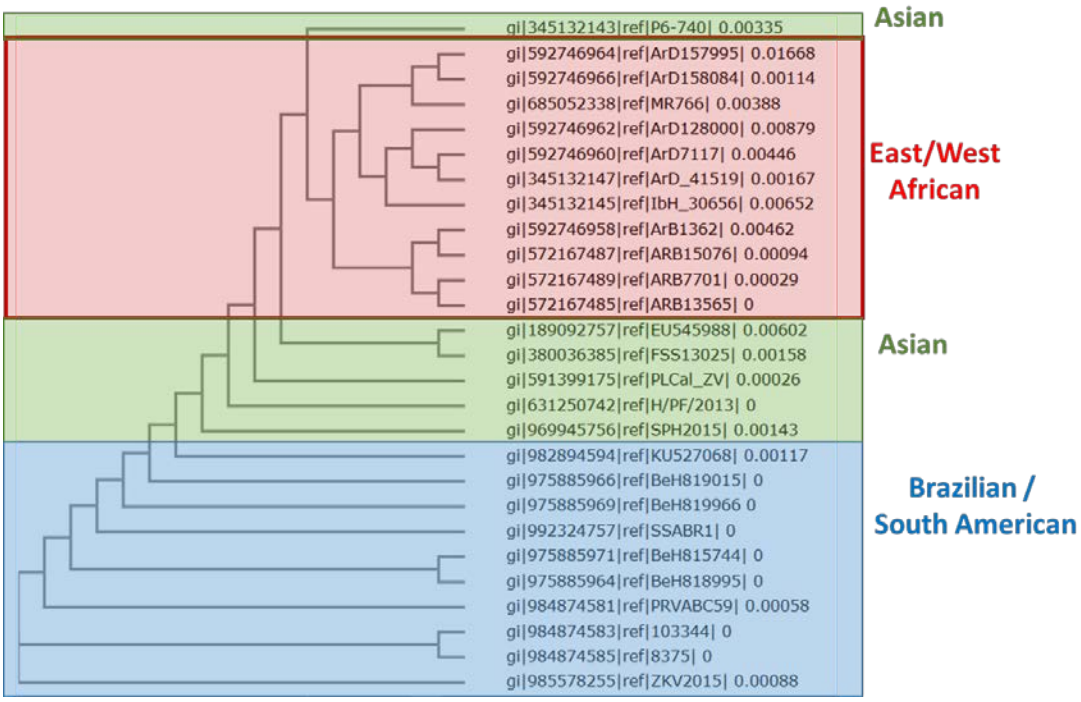
Supplementary Information

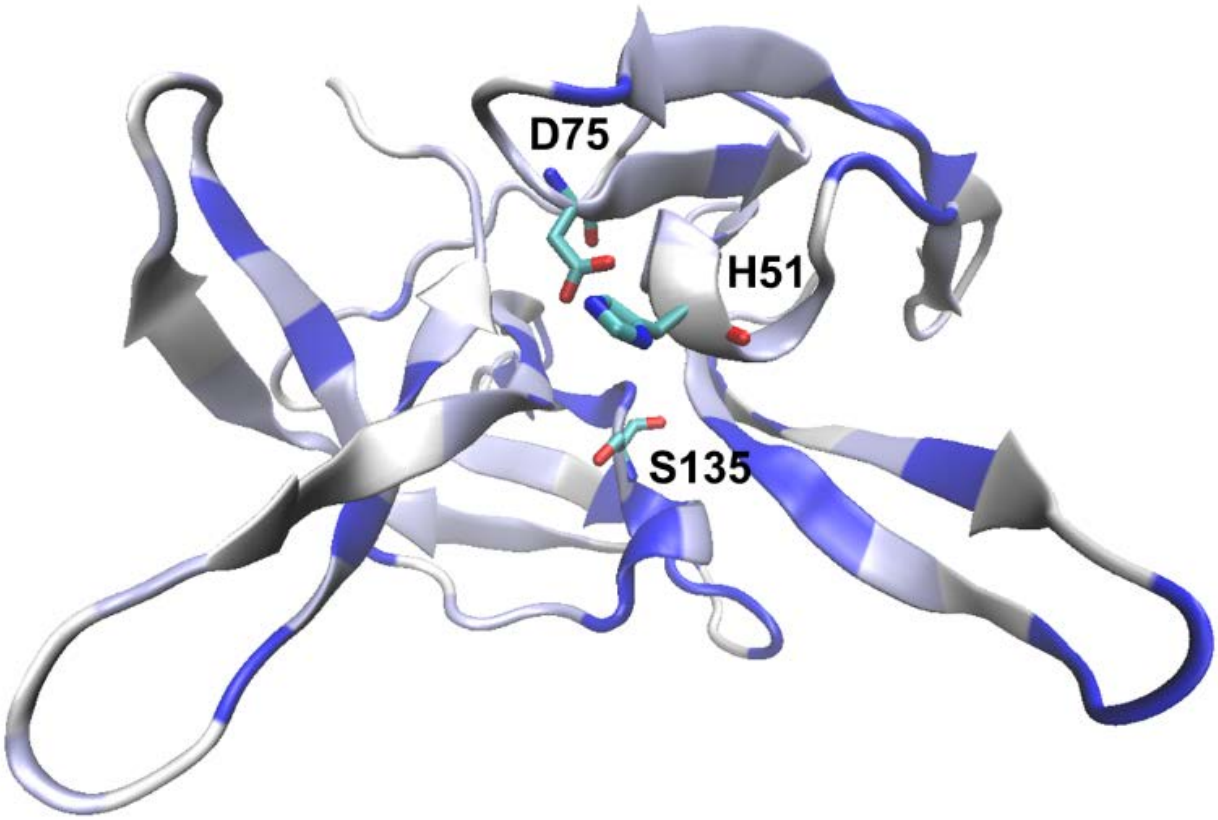
Predicting Zika Virus Structural Biology: Challenges and Opportunities for Intervention

Bryan D. Cox, Richard A. Stanton, and Raymond F. Schinazi

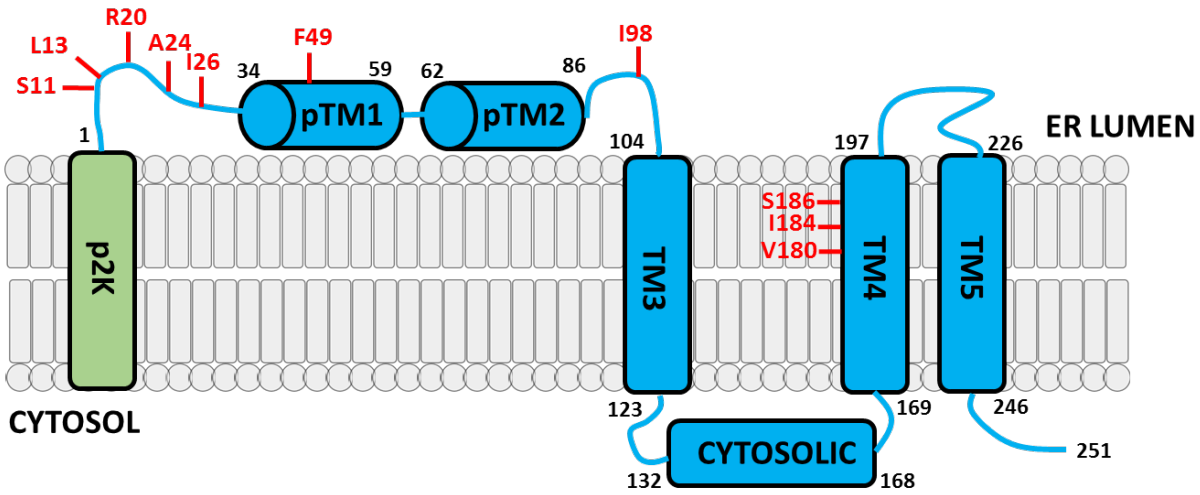
GI Numbers and codes for Zika Virus genomes and Polyprotein-Derived Phylogenic Tree

Asian 1966, P6-740 (gi 345132143); Asian 2007, EU545988 (gi 189092757); Asian 2010, FSS13025 (gi 380036385); Asian 2015, SPH2015 (gi 969945756); Brazil 2016, 103344 (gi 984874583); Asian 2015, 8375 (gi 984874585); Brazil 2016, PLCal_ZV (gi 591399175); Brazil 2016, KU527068 (gi 982894594); Brazil 2016, ZKV2015 (gi 985578255); Asian 2015, PRVABC59 (gi 984874581); Brazil 2016, BeH815744 (gi 975885971); Brazil 2016, BeH818995 (gi 975885964); Asian, H/PF/2013 (gi 631250742); Brazil 2016, BeH819015 (gi 975885966); Brazil 2016, BeH819966 (gi 975885969); Brazil 2016, SSABR1 (gi 992324757); E African 2001, ArD157995 (gi 592746964); W African 1997, ArD128000 (gi 592746962); W African 1968, IbH_30656 (gi 345132145); W African 1968, ArD7117 (gi 592746960); W African 1984, ArD_41519 gi (345132147); E African 1947, MR766 (gi 685052338); E African 2001, ArD158084 (gi 592746966); E African 1968, ArB1362 (gi 592746958); E African 1980, ARB15076 (gi 572167487); E African 1976, ARB7701 (gi 572167489); E African 1979, ARB13565 (gi 572167485)



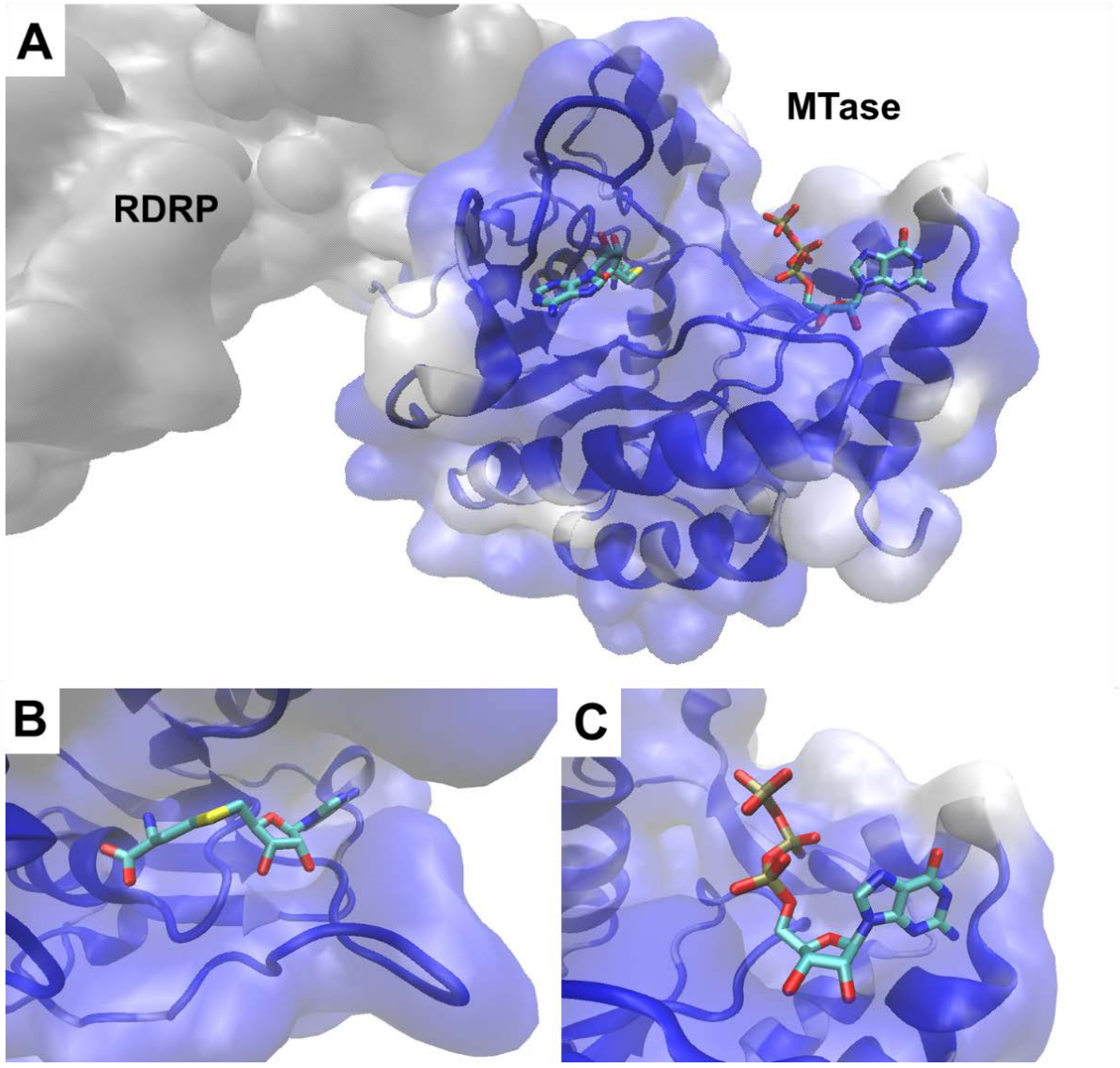


SI Figure 1. Structural comparison of ZIKV NS3 protease from homology model with HCV protease. The HCV protease was obtained from high-resolution crystal structure (PDBID 3M5L). The catalytic triad residues are indicated. Residues are colored from identical (blue) to dissimilar (white) according to sequence entropy in the VMD Multiseq tool.



SI Figure 2. Predicted membrane-spanning scheme of ZIKV p2K-NS4B. The domains were assigned by alignment to that proposed for JEV (GenBank D90195.1). Residues are numbered starting from the N-terminus of ZIKV NS4B. There are two membrane-interacting regions (pTM1 and 2) denoted along the ZIKV sequence with three segments spanning the membrane (TM3, 4 and 5) interrupted by a cytosolic domain between TM3 and TM4. Highlighted in red are residues not conserved among ZIKV strains.

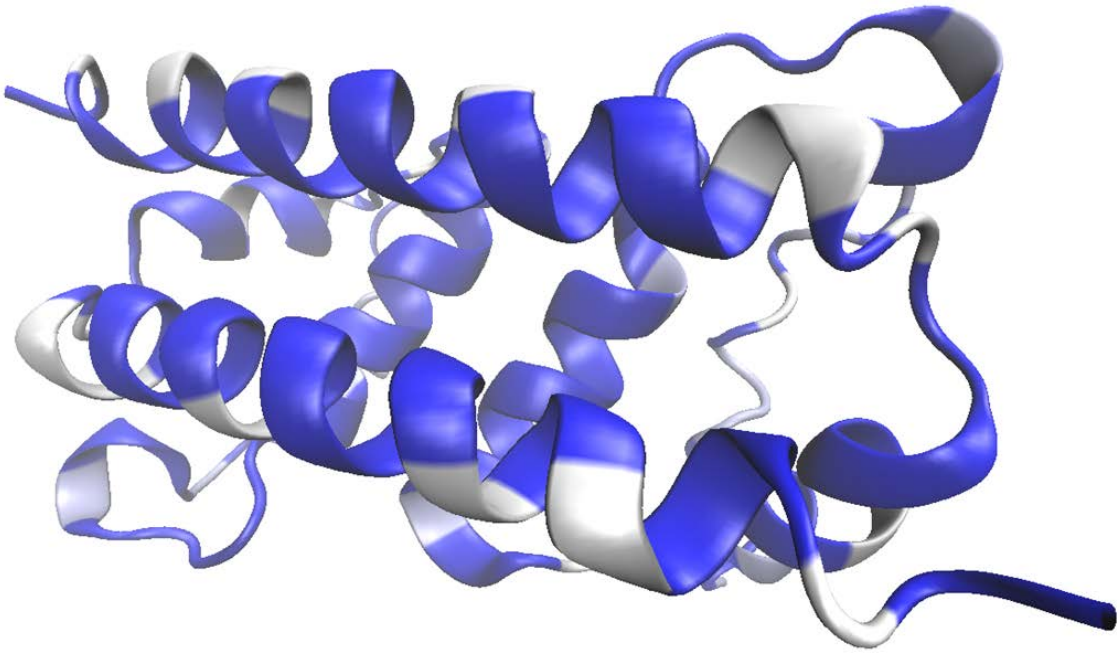
SI Figure 3. Residues at the RDRP-MTase interface predicted for ZIKV compared to those from fell-length DENV and JEV crystal structures. Colored in blue are interface residues shared by ZIKV and JEV, green are residues common between ZIVK and DENV-3, and red residues are ZIKV dissimilar to both templates.



SI Figure 4. Comparison of ZIKV MTase binding pockets. A) Homology model of full-length ZIKV NS5 derived from JEV structural template (PDBID 4K6M) with the indicated MTase and RDRP subunits. The MTase domain is aligned to that from DENV-3 crystal structure (PDBID 4V0R). The ZIKV model is colored by sequence entropy compared to the DENV structure. Shown in (B) and (C) are the identical MTase binding sites for S-adenosine methionine/homocysteine (SAM/SAH) and GTP, respectively.

SI Table 1. NS5 RDRP Active site residues between Flaviviridae.

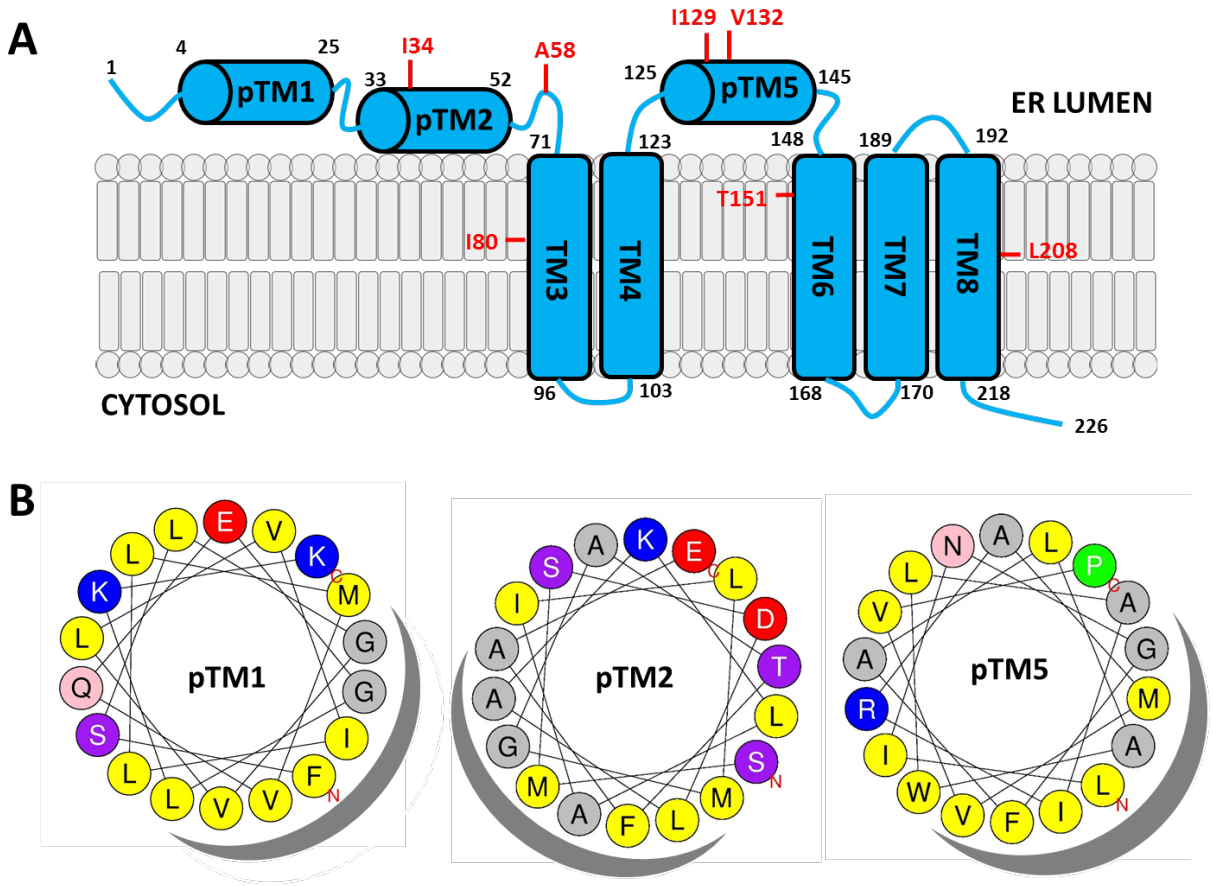
HCV Residue		JEV	DENV	Zika Model
48	Arg	Arg	Arg	Arg
141	Lys	Lys	*Lys*	Lys
158	Arg	Arg	*Arg*	Arg
159	Leu	Leu	*Leu*	Ala
160	Ile	Ile	*Ile*	Ile
220	Asp	Asp	Asp	Asp
221	Thr	Thr	Thr	Thr
222	Arg	Ala	Ala	Ala
223	His	Gly	Gly	Gly
224	Phe	Trp	Trp	Trp
225	Asp	Asp	Asp	Asp
226	Ser	Thr	Thr	Thr
280	Arg	Arg	Arg	Arg
281	Ala	Gly	Gly	Gly
282	Ser	Ser	Ser	Ser
287	Thr	Thr	Thr	Thr
291	Asn	Asn	Asn	Asn
316	Cys	Ser	Ser	Ser
317	Gly	Gly	Gly	Gly
318	Asp	Asp	Asp	Asp
319	Asp	Asp	Asp	Asp



SI Figure 5. Homology model of ZIKV anchC protein. The model was constructed based on an NMR-structure of the DENV C protein (PDBID 1R6R). Blue regions are structurally conserved residues between the model and the DENV template while white residues are dissimilar.

The anchC is 122 residues long and implicated in binding and stabilization of viral RNA. Sequential comparison reveals that the ZIKV anchC is 40% identical to the DENV C protein (PDBID 1R6R).[1] A homology model of ZIKV anchC protein (residues 25-100) was created using this NMR structure as a template. The DENV template predicts the ZIKV anchC is highly helical. The ZIKV and DENV sequences contain many positively charged residues to facilitate RNA interaction and stabilization. Though highly similar, the positions of basic residues in DENV are not conserved in the ZIKV model (K67N, K73G, K76E, R90A and R97A). To retain overall positive character, the ZIKV sequence substitutes Lys and Arg at other positions (N93R, G83K, P60K and S75K). These results suggest that the anchC protein maintains similar functional roles between ZIKV and DENV.

The anchC protein sequence is highly conserved among ZIKV strains. Two minor mutations occur outside the window of the homology model between the Asian/South American and African strains. These mutations are very similar (R101K and I110V) suggesting little impact on the overall function of ZIKV anchC.

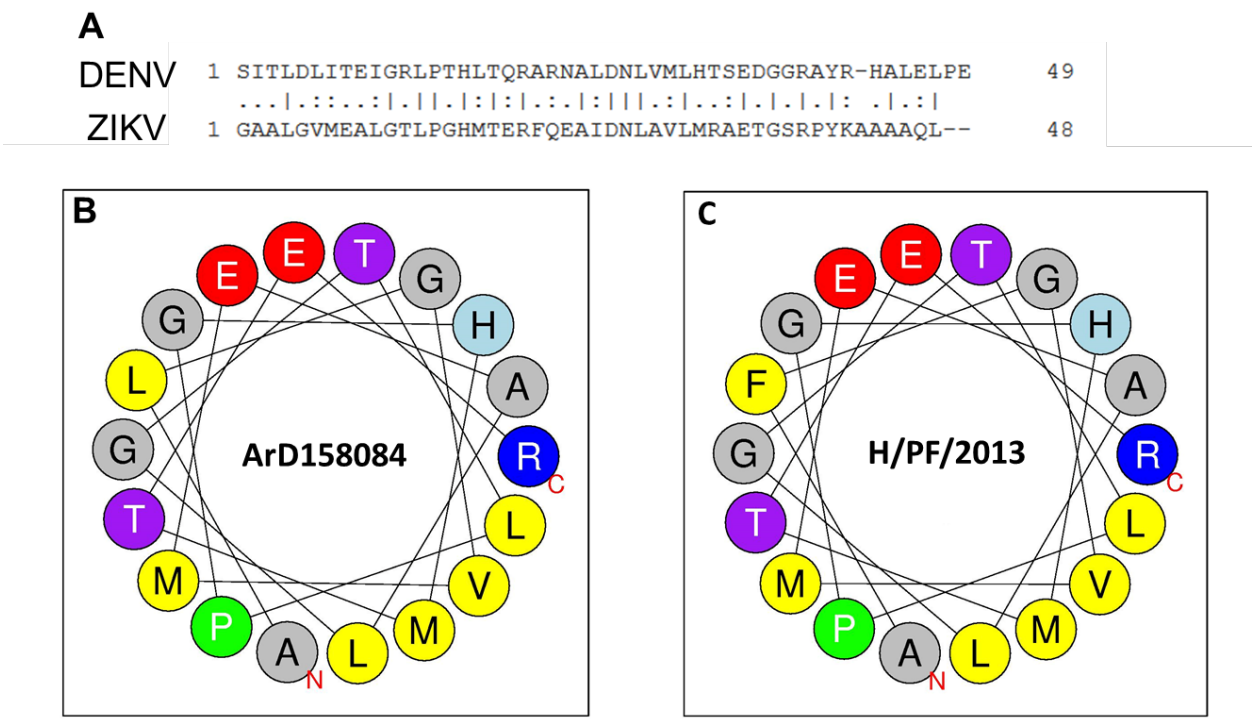


SI Figure 6. Predicted membrane-spanning scheme of ZIKV NS2A. A) The scheme was built by alignment to that proposed for DENV-2 (strain NGC). Residues are numbered starting from the N-terminus of ZIKV NS2A. There are eight membrane-interacting regions from DENV (pTM1-8) denoted along the ZIKV sequence with five segments spanning the membrane (TM3, 4, 6, 7 and 8). Highlighted in red are residues not conserved among ZIKV strains. B) Three membrane-interacting helices identified previously for DENV-2 NS2A are predicted to retain amphihelical character for ZIKV as predicted by HeliQuest. The predicted hydrophobic face is indicated with a grey swipe.

The flavivirus NS2A is responsible for multiple roles in viral replication and pathogenesis as a component to the replication complex. The N-terminal portion of the protein regulates NS1 cleavage and processing. NS2A also interacts with other viral proteins and host endoplasmic reticulum membrane to assemble and anchor the complex. The ZIKV NS2A protein is 226 residues long, but no X-ray or NMR structure is available to serve as a template for homology modeling. However, a recent study proposed the membrane topology of DENV NS2A.[2] The model proposed by Xie, et al. includes 8 membrane interacting regions numbered 1 to 8 along the sequence. Five of these regions were determined to span the membrane (TM3, 4, 6, 7 and 8), and the remaining three were membrane-interacting amphipathic helices (pTM1, pTM2 and pTM5). The ZIKV sequence was aligned to DENV-2 NS2A from this model and compared to the membrane-spanning topology. Though the DENV-2 and ZIKV NS2A proteins have low identity (27%) and modest similarity (48%), the amphipathic helices are predicted at the same loci supporting the ZIKV scheme. The NS2A has been probed for various flaviviruses to elucidate roles in viral replication, and ZIKV NS2A retains many of the residues required for viral production and

replication. A cluster at the N-terminus is required for NS1 processing in YFV (D50, K53, D70, E162, and R164),[3] and these residues are likewise in the ZIKV sequence (D52, K55, D71, and R166) though missing E162. The F113L mutant increased JEV production in Neuro-2a cells,[4] and by alignment ZIKV contains L111 at this position in TM4. The C-terminus is required for RNA replication in DENV.[5] This analysis predicts divergent sequences between DENV-2 and ZIKV in TM7 and TM8. Yet, the cytosolic C-terminal are similar (DENV-2: ^{210}TLSTYNKKR^{218}; ZIKV: ^{219}LLTRSGKR^{226}). These results suggest that ZIKV NS2A utilizes the same machinery as DENV and YFV to regulate NS1 cleavage, span the membrane, and promote viral replication.

Comparing the NS2A sequences across ZIKV strains reveals seven mutations between the African and Asian/Brazilian SIKV strains. Interestingly, these mutations occur solely in the endoplasmic reticulum (ER). Six of the mutants are replace similar residues in the transmembrane and ER lumen; M34I and A58 at pTM2, V80I in TM3, V129I and A132V in pTM5, and V/I208L in TM8. A single dissimilar mutation presents in TM6 (A151T), which may suggest different positioning of this region between ZIKV strains. Studies on Kunjin WNV support a single NS2A mutation in the ER lumen promotes apoptosis likely by a ER-stress mechanism.[6] These mutations merit additional studies in ZIKV NS2A modulation of virus-induced apoptosis.



SI Figure 7. Comparison of DENV and ZIKV NS4A N-terminus. A) Comparison of the consensus DENV and ZIKV NS4 revealing 36% identity and 58% similarity in this region. B and C) Helical wheel representation of the amphipathic helix predicted between residues 3 – 20 for ZIKV strains. The only major difference between the strains is an L4F mutation, and the N- and C-terminal residues are indicated.

The NS4 proteins are membrane-associated members of the flavivirus replication complex. Post-translational proteolytic cleavage results in the formation of NS4A (127 residues) and NS4B (251 residues) which contains an N-terminal protein 2K (p2K) motif (23 residues). NS4A modifies the host ER membrane to accommodate the replication complex. The DENV NS4A has a cytosolic amphipathic helix at the 48-residue N-terminal, and the biochemical nature of this structure is required for viral replication.[7] Comparison of DENV helix consensus sequence (residues 1 – 48) and ZIKV NS4A shows 36% identity and 58% similarity. Furthermore, the HeliQuest server predicts the ZIKV NS4 residues 1-48 likewise form an amphipathic helix. It is here that a L4F mutation occurs between ZIKV strains, but this mutation is not predicted to disrupt the secondary structure. These results suggest that, despite low identity, the DENV and ZIKV NS4A are similar at the N-terminus and likely perform similar roles.

REFERENCES

1. Ma, L.; Jones, C. T.; Groesch, T. D.; Kuhn, R. J.; Post, C. B., Solution structure of dengue virus capsid protein reveals another fold. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** *101* (10), 3414-3419.
2. Xie, X.; Gayen, S.; Kang, C.; Yuan, Z.; Shi, P.-Y., Membrane topology and function of dengue virus NS2A protein. *Journal of virology* **2013,** *87* (8), 4609-4622.
3. Voßmann, S.; Wieseler, J.; Kerber, R.; Kümmerer, B. M., A Basic Cluster in the N Terminus of Yellow Fever Virus NS2A Contributes to Infectious Particle Production. *Journal of virology* **2015,** *89* (9), 4951-4965.
4. Takamatsu, Y.; Morita, K.; Hayasaka, D., A Single Amino Acid Substitution in the NS2A Protein of Japanese Encephalitis Virus Affects Virus Propagation In Vitro but Not In Vivo. *Journal of virology* **2015,** *89* (11), 6126-6130.
5. Wu, R.-H.; Tsai, M.-H.; Chao, D.-Y.; Yueh, A., Scanning mutagenesis studies reveal a potential intramolecular interaction within the C-Terminal half of Dengue Virus NS2A involved in viral RNA replication and virus assembly and secretion. *Journal of virology* **2015,** *89* (8), 4281-4295.
6. Liu, I.-J.; Chiu, C.-Y.; Chen, Y.-C.; Wu, H.-C., Molecular mimicry of human endothelial cell antigen by autoantibodies to nonstructural protein 1 of dengue virus. *Journal of Biological Chemistry* **2011,** *286* (11), 9726-9736.
7. Stern, O.; Hung, Y.-F.; Valdau, O.; Yaffe, Y.; Harris, E.; Hoffmann, S.; Willbold, D.; Sklan, E. H., An N-terminal amphipathic helix in dengue virus nonstructural protein 4A mediates oligomerization and is essential for replication. *Journal of virology* **2013,** *87* (7), 4080-4085.