G bases at both positions, suggestive of a G●G pair and thus possible alternate conformations of this long-range base pair.

To explore this long-range interaction, we altered the A37-U51 pair by mutation and tested for Xrn1 resistance in vitro (fig. S4C). Substitution of either nucleotide individually had very little effect and substitution of both bases to convert the interaction to either a C-G or G●G resulted in only a moderate decrease in resistance. Overall, it is not clear why sequence conservation shows a preference for apparent Watson-Crick pairing partners; perhaps certain pairs improve Xrn1 resistance in the context of specific xrRNAs or they may contribute to functions other than Xrn1 resistance. Consistent with this, the preference for a purine at the first position of this interaction and a pyrimidine at the second position is reminiscent of a conserved trans R15:Y48 base pair observed in tRNAs (25). G●G pairs have been observed at this position in cysteine tRNAs, and tolerance for switching between G●G and R15:Y48 pairs has been shown to be dependent on the RNA sequence context (26).
The A37-U51 base pair defines the size of the ring and continues the progression of stacked bases from the P1 stem to the L3-S4 pseudoknot (fig. 3A&C). Specifically, A52 stacks on this pair but is not itself paired. This base is universally a purine (fig. S4A&B), likely to maximize stacking potential. The next base, A53, is extruded to form a crystal contact (fig. S5), but in solution it likely stacks under A52, supported by the presence of electron density where A53 could stack consistent with a 1,6-hexanediol molecule from the crystallization solution (fig. S5A). If stacked within the helix, A53 would continue the progression of stacked purines through S4 to the pseudoknot and could form an A•A pair with A35, providing additional stability (fig. 3C).

As predicted, the L3-S4 pseudoknot contains 4 consecutive G-C Watson-Crick base pairs (fig. 3C). The four G’s in L3 are made accessible by a U-turn, a motif found in other loops including tRNA anticodon loops. To test the functional importance of the ZIKV xrRNA1 pseudoknot we generated mutant RNAs (Pkmut1, Pkmut2) and tested them for Xrn1 resistance in vitro (fig. 3D). Disruption of the pseudoknot severely decreased Xrn1 resistance to the same degree as a previously-characterized mutation that disrupts the three-way junction (C22G) (19, 21). Restoring the pseudoknot (Comp) returned nearly wild-type Xrn1 resistance; thus the pseudoknot is critical for ZIKV xrRNA1 function.

The effects of structure-based mutations to xrRNA1 on Xrn1 resistance in vitro led us to predict that the same mutations would alter sfRNA formation during infection; therefore, we generated mutant ZIKV using an infectious clone based on strain FSS13025 (identical to PRVABC59 in xrRNA1) (27). Mutations in xrRNA1 analogous to C22G and Pkmut1 resulted in highly reduced sfRNA formation during infection, matching the in vitro Xrn1 resistance result (fig 3E). Interestingly, both sfRNAs were reduced, emphasizing the importance of xrRNA1 to sfRNA formation overall. A mutation analogous to A37C had no discernable effect on sfRNA formation, consistent with this mutant’s Xrn1 resistance in vitro (fig. S4C). We also assessed the effect of Xrn1 knockdown on ZIKV sfRNA production during infection of human cells (fig. S6); Xrn1 knockdown results in a reduction of one sfRNA species and a change in the overall sfRNA pattern, however the persistence of some sfRNA bands suggests Xrn1 could be redundant with other exonucleases in producing ZIKV sfRNAs. Importantly, Xrn1 resistance in vitro and ZIKV sfRNA production during infection are linked by the strong agreement between the effects of RNA mutations in both assays, demonstrating that the RNA structure is necessary for sfRNA formation.

One of the most functionally essential interactions is the L3-S4 pseudoknot, whose formation is accompanied by substantial structural differences compared to the MVE xrRNA2 structure (fig. 4A). The helical element created by the pseudoknot forms a continuous stack with the P4 helix (fig. S7A&B), placing the P4-L4 stem-loop in a different position compared to the MVE xrRNA2 structure (fig. 4B) and differing from that predicted by structural modeling (fig. S7C&D). P4’s stacking on the pseudoknot likely stabilizes the overall fold, but there are no obvious sequence-specific roles for this element in forming the structure. However, the P4-L4 stem-loop contains numerous conserved bases found in other MbFVs (fig. S8), suggesting it is important for some aspect of sfRNA formation or downstream function.
Using the fully-folded structure, we constructed a new model of the ZIKV xrRNA1 interacting with Xrn1 from D. melanogaster (with bound substrate analog) using previously reported biochemical information (19, 28), and the electrostatic charge distribution and shape of the enzyme’s surface (fig. 4C). The model shows extensive contacts between the enzyme and the RNA. A model of the partially folded MVE xrRNA with Xrn1 suggested that the RNA ring structure prevents Xrn1 from unwinding the structure (21), but in that form the P4 element projected away from the surface of Xrn1 (fig. 4C). In contrast, the fully folded ZIKV xrRNA1 closely matches the contours of an electropositive patch, with P4 contacting a conserved charged region on Xrn1’s winged helix domain. This putative P4-Xrn1 interaction could serve to stabilize the xrRNA’s pseudoknot interaction and thus enhance resistance to the enzyme, or P4 may form sequence-specific interactions with Xrn1 or with Xrn1-bound proteins. Also, because the winged helix domain is important for processive Xrn1 function (29), the bound ZIKV xrRNA may prevent conformational changes in the enzyme important for processivity. The new structure and derived hypotheses point the way to future studies to fully understand the formation and function of ZIKV sfRNAs, with implications for the development of interventions or vaccines.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the members of the Kieft Lab for discussions, and C. Musselman, and M. Stone for critical reading of this manuscript. We thank R. Soto from the M. Garcia-Blanco and S. Bradrick laboratories for sharing their protocol for sfRNA blots. The data presented in this manuscript are tabulated in the main paper and in the supplementary materials. Coordinates and structure factors have been deposited with PDB accession code 5TPY. The UC Denver X-ray Facility is supported by UC Cancer Center Support Grant P30CA046934 and NIH grant S10OD012033. “The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. B.M.A. is supported by National Institutes of Health fellowship F32GM117730. J.S.K. was an Early Career Scientist of the Howard Hughes Medical Institute and is supported by grants R35GM118070 and R01GM081346 from the National Institutes of Health. H.M.L. was supported by the Howard Hughes Medical Institute (HHMI) and Burroughs Wellcome Fund (BWF) as an HHMI-BWF Medical Research Fellow. A.R.M. and J.D.B. are supported by University of Colorado School of Medicine and Department of Medicine Institutional Pilot Funds. X.X. was awarded with a postdoctoral fellowship from Novartis Institutes for BioMedical Research. P.Y.S. is supported by a University of Texas Medical Branch startup award, a University of Texas STARs Award, and NIH grant R01AI087856.

References and Notes


47. (Schrödinger, LLC).


One Sentence Summary

ZIKV uses a convoluted RNA fold during infection to produce non-coding RNAs associated with pathogenic outcomes in flavivirus-infected fetal mice.
Fig. 1.
ZIKV produces sfRNAs by exonuclease resistance. A. Northern blot of purified RNA from mock (m)- and ZIKV (zv)-infected human A549, Vero, mosquito C6/36, and cultured primary mouse neuron cells. Probe was complementary to the putative dumbbell sequence in the ZIKV 3′ UTR. gRNA: genomic RNA. sfRNA1: band corresponding to largest sfRNA. Inset: high-contrast view of the region in the dashed box. Infections and northern blots were confirmed in three independent experiments. B. Secondary structure diagram with sequence conservation in known stem-loop (SL)-type xRNA from MbFVs. Universally conserved nucleotides and the position of tertiary interactions are shown. The P2, P3, and P4 stems and associated loops contain variable sequences (var). C. Sequence alignment of several known

*Science. Author manuscript; available in PMC 2017 June 19.*
or putative MbFV xrRNA sequences with regions colored as in panel B. D. Cartoon of the predicted ZIKV 3′ UTR secondary structure, with SL and dumbbell (DB) elements. In WNV _Kun_, the SL1 and SL2 RNAs are xrRNAs (4, 19, 21).

E. Ethidium-stained gel of _in vitro_ transcribed ZIKV 3′ UTR or a 3′ UTR lacking the first putative xrRNA1 element (ΔxrRNA1) treated with recombinant RppH (to generate a 5′ monophosphate) and recombinant Xrn1. The smaller band indicates Xrn1 resistance. The presence of Xrn1 activity in the absence of RppH is due to spontaneous loss of the 5′ pyrophosphate moiety that occurs at some level even in the absence of the RppH enzyme, as has been previously observed (19).
Fig. 2. Structure of the ZIKV xrRNA1. A. Secondary structure diagram of the crystallized RNA. Lower case letters represent sequences altered to facilitate RNA expression and crystallization. Colored lines indicate interactions discussed in the text. B. Xrn1 resistance assays of wild-type (Wt) ZIKV xrRNA1 and the crystallized RNA (Cryst) using $3'$ end-labeled RNA (yellow circle). The percent of the total RNA that formed an Xrn1-resistant band (listed below the gel) was quantified and reported as the average $\pm$ one standard deviation from the mean of three independent experiments. C. Ribbon representation of the structure of ZIKV xrRNA1, colored to match panel A. Magnesium ions are shown as yellow spheres. D. Detail of interactions at the $5'$ end of the RNA. Residue C22 (cyan) contacts the phosphate backbone of neighboring residues, setting up a kink in the RNA critical for folding. Residues U4, A24, and U42 (green) form a base triple interaction orienting the $5'$ end. Residue G3 forms a long-range base pairing interaction with residue C44. Residue G2
was mutated from a U to promote transcription; the wild-type sequence is predicted to form a base pair with residue A45 (predicted position change indicated by arrow) (21).
Fig. 3.
Details of the ZIKV xrRNA1 structure. A. Detailed view of the A37-U51 base pair (red) and intervening nucleotides (blue), which circle the 5′ end of the RNA. Other nucleotides discussed in the text are labeled. B. A37-U51 base pair (red) and intervening nucleotides (blue) are highlighted. The box shows the previously predicted secondary structure of the P3-L3 stem-loop. Leontis-Westhof nomenclature is used to indicate non-canonical pairing (30). Inset displays details of all three non-canonical base pairs; electron density is displayed at the 2 sigma contour level. C. The L3–S4 pseudoknot with the P4 stem co-axially stacked. Color matches figure 2A&C. D. Xrn1 resistance assays of pseudoknot mutants and a mutant known to disrupt xrRNA folding (C22G) (19, 21). Quantitation of resistance from three experiments is shown, determined as in figure 2B. E. Northern blot of viral RNA isolated from viral infection with wild-type virus and virus mutated in the xrRNA1 structure. The
mutants are labeled to match the analogous mutants in panel D and Figure S4C; corresponding positions in the viral RNA are provided in the box.
Fig. 4.
Model of ZIKV xRNA-Xrn1 interaction. A. Comparison of the S4 region (orange) and adjacent in the partially folded MVE (left) and fully folded ZIKV (right) xRNAs. B. Overlay of the MVE (cyan) and ZIKV (yellow) structures, showing the change in the position of the P4/L4 hairpin. C. Models of the MVE (top) and ZIKV (bottom) xRNAs docked onto the surface of Xrn1 colored by electrostatic potential (blue = positive, red = negative). Structural features are labeled.