DUAL FUNCTION OF ZIKA VIRUS NS2B-NS3 PROTEASE

Sergey A. Shiryaev[1, 3], Anton Cheltsov[2], Robert C. Liddington[1], and Alexey V. Terskikh[1, 3]

[1]Sanford-Burnham-Prebys Medical Discovery Institute, Infectious and Inflammatory Disease Center, La Jolla, CA 92037, USA;

[2]Q-mol LLC, San Diego, USA;

[3]Correspondence should be addressed to S.A.S. (shiryaev@sbpdiscovery.org) or A.V.T. (terskikh@sbpdiscovery.org)

ABSTRACT

Zika virus (ZIKV) serine protease, indispensable for viral polyprotein processing and replication, is composed of an NS2B polypeptide that associates with a proteolytic N-terminal fragment of NS3 polypeptide (NS3pro) to form NS2B-NS3pro. The larger C-terminal fragment of NS3 polypeptide contains helicase activity. In the present study, we discovered that ZIKV NS2B-NS3pro efficiently binds single-stranded (ss) RNA (K_d ~0.3 μM), suggesting that the protease may have a novel function. We tested an array of NS2B-NS3pro modifications and found that NS2B-NS3pro constructs that adopt the recently discovered "super-open" conformation could not bind ssRNA. Likewise, stabilization of NS2B-NS3pro in the "closed" (proteolytically active) conformation by substrate-like inhibitors abolished ssRNA binding. Therefore, we suggest that ssRNA binding occurs when ZIKV protease adopts the "open" conformation, which could be modeled using dengue NS2B-NS3pro in the open conformation. ssRNA binding competes with ZIKV NS2B-NS3pro protease activity, likely by shifting the complex into the open conformation. Modeling of ZIKV NS3 helicase activity based on homologous crystal structures suggests that the open conformation of NS3pro domains provides a positively charged surface contiguous with the NS3 helicase domain. Such a positively charged surface is well poised to bind ssRNA, providing an explanation for the previously observed requirement of NS3pro for RNA processivity by viral helicase. Our structure-function analyses suggest that binding of ssRNA by the protease domain of NS3 is likely to be a universal feature of *Flaviviridae*, given the high level of homology between NS3 protease-helicase proteins in this family.

INTRODUCTION

Zika virus (ZIKV) is a member of the *Flaviviridae* family that includes West Nile virus (WNV), dengue virus (DENV serotypes 1–4), Japanese encephalitis virus, hepatitis C virus (HCV), and tick-borne encephalitis virus) among many other human pathogens. Like other flaviviruses with

recurrent outbreaks, ZIKV is considered a major global health threat (1,2). ZIKV infection of pregnant mothers can cause microcephaly in the fetus (3-5) and infection of adults has been linked to the autoimmune neurodegenerative disease Guillain–Barré syndrome (6-8). Currently, no effective countermeasures, such as drugs or vaccines, have been approved by the United States Food and Drug Administration for the treatment or prevention of ZIKV infection (9-11).

Like other flaviviruses, ZIKV is a positive-sense, single-stranded RNA virus. The ~10.8 kb genome encodes a 3423-amino acid (aa) polyprotein precursor consisting of 3 structural and 7 nonstructural proteins. Once inserted into the host endoplasmic reticulum (ER) membrane, the viral polyprotein is cleaved by a viral protease at cytosol-exposed junctions between nonstructural proteins and within the capsid protein C. Host proteases, such as furin and signal peptidase, process the polyprotein within the ER **(Fig. 1A).**

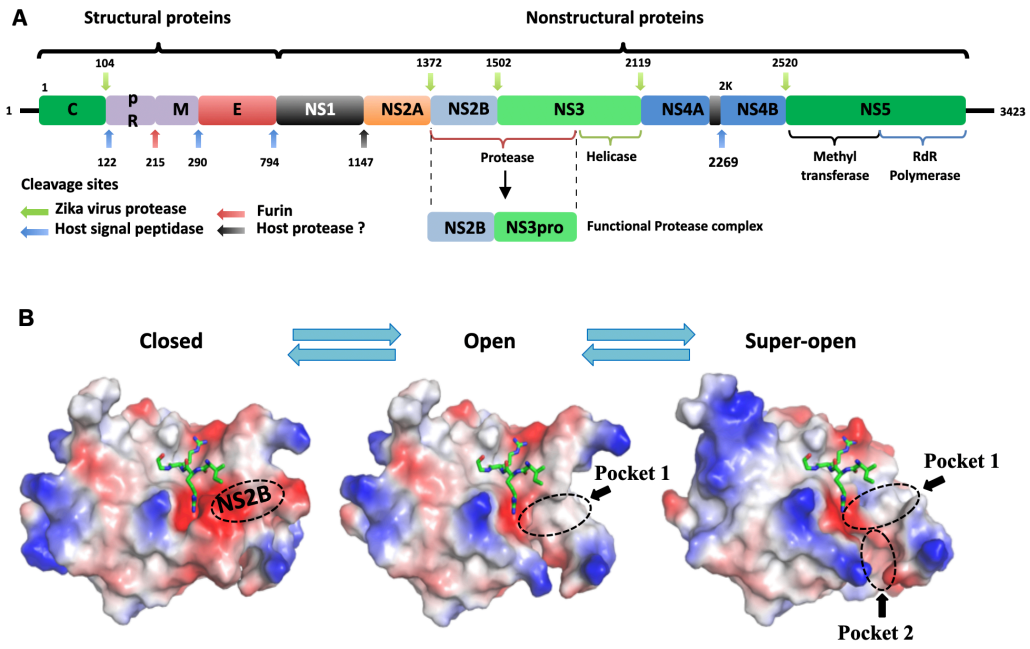


Figure 1. ZIKV polyprotein composition and processing by viral and host cell proteases. **(A)** Processing of ZIKV polyprotein. Positions of cleavage sites for host and viral proteases at the junctions between individual viral proteins are indicated by arrows. **(B)** Transitional equilibrium between closed, open, and super-open conformations of ZIKV NS2B-NS3pro, with pockets potentially involved in binding of ssRNA or allosteric inhibitors indicated. A peptide-based substrate (green ball and stick model) in the protease active center is modeled on the related structure of WNVpro + aprotinin (PDB ID 2IJO).

ZIKV NS3 protein is an ~615 aa multifunctional protein exhibiting endopeptidase, RNA helicase, RNA triphosphatase, and NTPase activities (12). The N-terminal 170-aa domain of NS3 (NS3pro) interacts with a cytosolic 48-aa region of the membrane-anchored NS2B to form a highly active serine protease termed NS2B-NS3pro **(Fig. 1A)** (13). The critical dependance of viral propagation on NS2B-NS3pro has made this polypeptide a promising antiviral drug target. However, the development of competitive inhibitors is challenging because of the high structural homology between the active centers of viral NS2B-NS3pro and multiple cellular serine proteases with basic functions (14). The development of allosteric anti-NS2B-NS3pro inhibitors may be one way to overcome this obstacle (14,15).

Recently, we determined the crystal structure of ZIKV NS3pro with the covalently linked NS2B cofactor with 2.5 Å and 3 Å resolution. These structures revealed that NS2B-NS3pro adopts a so-called “super-open” conformation (PDB IDs 5TFN, 5TFO, 6UM3, and 7M1V) that is clearly distinct from the previously reported proteolytically active “closed” conformation of ZIKV NS2B-NS3pro (PDB ID 5LC0) (16). The crystal structure of ZIKV protease in the “open” conformation has not been reported, but the structures of DENV and WNV NS2B-NS3pro in the open conformation (PDB IDs 2FOM and 2GGV) suggest that ZIKV protease open conformation can be modeled from the closed conformation by removing the C-terminal tail of the NS2B cofactor from a hydrophobic cleft of NS3pro (17). In addition to the NS2B dissociation, the super-open conformation imposes refolding of NS3pro C-terminal residues, which eliminates the substrate binding site. Given the apparent continuum of these structural changes, we propose that NS2B-NS3pro can dynamically switch between closed, open, and super-open conformations **(Fig. 1B)**.

The NS3-NS4 protease domain of HCV is distantly homologous to ZIKV NS2B-NS3pro, the major differences being the existence of HCV protease domain in a single conformation, the distinctive substrate specificities, and a strong positive charge around the active site of NS3-NS4 protease domain of HCV (18). Several independent studies have demonstrated that the protease domain of HCV NS3 binds to ssRNA, and full-length HCV NS3 protein possesses better dsRNA processivity than the helicase domain alone (19-21). The isoelectric point (pI) of ZIKV NS2B-NS3pro is lower than that of HCV protease, suggesting that it may not have the ability to bind RNA; however, the full-length NS3 of the homologous WNV and DENV2 viruses also demonstrate significantly better processivity compared with the corresponding helicase domains alone (22,23).

In the present study, we demonstrate that the protease domain of ZIKV NS3 binds ssRNA with a physiologically meaningful affinity. We also analyzed ssRNA binding to ZIKV NS2B-NS3pro in different conformational states, and based on these findings, we propose that ssRNA binds the open conformation of NS2B-NS3pro. Finally, we demonstrated that ssRNA binding competes with the proteolytic activity of NS2B-NS3pro and with binding of inhibitors, providing a novel approach for drug screening.

MATERIALS AND METHODS

Reagents. Routine laboratory reagents were purchased from Millipore Sigma (St. Louis, MO) unless indicated otherwise. Horseradish peroxidase-conjugated donkey anti-mouse IgGs and TMB/M substrate were from Jackson ImmunoResearch Laboratories (West Grove, PA), SuperSignal West Dura Extended Duration Substrate for ECL was from ThermoFisher (Carlsbad, CA) and SurModics IVD (Eden Prairie, MN), and oligonucleotides were synthesized by Integrated DNA Technologies (San Diego, CA).

Cloning, expression, and purification of ZIKV NS2B-NS3pro constructs. DNA sequences for the ZIKV constructs were synthesized and codon-optimized for efficient transcription in *E. coli.* The constructs were designed as single-chain two-component products lacking the hydrophobic transmembrane domain of the NS2B cofactor. To achieve this, the cytoplasmic portion of NS2B (residues 48–94) was linked to the NS3pro domain *via* a 9-residue linker (GGGGSGGGG). To block the proteolytic activity of the construct, where indicated, the catalytic Ser^{135} was substituted with Ala to give the inactive Ser135Ala construct.

ZIKV NS2B-NS3pro recombinant constructs with N-terminal His tag were used to transform competent *E.coli* BL21 (DE3) Codon Plus cells (Stratagene). Transformed cells were grown at 30°C in LB broth containing carbenicillin (0.1 mg/ml). Protein production was induced with 0.6 mM IPTG for 16 h at 18°C. Cells were collected by centrifugation at 5000 g at 4°C, and the cell pellet was resuspended in 20 mM Tris-HCl buffer, pH 8.0, containing 150 mM NaCl (TBS), and sonicated (eight 30 s pulses at 30 s intervals) on ice. The sample was then centrifuged at 40,000 g for 30 min at 4°C and the constructs were purified from the supernatant fraction using Ni-NTA Sepharose equilibrated with TBS containing 1 M NaCl. Impurities were removed by washing with the same buffer supplemented with 35 mM imidazole, and the column was equilibrated with standard TBS. The beads were co-incubated with thrombin (Sigma Aldrich) to cleave the His tag and release NS2B-NS3pro. Fractions containing recombinant protein were combined and purified by gel filtration on a S200 26/60 column (GE Healthcare) equilibrated with

standard TBS. Fractions containing ZIKV NS2B-NS3pro were concentrated to approximately 10 mg/ml using 10 kDa-cutoff concentrators (Millipore, Billerica, MA), and then flash frozen in small aliquots and stored at –80°C. Purity of the material was checked by SDS-PAGE (12% NuPAGE-MOPS, Invitrogen) followed by Coomassie staining.

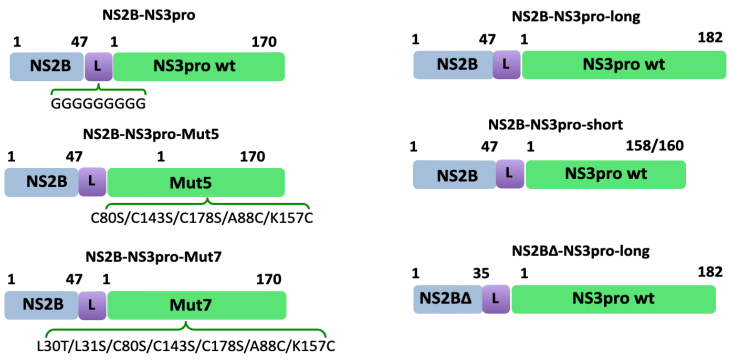
We generated a series of construct (Mut3, Mut5, and Mut6) that forced ZIKV NS2B-NS3pro to adopt the "super-open" conformation using the wild-type construct as a template. In the Mut3 construct, all three Cys residues in the original sequence were substituted with Ser residues (Cys80Ser, Cys143Ser, and Cys178Ser). In the construct with the "super-open" conformation only (Mut5), two additional Cys residues were inserted into the Mut3 construct (Ala88Cys and Lys157Cys). To obtain a catalytically inactive mutant with the "super-open" conformation (Mut6), an additional Ser135Ala was introduced into the Mut5 construct. For crystallization purposes, we also created a Mut7 construct containing Leu30Thr and Leu31Ser mutations in the Mut5 sequence. Mut7 was designed to minimize ZIKV protease dimerization during crystallization.

The DNA constructs were cloned into pGEX6P1 plasmid using BamH1 and EcoR1 cleavage sites, resulting in fusion of a GST tag at the N-terminus of the NS2B cofactor. *E. coli* were transformed with the individual recombinant constructs, and protein production was induced and the cells were disrupted as described above. The supernatant fraction containing the GST-tagged constructs were loaded onto Protino Glutathione Agarose 4B (Fisher Scientific) beads and impurities were removed by washing with TBS. To cleave the GST tag from the viral protease, the beads were co-incubated with 3C protease (Genscript). The NS2B-NS3pro constructs were then additionally purified by gel filtration on an S200 26/60 column (GE Healthcare) equilibrated with TBS. Purity of the purified construct was analyzed using SDS-PAGE followed by Coomassie staining. Purified protease was concentrated to approximately 10 mg/ml using 10 kDa-cutoff concentrators (Millipore, Billerica, MA), flash frozen in small aliquots, and stored at –80°C. To isolate NS2B-NS3pro complexed with aprotinin or WRPK3, the NS2B-NS3pro constructs were incubated with each inhibitor at an equimolar ratio and the protein-inhibitor complexes were purified by gel filtration on a S200 Superdex column.

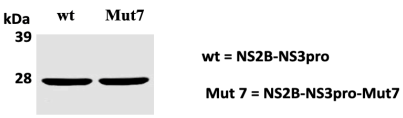
Fluorescent proteinase activity assay and IC_{50} determination. The peptide cleavage activity assays with purified ZIKV NS2B-NS3pro polypeptides were performed in 0.2 ml TBS containing 20% glycerol and 0.005% Brij 35 containing 20 μM of the cleavage peptide pyroglutamic acid Pyr-Arg-Thr-Lys-Arg-7-amino-4-methylcoumarin (Pyr-RTKR-AMC) and 10 nM enzyme. The reaction velocity was monitored continuously at λ_{ex}= 360 nm and λ_{em} = 465 nm on

a Tecan fluorescence spectrophotometer (Männedorf, Switzerland). To determine the IC_{50} values of the inhibitory compounds, ZIKV NS2B-NS3pro constructs (20 nM) were preincubated for 30 min at 20°C with various concentrations of compounds in 0.1 ml TBS containing 20% glycerol and 0.005% Brij 35. Pyr-RTKR-AMC substrate (20 µM) was then added in 0.1 ml of the same buffer. IC_{50} values were calculated by determining the compound concentration required to obtain 50% of the maximal inhibition of NS2B-NS3pro activity against Pyr-RTKR-AMC. GraphPad Prism was used as fitting software. All assays were performed in triplicate in 96-well plates.

A



B



C

	Capsid C	NS2A/NS2B	NS2B/NS3	NS3/NS4A	NS4B/NS5
ZIKV	V^{125}TRR↓GSAY^{132}	S^{1369}GKR↓SWPP^{1376}	T^{1499}GKR↓SGAL^{1506}	A^{2116}GKR↓GAAF^{2123}	G^{2766}PRR↓PVKY^{2773}
WNV	Q^{102}KKR↓GGTA^{109}	N^{1367}RKR↓GWPA^{1374}	Y^{1498}TKR↓GGVL^{1505}	S^{2117}GKR↓SQIG^{2124}	G^{2522}LKR↓GGAK^{2529}
DENV1	R^{97}RKR↓SVTM^{104}	W^{1341}GRK↓SWPL^{1348}	K^{1471}KQR↓SGVL^{1478}	A^{2090}GRR↓SVSG^{2097}	G^{2489}GRR↓GTGA^{2496}
DENV2	R^{97}RRR↓SAGV^{104}	S^{1342}KKR↓SWPL^{1349}	K^{1472}KQR↓AGVL^{1479}	A^{2090}GRK↓SLTL^{2097}	N^{2488}TRR↓GTGN^{2495}
DENV3	K^{97}RKK↓TSLC^{104}	L^{1340}KRR↓SWPL^{1347}	Q^{1470}TQR↓SGVL^{1477}	A^{2089}GRK↓SIAL^{2096}	T^{2487}GKR↓GTGS^{2494}
DENV4	G^{96}RKR↓STIT^{103}	A^{1341}SRR↓SWPL^{1348}	K^{1471}TQR↓SGAL^{1478}	S^{2089}GRK↓SITL^{2096}	T^{2484}PRR↓GTGT^{2491}
JEV	Q^{102}NKR↓GGNE^{109}	N^{1370}KKR↓GWPA^{1377}	T^{1501}TKR↓GGVF^{1508}	A^{2120}GKR↓SAVS^{2127}	S^{2524}LKR↓GRPG^{2531}
YFV	R^{98}KRR↓SHDV^{105}	F^{1351}GRR↓SIPV^{1358}	G^{1481}ARR↓SGDV^{1488}	E^{2104}GRR↓GAAE^{2111}	T^{2503}GRR↓GSAN^{2510}

Supplementary Figure 1. (A) NS2B-NS3pro constructs used in this study. All constructs were N-terminally fused with GST protein or HisTag for purification. NS2B central hydrophilic portion is shown in blue and the NS3 protease in green. L = GGGGSGGGG linker between NS2B and NS3pro. **(B)** Western blot analysis of purified wild-type and Mut7 NS2B-NS3pro proteins. **(C)** NS2B-NS3pro cleavage sequences in flaviviral polyproteins. Cleavage sites in the capsid protein C and at the NS2A/NS2B, NS2B/NS3, NS3/NS4A, NS4A/NS4B, and NS4B/NS5 boundaries are shown. ZIKV, Zika (GenBank AMB37295); WNV, West Nile virus (GenBank P06935); JEV, Japanese encephalitis (GenBank P19110); YFV, yellow fever (GenBank P19901); DENV1–4, dengue serotypes 1–4 (GenBank P33478, P29990, P27915, and P09866, respectively).

Fluorescence polarization assay of ssRNA/ssDNA binding to ZIKV NS2B-NS3 protease. Binding between purified recombinant ZIKV NS2B-NS3pro constructs and either ssRNA or ssDNA was assessed using a fluorescence polarization (FP) assay conducted in 0.1 ml of TBS containing 1 mM $MgCl_2$ at 25°C for 1 h. Samples of 10 nM 6-carboxyfluorescein-labeled 20-base ssRNA (poly-rA or -rU) or ssDNA (poly-dA or -dT) oligonucleotides were incubated with 50 nM to 50 μM of purified NS2B-NS3pro constructs. Polarization was monitored on a Bruker Daltonics fluorescence spectrophotometer (Fremont, CA). K_d values for ssRNA and ssDNA binding were calculated by determining the construct concentration needed to reach 50% polarization for 3′-fluorescein amidite-labeled ssRNA/DNA. All assays were performed in triplicate in 96-well plates. GraphPad Prism was used as fitting software.

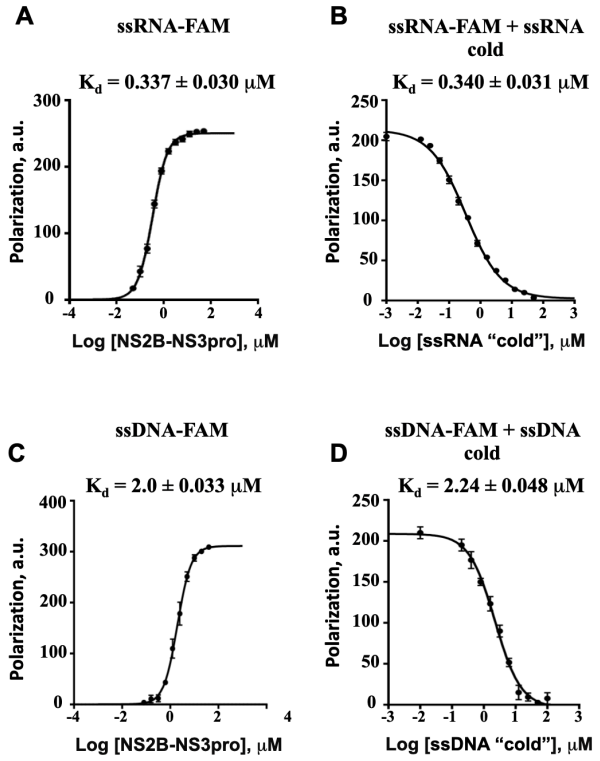


Figure 2. ZIKV NS2B-NS3pro binds ssRNA and ssDNA. Fluorescent polarization assays of **(A)** FAM-labeled ssRNA (20 poly-rU) binding to the indicated amounts of NS2B-NS3pro, **(B)** FAM-labeled ssRNA (20 poly-rU) binding to 1 μM of NS2B-NS3pro in the presence of the indicated amounts of unlabeled ssRNA (20 poly-rU). **(C, D)** As for A, B, except experiments were performed with FAM-labeled ssDNA (20 poly-dT).

RESULTS

ZIKV NS2B-NS3pro polypeptide binds ssRNA

To investigate ssRNA and ssDNA binding to ZIKV NS2B-NS3pro polypeptide, we performed classical FP assays in which FAM-labeled ssRNA or ssDNA was incubated with NS2B-NS3pro and binding was detected by a change (increase) in FP (24). We found that wild-type proteolytically competent NS2B-NS3pro, which is capable of adopting closed, open, or super-open conformations, bound robustly to ssRNA. The results of direct binding assays and competition assays performed in the presence of unlabeled ssRNA were in good agreement, both giving K_d values of 0.34 μM **(Fig. 2A, B)**. An increase in FP signal was also

obtained upon incubation of NS2B-NS3pro with FAM-labeled ssDNA **(Fig. 2C)**. However, both the direct binding and competition assays yielded K_d values for ssDNA that was ~10-fold lower than the K_d for binding of ssRNA (~2.0 μM *vs* 0.3 μM) **(Fig. 2C, D)**. The demonstration that unlabeled ssRNA and ssDNA efficiently competed with the FAM-labeled probes with similar kinetics **(Fig. 2B, D)** confirmed the specificity of binding.

To investigate ssRNA binding to ZIKV NS2B-NS3pro polypeptide in more detail, we generated a set of constructs aimed at probing some key structural and conformational requirements **(Supp. Fig. 1A)**. First, we tested a longer construct that contained 182 N-terminal residues of NS3 (NS2B-NS3-long) and a 12-aa linker region (^{171}EEETPVECFEPS^{182}) connecting the NS3pro domain with the NS3 helicase domain. NS2B-NS3-long was crystallized in the super-open conformation (PDB ID 5TFN) and should be capable of adopting the closed, open, and super-open conformations. Surprisingly, NS2B-NS3-long incubation with ssRNA or ssDNA did not increase the FP signal, indicating a lack of ssRNA or ssDNA binding **(Fig. 3A).** These observations suggests that the 12-aa linker region modulates ssRNA binding to NS2B-NS3pro.

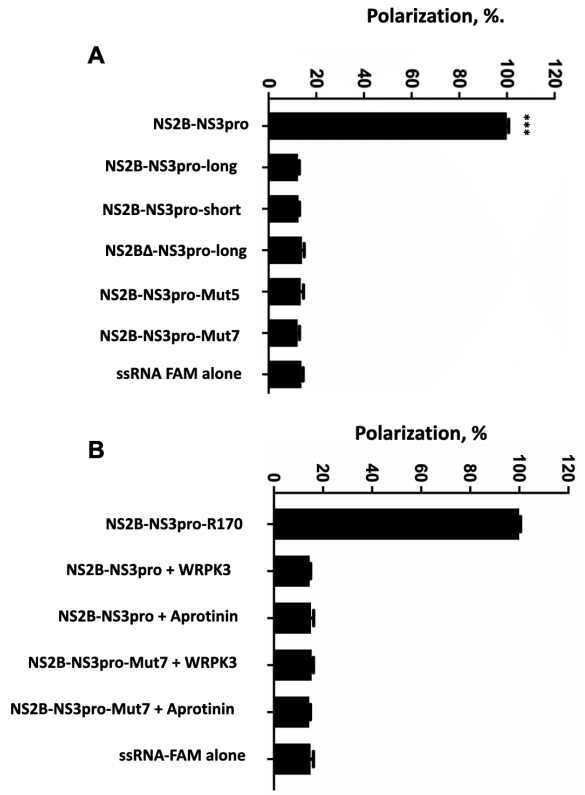


Figure 3. Structural and functional alterations in ZIKV NS2B-NS3pro preclude ssRNA binding. (A) Modifications of NS2B-NS3pro (see Supplementary Figure 1 for details) that prevent efficient binding of FAM-labeled ssRNA (20 poly-rU). **(B)** Inhibition of NS2B-NS3pro catalytic activity by aprotinin (~0.15 μM) and WRPK3 (covalent inhibitor) blocks ssRNA binding.

Next, we probed the conformational requirements of NS2B-NS3pro for ssRNA binding. The truncated NS2B-NS3-short construct contained 160 N-terminal residues of NS3 (NS2B-NS3-short) and can adopt only the super-open conformation due to the lack of aa 161–170, which are necessary to form the closed or open conformations (PDB 5LC0) (17). FP assays indicated that NS2B-NS3-short also failed to bind to

ssRNA or ssDNA **(Fig. 3A),** indicating a requirement for aa 161–170 in NS3pro to support ssRNA binding.

We next tested the NS2B-NS3pro-Mut5 and Mut7 constructs, which were designed to be stabilized in the super-open conformation *via* a disulfide bond between two cysteines introduced into the NS3 sequence. We solved the structures of Mut5 and Mut7 and confirmed that both mutants adopt the super-open conformation (PDB IDs 6UM3, 7M1V). However, neither of these mutants was able to bind ssRNA **(Fig. 3A),** suggesting that the super-open conformation is likely to be refractory to ssRNA binding.

We also designed an NS2B-NS3pro construct containing a truncated NS2B (NS2BΔ) which lacked ^{29}VALDESGDFSL^{39}, making the protein unable to adopt a closed conformation and thus form a proteolytically active complex with NS3 (17). The NS2BΔ-NS3-long construct was also unable to bind ssRNA **(Fig. 3A)**.

Taken together, these experiments suggest that full-length NS2B-NS3pro polypeptide efficiently binds ssRNA and that the tested structural modifications, particularly those that lock NS2B-NS3pro in the super-open conformation, eliminated ssRNA binding. Further, the C-terminal region of NS3pro and the linker between NS3pro and NS3 helicase appear to modulate ssRNA binding.

NS2B-NS3pro substrate-mimicking inhibitors compete with ssRNA binding

Several groups have demonstrated that binding of protein substrates, peptidomimetic peptides, or inhibitor compounds in the protease active center stimulates its transition to a proteolytically competent closed conformation (16,17,25-27). This phenomenon has been documented for proteases from several *Flaviviridae*, including WNV (PDB IDs 2IJO, 2YOL) (17,25), DENV-2 (PDB ID 3U1J) (26), and ZIKV (PDB IDs 5YOF, 5LCO) (16,27). To determine whether ssRNA can bind to ZIKV NS2B-NS3pro in the closed, proteolytically active state, we induced this conformation in NS2B-NS3pro using two substrate-mimicking inhibitors; aprotinin, a 70-aa serine protease inhibitor, and WRPK3, a synthetic peptide inhibitor that covalently binds to the NS2B-NS3pro active center (28). Previous studies demonstrated that these inhibitors bind only to the catalytically active NS2B-NS3pro protease in the closed conformation (17,28).

ZIKV NS2B-NS3pro was incubated with aprotinin or WRPK3 and the purified complexes were tested in the ssRNA FP assays. We observed no ssRNA binding to either of the complexes **(Fig. 3B)**. Since the active site of closed ZIKV protease is negatively charged **(Fig 1B),** direct binding of ssRNA to the same site as aprotinin or WRPK3 seems unlikely. Therefore, these results

suggest that substrate-mimicking inhibitors eliminate ssRNA binding by inducing and sustaining the closed conformation of NS2B-NS3pro, which is incompatible with ssRNA binding.

ssRNA binding inhibits the proteolytic activity of ZIKV NS2B-NS3pro

The experiments described thus far suggest that ssRNA cannot bind to NS2B-NS3pro constructs forced into closed or super-open conformations; by inference, this suggests that ssRNA binding occurs only when NS2B-NS3pro is in the open proteolytically inactive conformation. To test this, we investigated whether binding of ssRNA interferes with the proteolytic activity of NS2B-NS3pro. Indeed, using fluorogenic substrate cleavage assays, we found that binding of ssRNA or ssDNA inhibited the proteolytic activity of ZIKV NS2B-NS3 protease with IC_{50} values of 26.28 ± 1.12 μM for ssRNA and 64.75 ± 1.08 μM for ssDNA. These results suggest that binding of ssRNA or ssDNA to ZIKV NS2B-NS3 inhibits its proteolytic activity, presumably due to stabilization of the polypeptide in the open conformation, which is proteolytically inactive.

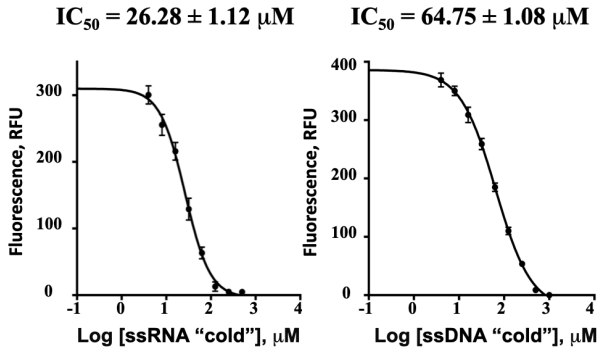


Figure 4. ssRNA inhibits the proteolytic activity of NS2B-NS3pro. Inhibition of ZIKV NS2B-NS3pro proteolytic activity by ssRNA (20-rA, left) and ssDNA (20-dA, right).

ssRNA binding offers new insights into the development of novel inhibitors of NS2B-NS3pro

The open conformation of NS2B-NS3pro is formed by rearrangement/dissociation of NS2B cofactor from the C-terminal half of NS3pro, leading to loss of proteolytic activity (29). NS2B dissociation uncovers a hydrophobic cleft in NS3pro, which presents a druggable pocket that is relatively conserved between the flaviviruses. Previously, we identified inhibitors of ZIKV, WNV, and DENV2 NS2B-NS3pro with sub- or low-micromolar IC_{50} values *in vitro* (13,14). These allosteric inhibitors were designed *in silico* to bind a region of NS2B-NS3pro distant from the proteolytic active site (14,30,31). In addition to demonstrating their potent inhibition of several flaviviral proteases, we showed that the inhibitors had no detectable effect on host serine proteases (furin and other proprotein convertases) that have a similar substrate specificity (13). Here, we tested ssRNA binding to ZIKV NS2B-NS3pro with a sub-library of the inhibitors that had previously been characterized for WNV and DENV2 NS2B-NS3pro. We

discovered that several of these compounds are also potent inhibitors of ssRNA binding to ZIKV NS2B-NS3pro **(Fig. 5A, bottom)** (13). One compound, NSC86314, very effectively inhibited binding to NS2B-NS3pro from ZIKV, WNV, and DENV *in vitro* with low to sub-micromolar IC_{50} values (13) (**Fig. 5A, bottom)**. NSC86314 was previously shown to inhibit WNV and DENV2 replicons in cell-based assays with IC_{50} values of <50 μM (13,30,31). A global search of docking locations on the ZIKV NS2B-NS3pro surface using the SWISSDOCK program (www.swissdock.ch) points to novel druggable pockets located outside the active site in the open and super-open conformations as the docking site in both WNV and ZIKV NS2B-NS3pro polypeptides **(Fig. 5B)**.

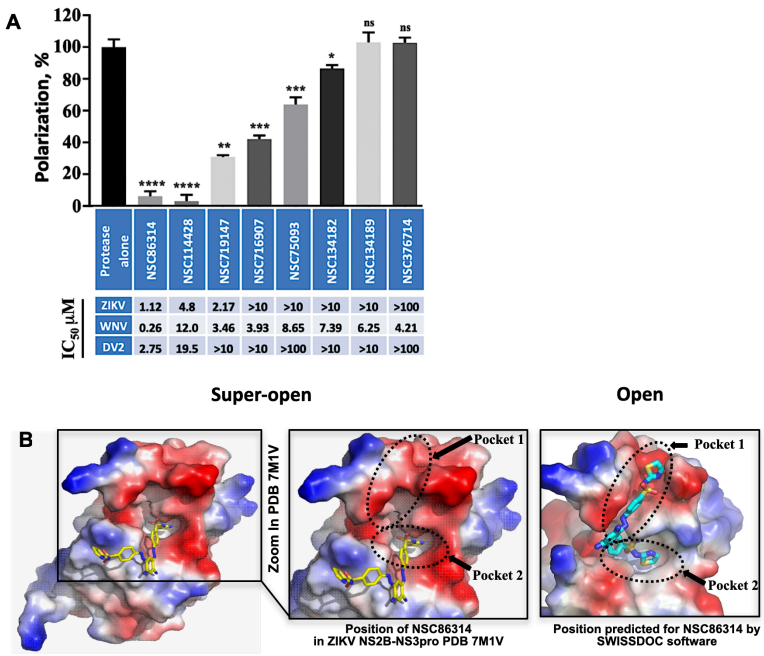


Figure 5. Allosteric inhibitors of ZIKV NS2B-NS3pro inhibit ssRNA binding. (A) Top, small molecule allosteric inhibitors (10 µM) targeting open/super-open conformations of ZIKV NS2B-NS3pro block ssRNA binding. Fluorescent polarization (%) using FAM labeled ssRNA (20 poly-rU). * p<0.05, ** p<0.05, *** p<0.0005, **** p<0.0005 by two-tailed unpaired t-test with Welch's correction. Bottom, IC_{50} for inhibition of ZIKV, WNV, and DENV2 NS2B-NS3pro proteolysis by the indicated inhibitors. **(B)** Crystal structure of ZIKV NS2B-NS3-Mut7 protease with the NSC86314 inhibitor. The position of NSC86314 differs from that predicted by docking modeling? with SwissDoc docking software due to dimerization of the NS2B-NS3 protease, leading to structural constraints in the crystal and inaccessibility of pocket 1.

To confirm the in silico docking results, we co-crystallized NSC86314 with the Mut7 ZIKV NS2B-NS3pro construct. Two additional mutations (Leu30Thr and Leu31Ser) were introduced into the membrane-binding loop to disrupt/modify the dimer interface of NS2B-NS3 protease. In our previously reported structures of NS2B-NS3pro in the super-open conformations, Leu30 and Lue31 were inserted into the hydrophobic cleft of the C-terminal β-barrel, preventing binding of inhibitors that target these areas (PDB ID 6UM3). The Mut7 crystal structure was solved and refined to a resolution of 1.8 Å (PDB ID

7M1V). The structure revealed that NSC83614 binds to the predicted hydrophobic pocket 2 near the C-terminus of NS3pro **(Fig. 5B)**. However, in contrast to the docking model, the dimerization of Mut7 and the structural constraints around pocket 1 made it inaccessible for a part of this inhibitor

We next performed Virtual Ligand Screening (VLS) of the NCI Chemotherapeutic Agents Repository to find compounds having a similar scaffold structure to NSC86314 that target the open and super-open conformations of NS2B-NS3 protease. One of the compounds identified, NSC114428, was able to block both the proteolytic activity of ZIKV protease and ssRNA binding **(Fig. 5A)** in the low μM range.

We next tested a series of flavivirus replication inhibitors with activity against ZIKV, WNV, and DENV proteases for their ability to inhibit ssRNA binding to ZIKV NS2B-NS3pro **(Fig. 5A, top)**. The most potent inhibitors of ZIKV and DENV2 protease activity tended to also be the most potent inhibitors of RNA binding. However, this association was not observed for inhibition of WNV protease activity, suggesting that the relationship between inhibition of flavivirus proteolytic activity and ssRNA binding may be complex (**Fig. 5A top)**. These results provide proof-of-principle for a new class of allosteric inhibitors that specifically target newly identified druggable pockets present in the open and super-open conformations of ZIKV NS2B-NS3pro. Our results suggest that such allosteric inhibitors might be able to block both the proteolytic and ssRNA-binding activities of NS2B-NS3pro from multiple flaviviruses.

Potential caveats and interpretations

The ssRNA-binding ZIKV NS2B-NS3pro construct was purified *via* the His tag prior to thrombin-mediated removal of the tag. In principle, other *E. coli* proteins could have been co-purified during affinity chromatography, and they may have affected our results with respect to ssRNA binding. For instance, the heat stable *E. coli* Hfq protein is known to harbor stretches of His residues and may have been co-purified during Ni affinity chromatography, resulting in false-positive RNA-binding activities (32). However, our purification scheme involved thrombin cleavage of the His tag to release ZIKV NS2B-NS3pro, and the thrombin site is absent in *E. coli* Hfq. Further, Hfq is an 11 kDa protein and has IC_{50} values for ssRNA binding in the low nanomolar range, which is inconsistent with our results demonstrating that purified NS2B-NS3pro binding to ssRNA occurred with an IC_{50} value in the submicromolar range and was inhibited by aprotinin, WRPK3, and several allosteric inhibitors. Additional experiments will be required to confirm that the open conformation of ZIKV NS2B-NS3pro is unable to bind to ssRNA. One approach will be to stabilize and crystallize

NS2B-NS3pro in the open conformation, and then confirm or refute ssRNA binding. Ultimately, co-crystallization of NS2B-NS3 with ssRNA will be required.

DISCUSSION

During viral propagation, the N-terminal proteolytic (NS3pro) and C-terminal helicase (NS3 helicase) domains function within the constraints of a single NS3 polypeptide. NS2B-NS3pro is responsible for autoproteolytic cleavage of junction regions between NS2B/NS3 and NS3/NS4A proteins bordering NS3 protein during viral polyprotein processing. After this step NS3 protein is associated with the ER membrane via NS2B cofactor (4 transmembrane domains) which interacts with NS3pro domain and NS4A cofactor (2 transmembrane domains) which interacts with NS3 helicase domain. Based on our experimental, structural, and modeling studies, we propose that NS3pro not only has proteolytic activity but also binds one ssRNA strand, thereby facilitating the unwinding of dsRNA generated by viral NS5 RNA-dependent RNA polymerase. We propose the existence of a dynamic equilibrium between the closed–open–super-open conformations of NS2B-NS3pro. The transition between conformations depends on the coordinated structural rearrangement of the C-terminal tail of NS3pro and the linker between NS2B-NS3pro and NS3 helicase, resulting in progressive loss of interactions between NS3pro and NS2B.

In support of our hypothesis, several studies have pointed to the importance of the flaviviral NS3 protease domain for proper functioning of NS3 helicase (19-21). The presence of the protease domain within full-length HCV NS3 protease-helicase constructs was shown to significantly enhance direct and functional RNA binding to NS3 helicase (20). Additionally, the presence of the NS3 protease domain contributed substantially to helicase translocation stepping efficiency by increasing the efficiency up to 93% per ATP hydrolysis event relative to 20% for the single NS3 helicase construct (19,23). Absence of the NS3 protease domain results in malfunctioning of NS3 helicase directional movement and the appearance of two opposing activities characterized by an ATP-dependent steady state between RNA unwinding and RNA annealing processes (21).

The linker between the NS3 protease and helicase domains has been demonstrated to play a critical role in their activities in the structurally and functionally similar DENV2 NS2B-NS3pro (22). Mutations in the linker affecting its flexibility and length were crucial for the ATPase and helicase activities and also led to significant reductions in viral genomic RNA synthesis (22). Given the high negative charge of these 12 aa (^{171}EEETPVECFEPS^{182}), we propose that the linker between NS3 protease and helicase domains competes with ssRNA for binding to NS3pro.

This may provide an explanation for the lack of ssRNA binding by the NS2B-NS3pro-long construct shown here.

In crystal structures, the ZIKV protease-helicase linker is disordered (PDB IDs 5TFN, 5TFO, 5T1V and 6UM3). However, in the closed and open structures of WNV and DENV2 proteases, the linkers are associated with a conserved region of the protease domain (17). We suggest that protease-helicase linker of ZIKV could dynamically interact with the protease domain, and its negative charge might weaken ssRNA binding (e.g. NS2B-NS3pro-long).

Published structures of ZIKV protease in the super-open state show that Ser^{160} is the last ordered NS3 residue, whereas this is Gly^{168} for proteases in the closed and open conformations.

The dissociation of NS2B from NS3pro during the transition from closed to open to super-open conformations changes the isoelectric point (pI) of the NS2B-NS3pro complex. Note that the progressive dissociation of NS2B from NS3pro removes the regions containing negatively charged residues in a conformation-dependent fashion, leading to an increased pI of the NS2B-NS3pro complex that should facilitate ssRNA binding.

Recent crystal structure of ZIKV NS2B-NS3pro in the super-open conformation with a small molecule compound occupying a new pocket (7M1V), suggests it may be feasible to specifically target the super-open conformation of NS2B-NS3pro. Molecular probes targeting the super-open conformation will help to elucidate the effects of the different NS2B-NS3 conformational states on the protease and the helicase activities.

Further structural and biochemical analyses will be required to elucidate how exactly ssRNA binds to NS3pro. Based on the model of dynamic switching between closed–open–super-open conformations, we propose that NS2B-NS3pro dynamically cycles between binding and releasing ssRNA, perhaps assisting NS3 helicase in pulling and unwinding dsRNA. In support of this hypothesis, the HCV NS3 protease has been shown to be important for directional movement of the NS3 helicase and translocation stepping efficiency in ATP hydrolysis during the viral dsRNA unwinding process (20).

The C-terminal regions of NS2B cofactors from various flaviviruses contain a conserved negative charge (33), which may explain why the productive binding of this fragment by protease in a closed state inhibits binding to negatively charged nucleotides. In the crystal structures of flavivirus proteases in the open state, the negatively charged fragment of NS2B is either disordered or loosely bound to NS3pro surface. Thus, differences in the location of the negatively charged NS2B fragment between the open and closed states is consistent with the preferred binding of ssRNA to the open conformation. Ongoing studies, including modeling of ssRNA

binding to ZIKV NS2B-NS3, will provide insights into the dynamic interactions of the protease and helicase domains.

The proposed binding of ssRNA to ZIKV NS2B-NS3pro in an open conformation provides new insights into the functional cycle of ZIKV RNA processing and viral replication. The appearance of pockets in NS2B-NS3pro accessible only in the open and super-open states may provide opportunities for the design of novel allosteric inhibitors that simultaneously affect ssRNA binding and protease catalytic activity.

We believe that these data will facilitate the development of new approaches for the design of allosteric inhibitors not only of ZIKV but also of other members of the *Flaviviridae* family, given their high (>70 %) primary sequence homology and high structural similarity.

ACKNOWLEDGMENTS

We thank Dr. Alexander Aleshin for critical discussions and suggestions and for providing early access to PDB ID 7M1V. We thank Dr. Laurie Bankston for designing the NS2BΔ-NS3-long construct, Dr. Alex Strongin for valuable advice and discussions, and Dr. Sumit Chanda for critical support to S.A.S. and valuable discussions. This work was supported by NIH grants 5R21AI134581 to R.C.L. and R01 NS105969-01 NOA to A.V.T.

AUTHOR CONTRIBUTIONS

S.A.S., R.C.L., and A.V.T. designed and supervised the study; S.A.S. and A.C. conducted the experiments; S.A.S., R.C.L., and A.V.T. analyzed the data and prepared the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

REFERENCES

1. Campos, G. S., Bandeira, A. C., and Sardi, S. I. (2015) Zika Virus Outbreak, Bahia, Brazil. *Emerg Infect Dis* **21**, 1885-1886
2. Mackenzie, J. S., Gubler, D. J., and Petersen, L. R. (2004) Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses. *Nat Med* **10**, S98-109
3. Vogel, G. (2016) INFECTIOUS DISEASE. Evidence grows for Zika virus as pregnancy danger. *Science* **351**, 1123-1124

4. Solomon, I. H., Milner, D. A., and Folkerth, R. D. (2016) Neuropathology of Zika Virus Infection. *J Neuroinfect Dis* **7**
5. Rasmussen, S. A., Jamieson, D. J., Honein, M. A., and Petersen, L. R. (2016) Zika Virus and Birth Defects--Reviewing the Evidence for Causality. *N Engl J Med* **374**, 1981-1987
6. Enfissi, A., Codrington, J., Roosblad, J., Kazanji, M., and Rousset, D. (2016) Zika virus genome from the Americas. *Lancet* **387**, 227-228
7. Cao-Lormeau, V. M., Blake, A., Mons, S., Lastere, S., Roche, C., Vanhomwegen, J., Dub, T., Baudouin, L., Teissier, A., Larre, P., Vial, A. L., Decam, C., Choumet, V., Halstead, S. K., Willison, H. J., Musset, L., Manuguerra, J. C., Despres, P., Fournier, E., Mallet, H. P., Musso, D., Fontanet, A., Neil, J., and Ghawche, F. (2016) Guillain-Barre Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet* **387**, 1531-1539
8. Lucchese, G., and Kanduc, D. (2016) Zika virus and autoimmunity: From microcephaly to Guillain-Barré syndrome, and beyond. *Autoimmun Rev* **15**, 801-808
9. Gorshkov, K., Shiryaev, S. A., Fertel, S., Lin, Y. W., Huang, C. T., Pinto, A., Farhy, C., Strongin, A. Y., Zheng, W., and Terskikh, A. V. (2018) Zika Virus: Origins, Pathological Action, and Treatment Strategies. *Front Microbiol* **9**, 3252
10. Song, B. H., Yun, S. I., Woolley, M., and Lee, Y. M. (2017) Zika virus: History, epidemiology, transmission, and clinical presentation. *J Neuroimmunol* **308**, 50-64
11. Baud, D., Gubler, D. J., Schaub, B., Lanteri, M. C., and Musso, D. (2017) An update on Zika virus infection. *Lancet* **390**, 2099-2109
12. Padmanabhan, R., Mueller, N., Reichert, E., Yon, C., Teramoto, T., Kono, Y., Takhampunya, R., Ubol, S., Pattabiraman, N., Falgout, B., Ganesh, V. K., and Murthy, K. (2006) Multiple enzyme activities of flavivirus proteins. *Novartis Found Symp* **277**, 74-84; discussion 84-76, 251-253
13. Shiryaev, S. A., Farhy, C., Pinto, A., Huang, C. T., Simonetti, N., Elong Ngono, A., Dewing, A., Shresta, S., Pinkerton, A. B., Cieplak, P., Strongin, A. Y., and Terskikh, A. V. (2017) Characterization of the Zika virus two-component NS2B-NS3 protease and structure-assisted identification of allosteric small-molecule antagonists. *Antiviral Res* **143**, 218-229
14. Shiryaev, S. A., Cheltsov, A. V., Gawlik, K., Ratnikov, B. I., and Strongin, A. Y. (2011) Virtual ligand screening of the National Cancer Institute (NCI) compound library leads to the allosteric inhibitory scaffolds of the West Nile Virus NS3 proteinase. *Assay Drug Dev Technol* **9**, 69-78
15. Shiryaev, S. A., Cheltsov, A. V., and Strongin, A. Y. (2012) Probing of exosites leads to novel inhibitor scaffolds of HCV NS3/4A proteinase. *PLoS One* **7**, e40029
16. Lei, J., Hansen, G., Nitsche, C., Klein, C. D., Zhang, L., and Hilgenfeld, R. (2016) Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. *Science* **353**, 503-505
17. Aleshin, A. E., Shiryaev, S. A., Strongin, A. Y., and Liddington, R. C. (2007) Structural evidence for regulation and specificity of flaviviral proteases and evolution of the Flaviviridae fold. *Protein Sci* **16**, 795-806
18. Shiryaev, S. A., Chernov, A. V., Shiryaeva, T. N., Aleshin, A. E., and Strongin, A. Y. (2011) The acidic sequence of the NS4A cofactor regulates ATP hydrolysis by the HCV NS3 helicase. *Arch Virol* **156**, 313-318
19. Rajagopal, V., Gurjar, M., Levin, M. K., and Patel, S. S. (2010) The protease domain increases the translocation stepping efficiency of the hepatitis C virus NS3-4A helicase. *J Biol Chem* **285**, 17821-17832
20. Beran, R. K., Serebrov, V., and Pyle, A. M. (2007) The serine protease domain of hepatitis C viral NS3 activates RNA helicase activity by promoting the binding of RNA substrate. *J Biol Chem* **282**, 34913-34920

21. Gebhard, L. G., Kaufman, S. B., and Gamarnik, A. V. (2012) Novel ATP-independent RNA annealing activity of the dengue virus NS3 helicase. *PLoS One* **7**, e36244
22. Luo, D., Wei, N., Doan, D. N., Paradkar, P. N., Chong, Y., Davidson, A. D., Kotaka, M., Lescar, J., and Vasudevan, S. G. (2010) Flexibility between the protease and helicase domains of the dengue virus NS3 protein conferred by the linker region and its functional implications. *J Biol Chem* **285**, 18817-18827
23. Chernov, A. V., Shiryaev, S. A., Aleshin, A. E., Ratnikov, B. I., Smith, J. W., Liddington, R. C., and Strongin, A. Y. (2008) The two-component NS2B-NS3 proteinase represses DNA unwinding activity of the West Nile virus NS3 helicase. *J Biol Chem* **283**, 17270-17278
24. Pagano, J. M., Clingman, C. C., and Ryder, S. P. (2011) Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA* **17**, 14-20
25. Hammamy, M. Z., Haase, C., Hammami, M., Hilgenfeld, R., and Steinmetzer, T. (2013) Development and characterization of new peptidomimetic inhibitors of the West Nile virus NS2B-NS3 protease. *ChemMedChem* **8**, 231-241
26. Lin, K. H., Ali, A., Rusere, L., Soumana, D. I., Kurt Yilmaz, N., and Schiffer, C. A. (2017) Dengue Virus NS2B/NS3 Protease Inhibitors Exploiting the Prime Side. *J Virol* **91**
27. Nitsche, C. (2019) Proteases from dengue, West Nile and Zika viruses as drug targets. *Biophys Rev* **11**, 157-165
28. Rut, W., Zhang, L., Kasperkiewicz, P., Poreba, M., Hilgenfeld, R., and Drag, M. (2017) Extended substrate specificity and first potent irreversible inhibitor/activity-based probe design for Zika virus NS2B-NS3 protease. *Antiviral Res* **139**, 88-94
29. Shiryaev, S. A., Aleshin, A. E., Muranaka, N., Kukreja, M., Routenberg, D. A., Remacle, A. G., Liddington, R. C., Cieplak, P., Kozlov, I. A., and Strongin, A. Y. (2014) Structural and functional diversity of metalloproteinases encoded by the Bacteroides fragilis pathogenicity island. *FEBS J* **281**, 2487-2502
30. Johnston, P. A., Phillips, J., Shun, T. Y., Shinde, S., Lazo, J. S., Huryn, D. M., Myers, M. C., Ratnikov, B. I., Smith, J. W., Su, Y., Dahl, R., Cosford, N. D., Shiryaev, S. A., and Strongin, A. Y. (2007) HTS identifies novel and specific uncompetitive inhibitors of the two-component NS2B-NS3 proteinase of West Nile virus. *Assay Drug Dev Technol* **5**, 737-750
31. Sidique, S., Shiryaev, S. A., Ratnikov, B. I., Herath, A., Su, Y., Strongin, A. Y., and Cosford, N. D. (2009) Structure-activity relationship and improved hydrolytic stability of pyrazole derivatives that are allosteric inhibitors of West Nile Virus NS2B-NS3 proteinase. *Bioorg Med Chem Lett* **19**, 5773-5777
32. Milojevic, T., Sonnleitner, E., Romeo, A., Djinovic-Carugo, K., and Blasi, U. (2013) False positive RNA binding activities after Ni-affinity purification from Escherichia coli. *RNA Biol* **10**, 1066-1069
33. Radichev, I., Shiryaev, S. A., Aleshin, A. E., Ratnikov, B. I., Smith, J. W., Liddington, R. C., and Strongin, A. Y. (2008) Structure-based mutagenesis identifies important novel determinants of the NS2B cofactor of the West Nile virus two-component NS2B-NS3 proteinase. *J Gen Virol* **89**, 636-641