Programmed -2/-1 Ribosomal Frameshifting in Simarteriviruses: An Evolutionarily Conserved Mechanism

Yanhua Li[a,b], Andrew E. Firth[b], Ian Brierley[b], Yingyun Cai[c], Sawsan Napthine[b], Tao Wang[a,d], Xingyu Yan[a], Jens H. Kuhn[c], Ying Fang[a]#

[a]Department of Diagnostic Medicine and Pathobiology, Kansas State University, Manhattan, Kansas, USA

[b]Division of Virology, Department of Pathology, University of Cambridge, Cambridge CB2 1QP, United Kingdom

[c]Integrated Research Facility at Fort Detrick, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Fort Detrick, Frederick, Maryland, USA

[d]Yangzhou University, Yangzhou, P. R. China

Running title: -2/-1 PRF in Simarteriviruses

#Address correspondence to Ying Fang, yfang@vet.k-state.edu.

ABSRACT

The -2/-1 programmed ribosomal frameshifting (-2/-1 PRF) mechanism in porcine reproductive and respiratory syndrome virus (PRRSV) leads to the translation of two additional viral proteins, nsp2TF and nsp2N. This -2/-1 PRF mechanism is transactivated by a viral protein nsp1β and cellular poly(rC) binding proteins (PCBPs). Critical elements for -2/-1 PRF, including a slippery sequence and a downstream C-rich motif, were also identified in eleven simarteriviruses. However, the slippery sequences (XXXUCUCU instead of XXXUUUUU) in seven simarteriviruses can only facilitate -2 PRF to generate nsp2TF. The nsp1β of simian hemorrhagic fever virus (SHFV) was identified as a key factor that transactivates both -2 and -1 PRF, and universally conserved Tyr111 and Arg114 in nsp1β are essential for this activity. *In vitro* translation experiments demonstrated the involvement of PCBPs in simarterivirus -2/-1 PRF. Using SHFV reverse genetics, we confirmed critical roles of nsp1β, slippery sequence and C-rich motif in -2/-1 PRF in SHFV-infected cells. Attenuated virus growth ability was observed in SHFV mutants with impaired expression of nsp2TF and nsp2N. Comparative genomic sequence analysis showed that key elements of -2/-1 PRF are highly conserved in all known arteriviruses, except equine arteritis virus (EAV) and wobbly possum disease virus (WPDV). Furthermore, -2/-1 PRF with SHFV PRF signal RNA can be stimulated by heterotypic nsp1βs of all tested non-EAV arteriviruses. Taken together, these data suggest that -2/-1 PRF is an evolutionarily conserved mechanism employed in non-EAV/WPDV arteriviruses for the expression of additional viral proteins that are important for viral replication.

IMPORTANCE

Simarteriviruses are a group of arteriviruses infecting nonhuman primates, and a number of new species have been established in recent years. Although these arteriviruses are widely distributed among African nonhuman primates of different species and some of them caused lethal hemorrhagic fever disease, this group of viruses has been under-characterized. Since wild nonhuman primates are historically important sources or reservoirs of human pathogens, there is concern that simarteriviruses may be “preemergent” zoonotic pathogens. Thus, molecular characterization of simarteriviruses is becoming a priority in arterivirology. In this study, we demonstrated that an evolutionarily conserved ribosomal frameshifting mechanism is used by simarteriviruses and other distantly related arteriviruses for the expression of additional viral proteins. This mechanism is unprecedented in eukaryotic systems. Given the crucial role of ribosome function in all living systems, the potential impact of the in-depth characterization of this novel mechanism reaches beyond the field of virology.

INTRODUCTION

The family *Arteriviridae* within the order *Nidovirales* initially contained four species of positive-stranded RNA viruses, namely porcine reproductive and respiratory syndrome virus (PRRSV), mouse lactate dehydrogenase-elevating virus (LDV), equine arteritis virus (EAV), and simian hemorrhagic fever virus (SHFV), which were assigned to four species (1). Currently, this family has been reclassified into six subfamilies with 19 species (2), in which the two genotypes of PRRSV currently belong to two different species (PRRSV-1 and PRRSV-2), and the newly identified wobbly possum disease virus [WPDV; (3)], Chinese rat arterivirus (RatAV) and Ningxia rat arterivirus (RatAV_Ningxia2015) (2, 4), African pouched rat arterivirus [APRAV; (5)] and new simarteriviruses were added as members of new species (6). Among these arteriviruses, equine arteritis virus (EAV; *Equarterivirinae*) and PRRSV-1 and PRRSV-2 (*Variarterivirinae*) are economically important veterinary pathogens (7). Simarteriviruses comprise a rapidly expanding subfamily of under-characterized arteriviruses now known to infect a wide range of nonhuman primates. The most notorious simarteriviruses, Pebjah virus (PBJV), simian hemorrhagic encephalitis virus (SHEV) and SHFV, have repeatedly caused highly lethal hemorrhagic fever epizootics among captive macaques in the UK, US, and USSR throughout the 1960s to 1990s [reviewed in (8)], but their natural host reservoirs are unknown. In contrast, other simarteriviruses found in nature, in specific sub-clinically infected African nonhuman primates, have not been implicated in epizootics. These viruses include DeBrazza's monkey virus 1 (DeBMV-1) in DeBrazza's monkeys (*Cercopithecus neglectus*); Drakensberg

Mountain vervet virus (DMVV-1) in vervet monkeys (*Chlorocebus pygerythrus*) (6); Kafue kinda-chacma baboon virus (KKCBV) in Kinda baboons (*Papio kindae*); Kibale red colobus viruses 1 and 2 (KRCV-1 and KRCV-2) in Ugandan red colobus (*Procolobus rufomitratus tephrosceles*) (9); Kibale red-tailed guenon viruses 1 and 2 (KRTGV-1/2) in red-tailed monkeys (*Cercopithecus ascanius*) (10); Mikumi yellow baboon virus 1 (MYBV-1) in yellow baboons (*Papio cynocephalus*) (11); SWBV-1 in olive baboons (Papio anubis) (11); and Zambian malbrouck virus 1 (ZMbV-1) in malbrouck monkeys (*Chlorocebus cynosuros*) (6). In experimental settings, KRCV-1 can cause mild disease in crab-eating macaques (*Macaca fascicularis*) (12), whereas SWBV-1 infects, but does not appear to cause disease in rhesus monkeys (*Macaca mulatta*) (13). Since wild nonhuman primates are historically important sources/reservoirs of human pathogens, and because of their broad and diverse distribution among African monkeys, there is concern that some simian arteriviruses may be "preemergent" zoonotic pathogens (14). Hence, increased molecular characterization of simarteriviruses is becoming a priority in arterivirology.

Most RNA viruses have polycistronic genomes and have evolved strategies to overcome a limitation of the eukaryotic translation apparatus, namely that in general, only the 5'-most open reading frame (ORF) on an mRNA is translated. These include non-canonical translation mechanisms such as programmed ribosomal frameshifting [PRF] and alternative initiation, and in addition, the expression of polyproteins that are subsequently cleaved by virus or host proteases. Viruses may also produce subgenomic mRNAs that are functionally monocistronic.

Arteriviruses employ several of these strategies to co-ordinate their complex replication cycle (15, 16).

Arterivirus genomes vary in length between 12.5-15.5 kb and contain 10-15 known ORFs. All but two are located towards the 3' end of the genome and encode viral structural proteins that are translated from a nested set of subgenomic mRNAs (17). ORF1a and ORF1b, at the 5' end of the genome, comprise some three-quarters of the coding capacity and encode replicase-associated proteins. Translation of the genomic RNA yields the ORF1a polyprotein and in addition, an ORF1ab fusion polyprotein following -1 PRF at the ORF1a/ORF1b junction (18, 19). The two polyproteins are processed into individual functional nonstructural proteins (nsps) by ORF1a-encoded protease domains. In PRRSV, these comprise two papain-like proteases, PLP1α and PLP1β, located in nsp1α and nsp1β, respectively, a papain-like protease (PLP2) domain situated in the N-terminus of nsp2, and a serine protease domain resided in nsp4. The rapid release of nsp1α, nsp1β and nsp2 from the N-terminus of the polyprotein is mediated by autocatalytic cleavage with PLP1α (between nsp1α/1β), PLP1β (nsp1β/2) and PLP2 (nsp2/3) (20). In the SHFV nsp1 region, three papain-like proteases (PLPα, PLPβ and PLPγ) are present within nsp1 subunits nsp1α, nsp1β, and nsp1γ. Similar to PRRSV, these nsp1 subunits are released from the N-terminus of SHFV polyproteins by autocleavages (21). Nsp2, the largest replicase subunit of arteriviruses, is a multifunctional protein, which plays important roles in viral replication and virus-host interaction (20, 22-30). In addition to cleavage of the nsp2/3 site, the PLP2 domain functions as a co-factor for the serine protease during proteolytic processing of

the C-terminal region of pp1a and pp1ab (20, 31). The C-terminus of nsp2 contains a highly conserved Cys-rich domain of unknown function, and a multispanning transmembrane domain that plays critical role in the formation of membranous structures (32).

PRRSV uses an unusual -2 programmed ribosomal frameshifting (PRF) signal to direct efficient expression of an additional protein from the +1 reading frame overlapping the nsp2-encoding region. The -2 PRF generates a transframe (TF) fusion protein, nsp2TF. It consists of the N-terminal two thirds of nsp2, followed by a unique C-terminal domain that is encoded by a novel overlapping TF ORF (33). At the same frameshifting site, an immediate stop codon was generated by -1 PRF, which leads to the expression of a truncated nsp2, designated as nsp2N (34). Remarkably, both cellular poly(rC) binding proteins (PCBPs) and viral nsp1β are required for efficient -2/-1 PRF (35). Sequence analysis showed that the signals for -2/-1 PRF, including a slippery sequence and downstream C-rich RNA motif, are highly conserved in all arterivirus genomes except that of EAV (36). However, variations in slippery sequences were identified in several newly identified simarteriviruses. In this study, we demonstrated that -2/-1 PRF identified in PRRSV is evolutionarily conserved in non-EAV/WPDV arteriviruses, with particular emphasis on simarteriviruses. This study provides additional insights into the biological characteristics of arteriviruses and advances our knowledge of non-canonical translation mechanisms in both virus infection and cellular systems.

MATERIALS AND METHODS

Cells and viruses: Embryonic grivet (*Chlorocebus aethiops*) kidney epithelial cells (MARC-145; ATCC CRL-12231) were used for SHFV propagation and subsequent experiments. Human embryonic kidney 293T (HEK-293T; ATCC CRL-3216) cells were used for ectopic protein expression. These cells were maintained in minimum essential medium (MEM; Thermo Fisher Scientific, Waltham, MA) supplemented with 10% heat-inactivated fetal bovine serum (FBS; Sigma-Aldrich, St. Louis, MO), 100 U/ml of Penicillin-Streptomycin (Thermo Fisher Scientific, Waltham, MA), and 0.25 μg/mL of Gibco™ Amphotericin B (Thermo Fisher Scientific, Waltham, MA) at 37 °C with 5% CO_2. SHFV isolate KS-06_17_11 [strain LVR 42-0/M6941; GenBank accession No. KM373784.1 (37)] and mutants thereof were used for experiments. These mutants were rescued with SHFV reverse genetics and designated as vY111A, vR114A, vSS1 and vSS2, and vCC1.

Plasmids: Based on a previous study (21), SHFV nsp1β is predicted to contain amino acids (aa) 166-350 of SHFV ORF1a, while nsp2 contains aa 486-1236 of SHFV ORF1a. In this study, the coding sequences of nsp1β and nsp2 were inserted into the pcDNA^{TM}3.1$^{(+)}$ vector (Thermo Fisher Scientific, Waltham, MA) under the control of the CMV promoter and also N-terminally tagged with FLAG (DYKDDDDK) epitope or hemagglutinin (HA) epitope, designated as pFLAG-SHFV-nsp1β and pHA-SHFV-nsp2, respectively. Mutations in nsp1β were generated using the QuikChange® Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions. Wild type (wt) nsp1β and mutants thereof were also

cloned into the pGEX-6p-2 vector (GE Healthcare, Chicago, IL) for expression as glutathione-S-transferase (GST)-tagged proteins. The coding regions of nsp1β and PLP2 domain were also cloned into pET28a (+) vector (MilliporeSigma, Burlington, MA). Plasmids expressing nsp1β and nsp2 proteins of PRRSV were described elsewhere (34). The predicted nsp1β gene sequence of KRCV-1 (GenBank accession No. HQ845737.1) were cloned into the p3xFLAG-Myc-CMV™-24 expression vector (Sigma-Aldrich, St. Louis, MO), while the predicted nsp2 gene sequence of KRCV-1 (GenBank accession No. HQ845737.1) tagged with hemagglutinin (HA) epitope at the N-terminus was cloned into the pcDNA^{TM}3.1$^{(+)}$ vector (Thermo Fisher Scientific, Waltham, MA). The LDV Plagemann strain (LDV-P) nsp1β gene was codon-optimized for expression in human cells and cloned into the p3xFLAG-Myc-CMV™-24 expression vector (Sigma-Aldrich, St. Louis, MO). The plasmid was designated as pFLAG-LDV-nsp1β. The plasmid expressing nsp1 of EAV Bucyrus strain was described previously (38). A dual luciferase reporter plasmid, pDluc (39), was used for evaluation of *in vitro* programmed ribosomal frameshifting (PRF) efficiencies. The pDluc-SHFV/WT plasmid and mutants thereof were generated by inserting the SHFV -2/-1 PRF signal into pDluc between *Renilla* and firefly luciferase reporter genes using the method we described previously (34). The pDluc-PRRSV/WT plasmid containing -2/-1 PRF signals of PRRSV-1 was described in our previous study (34). Human PCBP2 (NM_005016.5) cloned in pcDNA^{TM}3.1$^{(+)}$ was kindly provided by Professor Asit K. Pattnaik, University of Nebraska-Lincoln, Nebraska. For expression of the hexahistidine (His)-tagged recombinant protein, the PCBP2 gene was cloned

into the pET28a (+) vector (MilliporeSigma, Burlington, MA) with a His tag on the N-terminus.

Antibodies: Monoclonal antibodies (mAbs) against SHFV nsp1β and the nsp2 N-terminal PLP2 domain were generated using the method that we described previously (40). The mouse experiment was performed according to protocols approved by the Institutional Animal Care and Use Committee (IACUC) of Kansas State University. Briefly, BALB/c laboratory mice (Jackson laboratory, Bar Harbor, ME) were immunized 3 times with 50 μg purified protein mixed with Freund's incomplete adjuvant (Sigma-Aldrich, St. Louis, MO) at 2-week intervals. Mouse splenocytes were fused with NS-1 myeloma cells. Specific anti-nsp1β and anti-nsp2 mAbs were obtained by screening with an immunofluorescence assay. MAbs 133-243 and 134-260 against the SHFV nsp2 N-terminal PLP2 domain and mAb 76-69 against SHFV nsp1β were used in this study. MAbs 36-19 and 140-68 were used for specifically recognizing nsp2 of PRRSV-1 and PRRSV-2, respectively (34). A polyclonal antibody (pAb) against the SHFV nsp2TF C-terminal peptide (RLDSTVVFEETTPLLDQVPVC; nsp2TFC) was produced in rabbits by GenScript (Piscataway, NJ). The following antibodies from commercial resources were also used in this study: anti-HA tag mAb 16B12 (Biolegend, San Diego, CA), anti-FLAG tag mAb M2 (Sigma-Aldrich, St. Louis, MO), anti-human PCBP2 mAb 23-G (Santa Cruz Biotechnology, Dallas, TX), anti-human PCBP1 rabbit pAb (Sigma-Aldrich, St. Louis, MO), anti-Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) pAb (Santa Cruz Biotechnology, Dallas, TX), and anti-Renilla luciferase mAb clone 1D5.2 (MilliporeSigma, Burlington, MA). The mAb 12A4 against EAV nsp1 (41) was generously provided by Dr. Udeni Balasuriya at

Louisiana State University.

Protein expression and purification: Recombinant proteins were expressed and purified using the methods described previously (35). Briefly, His-tagged nsp1β and the PLP2 domain of nsp2 were expressed in *E. coli* BL21/DE3/pLysS (Thermo Fisher Scientific, Waltham, MA) and purified using Ni-nitrilotriacetic acid agarose resin (Ni-NTA; Qiagen, Germantown, MD). GST-tagged nsp1β and mutants thereof were expressed in *E. coli* BL21/DE3/pLysS (Thermo Fisher Scientific, Waltham, MA) and purified using glutathione agarose resin (Thermo Fisher Scientific, Waltham, MA). His-tagged PCBP2 were expressed in *E. coli* BL21/DE3/pLysS (Thermo Fisher Scientific, Waltham, MA) and purified using Ni-nitrilotriacetic acid agarose resin (Ni-NTA; Qiagen, Germantown, MD). Proteins were dialyzed, quantified with Pierce BCA Protein Assay Kit (Thermo Fisher Scientific, Waltham, MA), and stored at −80 ^{o}C.

***In vitro* translation:** *In vitro* translation was performed using nuclease-treated rabbit reticulocyte lysate (RRL; Promega, Madison, WI) or wheat germ extract (WGE; Promega, Madison, WI) as described previously (35). Briefly, 5'-capped messenger RNAs were generated using T7 RNA polymerase (New England Biolabs, Ipswich, MA) with *Fsp*I-linearized pDluc reporter plasmid. *In vitro* translation reactions (10 µl) were reconstituted with mRNA template (50 µg/ml), RRL (9 µl) or WGE (9 µl) containing 20 µM amino acids without methionine (Promega, Madison, WI), 0.2 MBq [^{35}S]-methionine (PerkinElmer, Waltham, MA), and purified viral protein and/or cellular protein. After 1 h incubation at 30 ºC, proteins were separated in 12% SDS–PAGE gels. The protein gels were dried and exposed to X-ray film or to a Cyclone Plus Storage Phosphor

Screen (PerkinElmer, Waltham, MA). Images were developed using a Typhoon TRIO Variable Mode Imager (GE Healthcare, Chicago, IL), and protein bands were quantified using ImageQuant^{TM} TL software (GE Healthcare, Chicago, IL). The formula [IFS1/MetFS1]/[IS/MetS + IFS1/MetFS1 + IFS2/MetFS2] was used to calculate -1 PRF efficiency. In this formula, the number of methionines in the product without a frameshift (stop) and the -1/-2 frameshift (FS) products are denoted by MetS, MetFS1 and MetFS2, respectively, and the densitometry values for the same products are denoted by IS, IFS1 and IFS2, respectively. The -2 PRF efficiency was calculated similarly. All frameshifting assays were repeated at least three times.

Immunoprecipitation and western blots: Expression of PRRSV nsp2TF and nsp2N was detected by immunoprecipitation (IP) and western blot analysis as described previously (34). Similar methods were used for detecting the expression of the SHFV nsp2-related proteins. In the ectopic expression system, HEK-293T cells in 6-well plates were co-transfected with plasmid DNA (1 μg) expressing SHFV nsp2 and plasmid DNA (0.5 μg) expressing SHFV nsp1β or its mutants. The empty vector was included as a control. At 36 h post transfection (hpt), cell lysates were harvested with 300 μl IP lysis wash buffer per well (Thermo Fisher Scientific, Waltham, MA) supplemented with Protease Inhibitor Cocktail (Sigma-Aldrich, St. Louis, MO). For the SHFV infection system, MARC-145 cells seeded in 6-well plates were infected with the parental virus or mutants thereof at a multiplicity of infection (MOI) of 0.01; cells were mock-inoculated to serve as a negative control. At 36 hpi, cell lysates were harvested with 300 μl IP lysis wash buffer per well (Thermo Fisher Scientific, Waltham, MA), supplemented with Protease Inhibitor

Cocktail (Sigma-Aldrich, St. Louis, MO). IP was performed using the Pierce™ Classic Magnetic IP/Co-IP Kit (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's instructions. SHFV nsp2-related proteins were immunoprecipitated with mAb 133-243 against the SHFV PLP2 domain. Western blot analysis was conducted with mAb 134-260 against the SHFV PLP2 domain and rabbit anti-sera against the nsp2TF C-terminal peptide. IRDye® 800CW Goat anti-Mouse IgG (H + L) or/and IRDye® 680RD Goat anti- Rabbit IgG (H + L) (LI-COR Biosciences, Lincoln, NE) were used as secondary antibodies. The target proteins were visualized using a digital imaging system (Odyssey infrared imaging system; LI-COR Biosciences, Lincoln, NE).

IP and western blots were also performed to determine the interactions between SHFV nsp1β and PCBP1/2 as described previously (42). In the ectopic expression system, HEK-293T cells seeded in 6-cm dishes were transfected with plasmid DNA (3 μg) expressing SHFV nsp1β. An empty vector was included as a control. At 36 hpt, cell lysates were harvested with 500 μl IP lysis wash buffer per well (Thermo Fisher Scientific, Waltham, MA) supplemented with Protease Inhibitor Cocktail (Sigma-Aldrich, St. Louis, MO). IP was conducted using FLAG mAb M2 to precipitate the FLAG-tagged SHFV-nsp1β, and western blot analysis was performed to detect SHFV nsp1β and PCBP1/2 using specific mAbs. For the SHFV infection system, MARC-145 cells seeded in 10-cm dishes were infected with SHFV at an MOI of 0.01. Mock-infected cells were included in the analysis as a negative control. At 36 hpi, cell lysates were harvested with 1 ml IP lysis wash buffer per dish (Thermo Fisher Scientific, Waltham, MA) supplemented with

Protease Inhibitor Cocktail (Sigma-Aldrich, St. Louis, MO). IP was conducted using PCBP2 mAb 23-G, and western blot analysis was performed to detect SHFV nsp1β and PCBP2 with specific mAbs.

SHFV reverse genetics system: The SHFV infectious clone (pCMV-SHFV) consists of a commercially synthesized cDNA covering the full-length genome of SHFV (GenBank #KM373784.1) with a CMV immediate-early enhancer and promoter at the 5' end and a hepatitis delta virus ribozyme sequence (pCMV-SHFV) at the 3' end. This cassette was assembled into a pACYC177 vector backbone (GenBank #X06402). The full-length cDNA clones containing mutations at the region of nsp1β or the -2/-1 PRF signal were created using the QuikChange® Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions. Viruses were rescued by transfection of 70–80% confluent MARC-145 cells with 2 μg of pCMV-SHFV or mutants thereof using Transit-LT1 transfection reagent (Mirus Bi LLC, Madison, WI). Viability of the recombinant viruses was determined by observing the cytopathic effect (CPE) and by immunofluorescence assay using SHFV PLP2 mAb 134-260. The cell culture supernatant was harvested at 80% CPE. The recombinant viruses were serially passaged 5 times on MARC-145 cells. Passage 5 (P5) viruses were used for subsequent study. To test the stability of mutants, substitutions in mutants of P5 viruses were verified by sequencing -2/-1 PRF signal and nsp1β coding regions.

***In vitro* characterization of recombinant viruses:** Recombinant viruses were characterized by determining the viral growth kinetics. For multi-step growth curves, MARC-145 cells were

seeded in 24-well plates. When the cells reached 100% confluence, they were infected with parental virus or mutants thereof at an MOI of 0.01. Cell culture supernatants were collected at 12, 24, 36, 48, 60 and 72 h post-infection (hpi). Virus titer was measured by titration on MARC-145 cells and calculated as $TCID_{50}$/mL according to the Reed and Muench method (43). To determine the plaque morphology of the parental virus and corresponding mutants, a plaque assay was performed in MARC-145 cells using the method described previously (33).

RESULTS

Key elements of a -2/-1 programmed ribosomal frameshifting (PRF) mechanism are conserved in non-EAV/WPDV arteriviruses

Previously, we identified the frameshift RNA signal of PRRSV, which consists of a slippery sequence (G_GUU_UUU) and a downstream C-rich motif (CCCANCUCC) separated from the slippery sequence by a 10-nt "spacer" (34, 35). In addition, the PRRSV replicase subunit nsp1β functions as a transactivator for both -2 and -1 PRF; in particular, a highly conserved α-helix motif in nsp1β is critical for PRF transactivation (34). In this study, comparative genomic sequence analysis of arteriviruses was performed to identify the key elements of -2/-1 PRF in viruses of all relevant species in the *Arteriviridae* family. The slippery sequence and C-rich motif of the PRF signal were identified in all simarteriviruses, although a few substitutions were observed for viruses of different species, including U_GUU_UUU (KRTGV, PBJV, and DeMAV), G_GUC_UCU (KRCV-1, KRCV-2, MYBV-1, and KKCBV) and U_UUC_UCU

(FSVV, SHEV, and ZMbV-1) (Fig. 1A). These signals all allow anticodon:codon re-pairing in the A-site following a -2 PRF but, as that in PRRSV, the potential for re-pairing in the P-site is more limited. Of note, these PRF RNA signals were found in all known arteriviruses except EAV. Additionally, the distance between slippery sequence and C-rich motif is 9-10 nucleotides (nt) except in WPDV where the closest C-rich sequence was at a distance of 19 nt. The highly conserved motif in PRRSV-2 nsp1β is also found in all known arteriviruses except WPDV and EAV (Fig.1B). In line with our previous studies in PRRSV (44, 45), two arginine residues in this motif are conserved among most arteriviruses (Arg114 and Arg115 in SHFV). In APRAV-1, the second arginine is replaced with a leucine residue (Fig. 1B). The 3D structures of nsp1β of arteriviruses except WPDV and EAV were predicted with the I-TASSER server and RaptorX server guided by the known crystal structure of PRRSV-2 (XH-GD strain) nsp1β (46-48) (Fig. 1C). The highly conserved motif consistently forms an α-helix. The longer distance between slippery sequence and C-rich sequence in WPDV and its different nsp1β protein conformation compared to other arteriviruses suggest that WPDV likely does not utilize frameshifting at this site. Taken together, these data suggest that -2/-1 PRF might be an evolutionarily conserved mechanism for the expression of additional viral proteins in non-EAV/WPDV arteriviruses.

SHFV nsp1β transactivates -2/-1 ribosomal frameshifting

In PRRSV, -2/-1 PRF results in the translation of two novel proteins, nsp2TF and nsp2N (34). To test our hypothesis that -2/-1 PRF is utilized to translate nsp2TF and nsp2N proteins in

simarteriviruses, we initially characterized nsp2-related proteins in SHFV-infected MARC-145 cells. For the detection of nsp2-related proteins, we generated two monoclonal antibodies (mAb 133-243 and mAb 134-260) against the PLP2 domain, which is shared by SHFV nsp2 and the predicted nsp2TF and nsp2N. A rabbit polyclonal antibody against the unique epitope (RLDSTVVFEETTPL) at the C-terminus of nsp2TF was also generated for specific identification of nsp2TF. The N-terminus of SHFV nsp2 (Gly484/Gly485/Gly486; the residue positions based on SHFV pp1a sequence) was determined by mass spectrometry sequence analysis previously (21). We further predicted the C-terminus of SHFV nsp2 based on the putative cleavage site between nsp2 and nsp3. As shown in Fig. 2A, the cleavage site of SHFV nsp2/3 was predicted to be between Gly1236 and Gly1237 (the residue positions based on SHFV pp1a sequence) by sequence alignment analysis with the nsp2/3 cleavage site in PRRSVs (20). Thus, the SHFV nsp2 gene is predicted to be nt 1649-3901 of the SHFV genome, and the predicted protein size is 81.2 kDa. If -2/-1 PRF occurs, nsp2TF would be translated from SHFV genomic nucleotides 1,649–2,860 fused to nucleotides 2,859–3,536, which results in a product with a predicted size of 68.7 kDa; nsp2N would be translated from SHFV genomic nucleotides 1,649–2,860 fused to 2,860–3,093, resulting in a product with a predicted size of 52 kDa (Fig. 2B). Western blot analysis was performed to identify nsp2-related proteins in SHFV-infected MARC-145 cells. As shown in Figure 2, four major virus-specific bands were detected by mAb 134-260 against SHFV PLP2 (Fig. 2C, left panel). The top three largest proteins appear to be nsp2, nsp2TF and nsp2N, although their masses are larger than those that were predicted - a

phenomenon that was previously reported for nsp2-related proteins of PRRSV (33). The rabbit polyclonal antibody (pAb) specifically designed to detect SHFV nsp2TF recognized two protein bands on western blot. The top band appeared to be a cellular protein, since this band was also detected in mock-infected cells, while the second large protein band was specific to nsp2TF (Fig. 2C, right panel), indicating that SHFV indeed expresses nsp2TF via -2 PRF. The fourth band with a size less than 50 kDa is a C-terminal truncated isoform of nsp2, which may be generated by a proteolytic cleavage of PLP2 or some other proteinase. The identity of this product needs to be further studied in the future. To confirm the identity of these nsp2-related proteins, immunoprecipitation (IP) was performed with anti-PLP2 mAb 133-243 using cell lysate of SHFV-infected MARC-145 cells, while cell lysate of mock-infected MARC-145 cells was used as a negative control. As expected, four major bands were detected in the IP product using mAb 134-260 (Fig. 2D, left panel), and the second largest protein band was also recognized by the nsp2TF-specific pAb (Fig. 2D, right).

An ectopic expression system was further employed to investigate whether the expression of SHFV nsp2TF and nsp2N requires nsp1β as a PRF transactivator. In HEK-293T cells transfected with a plasmid expressing SHFV nsp2 alone, the PRF products, nsp2TF and nsp2N, were not detected (Fig. 3A, lane “EV”). When HEK-293T cells were co-transfected with plasmids expressing nsp2 and nsp1β, both PRF products, nsp2TF and nsp2N, were detected (Figure 3A, lane “nsp1β”). These results indicate that SHFV nsp1β is critical for the expression of nsp2TF and nsp2N. To further identify the key residues in SHFV nsp1β, amino acids Gly109, Lys110,

Tyr111, Arg114 and Arg115 in the conserved motif (Fig. 1A) were targeted for mutagenesis with alanine replacement. Consistent with our previous findings on PRRSV (34, 45), mutation of R114A completely abolished the expression of nsp2TF and nsp2N, whereas the G109A and K110A mutations had no obvious effects on the expression of nsp2TF and nsp2N. The R115A mutant stimulated lower expression levels of nsp2TF and nsp2N, which may be due to the lower expression level of this nsp1β mutant. The Y111A mutant was also unable to stimulate the expression of nsp2TF and nsp2N, indicating that Tyr111 is another key residue for nsp1β's function in transactivation of -2/-1 PRF (Fig. 3A). This result was further confirmed with PRRSV: nsp2TF and nsp2N were not detectable in HEK-293T cells expressing PRRSV nsp1β mutants that contained the corresponding alanine substitution at Y125 (PRRSV-2) or Y131 (PRRSV-1) (Fig. 3B). Similar to the R115A mutant, a lower expression level was observed for the SHFV Y111A mutant (Fig. 3A). However, this lower expression should not be the direct reason leading to the loss of expression of nsp2TF and nsp2N, since the Y111A mutant with similar expression level to wild type nsp1β in vY111A-infected cells was unable to stimulate the expression of nsp2TF and nsp2N (see Fig. 5 below).

A dual luciferase reporter system was employed to confirm the activity of SHFV nsp1β in stimulating -2/-1 PRF at the predicted PRF signal in the SHFV genome. A dual luciferase reporter plasmid, pDluc-SHFV/WT, was generated using the approach that we described previously (34), in which 79 nt containing the SHFV PRF signal are inserted in the plasmid between *Renilla* luciferase (Rluc) and firefly luciferase (Fluc) ORFs. In this construct, Fluc is

encoded in the -2 frame relative to the upstream Rluc ORF. Besides an in-frame Rluc product (stop), two frameshifting products (-2FS and -1FS) can be expressed through -2/-1 PRF (Fig. 4A). The predicted sizes of the stop, -1FS and -2FS products are 40.4 kDa, 43.4 kDa and 100.1 kDa, respectively. In HEK-293T cells co-transfected with the reporter plasmid and a plasmid expressing SHFV nsp1β, all three nsp2-related proteins were detected at the predicted sizes by a monoclonal antibody against their shared N-terminal Rluc (Fig. 4B lane "nsp1β"). Only the in-frame translation product (Stop) was detected in cells without the expression of SHFV nsp1β (Fig. 4B lane "EV" in the "pDluc-SHFV/WT" transfected cells). The ability of nsp1β mutants to activate -2/-1 PRF was also tested using this reporter system. In cells expressing the SHFV nsp1β Y111A or R114A mutants, -2FS and -1FS expression products were undetectable, indicating that Tyr111 and Arg114 are essential to nsp1β's ability to stimulate frameshifting (Fig. 4B). In cells expressing the SHFV nsp1β R115A mutant, reduced expression of the -2FS product was observed while the -1FS product was undetectable. Compared to the reduced expression level of nsp2TF and nsp2N stimulated by the R115A mutant in Fig. 3A, the result suggests that the SHFV nsp1β R115A mutant in the luciferase reporter system may stimulate a much lower level of -1FS that was under the detection limit of western blot analysis.

These findings were further confirmed in an *in vitro* translation assay using rabbit reticulocyte lysate (RRL) programmed with an *in vitro* transcribed reporter mRNA and purified nsp1β protein (Fig. 4C). In this system, the predicted masses of stop, -1FS and -2FS products are 40.4 kDa, 43.4 kDa and 70.8 kDa. When *in vitro* translation was performed with only reporter

mRNA from pDluc-SHFV/WT, the -2FS product was not detected and only a trace amount of -1FS was observed (Fig. 4C, lane “DB”). In contrast, with the addition of purified nsp1β, -1FS and -2FS were detected at the predicted molecular masses. Furthermore, within the range of 0~1 μM of nsp1β, expression levels of the -1 and -2 PRF products were observed to be dose-dependent (Fig. 4C-D). When the concentration of nsp1β was higher than 1 μM, both frameshifts reached their maximum efficiencies, which were 11.7% (-2 PRF) and 3.5% (-1 PRF) (Fig. 4D). Furthermore, purified nsp1β mutant R114A was impaired with its ability to stimulate -2/-1 PRF, as demonstrated by the disappearance of the -2FS product and the amount of -1FS product was reduced to the background level (Fig. 4E).

To further confirm these results in the context of SHFV infection, two recombinant viruses, vY111A and vR114A, were rescued using a SHFV reverse genetics system (constructed from variant strain NIH *LVR*42-0/*M*6941). The ability of vY111A and vR114A to express nsp2TF and nsp2N was evaluated by western blot analysis using cell lysates of SHFV-infected MARC-145 cells. In MARC-145 cells infected with either of the two nsp1β mutants, nsp2TF was not detected and only a low expression level of nsp2N was detected (Fig. 5), which is consistent with the results generated in the *in vitro* expression systems.

Both slippery sequence and C-rich motif are required for efficient -2/-1 PRF in simian arteriviruses

As described above (Fig. 1A), the -2/-1 PRF signals (slippery sequence and C-rich RNA motif)

were found in all simarteriviruses. However, four types of slippery sequence were observed in simarteriviruses of different species, namely G_GUU_UUU (SHFV), U_GUU_UUU (DeMAV, KRTGV, and PBJV), G_GUC_UCU (KKCBV, KRCV-1, KRCV-2, and MYBV-1) and U_UUC_UCU (FSVV, SHEV, ZMbV-1). To test whether some of these variant sequences could support -2/-1 PRF in the context of the SHFV 3' stimulatory sequence and SHFV nsp1β, plasmids pDluc-SHFV/SS1 and pDluc-SHFV/SS2 were created by introducing mutations at the SHFV slippery sequence in the dual luciferase reporter plasmid pDluc-SHFV/WT to mimic the PRF signals of distinct simarteriviruses (Fig. 6A). In the pDluc-SHFV/SS1 construct, the slippery sequence (G_GUU_UUU) was changed to G_GUC_UCU, whereas in the pDluc-SHFV/SS2 construct, the slippery sequence was changed to U_UUC_UCU (Fig. 6A). In HEK-293T cells expressing SHFV nsp1β, the SHFV/WT slippery sequence permits -2/-1 PRF, as evidenced by the detection of -1FS and -2FS expression products with mAb against *Renilla* luciferase in western blots (Fig. 6B). In contrast, only the -2FS product was detected when the SHFV/SS1 or SHFV/SS2 constructs were used. We further confirmed this result using an *in vitro* translation assay in RRL. The reporter mRNAs transcribed from pDluc-SHFV/SS1 or pDluc-SHFV/SS2 constructs and translated in RRL only generated stop and -2FS products in the presence of SHFV nsp1β (Fig. 6C).

We further confirmed the results using simarterivirus KRCV-1. As indicated above, the genome of this virus contains the slippery sequence (G_GUC_UCU). The predicted coding regions of KRCV-1 nsp1β and nsp2 were cloned into eukaryotic expression vectors. The

predicted sizes of KRCV-1 nsp2 and nsp2TF are 77.9 kDa and 64.7 kDa. In HEK-293T cells co-transfected with plasmids containing nsp1β and nsp2, the expression of nsp2TF was detected but no nsp2N was detected. As we expected, no frameshifting products were detected in HEK-293T cells that were not transfected with the plasmid expressing nsp1β (Fig. 6D). These data indicate that the slippery sequence of KRCV-1 lacks the ability to support -1 PRF, which is consistent with the results generated with the luciferase reporter system.

Subsequently, we investigated the function of the C-rich RNA motif in SHFV -2/-1 PRF. The plasmid pDluc-SHFV/CC1 was constructed by introducing synonymous mutations to disrupt two C-rich patches within the PRF signal (Fig. 6A). HEK-293T cells were co-transfected with pDluc-SHFV/CC1 and the plasmid expressing SHFV nsp1β. No frameshifting products were detected (Fig. 6B). This result was further confirmed by *in vitro* translation assay using RRL. Again, no frameshifting products were detected in *in vitro* translation reactions when using the reporter mRNAs transcribed from the pDluc-SHFV/CC1 construct (Fig. 6C), thus indicating that the substitutions of C residues in the C-rich RNA motif knock out both -2 and -1 PRF in the case of SHFV.

Next, we confirmed the data from *in vitro* analysis using recombinant viruses carrying the mutations designed in pDluc-SHFV/SS1, pDluc-SHFV/SS2 and pDluc-SHFV/CC1. Using the SHFV reverse genetics system, three recombinant viruses were generated and designated as vSS1, vSS2, and vCC1. Western blot analysis was performed to assess the expression of nsp2TF and nsp2N in WT- and mutant virus-infected MARC-145 cells. The results were consistent with

454 the data generated in the *in vitro* expression system, in which the -1 PRF product (nsp2N) was
455 not detected in vSS1 and vSS2, whereas expression of both -2 PRF and -1 PRF products was not
456 detected in vCC1 (Fig. 6E). These data indicate that the C-rich RNA motif is required for
457 efficient -2/-1 PRF in simarteriviruses, and that X_XUC_UCU variants of the slippery sequence
458 lack the ability to facilitate -1 PRF.
459
460 **Poly(rC) binding proteins are important for efficient -2/-1 PRF in simarteriviruses**
461 In our previous study, poly(rC) binding proteins (PCBPs) were demonstrated to be critical for
462 -2/-1 PRF in PRRSV (35). To test whether PCBPs are also involved in stimulating -2/-1 PRF in
463 simarteriviruses, *in vitro* translations were performed with SHFV reporter mRNA and the
464 addition of nsp1β and/or PCBP2 in wheat germ extract (WGE). *In vitro* translation using RRL
465 was included as the control. As expected, in RRL, both -2FS and -1FS were stimulated by the
466 presence of nsp1β (Fig. 7A). In the WGE system, although there was some expression of -2FS
467 and -1FS products in the absence of any exogenous protein, these levels were greatly stimulated
468 only upon the addition of both nsp1β and PCBP2. The presence of nsp1β or PCBP2 alone
469 showed no stimulatory effect on the translation of -2FS and -1FS products. (Note that, in contrast
470 to RRL, WGE likely contains endogenous PCBPs that are divergent from those of mammalian
471 cells and not active in the stimulation of PRF.) Interestingly, the frameshifting efficiencies for -1
472 PRF in WGE were much higher than those observed using RRL (Fig. 7B). With the addition of
473 nsp1β and PCBP2 in WGE, the efficiency of -2 PRF increased from 0.9% to 9.8%, whereas the

efficiency of -1 PRF increased from 8% to 29%. These data suggest that PCBPs are required for efficient -2/-1 PRF in simarteriviruses.

Previously, the interaction between PCBPs and PRRSV-1 nsp1β was determined to be required for nsp1β's ability to bind the PRRSV-1 -2/-1 PRF RNA signal (35). In this study, we further analyzed the interactions between PCBP1/2 and SHFV nsp1β. In HEK-293T cells transfected with plasmids expressing FLAG-tagged SHFV nsp1β and PCBP2, the protein complex of nsp1β and PCBP2 was immunoprecipitated by anti-FLAG mAb and subsequently detected by western blot analysis using anti-PCBP2 mAb (Fig. 8A). Consistently, in SHFV-infected MARC-145 cells, nsp1β was pulled down by the mAb against PCBP2 (Fig. 8B). Similarly, the interaction between SHFV nsp1β and PCBP1 was also demonstrated by immunoprecipitation and western blot analysis (Fig. 8C).

The frameshift products play a role in SHFV replication *in vitro*

As described above, a panel of recombinant viruses containing mutations in nsp1β or the -2/-1 PRF signal regions was rescued using SHFV reverse genetics. Five recombinant viruses with deficiencies in the expression of nsp2N and/or nsp2TF were further passaged five times in MARC-145 cells, and the introduced mutations were verified for P5 viruses (data not shown). The P3 recombinant viruses were used to evaluate viral growth kinetics *in vitro*, and wild type (WT) SHFV recovered by the reverse genetics was included as the control. Growth kinetic analysis showed that vY111A, vR114A and vCC1 had attenuated viral growth in MARC-145

cells, whereas vSS1 and vSS2 displayed similar growth kinetics as that of WT virus (Fig. 9A). Before 48 hpi, virus titers of vCC1 were reduced by about 1-log compared to WT. In contrast, growth of the vY111A and vR114A mutants was more significantly reduced, with a 1.5 to 3-log decrease in virus titers throughout the time course of the study (Fig. 9A). The plaque assay results consistently showed that the vCC1, vY111A and vR114A mutants developed smaller plaques than those caused by WT virus (Fig. 9B). All three mutants displaying attenuated growth have lost the ability to express nsp2TF and nsp2N (Fig. 5 and 6E). On the other hand, vSS1 and vSS2 express nsp2TF but not nsp2N and had growth kinetics more similar to that of WT virus. This suggests that nsp2TF plays a role in SHFV replication whereas nsp2N appears not to be important for viral growth in MARC-145 cells. Nonetheless, the relatively mild attenuation of vCC1, which expresses neither nsp2TF nor nps2N, confirms that neither protein is essential for viral growth in cell culture, but they may be important for maintaining maximal virus fitness.

Heterotypic arterivirus nsp1βs stimulate ribosomal frameshifting on the SHFV -2/-1 PRF signal

The genomes of most arteriviruses (with the exception of EAV and WPDV) share a highly conserved RNA-binding motif in nsp1β, which is critical for nsp1β's function in -2/-1 PRF transactivation (Fig. 1B). To further determine whether this is an evolutionarily conserved mechanism in the *Arteriviridae* family, we evaluated the ability of PRRSV-1 nsp1β to transactivate frameshifting on the -2/-1 PRF signal from SHFV, and *vice versa*. *In vitro*

translation was performed in the RRL system with reporter mRNA of SHFV and nsp1β protein of PRRSV-1. As expected, -2 and -1 frameshift products were detected. Increasing concentrations of PRRSV-1 nsp1β led to dose-dependent expression of the two PRF products (Fig. 10A left panel and 10B). Consistently, SHFV nsp1β displayed similar activity to stimulate -2/-1 PRF on the reporter mRNA of PRRSV-1 (Figure 10A, right panel).

To confirm that the PRF transactivation mechanism of nsp1β is also conserved among other arteriviruses, we included additional heterotypic nsp1βs in the assay. The -2FS product was detected in HEK-293T cells transfected with pDluc-SHFV/WT and a plasmid expressing nsp1β from arterivirus of other species, including KRCV-1, PRRSV-1, PRRSV-2 and LDV. Again, the -2FS product was detected; however, the -1FS product was not observed, which may be due to the low efficiency of -1 PRF (Fig. 10C). No frameshifting product was detected in cells expressing EAV nsp1. Since the canonical -2/-1 PRF signal was identified in viruses of all known arteriviral species except EAV and WPDV, these data demonstrate that the transactivation function of nsp1β on -2/-1 PRF is evolutionarily conserved in non-EAV/WPDV arteriviruses.

DISCUSSION

Arteriviruses are a group of mammalian positive-sense RNA viruses. Although most of them have not been associated with overt disease, some arteriviruses cause acute respiratory syndrome, abortion, lethal hemorrhagic fever, or neurological impairment (49-51). EAV, PRRSV-1, and PRRSV-2 are veterinary pathogens with significant economic impact (7). PBJV, SHEV, and

SHFV are known etiologic agents of almost uniformly lethal viral hemorrhagic fever in macaques (8). Most related simarteriviruses have not been identified as pathogens but their infectivity for and transmission ability among nonhuman primates cause concern regarding zoonotic transmission (14). Improved understanding of the biological characteristics of arteriviruses would facilitate the development of disease control strategies and may also advance our knowledge of the factors that drive zoonotic transmission of RNA viruses.

The arterivirus -2/-1 PRF and the involvement of a *trans*-activating viral protein and host factors in ribosomal frameshifting are unprecedented in eukaryotic systems. Our recent studies revealed that the PRF products of PRRSV, nsp2TF and nsp2N, are important for viral replication. On the other hand, these proteins function as innate immune antagonists, suggesting that recombinant PRRSV with impaired nsp2TF/nsp2N expression could be developed as candidate vaccines (52). In a recent comparative genomic study, conserved +1 and -2 PRF signals were identified in the additional sex combs-like (ASXL) genes 1 and 2, respectively, and hypothesized to be utilized for the expression of a conserved overlapping ORF via PRF (53). ASXL genes encode important epigenetic and transcriptional regulatory proteins, and truncation or frameshift mutants of ASXL are linked to myeloid malignancies and genetic diseases. This study highlights the significance of the -2 PRF mechanism, suggesting that the mechanism could be more widely employed in regulating viral/host gene expression.

Comparative genomic analysis of 19 arteriviruses revealed that the key elements for -2/-1 PRF, including the slippery sequence and downstream C-rich RNA motif, are highly conserved

554 in all known arteriviruses except EAV. Of note, the distance between the slippery sequence and
555 downstream C-rich RNA motif is consistently 9-10 nt in all non-EAV arteriviruses with the
556 exception of that of WPDV, which has a stretch that is 9 nucleotides longer. WPDV is
557 phylogenetically distant from other arteriviruses (54). It also lacks a long overlapping TF ORF
558 (36); thus, like EAV, WPDV most likely does not utilize frameshifting in this region of the nsp2
559 gene. Our experimental data indicate that a -2/-1 PRF mechanism similar to that used by PRRSV
560 is employed by simarteriviruses to express nsp2TF and/or nsp2N analogs. In addition, we also
561 experimentally demonstrated that nsp1β proteins of other arteriviruses (KRCV-1, PRRSV-1,
562 PRRSV-2 and LDV, but not EAV) are able to stimulate ribosomal frameshifting on the SHFV
563 -2/-1 PRF RNA signal. These results indicate that -2/-1 PRF is an evolutionarily conserved
564 mechanism used by most arteriviruses for the expression of additional viral proteins.
565 Furthermore, in line with our previous study on PRRSV, at least nsp2TF plays an important role
566 in SHFV replication. Previous studies suggest that the N-terminal PLP2 domain shared by
567 nsp2-related proteins of arteriviruses (PRRSV-1, PRRSV-2, LDV, and SHFV) suppress host type
568 I IFN response through its deubiquitylation activity (30). In comparison with nsp2, PRRSV-2
569 nsp2TF and nsp2N displayed a stronger inhibitory effect on host innate immune responses (52).
570 Therefore, we suspect that SHFV nsp2TF and nsp2N may also play important roles in SHFV
571 infection and virus-host interaction.
572 Three variants of the G_GUU_UUU slippery sequence were identified in simarteriviruses,
573 namely U_GUU_UUU (DeMAV, KRTGV-1, and PBJV), G_GUC_UCU (KKCBV, KRCV-1,

574 KRCV-2, and MYBV-1) and U_UUC_UCU (FSVV, SHEV, and ZMbV-1). In the case of PRRSV,
575 we proposed that tandem slippage of ribosome-bound tRNAs on G_GUU_UUU allows complete
576 A-site re-pairing in both −1 and −2 frames (tRNA anticodon:mRNA codon pairing in 0-frame is
577 3′-AAG-5′:5′- UUU-3′; single $tRNA^{Phe}$ isoacceptor AAG). In contrast, however, tandem slippage
578 of ribosome-bound tRNAs on G_GUC_UCU and U_UUC_UCU does not allow A-site re-pairing
579 in the -1 frame. Consistently, our results showed that the SHFV SS1 and SS2 mutants that mimic
580 these slippery sequence variants only permit -2 PRF. Thus, those simarteriviruses carrying the
581 slippery sequence of G_GUC_UCU or U_UUC_UCU are unlikely to be able to express nsp2N.
582 As indicated above, the SHFV nsp2N is predicted to be an innate immune antagonist. It will be
583 interesting to compare the pathogenicity of this group of simarteriviruses (lack of nsp2N
584 expression) with that of SHFV (expression of nsp2N).
585 PRRSV nsp1β transactivates ribosomal frameshifting through a highly conserved α-helix
586 motif, in which a universally conserved arginine (Arg128 in PRRSV-2 and Arg134 in PRRSV-1)
587 is a key residue for nsp1β activity (34, 45). Comparative sequence analysis of nsp1β of 19
588 arteriviruses showed that the highly conserved α-helix motif to be present in all of them except
589 EAV and WPDV. In our study, we observed that SHFV nsp1β stimulates -2/-1 PRF to express
590 nsp2TF and nsp2N, and an alanine substitution at Arg114 completely impairs this process.
591 Consistently, nsp1β of KRCV-1, a newly isolated simarterivirus, is also capable of stimulating -2
592 PRF. Based on sequence alignment, Arg114 in SHFV nsp1β is the corresponding residue to
593 Arg128 in PRRSV-2 and Arg134 in PRRSV-1, and this arginine is universally conserved among

all arteriviruses except EAV and WPDV. Besides this arginine, Tyr111 in SHFV is also highly conserved among all arteriviruses except EAV and WPDV. Of note, alanine substitution introduced at this residue in SHFV and at analogous residues in PRRSV-1 (Tyr131) and 2 (Tyr125) also impaired nsp1β's ability to stimulate -2/-1 PRF of both PRRSV-1 and PRRSV-2. Remarkably, nsp1βs from PRRSV-1, PRRSV-2, KRCV-1 and LDV are capable of stimulating ribosomal frameshifting on the SHFV -2/-1 PRF signal. The amino acid sequence identity of nsp1βs among these arteriviruses ranges from 21.4% to 57.7%. However, *in silico* structure prediction showed that, with the exception of EAV nsp1 and WPDV nsp1β, all arteriviral nsp1βs share a similar 3D structure, and the conserved α-helix motif was also found in these predicted structures. These data suggest that nsp1β protein structure, especially the α-helix region, is essential for its PRF transactivation function.

Cellular PCBPs were initially identified as interacting partners of PRRSV nsp1β, nsp9 and the genomic 5' untranslated region (5'-UTR) (55). Our recent study further explored the function of PCBPs in enhancing -2/-1 PRF together with viral protein nsp1β (35). We also observed that the enhancement of -2/-1 PRF in SHFV is dependent on the addition of both viral nsp1β and PCBP2 in WGE (Fig. 7A). Interestingly, the frameshifting efficiencies for -1 PRF in WGE were much higher than those observed using RRL. In our previous study on the PRRSV frameshift signal (Napthine et al., 2016), we found efficient -2 PRF in RRL upon the addition of PRRSV nsp1β only. Additional supplementation with PCBP2 maintained -2 PRF but also led to an increase in -1 PRF. We reasoned that RRL may already contain endogenous PCBPs, with a

balance of paralogs that favored -2 PRF over -1 PRF. Turning to WGE translations, we found that supplementation with PCBP2 led preferentially to -1 PRF whereas supplementation with PCPB1 led preferentially to -2 PRF. This suggests that PCBP1 is the abundant form in RRL. Knockdown of either PCBP1 or PCBP2 in mammalian cells using siRNAs produced consistent results. In the current SHFV study, supplementation with PCBP2 could particularly promote -1 PRF in WGE due to the absence of competing endogenous mammalian PCBP1 (compare Fig. 7B with Fig. 4D). As we expected, the interaction between PCBP1/2 and nsp1β was detected by coimmunoprecipitation in HEK-293T cells (Fig. 8A and 8C). In SHFV-infected MARC-145 cells, SHFV nsp1β was also coimmunoprecipitated with endogenous PCBP2 (Fig. 8B). These data suggest that PCBPs are required for efficient -2/-1 PRF in arteriviruses, and the interaction between nsp1β and PCBPs may be also required for the mechanism.

In conclusion, our study demonstrates that -2/-1 PRF is an evolutionarily conserved mechanism used by distantly related arteriviruses for the expression of additional viral proteins; and the PRF products are important for viral replication. This mechanism is unprecedented in eukaryotic systems, not only with the efficient shift to the -2 frame, but also with the involvement of a *trans*-activating viral protein factor (nsp1β) and host cellular protein (PCBPs). Given the crucial role of ribosome function in all living systems, the potential impact of the in-depth characterization of this novel mechanism reaches beyond the field of virology.

ACKNOWLEDGEMENT

We thank Dr. Eric J. Snijder (Leiden University Medical Center, Leiden, The Netherlands) for helpful discussion. Special thanks to Dr. Asit K. Pattnaik (University of Nebraska-Lincoln, Lincoln, Nebraska) for providing pcDNA^{TM}3.1$^{(+)}$ plasmid containing human PCBP2 gene, and Dr. Udeni Balasuriya (Louisiana State University, Baton Rouge, Louisiana) for providing mAb 12A4 against EAV nsp1. We would also like to thank Laura Bollinger of the IRF-Frederick for critically editing the manuscript. This project was supported by Agriculture and Food Research Initiative Competitive Grant no. 2015-67015-22969 from the USDA National Institute of Food and Agriculture (to Y.F.), and Wellcome Trust grant [106207] and European Research Council grant [646891] (to A.E.F.). This work was also funded in part through Battelle Memorial Institute's prime contract with the US National Institute of Allergy and Infectious Diseases (NIAID) under Contract No. HHSN272200700016I (to Y.C., J.H.K). Tao Wang is a visiting scholar from Yangzhou University, Yangzhou, China, who is supported by China Scholarship Council. The content of this publication does not necessarily reflect the views or policies of the US Department of Health and Human Services, or the institutions and companies affiliated with the authors.

REFERNCES

1. Snijder EJ, Kikkert M, Fang Y. 2013. Arterivirus molecular biology and pathogenesis. J Gen Virol 94:2141-63.

2. International Committee on Taxonomy of Viruses (ICTV). 2018. Virus Taxonomy: 2018b Release. https://talkictvonlineorg/taxonomy/.

3. Dunowska M, Biggs PJ, Zheng T, Perrott MR. 2012. Identification of a novel nidovirus associated with a neurological disease of the Australian brushtail possum (Trichosurus vulpecula). Vet Microbiol 156:418-24.

4. Wu Z, Lu L, Du J, Yang L, Ren X, Liu B, Jiang J, Yang J, Dong J, Sun L, Zhu Y, Li Y, Zheng D, Zhang C, Su H, Zheng Y, Zhou H, Zhu G, Li H, Chmura A, Yang F, Daszak P, Wang J, Liu Q, Jin Q. 2018. Comparative analysis of rodent and small mammal viromes to better understand the wildlife origin of emerging infectious diseases. Microbiome 6:178.

5. Kuhn JH, Lauck M, Bailey AL, Shchetinin AM, Vishnevskaya TV, Bao Y, Ng TF, LeBreton M, Schneider BS, Gillis A, Tamoufe U, Diffo Jle D, Takuo JM, Kondov NO, Coffey LL, Wolfe ND, Delwart E, Clawson AN, Postnikova E, Bollinger L, Lackemeyer MG, Radoshitzky SR, Palacios G, Wada J, Shevtsova ZV, Jahrling PB, Lapin BA, Deriabin PG, Dunowska M, Alkhovsky SV, Rogers J, Friedrich TC, O'Connor DH, Goldberg TL. 2016. Reorganization and expansion of the nidoviral family Arteriviridae. Arch Virol 161:755-68.

6. Bailey AL, Lauck M, Ghai RR, Nelson CW, Heimbruch K, Hughes AL, Goldberg TL, Kuhn JH, Jasinska AJ, Freimer NB, Apetrei C, O'Connor DH. 2016. Arteriviruses, Pegiviruses, and Lentiviruses Are Common among Wild African Monkeys. J Virol 90:6724-6737.

7. MacLachlan NJ, Dubovi EJ. 2017. *Arteriviridae* and *Roniviridae*, p 463-476, Fenner's veterinary virology, 5th ed doi:https://doi.org/10.1016/B978-0-12-800946-8.00025-8. Academic Press, Boston.

8. Lauck M, Alkhovsky SV, Bào Y, Bailey AL, Shevtsova ZV, Shchetinin AM, Vishnevskaya TV, Lackemeyer MG, Postnikova E, Mazur S, Wada J, Radoshitzky SR, Friedrich TC, Lapin BA, Deriabin PG, Jahrling PB, Goldberg TL, O'Connor DH, Kuhn JH. 2015. Historical outbreaks of simian hemorrhagic fever in captive macaques were caused by distinct arteriviruses. J Virol 89:8082-7.

9. Lauck M, Hyeroba D, Tumukunde A, Weny G, Lank SM, Chapman CA, O'Connor DH, Friedrich TC, Goldberg TL. 2011. Novel, divergent simian hemorrhagic fever viruses in a wild Ugandan red colobus monkey discovered using direct pyrosequencing. PLoS One 6:e19056.

10. Lauck M, Sibley SD, Hyeroba D, Tumukunde A, Weny G, Chapman CA, Ting N, Switzer WM, Kuhn JH, Friedrich TC, O'Connor DH, Goldberg TL. 2013. Exceptional simian hemorrhagic fever virus diversity in a wild African primate community. J Virol 87:688-91.

11. Bailey AL, Lauck M, Sibley SD, Pecotte J, Rice K, Weny G, Tumukunde A, Hyeroba D, Greene J, Correll M, Gleicher M, Friedrich TC, Jahrling PB, Kuhn JH, Goldberg TL, Rogers J, O'Connor DH. 2014. Two novel simian arteriviruses in captive and wild baboons (Papio spp.). J Virol 88:13231-9.

12. Wahl-Jensen V, Johnson JC, Lauck M, Weinfurter JT, Moncla LH, Weiler AM, Charlier O, Rojas O, Byrum R, Ragland DR, Huzella L, Zommer E, Cohen M, Bernbaum JG, Cai Y, Sanford HB, Mazur S, Johnson RF, Qin J, Palacios GF, Bailey AL, Jahrling PB, Goldberg TL, O'Connor DH, Friedrich TC, Kuhn JH. 2016. Divergent Simian Arteriviruses Cause Simian Hemorrhagic Fever of Differing Severities in Macaques. MBio 7:e02009-15.

13. Buechler C, Semler M, Baker DA, Newman C, Cornish JP, Chavez D, Guerra B, Lanford R, Brasky K, Kuhn JH, Johnson RF, O'Connor DH, Bailey AL. 2018. Subclinical infection of macaques and baboons with a naboon simarterivirus. Viruses 10:701.

14. Bailey AL, Lauck M, Sibley SD, Friedrich TC, Kuhn JH, Freimer NB, Jasinska AJ, Phillips-Conroy JE, Jolly CJ, Marx PA, Apetrei C, Rogers J, Goldberg TL, O'Connor DH. 2016. Zoonotic Potential of Simian Arteriviruses. J Virol 90:630-5.

15. Gorbalenya AE, Enjuanes L, Ziebuhr J, Snijder EJ. 2006. Nidovirales: evolving the largest RNA virus genome. Virus Res 117:17-37.

16. Firth AE, Brierley I. 2012. Non-canonical translation in RNA viruses. J Gen Virol 93:1385-409.

17. Pasternak AO, Spaan WJ, Snijder EJ. 2006. Nidovirus transcription: how to make sense...?

712 J Gen Virol 87:1403-21.
713 18. Brierley I, Digard P, Inglis SC. 1989. Characterization of an efficient coronavirus
714 ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell 57:537-47.
715 19. den Boon JA, Snijder EJ, Chirnside ED, de Vries AA, Horzinek MC, Spaan WJ. 1991.
716 Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. J
717 Virol 65:2910-20.
718 20. Li Y, Tas A, Sun Z, Snijder EJ, Fang Y. 2015. Proteolytic processing of the porcine
719 reproductive and respiratory syndrome virus replicase. Virus Res 202:48-59.
720 21. Vatter HA, Di H, Donaldson EF, Radu GU, Maines TR, Brinton MA. 2014. Functional
721 analyses of the three simian hemorrhagic fever virus nonstructural protein 1 papain-like
722 proteases. J Virol 88:9129-40.
723 22. Song T, Fang L, Wang D, Zhang R, Zeng S, An K, Chen H, Xiao S. 2016. Quantitative
724 interactome reveals that porcine reproductive and respiratory syndrome virus
725 nonstructural protein 2 forms a complex with viral nucleocapsid protein and cellular
726 vimentin. J Proteomics 142:70-81.
727 23. Wang L, Zhou L, Zhang H, Li Y, Ge X, Guo X, Yu K, Yang H. 2014. Interactome profile
728 of the host cellular proteins and the nonstructural protein 2 of porcine reproductive and
729 respiratory syndrome virus. PLoS One 9:e99176.
730 24. Xiao Y, Wu W, Gao J, Smith N, Burkard C, Xia D, Zhang M, Wang C, Archibald A,
731 Digard P, Zhou EM, Hiscox JA. 2016. Characterization of the Interactome of the Porcine

732 Reproductive and Respiratory Syndrome Virus Nonstructural Protein 2 Reveals the
733 Hyper Variable Region as a Binding Platform for Association with 14-3-3 Proteins. J
734 Proteome Res 15:1388-401.
735 25. Fang Y, Fang L, Wang Y, Lei Y, Luo R, Wang D, Chen H, Xiao S. 2012. Porcine
736 reproductive and respiratory syndrome virus nonstructural protein 2 contributes to
737 NF-kappaB activation. Virol J 9:83.
738 26. Sun Z, Li Y, Ransburgh R, Snijder EJ, Fang Y. 2012. Nonstructural protein 2 of porcine
739 reproductive and respiratory syndrome virus inhibits the antiviral function of
740 interferon-stimulated gene 15. J Virol 86:3839-50.
741 27. Faaberg KS, Kehrli ME, Jr., Lager KM, Guo B, Han J. 2010. In vivo growth of porcine
742 reproductive and respiratory syndrome virus engineered nsp2 deletion mutants. Virus Res
743 154:77-85.
744 28. Sun Z, Chen Z, Lawson SR, Fang Y. 2010. The cysteine protease domain of porcine
745 reproductive and respiratory syndrome virus nonstructural protein 2 possesses
746 deubiquitinating and interferon antagonism functions. J Virol 84:7832-46.
747 29. van Kasteren PB, Bailey-Elkin BA, James TW, Ninaber DK, Beugeling C, Khajehpour M,
748 Snijder EJ, Mark BL, Kikkert M. 2013. Deubiquitinase function of arterivirus papain-like
749 protease 2 suppresses the innate immune response in infected host cells. Proc Natl Acad
750 Sci U S A 110:E838-47.
751 30. van Kasteren PB, Beugeling C, Ninaber DK, Frias-Staheli N, van Boheemen S,

752 Garcia-Sastre A, Snijder EJ, Kikkert M. 2012. Arterivirus and nairovirus ovarian tumor
753 domain-containing Deubiquitinases target activated RIG-I to control innate immune
754 signaling. J Virol 86:773-85.
755 31. Wassenaar AL, Spaan WJ, Gorbalenya AE, Snijder EJ. 1997. Alternative proteolytic
756 processing of the arterivirus replicase ORF1a polyprotein: evidence that NSP2 acts as a
757 cofactor for the NSP4 serine protease. J Virol 71:9313-22.
758 32. Snijder EJ, van Tol H, Roos N, Pedersen KW. 2001. Non-structural proteins 2 and 3
759 interact to modify host cell membranes during the formation of the arterivirus replication
760 complex. J Gen Virol 82:985-94.
761 33. Fang Y, Treffers EE, Li Y, Tas A, Sun Z, van der Meer Y, de Ru AH, van Veelen PA,
762 Atkins JF, Snijder EJ, Firth AE. 2012. Efficient -2 frameshifting by mammalian
763 ribosomes to synthesize an additional arterivirus protein. Proc Natl Acad Sci U S A
764 109:E2920-8.
765 34. Li Y, Treffers EE, Napthine S, Tas A, Zhu L, Sun Z, Bell S, Mark BL, van Veelen PA, van
766 Hemert MJ, Firth AE, Brierley I, Snijder EJ, Fang Y. 2014. Transactivation of
767 programmed ribosomal frameshifting by a viral protein. Proc Natl Acad Sci U S A
768 111:E2172-81.
769 35. Napthine S, Treffers EE, Bell S, Goodfellow I, Fang Y, Firth AE, Snijder EJ, Brierley I.
770 2016. A novel role for poly(C) binding proteins in programmed ribosomal frameshifting.
771 Nucleic Acids Res 44:5491-503.

772 36. Gulyaeva A, Dunowska M, Hoogendoorn E, Giles J, Samborskiy D, Gorbalenya AE.
773 2017. Domain Organization and Evolution of the Highly Divergent 5' Coding Region of
774 Genomes of Arteriviruses, Including the Novel Possum Nidovirus. J Virol 91.
775 37. Lauck M, Palacios G, Wiley MR, L Y, Fang Y, Lackemeyer MG, Cai Y, Bailey AL,
776 Postnikova E, Radoshitzky SR, Johnson RF, Alkhovsky SV, Deriabin PG, Friedrich TC,
777 Goldberg TL, Jahrling PB, O'Connor DH, Kuhn JH. 2014. Genome Sequences of Simian
778 Hemorrhagic Fever Virus Variant NIH LVR42-0/M6941 Isolates (Arteriviridae:
779 Arterivirus). Genome Announc. 9:2(5).
780 38. Go YY, Li Y, Chen Z, Han M, Yoo D, Fang Y, Balasuriya UB. 2014. Equine arteritis virus
781 does not induce interferon production in equine endothelial cells: identification of
782 nonstructural protein 1 as a main interferon antagonist. Biomed Res Int 2014:420658.
783 39. Fixsen SM, Howard MT. 2010. Processive selenocysteine incorporation during synthesis
784 of eukaryotic selenoproteins. J Mol Biol 399:385-96.
785 40. Li Y, Tas A, Snijder EJ, Fang Y. 2012. Identification of porcine reproductive and
786 respiratory syndrome virus ORF1a-encoded non-structural proteins in virus-infected cells.
787 J Gen Virol 93:829-39.
788 41. Wagner HM, Balasuriya UB, James MacLachlan N. 2003. The serologic response of
789 horses to equine arteritis virus as determined by competitive enzyme-linked
790 immunosorbent assays (c-ELISAs) to structural and non-structural viral proteins. Comp
791 Immunol Microbiol Infect Dis 26:251-60.

792 42. Guo R, Katz BB, Tomich JM, Gallagher T, Fang Y. 2016. Porcine Reproductive and
793 Respiratory Syndrome Virus Utilizes Nanotubes for Intercellular Spread. J Virol
794 90:5163-5175.
795 43. Reed LJ, Muench H. 1938. A simple method of estimating fifty per cent endpoints.
796 American journal of epidemiology 27:493-497.
797 44. Li Y, Zhu L, Lawson SR, Fang Y. 2013. Targeted mutations in a highly conserved motif
798 of the nsp1beta protein impair the interferon antagonizing activity of porcine
799 reproductive and respiratory syndrome virus. J Gen Virol 94:1972-83.
800 45. Li Y, Shyu DL, Shang P, Bai J, Ouyang K, Dhakal S, Hiremath J, Binjawadagi B,
801 Renukaradhya GJ, Fang Y. 2016. Mutations in a Highly Conserved Motif of nsp1beta
802 Protein Attenuate the Innate Immune Suppression Function of Porcine Reproductive and
803 Respiratory Syndrome Virus. J Virol 90:3584-99.
804 46. Xue F, Sun Y, Yan L, Zhao C, Chen J, Bartlam M, Li X, Lou Z, Rao Z. 2010. The crystal
805 structure of porcine reproductive and respiratory syndrome virus nonstructural protein
806 Nsp1beta reveals a novel metal-dependent nuclease. J Virol 84:6461-71.
807 47. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. 2015. The I-TASSER Suite: protein
808 structure and function prediction. Nat Methods 12:7-8.
809 48. Wang S, Li W, Liu S, Xu J. 2016. RaptorX-Property: a web server for protein structure
810 property prediction. Nucleic Acids Res 44:W430-5.
811 49. Shevtsova ZV, Krylova RI. 1971. [A comparative study of 2 strains of simian

812 hemorrhagic fever virus]. Vopr Virusol 16:686-8.
813 50. Wensvoort G. 1993. Lelystad virus and the porcine epidemic abortion and respiratory
814 syndrome. Vet Res 24:117-24.
815 51. Elazhary Y, Weber J, Bikour H, Morin M, Girard C. 1991. 'Mystery swine disease' in
816 Canada. Vet Rec 129:495-6.
817 52. Li Y, Shang P, Shyu D, Carrillo C, Naraghi-Arani P, Jaing CJ, Renukaradhya GJ, Firth AE,
818 Snijder EJ, Fang Y. 2018. Nonstructural proteins nsp2TF and nsp2N of porcine
819 reproductive and respiratory syndrome virus (PRRSV) play important roles in
820 suppressing host innate immune responses. Virology 517:164-176.
821 53. Dinan AM, Atkins JF, Firth AE. 2017. ASXL gain-of-function truncation mutants:
822 defective and dysregulated forms of a natural ribosomal frameshifting product? Biol
823 Direct 12:24.
824 54. Vanmechelen B, Vergote V, Laenen L, Koundouno FR, Bore JA, Wada J, Kuhn JH,
825 Carroll MW, Maes P. 2018. Expanding the Arterivirus Host Spectrum: Olivier's Shrew
826 Virus 1, A Novel Arterivirus Discovered in African Giant Shrews. Sci Rep 8:11171.
827 55. Beura LK, Dinh PX, Osorio FA, Pattnaik AK. 2011. Cellular poly(c) binding proteins 1
828 and 2 interact with porcine reproductive and respiratory syndrome virus nonstructural
829 protein 1beta and support viral replication. J Virol 85:12939-49.
830 56. Sievers F, Higgins DG. 2018. Clustal Omega for making accurate alignments of many
831 protein sequences. Protein Sci 27:135-145.

832 57. Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new
833 ENDscript server. Nucleic Acids Res 42:W320-4.
834
835

FIGURE LEGENDS

Figure 1. Bioinformatic analysis of -2/-1 programmed ribosomal frameshifting (PRF) RNA signals and nsp1β of arteriviruses. (**A**) The -2/-1 PRF signal, including slippery sequence (orange) and downstream C-rich RNA motif (blue), identified in the nsp2 coding region of arteriviruses. For each sequence, the genome coordinate of the first nucleotide in the alignment is specified. (**B**) A highly conserved α-helix motif in the papain-like cysteine protease β domain (PCPβ) of arterivirus nsp1β. The protein sequences of arterivirus nsp1βs were predicted as described previously (36). Sequence alignment of arterivirus nsp1βs was performed with Clustal Omega (56), and the figure was created with ESPript 3 (57). The highly conserved α-helix motif is indicated with a rectangle in black, and the residues in this motif targeted for mutagenesis are marked with "#". For each sequence, the nsp1β or nsp1 coordinate of the first amino acid in the alignment is specified. (**C**) The 3D structures of arterivirus nsp1βs predicted by the I-TASSER server (47). The α-helix secondary structure of the highly conserved motif in ZMbV-1 nsp1β was also predicted with RaptorX server (48), shown as ZMbV-1_RaptorX. The representative virus strains from 19 arteriviruses were included in the above bioinformatic analysis (2); their GenBank accession numbers are: XH-GD (EU624117.1), PRRSV-2 (SD95-21; KC469618.1), PRRSV-1 (SD01-08; DQ489311.1), RatAV (KP280006.1), RatAV_Ningxia2015 (KU302440.1), LDV-P (U15146.1), APRAV (NC_026439.1), SHFV (KM373784.1), KRTGV (JX473847.1), PBJV (KR139838.1), DeMAV (KP126831.1), KRCV-1 (HQ845737.1), KRCV-2 (KC787631.1), MYBV-1 (KM110935.1), KKCBV (KT447550.1), FSVV (KR862306.1), SHEV (KM677927.1),

ZMbV-1 (KT166441.1), EAV (X53459.3), WPDV (JN116253.3). The highly conserved α-helix motif is highlighted in red.

Figure 2. SHFV nsp2-related proteins expressed in virus-infected MARC-145 cells. (**A**) The C-terminal end of SHFV nsp2 predicted by sequence alignment of polyprotein 1a (pp1a) from SD01-08 (PRRSV-1; DQ489311.1), SD95-21 (PRRSV-2; KC469618.1) and SHFV (KM373784.1). SHFV nsp2 is predicted to be released from pp1a through a proteolytic cleavage by the PLP2 domain. The cleavage site between SHFV nsp2 and nsp3 was predicted to be at Gly1236/Gly1237. For each sequence, the pp1a coordinate of the last amino acid in the alignment is specified. (**B**) A schematic diagram of putative SHFV nsp2, nsp2TF and nsp2N. SHFV nsp2 is encoded by SHFV genomic region of nucleotide (nt) 1649~3901, whereas nsp2TF (nt 1649~2860 + 2859~3536) and nsp2N (nt 1649~2860 + 2860~3093) are translated through -2 PRF and -1 PRF, respectively. The epitopes recognized by antibodies are pointed with black arrows. (**C**) SHFV nsp2-related proteins detected by western blot analysis. MARC-145 cells were infected with SHFV at an MOI of 0.01, and cell lysates were harvested at 36 hpi. The nsp2-related products were detected by anti-PLP2 mAb 134-260, including nsp2, nsp2TF, nsp2N and an unknown nsp2-related protein marked with "*". SHFV nsp2TF was also recognized by rabbit pAb against the nsp2TF C-terminal peptide (nsp2TFC). (**D**) SHFV nsp2-related proteins detected by immunoprecipitation and western blot analysis. MARC-145 cells were infected with SHFV at an MOI of 0.01, and cell lysates were harvested at 36 hpi. The nsp2-related products were immunoprecipitated (IP) using anti-PLP2 mAb 133-243 and probed by western blot (WB)

with anti-PLP2 mAb 134-260 and rabbit pAb anti-nsp2TFC. An unknown nsp2-related protein is marked with "*".

Figure 3. **Identification of key residues involved in PRF transactivation function of nsp1β from SHFV and PRRSV.** (**A**) Immunoprecipitation and western blot analysis of amino acid residues critical for the transactivation function of SHFV nsp1β. HEK-293T cells were co-transfected with a plasmid expressing HA-tagged SHFV nsp2 and a plasmid expressing FLAG-tagged SHFV nsp1β or mutants thereof. An empty vector (EV) was used as the control. Whole cell lysates (WCLs) were harvested at 36 h post transfection (hpt). The expression of nsp2, nsp2TF and nsp2N was detected by western blot (WB) using anti-HA mAb, while FLAG-tagged nsp1β was detected using anti-FLAG mAb M2. The expression of HA-tagged nsp2TF was further confirmed by IP using anti-HA mAb and WB detection using anti-nsp2TFC pAb. (**B**) Immunoprecipitation and western blot analysis of amino acid residues critical for the transactivation function of nsp1β from PRRSV-1 and PRRSV-2. HEK-293T cells were co-transfected with a plasmid expressing PRRSV nsp2 and a plasmid expressing FLAG-tagged PRRSV nsp1β or mutants thereof. An empty vector (EV) was used as a control. PRRSV nsp2-related proteins were immunoprecipitated with anti-PRRSV PLP2 mAb. The expression of nsp2, nsp2TF and nsp2N was evaluated by western blot using anti-PRRSV PLP2 domain mAb, and FLAG-tagged nsp1β was detected with anti-FLAG mAb M2. GAPDH was monitored as a loading control.

Figure 4. **SHFV nsp1β stimulates -2/-1 PRF in pDluc reporter systems. (A)** A schematic

diagram of dual luciferase (pDluc) constructs. A 79-nt sequence (WT) containing the SHFV putative slippery sequence and downstream C-rich motif was inserted between *Renilla* luciferase (Rluc) and firefly luciferase (Fluc; in the -2 frame relative to Rluc) ORFs. The in-frame control (IFC) construct was generated by inserting two Us after the slippery sequence. In addition to the "stop" product translated without frameshifting, -2FS and -1FS products containing *Renilla* luciferase could be translated via programmed -2 PRF and -1 PRF, respectively. **(B)** Identification of critical residues in SHFV nsp1β involved in -2/-1 PRF. HEK-293T cells were co-transfected with plasmid pDluc-SHFV/WT containing wild type SHFV -2/-1 PRF signal and a plasmid expressing SHFV nsp1β or mutants thereof; empty vector (EV) was used as the control. Non-frameshift, -1 PRF and -2 PRF products were detected by western blot using anti-*Renilla* luciferase (Rluc) mAb, and are indicated as stop, -1FS and -2FS, respectively. FLAG-tagged nsp1β was detected with mAb M2 and GAPDH was detected as a loading control. **(C)** SHFV/WT mRNA transcribed from plasmid pDluc-SHFV/WT was translated in RRL in the presence of different concentrations of GST-tagged nsp1β (μM) or with dilution buffer (DB) as a control. **(D)** Efficiencies of -2 PRF and -1 PRF were calculated based on the quantified bands using ImageQuant^{TM} TL software (GE Healthcare). **(E)** *In vitro* translation of SHFV/WT mRNA was performed with 1 μM GST-nsp1β or its mutant. (**C and E**) *In vitro* translation products were resolved in 12% SDS-PAGE and visualized by autoradiography. Products generated without frameshifting or from −1 or −2 PRF are indicated as stop, -1FS and -2FS, respectively.

Figure 5. Tyr111 and Arg114 residues on SHFV nsp1β are critical for -2/-1 PRF during

SHFV infection. MARC-145 cells were infected with SHFV or mutants (vY111A and vR114A) at an MOI of 0.01, and cell lysates were harvested at 36 hpi. The nsp2-related products were detected by anti-PLP2 mAb 134-260. SHFV nsp2TF was also recognized by rabbit pAb against the nsp2TF C-terminal peptide (nsp2TFC). The expression of SHFV nsp1β and mutants thereof was evaluated with mAb 76-69. GAPDH was monitored as a loading control.

Figure 6. The slippery sequence and downstream C-rich RNA motif are required for SHFV −2/−1 PRF. (A) A schematic diagram of dual luciferase (pDluc) constructs. A 79-nt sequence (WT) containing the SHFV putative slippery sequence and downstream C-rich motif assumed to be necessary for -2/-1 PRF was inserted between *Renilla* luciferase (Rluc) and firefly luciferase (Fluc; in the -2 frame relative to Rluc) ORFs. Constructs SS1 and SS2 contain substitutions (red) at the slippery sequence to mimic natural PRF signal variations identified in simarteriviruses other than SHFV. The CC1 construct contains substitutions (red) to disrupt the C-rich motif, while the in-frame control (IFC) construct was generated by inserting two Us after the slippery sequence. **(B)** Effect of mutations in the slippery sequence (SS1 and SS2) or C-rich motif (CC1) on SHFV -2/-1 PRF in the pDluc reporter system. HEK-293T cells were co-transfected with pDluc-SHFV/WT or mutants thereof and a plasmid expressing SHFV nsp1β; empty vector (EV) was used as a control. Non-frameshift, -1 PRF and -2 PRF products (indicated as stop, -1FS and -2FS, respectively) were detected by western blot using anti-*Renilla* luciferase (Rluc) mAb. FLAG-tagged nsp1β was detected with anti-FLAG mAb, and GAPDH was detected as a loading control. **(C)** *In vitro* translation of mRNA derived from the pDluc-SHFV/WT construct or its

mutant in the presence of GST-nsp1β. *In vitro* translation products were resolved in 12% SDS-PAGE and visualized by autoradiography. Products generated without frameshifting or from −1 or −2 PRF are indicated as stop, -1FS, and -2FS respectively. **(D)** The expression of nsp2TF in KRCV-1 stimulated by nsp1β. HEK-293T cells were co-transfected with a plasmid expressing HA-tagged KRCV-1 nsp2 and a plasmid expressing FLAG-tagged wild type nsp1β; empty vector (EV) was used as a control. The expression of nsp2 and nsp2TF was detected by western blot analysis using a mAb against HA-tag, and FLAG-tagged nsp1β was detected with FLAG M2 antibody. GAPDH was monitored as a loading control. (**E**) MARC-145 cells were infected with SHFV or mutants (vSS1, vSS2, and vCC1) at an MOI of 0.01, and cell lysates were harvested at 36 hpi. The nsp2-related products were detected by anti-PLP2 mAb 134-260. SHFV nsp2TF was also detected by rabbit pAb against the nsp2TF C-terminal peptide (nsp2TFC). GAPDH was monitored as a loading control.

Figure 7. Poly(rC) binding protein 2 enhances SHFV -2/-1 PRF in *in vitro* translation systems. (A) *In vitro* translation of mRNA from plasmid pDluc-SHFV/WT in wheat germ extract (WGE) in the presence of nsp1β, PCBP2 or both proteins. Translation of SHFV/WT mRNA in RRL is also shown as a control. Products generated without frameshifting or from -1 or -2 PRF are indicated as stop, -1FS and -2FS, respectively. **(B)** Efficiencies of -2 PRF and -1 PRF were calculated based on the quantified bands using ImageQuant™ TL software (GE Healthcare).

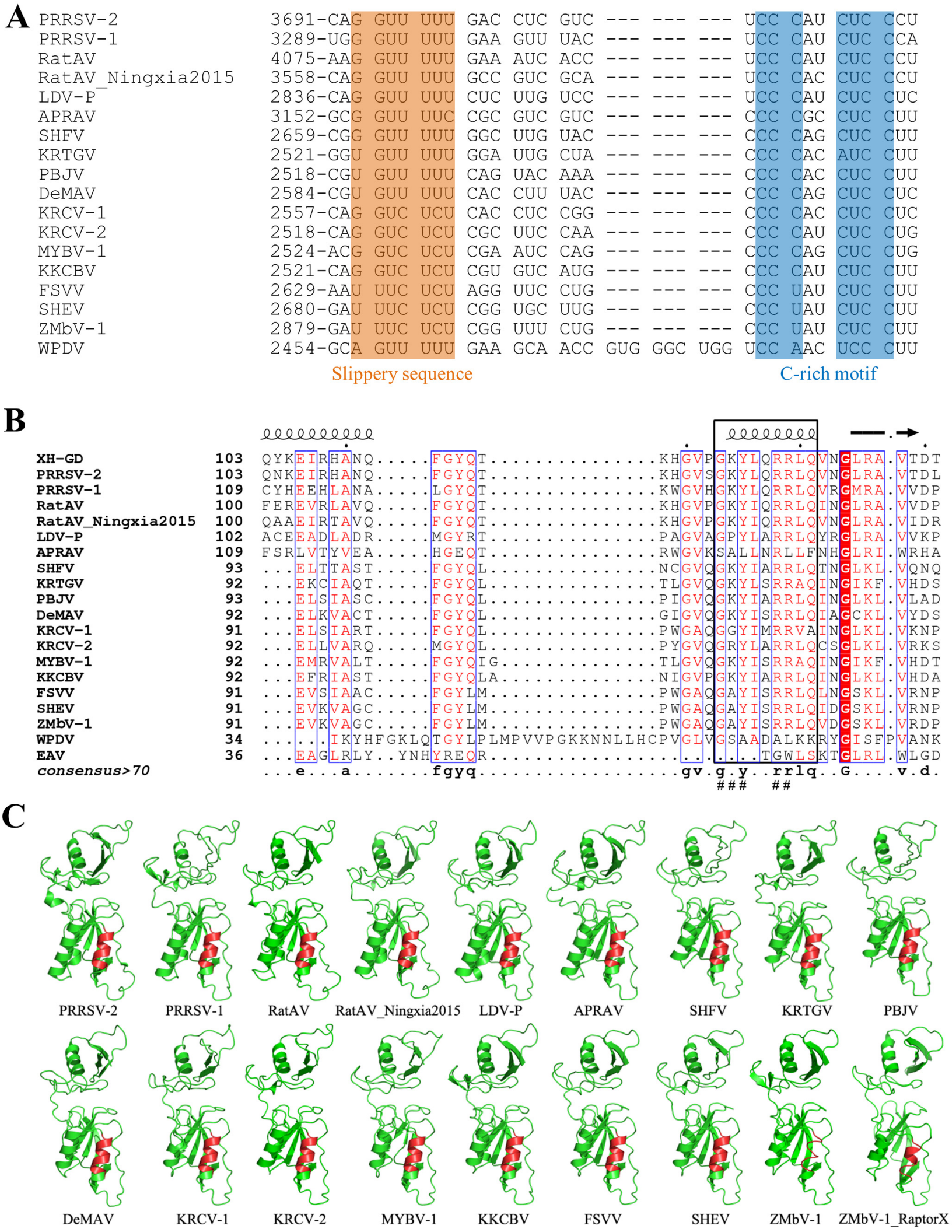
Figure 8. SHFV nsp1β binds to poly(rC) binding protein (PCBP) 1 and PCBP2. (A and C) Ectopically expressed SHFV nsp1β interacts with PCBP1 and PCBP2 in HEK-293T cells.

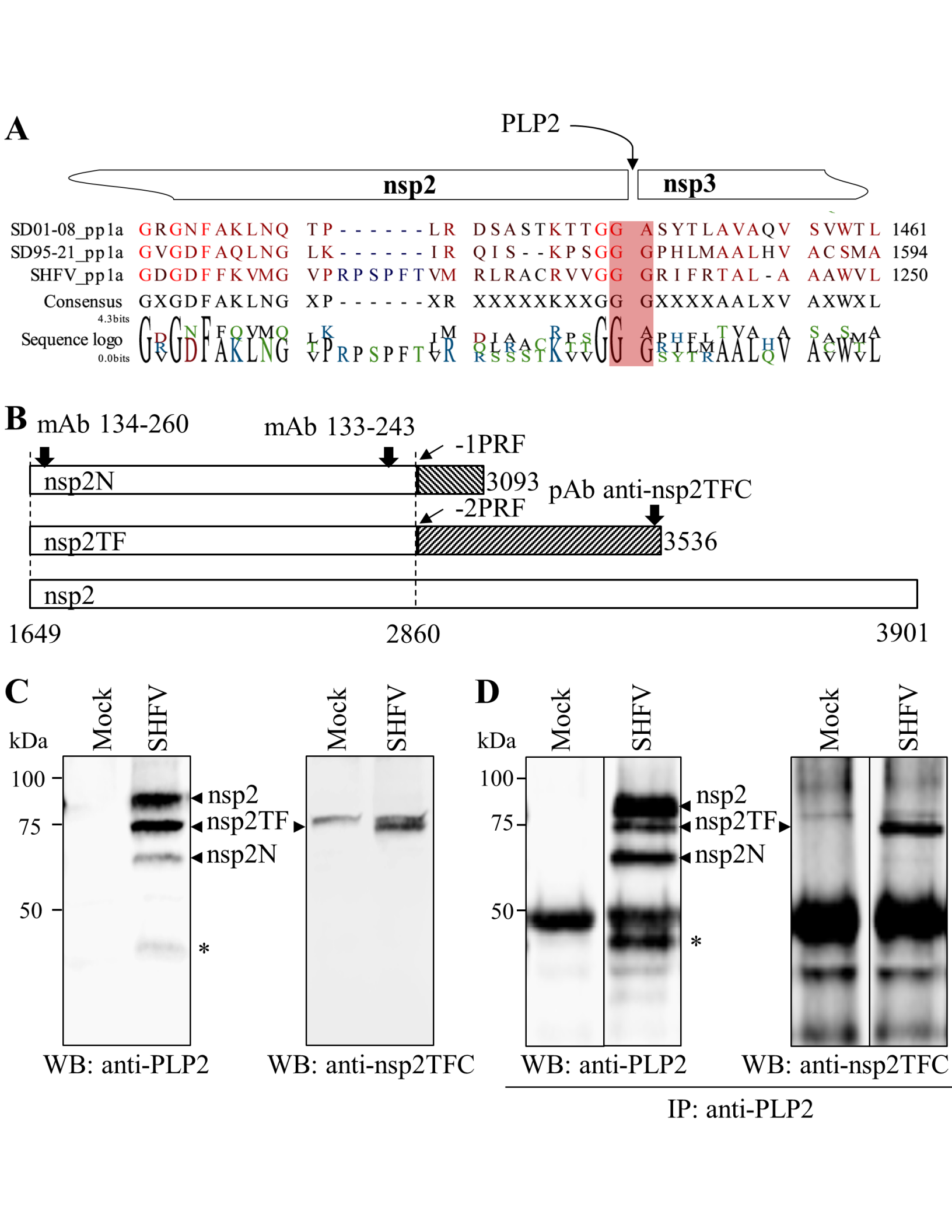
HEK-293T cells were transfected with a plasmid expressing FLAG-tagged SHFV nsp1β or empty vector. The interaction between nsp1β and PCBP1/2 was determined by immunoprecipitation (IP) using anti-FLAG M2 antibody and western blot (WB) analysis using antibodies against PCBP1/2. **(B)** SHFV nsp1β interacts with PCBP2 in virus-infected MARC-145 cells. MARC-145 cells were infected with SHFV at an MOI of 0.01 and harvested at 36 hpi. The interaction between nsp1β and PCBP2 was determined by immunoprecipitation (IP) with anti-PCBP2 mAb and western blot (WB) analysis with mAb 76-69 against SHFV nsp1β. The expression of nsp1β and PCBP1/2 was monitored by western blot analysis with specific antibodies. WCL: whole cell lysate.

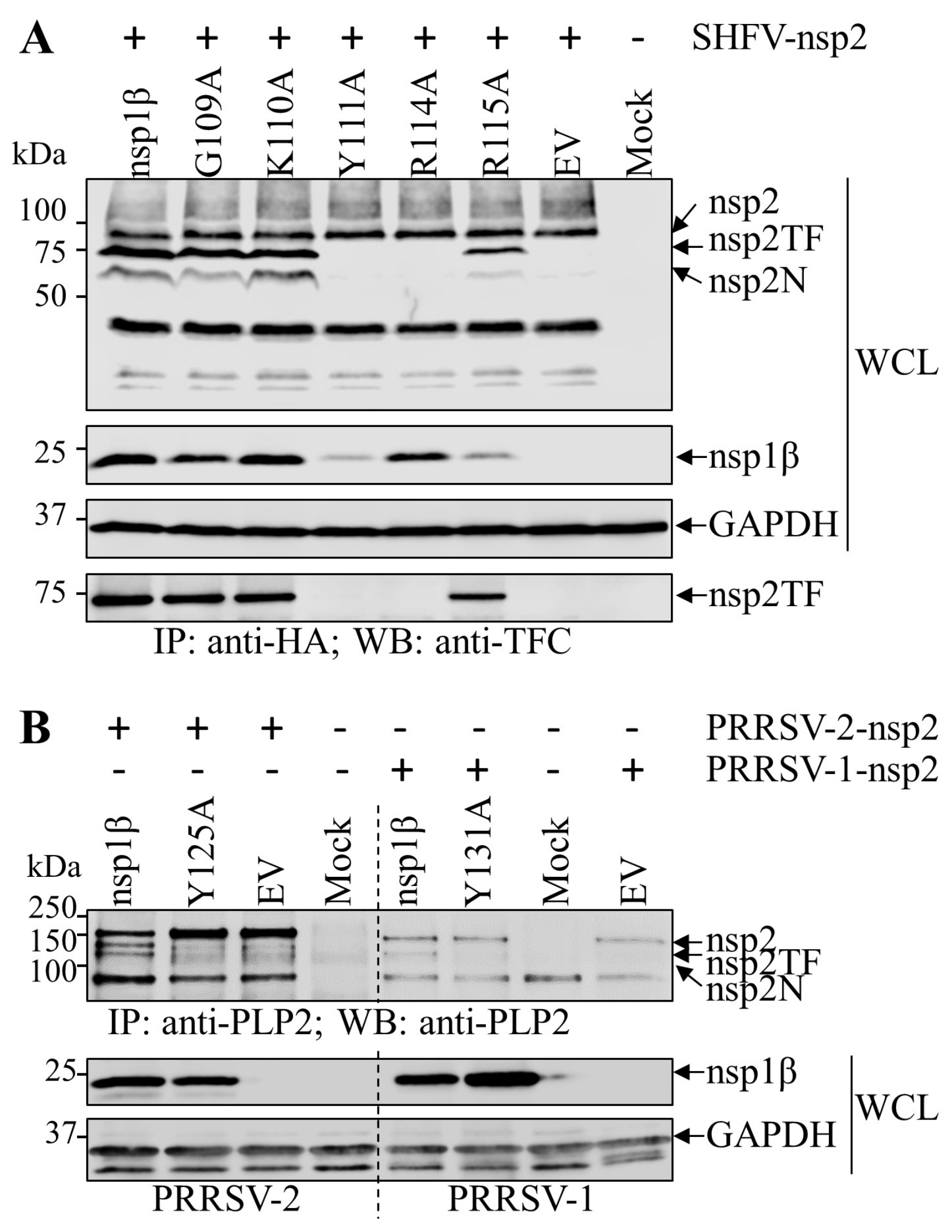
Figure 9. *In vitro* growth characterization of SHFV wild type and mutant viruses. (A) Multiple-step virus growth curve of wild type (WT) and mutant viruses. Each data point shown represents the mean value from duplicated treatments. Error bars show standard errors of the mean (SEM). **(B)** Plaque morphology of wild type (WT) SHFV and mutants thereof. Confluent cell monolayers were infected with 10-fold serial dilutions of the virus suspension. After 2 h incubation, an agar overlay was added on top of the infected cells. Plaques were observed after 3 d of incubation at 37 °C. Cells were stained with 0.1% crystal violet.

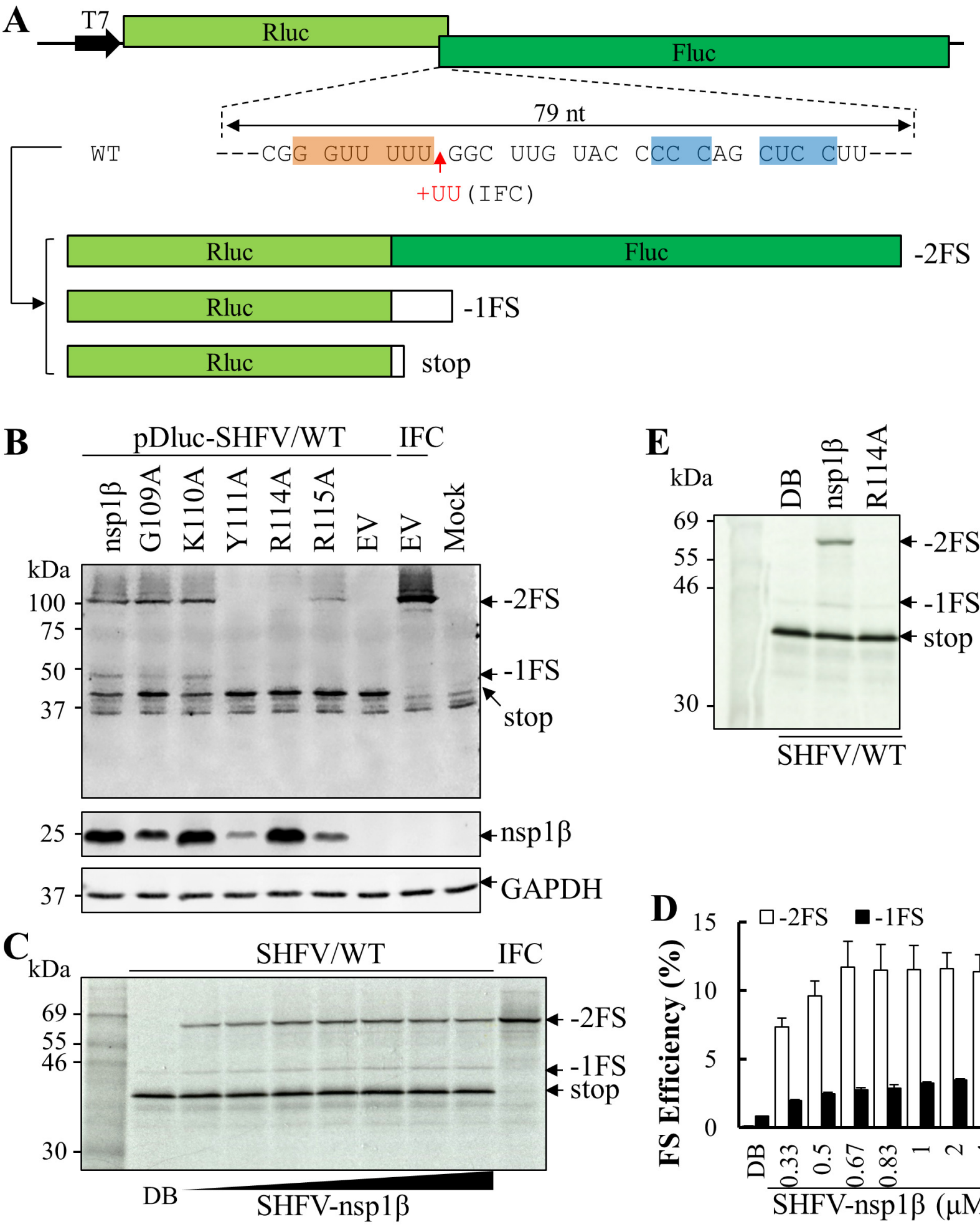
Figure 10. Heterotypic arterivirus nsp1βs stimulate -2/-1 PRF at signals of divergent arteriviruses. (A) Left panel: *In vitro* translation of mRNA from plasmid pDluc-SHFV/WT in the presence of SHFV-nsp1β (1 μM), different concentrations of PRRSV-nsp1β (from 0.5 μM to 1.5 μM), or dilution buffer (DB). Right panel: *In vitro* translation of mRNA from plasmid

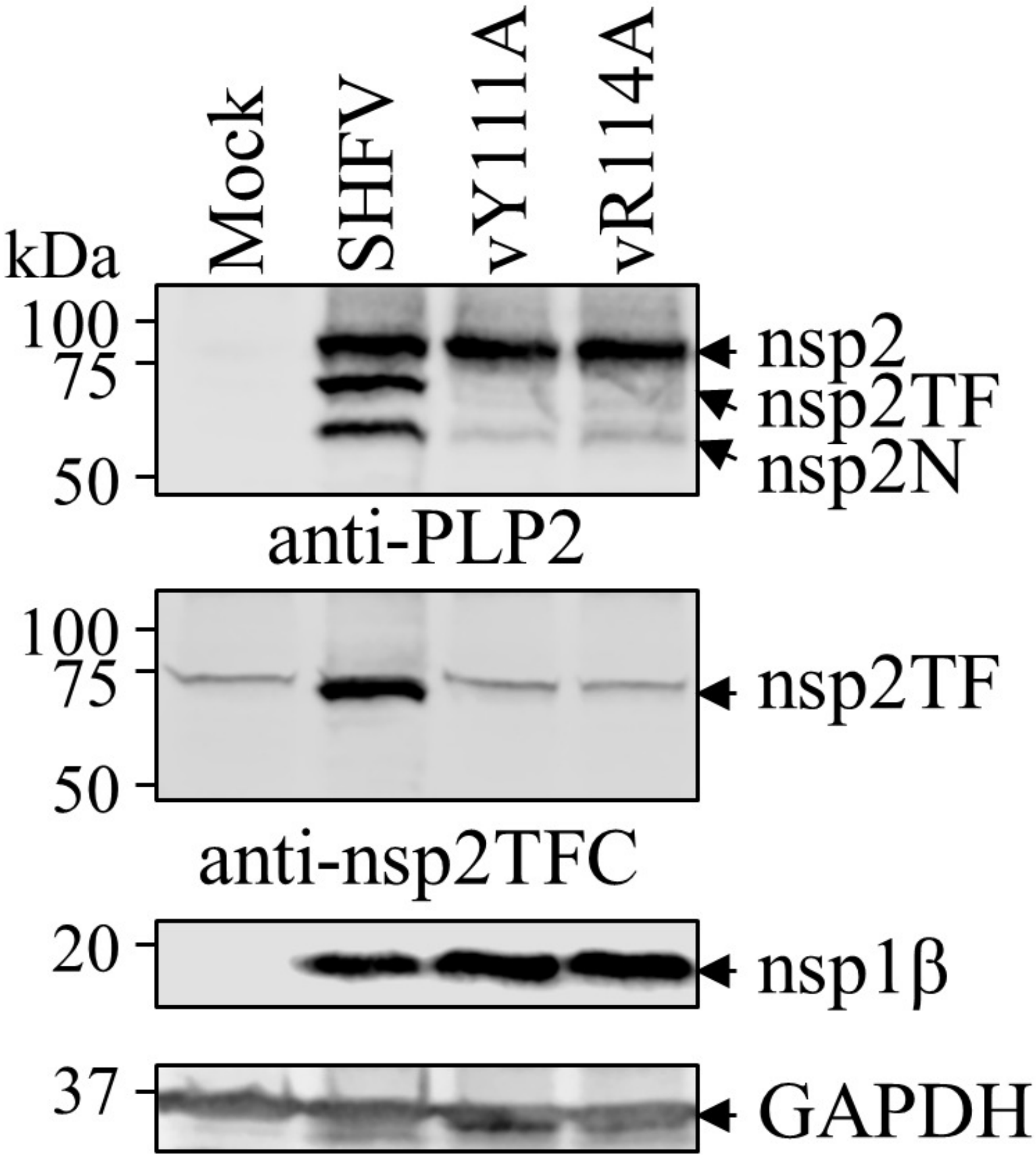
pDluc-PRRSV-1/WT in the presence of PRRSV-nsp1β (1 μM), different concentrations of SHFV-nsp1β (from 0.5 μM to 1.5 μM), or dilution buffer (DB). Products generated without PRF or from −1 or −2 PRF are indicated as stop, -1FS and -2FS, respectively. **(B)** Quantification of SHFV -2 PRF and -1 PRF efficiencies when stimulated by SHFV nsp1β or PRRSV nsp1β. FS efficiencies were calculated based on the protein bands quantified by ImageQuant™ TL software (GE Healthcare). Each data point shown represents the mean value from two independent experiments, and error bars show standard errors of the mean (SEM). (**C**) Analysis of -2/-1 frameshifting at the SHFV PRF signal when stimulated by heterotypic arterivirus nsp1βs. HEK-293T cells were co-transfected with the plasmid pDluc-SHFV/WT and a plasmid expressing heterotypic nsp1β. Empty vector (EV) was used as a control. Non-frameshift, -1 PRF and -2 PRF products (indicated as stop, -1FS and -2FS, respectively) were detected by western blot using anti-*Renilla* luciferase mAb. FLAG-tagged nsp1β was detected with mAb M2, and EAV nsp1 was probed with mAb 12A4. GAPDH was detected as a loading control.

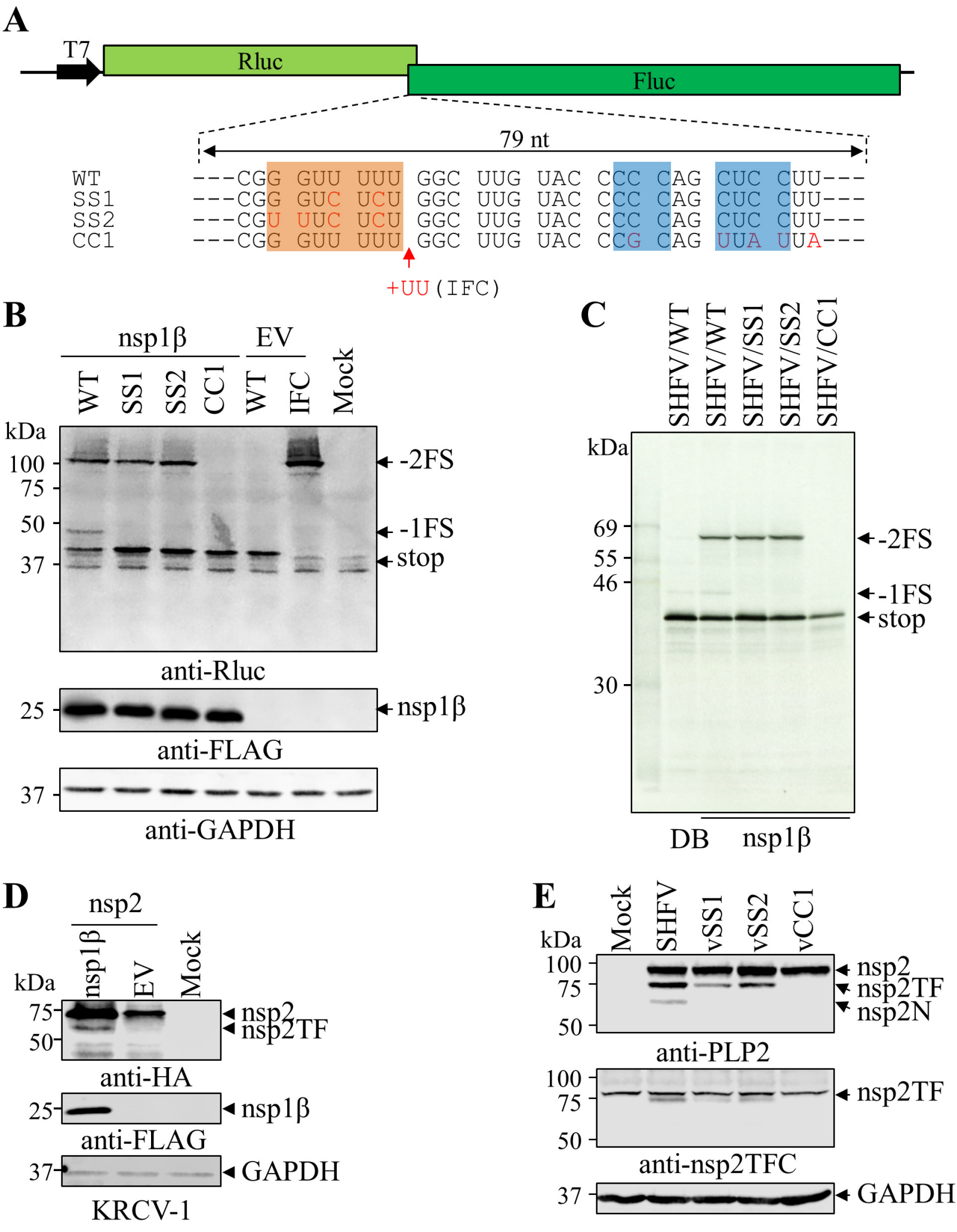

A
PRRSV-2 3691-CAG GUU UUU GAC CUC GUC --- --- --- UCC CAU CUC CCU
PRRSV-1 3289-UGG GUU UUU GAA GUU UAC --- --- --- UCC CAU CUC CCA
RatAV 4075-AAG GUU UUU GAA AUC ACC --- --- --- UCC CAU CUC CCU
RatAV_Ningxia2015 3558-CAG GUU UUU GCC GUC GCA --- --- --- UCC CAC CUC CCU
LDV-P 2836-CAG GUU UUU CUC UUG UCC --- --- --- UCC CAU CUC CUC
APRAV 3152-GCG GUU UUC CGC GUC GUC --- --- --- CCC CGC CUC CUU
SHFV 2659-CGG GUU UUU GGC UUG UAC --- --- --- CCC CAG CUC CUU
KRTGV 2521-GGU GUU UUU GGA UUG CUA --- --- --- CCC CAC AUC CUU
PBJV 2518-CGU GUU UUU CAG UAC AAA --- --- --- CCC CAC CUC CUU
DeMAV 2584-CGU GUU UUU CAC CUU UAC --- --- --- CCC CAG CUC CUC
KRCV-1 2557-CAG GUC UCU CAC CUC CGG --- --- --- CCC CAC CUC CUC
KRCV-2 2518-CAG GUC UCU CGC UUC CAA --- --- --- CCC CAU CUC CUG
MYBV-1 2524-ACG GUC UCU CGA AUC CAG --- --- --- CCC CAG CUC CUG
KKCBV 2521-CAG GUC UCU CGU GUC AUG --- --- --- CCC CAU CUC CUU
FSVV 2629-AAU UUC UCU AGG UUC CUG --- --- --- CCC UAU CUC CUU
SHEV 2680-GAU UUC UCU CGG UGC UUG --- --- --- CCC UAU CUC CUU
ZMbV-1 2879-GAU UUC UCU CGG UUU CUG --- --- --- CCC UAU CUC CUU
WPDV 2454-GCA GUU UUU GAA GCA ACC GUG GGC UGG UCC AAC UCC CUU
Slippery sequence
C-rich motif
B
XH-GD 103 QYKEIRHANQ.....FGYQT.....KHGVPGKYLQRRLQVNGLRA.VTDT
PRRSV-2 103 QNKEIRHANQ.....FGYQT.....KHGVSGKYLQRRLQVNGLRA.VTDL
PRRSV-1 109 CYHEEHLANA.....LGYQT.....KWGVHGKYLQRRLQVRGMRA.VVDP
RatAV 100 FEREVRLAVQ.....FGYQT.....KHGVPGKYIQRRLQINGLRA.VVDP
RatAV_Ningxia2015 100 QAAEIRTAVQ.....FGYQT.....KHGVPGKYIQRRLQVNGLRA.VIDR
LDV-P 102 ACEEADLADR.....MGYRT.....PAGVAGPYLARRLQYRGLRA.VVKP
APRAV 109 FSRLVTYVEA.....HGEQT.....RWGVKSALLNRLLFNHGLRI.WRHA
SHFV 93 ...ELTTAST.....FGYQL.....NCGVQGKYIARRLQTNGLKL.VQNQ
KRTGV 92 ...EKCIAQT.....FGYQL.....TLGVQGKYLSRRAQINGIKF.VHDS
PBJV 93 ...ELSIASC.....FGYQL.....PIGVQGKYIARRLQINGLKL.VLAD
DeMAV 92 ...ELKVACT.....FGYQL.....GIGVQGKYISRRLQIAGCKL.VYDS
KRCV-1 91 ...ELSIART.....FGYQL.....PWGAQGGYIMRRVAINGLKL.VKNP
KRCV-2 92 ...ELLVARQ.....MGYQL.....PYGVQGRYLARRLQCSGLKL.VRKS
MYBV-1 92 ...EMRVALT.....FGYQIG.....TLGVQGKYISRRAQINGIKF.VHDT
KKCBV 92 ...EFRIAST.....FGYQLA.....NIGVPGKYIARRLQINGLKL.VHDA
FSVV 91 ...EVSIAAC.....FGYLM.....PWGAQGAYISRRLQLNGSKL.VRNP
SHEV 91 ...EVKVAGC.....FGYLM.....PWGAQGAYISRRLQIDGSKL.VRNP
ZMbV-1 91 ...EVKVAGC.....FGYLM.....PWGAQGAYISRRLQVDGSKL.VRDP
WPDV 34 .....IKYHFGKLQTGYLPLMPVVPGKKNNLLHCPVGLVGSAADALKKRYGISFPVANK
EAV 36 ...EAGLRLY..YNHYREQR.....TGWLSKTGLRL.WLGD
consensus>70 ...e...a.....fgyq.....gv.g.y..rrlq..G....v.d.
###
##
C
PRRSV-2
PRRSV-1
RatAV
RatAV_Ningxia2015
LDV-P
APRAV
SHFV
KRTGV
PBJV
DeMAV
KRCV-1
KRCV-2
MYBV-1
KKCBV
FSVV
SHEV
ZMbV-1
ZMbV-1_RaptorX

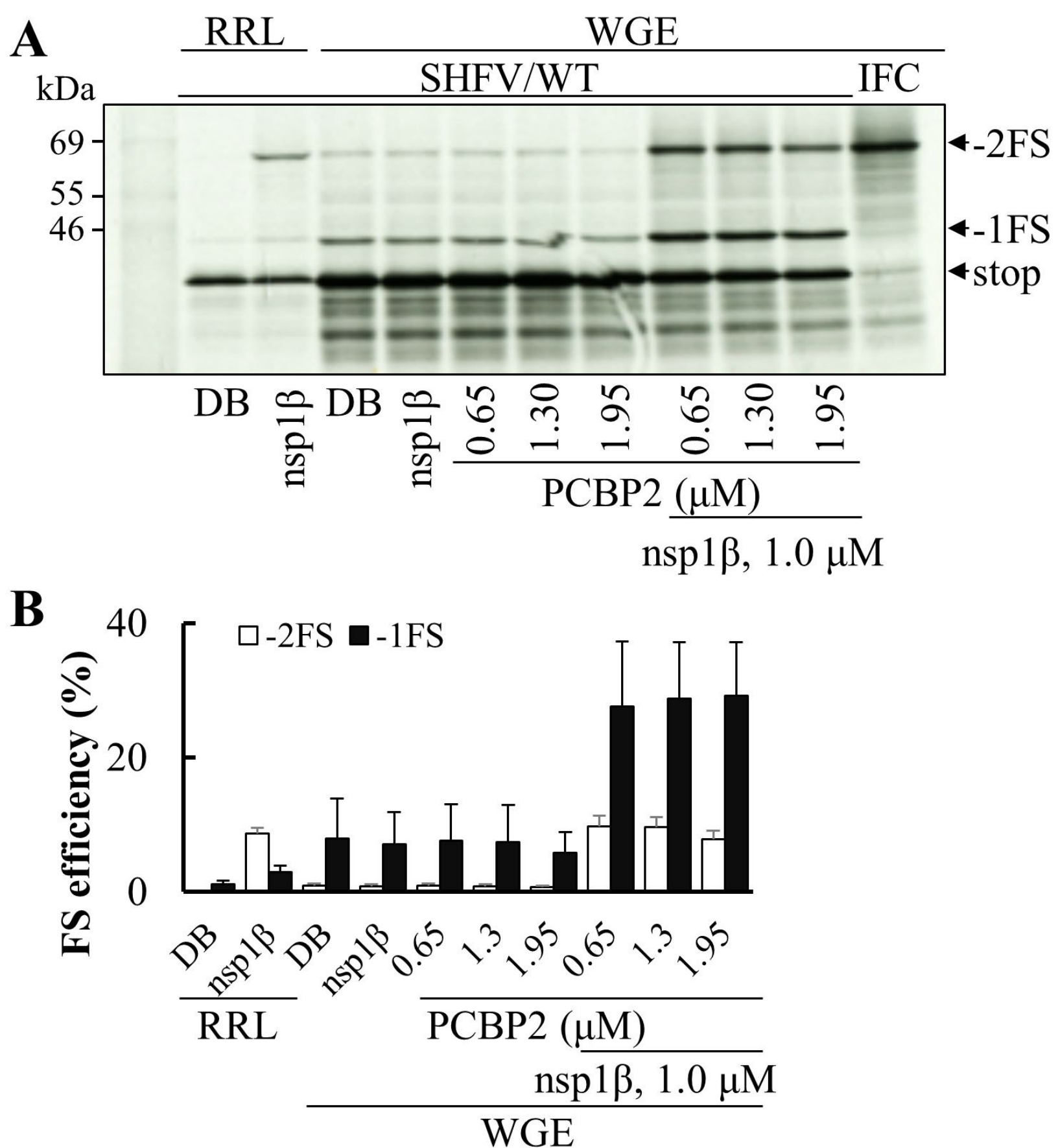

A
PLP2
nsp2
nsp3
SD01-08_pp1a GRGNFAKLNQ TP------LR DSASTKTTGG ASYTLAVAQV SVWTL 1461
SD95-21_pp1a GVGDFAQLNG LK------IR QIS--KPSGG GPHLMAALHV ACSMA 1594
SHFV_pp1a GDGDFFKVMG VPRPSPFTVM RLRACRVVGG GRIFRTAL-A AAWVL 1250
Consensus GXGDFAKLNG XP------XR XXXXXKXXGG GXXXXAALXV AXWXL
4.3bits
Sequence logo
0.0bits
B
mAb 134-260
mAb 133-243
-1PRF
nsp2N
3093
pAb anti-nsp2TFC
-2PRF
nsp2TF
3536
nsp2
1649
2860
3901
C
Mock
SHFV
kDa
100
75
50
nsp2
nsp2TF
nsp2N
*
WB: anti-PLP2
Mock
SHFV
WB: anti-nsp2TFC
D
Mock
SHFV
kDa
100
75
50
nsp2
nsp2TF
nsp2N
*
WB: anti-PLP2
Mock
SHFV
WB: anti-nsp2TFC
IP: anti-PLP2

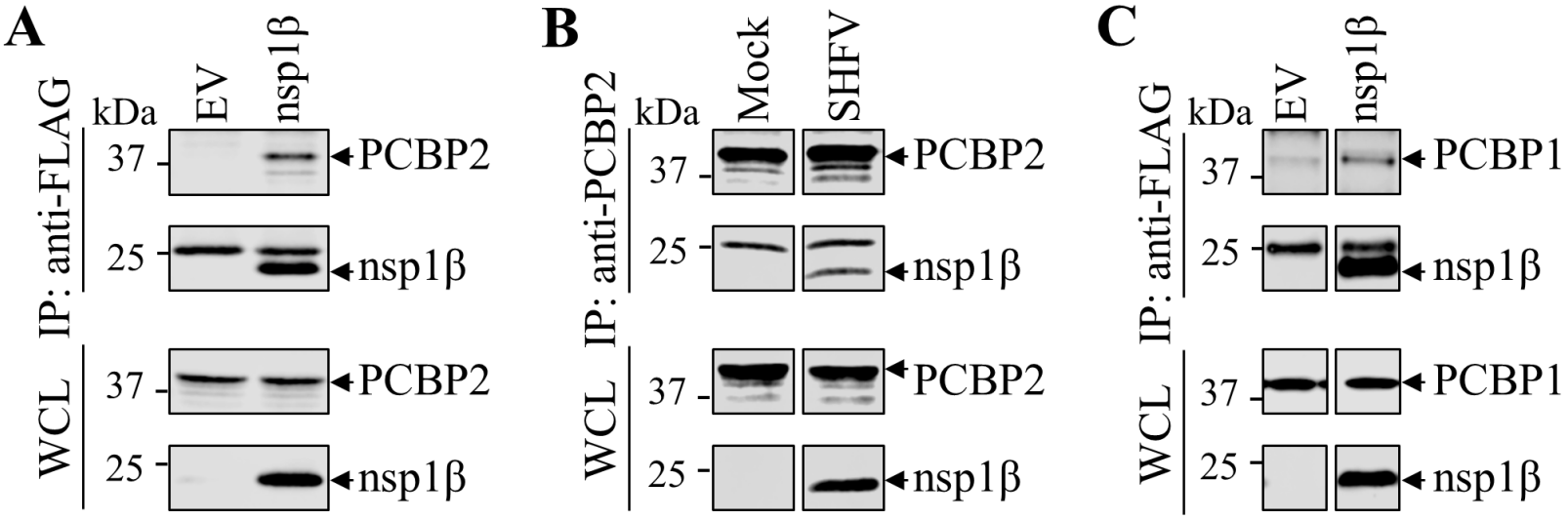

A
+ + + + + + + - SHFV-nsp2
nsp1β G109A K110A Y111A R114A R115A EV Mock
kDa
100
75
50
nsp2
nsp2TF
nsp2N
WCL
25
nsp1β
37
GAPDH
75
nsp2TF
IP: anti-HA; WB: anti-TFC
B
+ + + - - - - - PRRSV-2-nsp2
- - - - + + - + PRRSV-1-nsp2
nsp1β Y125A EV Mock nsp1β Y131A Mock EV
kDa
250
150
100
nsp2
nsp2TF
nsp2N
IP: anti-PLP2; WB: anti-PLP2
25
nsp1β
WCL
37
GAPDH
PRRSV-2
PRRSV-1

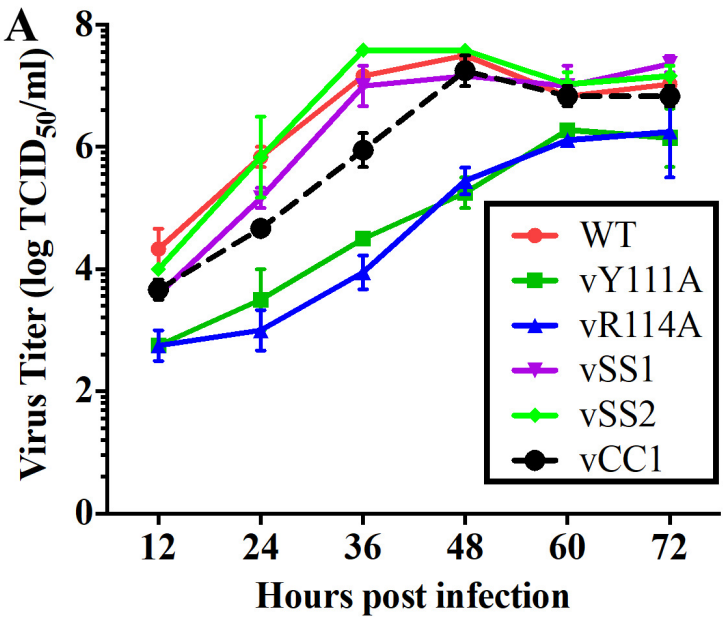

A
T7
Rluc
Fluc
79 nt
WT
---CGG GUU UUU GGC UUG UAC CCC CAG CUC CUU---
+UU (IFC)
Rluc
Fluc
-2FS
Rluc
-1FS
Rluc
stop
B
pDluc-SHFV/WT
IFC
nsp1β
G109A
K110A
Y111A
R114A
R115A
EV
EV
Mock
kDa
100
75
50
37
-2FS
-1FS
stop
25
nsp1β
37
GAPDH
C
kDa
SHFV/WT
IFC
69
55
46
30
-2FS
-1FS
stop
DB
SHFV-nsp1β
D
FS Efficiency (%)
15
10
5
0
-2FS
-1FS
DB
0.33
0.5
0.67
0.83
1
2
4
SHFV-nsp1β (μM)
E
kDa
DB
nsp1β
R114A
69
55
46
30
-2FS
-1FS
stop
SHFV/WT

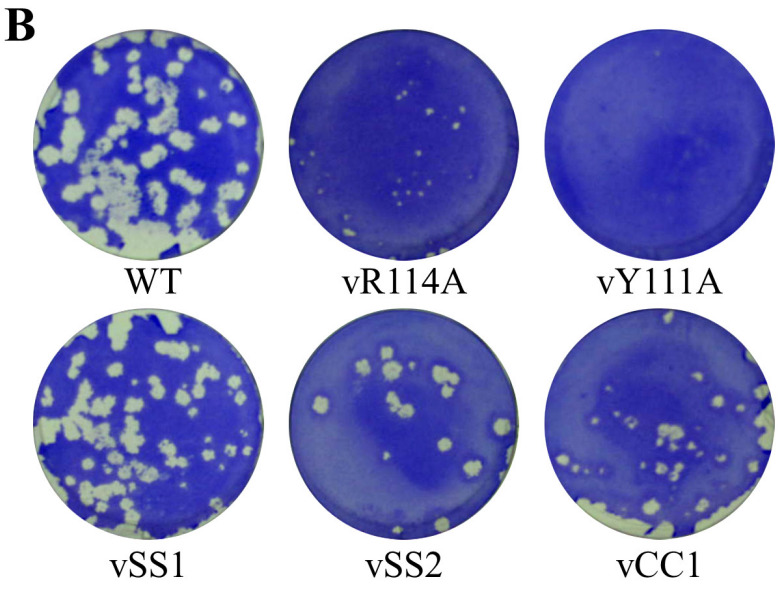

Mock
SHFV
vY111A
vR114A
kDa
100
75
50
nsp2
nsp2TF
nsp2N
anti-PLP2
100
75
50
nsp2TF
anti-nsp2TFC
20
nsp1β
37
GAPDH

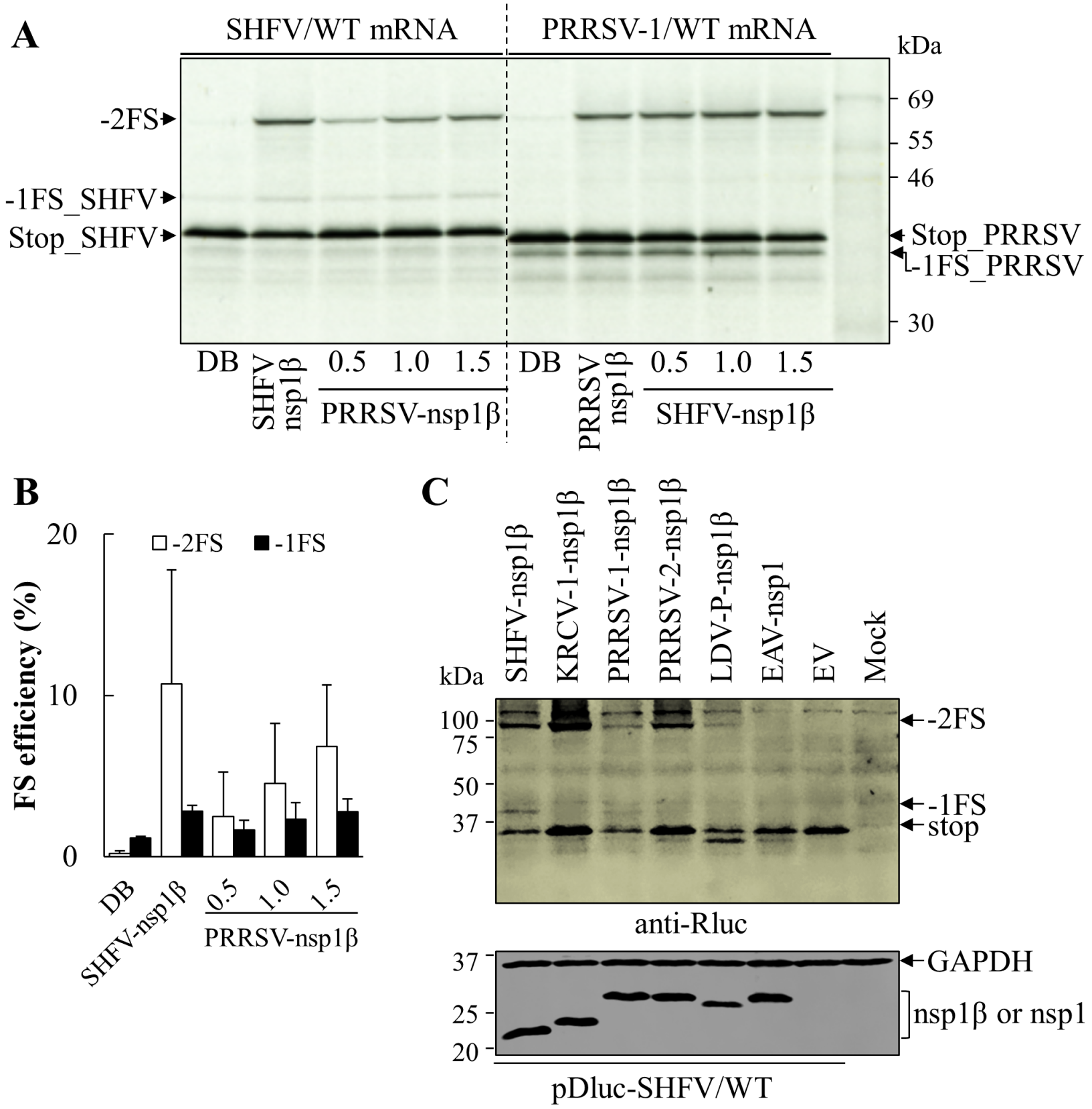

A
T7
Rluc
Fluc
79 nt
WT ---CGG GUU UUU GGC UUG UAC CCC CAG CUC CUU---
SS1 ---CGG GUC UCU GGC UUG UAC CCC CAG CUC CUU---
SS2 ---CGU UUC UCU GGC UUG UAC CCC CAG CUC CUU---
CC1 ---CGG GUU UUU GGC UUG UAC CCG CAG UUA UUA---
+UU (IFC)
B
nsp1β
EV
WT
SS1
SS2
CC1
WT
IFC
Mock
kDa
100
75
50
37
-2FS
-1FS
stop
anti-Rluc
25
nsp1β
anti-FLAG
37
anti-GAPDH
C
SHFV/WT
SHFV/WT
SHFV/SS1
SHFV/SS2
SHFV/CC1
kDa
69
55
46
30
-2FS
-1FS
stop
DB
nsp1β
D
nsp2
nsp1β
EV
Mock
kDa
75
50
nsp2
nsp2TF
anti-HA
25
nsp1β
anti-FLAG
37
GAPDH
KRCV-1
E
Mock
SHFV
vSS1
vSS2
vCC1
kDa
100
75
50
nsp2
nsp2TF
nsp2N
anti-PLP2
100
75
50
nsp2TF
anti-nsp2TFC
37
GAPDH


A
RRL
WGE
SHFV/WT
IFC
kDa
69
55
46
-2FS
-1FS
stop
DB
nsp1β
DB
nsp1β
0.65
1.30
1.95
0.65
1.30
1.95
PCBP2 (μM)
nsp1β, 1.0 μM
B
40
20
0
FS efficiency (%)
-2FS
-1FS
DB
nsp1β
DB
nsp1β
0.65
1.3
1.95
0.65
1.3
1.95
RRL
PCBP2 (μM)
nsp1β, 1.0 μM
WGE


A
kDa
EV
nsp1β
IP: anti-FLAG
37
PCBP2
25
nsp1β
WCL
37
PCBP2
25
nsp1β
B
kDa
Mock
SHFV
IP: anti-PCBP2
37
PCBP2
25
nsp1β
WCL
37
PCBP2
25
nsp1β
C
kDa
EV
nsp1β
IP: anti-FLAG
37
PCBP1
25
nsp1β
WCL
37
PCBP1
25
nsp1β


A
Virus Titer (log TCID50/ml)
0
2
4
6
8
12
24
36
48
60
72
Hours post infection
WT
vY111A
vR114A
vSS1
vSS2
vCC1


B
WT
vR114A
vY111A
vSS1
vSS2
vCC1


A
SHFV/WT mRNA
PRRSV-1/WT mRNA
kDa
-2FS
-1FS_SHFV
Stop_SHFV
69
55
46
30
Stop_PRRSV
-1FS_PRRSV
DB
SHFV nsp1β
0.5
1.0
1.5
PRRSV-nsp1β
DB
PRRSV nsp1β
0.5
1.0
1.5
SHFV-nsp1β
B
FS efficiency (%)
20
10
0
-2FS
-1FS
DB
SHFV-nsp1β
0.5
1.0
1.5
PRRSV-nsp1β
C
kDa
SHFV-nsp1β
KRCV-1-nsp1β
PRRSV-1-nsp1β
PRRSV-2-nsp1β
LDV-P-nsp1β
EAV-nsp1
EV
Mock
100
75
50
37
-2FS
-1FS
stop
anti-Rluc
37
25
20
GAPDH
nsp1β or nsp1
pDluc-SHFV/WT