THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 25, pp. 22521–22534, June 24, 2011
Printed in the U.S.A.

Identification of *Cis*-Acting Elements in the 3′-Untranslated Region of the Dengue Virus Type 2 RNA That Modulate Translation and Replication*[S]

Received for publication, February 25, 2011, and in revised form, April 21, 2011 Published, JBC Papers in Press, April 22, 2011, DOI 10.1074/jbc.M111.234302

Mark Manzano[‡], Erin D. Reichert[‡], Stephanie Polo[§], Barry Falgout[§], Wojciech Kasprzak[¶], Bruce A. Shapiro[‖1], and Radhakrishnan Padmanabhan[‡2]

From the [‡]Department of Microbiology and Immunology, Georgetown University School of Medicine, Washington, D. C. 20057, the [§]Center for Biologics Evaluation and Review, Food and Drug Administration, Bethesda, Maryland 20892, the [¶]Basic Science Program, SAIC-Frederick, Inc., and the [‖]Center for Cancer Research Nanobiology Program, NCI-Frederick, National Institutes of Health, Frederick, Maryland 21702

Using the massively parallel genetic algorithm for RNA folding, we show that the core region of the 3′-untranslated region of the dengue virus (DENV) RNA can form two dumbbell structures (5′- and 3′-DBs) of unequal frequencies of occurrence. These structures have the propensity to form two potential pseudoknots between identical five-nucleotide terminal loops 1 and 2 (TL1 and TL2) and their complementary pseudoknot motifs, PK2 and PK1. Mutagenesis using a DENV2 replicon RNA encoding the *Renilla* luciferase reporter indicated that all four motifs and the conserved sequence 2 (CS2) element within the 3′-DB are important for replication. However, for translation, mutation of TL1 alone does not have any effect; TL2 mutation has only a modest effect in translation, but translation is reduced by ~60% in the TL1/TL2 double mutant, indicating that TL1 exhibits a cooperative synergy with TL2 in translation. Despite the variable contributions of individual TL and PK motifs in translation, WT levels are achieved when the complementarity between TL1/PK2 and TL2/PK1 is maintained even under conditions of inhibition of the translation initiation factor 4E function mediated by LY294002 via a noncanonical pathway. Taken together, our results indicate that the *cis*-acting RNA elements in the core region of DENV2 RNA that include two DB structures are required not only for RNA replication but also for optimal translation.

The dengue virus (DENV)[3] is a mosquito-borne flavivirus (MBFV) in the *Flaviviridae* family that consists of over 70 members, many of which are significant human pathogens (1). The MBFV members are classified into three subgroups: DENV, yellow fever virus, and Japanese encephalitis virus (JEV) (2, 3). The four serotypes of DENV (DENV1 to -4) cause an estimated 50 million cases of infections, with ~10% of those leading to severe forms of the disease, dengue hemorrhagic fever and dengue shock syndrome (4–6).

The viral genome is a single-stranded RNA of positive (+) polarity containing ~11 kilobases (10,723 nt for DENV2 New Guinea C strain, GenBank™ accession number M29095 (7)). The 3′-end is non-polyadenylated, and the 5′-end has a type I cap structure (for a review, see Ref. 8). Flanking the single long open reading frame are the 5′- and 3′-UTRs, which contain conserved *cis*-acting RNA secondary structure elements required for translation and replication (9–18). The viral RNA is translated to form a polyprotein precursor, which is processed by host and viral proteases in the endoplasmic reticulum membrane. Protein processing gives rise to three structural (C, prM, and E) and seven nonstructural (NS) proteins: NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5 in that order (for reviews, see Refs. 8, 19, and 20) and references therein). According to the current model for replication, assembly of the viral replicase complex occurs in a cytoplasmic membrane organelle, followed by negative (−)-strand RNA synthesis starting at the 3′-end of the viral genome, resulting in a double-stranded replicative form. NS3 and NS5 are multifunctional proteins, known to physically and functionally interact as a complex and thought to be components of the replicase complex in flavivirus-infected cells (21–23). The (−)-strand is used as a template for synthesis of progeny (+)-RNA (24) (reviewed in Refs. 1 and 25).

The 5′-UTRs of flavivirus RNAs are ~100 nt in length and fold to form stable secondary structures (16, 26, 27). The 5′-terminal region within ~160 nt forms two stem-loop (SL) structures, SLA and SLB (Fig. 1*A*), and is important for binding the NS5 protein, the viral RNA-dependent RNA polymerase (RdRP), *in vitro* (16). Mutations that disrupt this interaction

* This work was supported, in whole or in part, by National Institutes of Health (NIH) Grants AI 57705 and AI 70791 (to R. P.); by NCI, NIH, Contracts NO1-CO-12400 and HHSN261200800001E (to W. K.); and by the Intramural Research Program of the NCI, NIH, Center for Cancer Research. This work constitutes a partial fulfillment of the requirements for the degree of Doctor of Philosophy at Georgetown University School of Medicine for E. R. and M. M.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables 1 and 2 and Figs. 1–3.

[1] To whom correspondence may be addressed. E-mail: shapirbr@mail.nih.gov.

[2] To whom correspondence may be addressed. E-mail: rp55@georgetown.edu.

[3] The abbreviations used are: DENV, dengue virus; EMCV, encephalomyocarditis virus; IRES, internal ribosome entry site; MBFV, mosquito-borne flavivirus; JEV, Japanese encephalitis virus; nt, nucleotide(s); RdRP, RNA-dependent RNA polymerase; UAR, upstream of AUG region; CS, conserved sequence; RCS, repeat conserved sequence; DB, dumbbell structure; Rluc, *Renilla* luciferase; MPGAfold, massively parallel genetic algorithm; hpt, h post-transfection; qPCR, quantitative PCR; WNV, West Nile virus; cHP, capsid-coding region hairpin; TL, terminal loop; EMEM, Earle's minimum essential medium; PK, pseudoknot.

severely affected viral replication (16, 28). Downstream of the 5′-UTR in the capsid-coding region is the capsid-coding region hairpin RNA (cHP) (Fig. 1*A*), which was shown to play a role for the efficient selection of the start codon for initiation of translation of the single polyprotein (29, 30). Moreover, it is also required for DENV and WNV replication (29, 30). Recently, another base-paired interaction between the motifs known as the "5′-3′ downstream of AUG region," or DAR, was also identified to be required for viral RNA replication (31, 32).

The 3′-UTR of flaviviruses exhibits a great deal of sequence divergence and size heterogeneity but also contains conserved sequences that are required for replication (9–14, 16, 17, 33). It consists of the variable region, core region (CR), and terminal stem-loop (3′-SL) region (Fig. 1*A*).

The 3′-SL region is formed within ~100 nt from the 3′-end (34, 35), which has been shown to play a significant role in replication (33, 36–39). The CR includes the 3′-cyclization sequence (3′-CS1) containing a short motif that is complementary to the 5′ cyclization sequence (5′-CS) located in the capsid coding region (Fig. 1*A*). A long range RNA-RNA interaction between these complementary motifs resulting in circularization of the genome was postulated (40) and demonstrated by atomic force microscopy (13). The physical and functional RNA-RNA interaction between 5′- and 3′-CS1 was shown to be essential for RNA synthesis *in vitro* (10, 41) and in replication of subgenomic replicons or infectious clones in cultured mammalian cells (11, 12, 14, 42, 43). In addition to the 5′- and 3′-CS1, two regions of complementarity in the 5′- and 3′-terminal sequences, termed the upstream of AUG region (UAR), were also identified as a requirement for cyclization and RNA replication (13, 17, 44). The 5′-UAR is located upstream of the 5′-CS, and the 3′-UAR is downstream of 3′-CS1 (Fig. 1*A*).

The CR exhibits relatively high sequence conservation among members of MBFV and is believed to fold into well defined secondary structures independent of other regions of the 3′-UTR (45, 46). Within this region, upstream from the 3′-CS1, the CS2 is present in all three subgroups in the MBFV, whereas the repeat conserved sequence 2 (RCS2) is present only in JEV and DENV subgroups (40) (Fig. 1*B*).

The regions containing CS2 and RCS2 are part of two almost identical dumbbell-like (DB) secondary structures that have been postulated to form in DENV (Fig. 1*B*). It is thought that this duplication of dumbbell-like structures arose because of the repetition of CS2 and RCS2 sequences in the 3′- and 5′-DB, respectively. The leftmost loops of the 5′- and 3′-DBs contain a set of identical sequences (underlined in 5′-GAAGCUGUA-3′) (45). Sequence analysis of different serotypes of DENV showed that the internal five nucleotides in these loops (referred to as TL1 and TL2) and the two complementary pentanucleotide sequences downstream of each DB, PK2 (5′-GCAGC-3′) and PK1 (5′-ACAGC-3′), are highly conserved, suggesting possible base pair interactions between TL1/PK2 and TL2/PK1 to form two pseudoknots (5′-ψ and 3′-ψ) (45).

The role of the CR region in flavivirus replication or translation is not understood at present. In this study, we sought to examine the functional roles of the TLs and PKs by using a *Renilla* luciferase (Rluc) reporter replicon in BHK21 cells. We show here that all four motifs played a crucial role in replication in a sequence-dependent manner. In contrast, there seems to be a differential role for TL1/PK2 and TL2/PK1 in translation. The lack of functional similarity between TL1 and TL2 despite their identical sequences within two DBs could be explained by results of our analysis of RNA secondary structures, their stabilities, and frequency of occurrence computed by the massively parallel genetic algorithm (MPGAfold) (47–51). The results of our study, taken together, provide a new insight into the role of the *cis*-acting elements in the CR region in translation and replication.

EXPERIMENTAL PROCEDURES

Cells and Primers

BHK-21 cells (ATCC, Manassas, VA) were maintained at 37 °C with 5% CO_2 in complete Earle's minimum essential medium (EMEM) (CellGro, Manassas, VA), containing high glucose (1 g/liter) and L-glutamine, and supplemented with 10% fetal bovine serum and antibiotics (Invitrogen). All primers were obtained from Integrated DNA Technologies (Coralville, IA) and resuspended in H_2O. Primer sequences are listed in supplemental Table 1.

Replicon Construction

DENV2 Rluc Replicon—The Rluc gene was amplified from phRL-SV40 (Promega, Madison, WI) placing SalI sites on both the 5′- and 3′-ends using RlucF and RlucR primers (supplemental Table 1). The resulting fragment was cloned into pGEMTEasy (Promega). The resulting insert vector was used as a template using RlucF and RlucIresR primers (supplemental Table 1) to create a fragment containing Rluc with a SalI site at the 5′-end and the 5′-terminus of EMCV IRES at the 3′-end. Next, using the GFP-expressing replicon, pRS424GFPIresDEN2 (52), as a template and primers IresF and NS1BstEIIR, a fragment containing the entire sequence of EMCV IRES followed by the 5′-end of NS1 up to the BstEII site was created. Overlap PCR was performed using the products from the first two PCRs and primers RlucF and NS1BstEIIR. The entire fragment was cloned into pGEMTEasy. The insert was subcloned into the replicon plasmid pRS424GFPIresDEN2 using SalI and BstEII (New England Biolabs, Ipswich, MA) to create the DENV2 Rluc replicon.

GND Replicon—A replication-defective replicon clone was created by engineering a point mutation in the conserved GDD motif of the RdRP domain in *NS5* (nucleotides 9553–9561). First, an AatII/ApaI restriction fragment from the C-terminal region of *NS5* to the ApaI site at the 3′-UTR was isolated from pRS424GFPIresDEN2 and subcloned into pBR322 vector. The GDD → GND mutation was introduced by site-directed mutagenesis using primers GNDF and GNDR (supplemental Table 1). The fragment was released with PmlI and ApaI and subcloned into XhoI-SacI sites of pBR322, which contained the replicon sequence from the XhoI site in NS3 to the SacI site at the end of the 3′-UTR. The sequence between XhoI and SacI was cloned into the WT DENV2 Rluc replicon to yield the GND mutant.

Replicon Mutants

The 5′- or 3′-UTR fragments containing the desired mutations were first amplified by overlap extension site-directed

mutagenesis PCR using primers containing desired mutations (supplemental Table 1) and were purified from agarose gel using the Zymoclean gel DNA recovery kit (Zymo Research, Irvine, CA). Three purified PCRs were combined in a single tube and vacuum-dried. Twenty micrograms of WT replicon were digested overnight with ~40 units of SalI for 5′-UTR mutations or BbvCI for 3′-UTR mutations. These were purified by phenol/chloroform/isoamyl alcohol (25:24:1), ethanol-precipitated, and resuspended in 5 μl of water. The linearized WT or the mutant replicon plasmid was mixed with the vacuum-dried mutant PCR fragments, and the mixture was introduced into *Saccharomyces cerevisiae* YPH857 (50 μl) made chemically competent using the S.c. EasyComp Transformation Kit (Invitrogen) (53). The 5′- or 3′-UTRs of mutant replicons were amplified from the colonies, and their sequences were verified (MC Lab, South San Francisco, CA). Plasmid DNAs harboring the mutations were recovered using the Zymoprep II yeast plasmid minipreparation kit (Zymo Research), and 2 μl were used to transform 20 μl of MAX Efficiency Stbl2-competent cells (Invitrogen). Replicon plasmids were propagated in *Escherichia coli* from single colonies, and their sequences within 5′- and 3′-UTRs were again verified. Details of sequences for mutant replicons used in this study are shown in supplemental Table 1 and Figs. 3*A*, 4*A*, and 5*A*.

Templates for in Vitro RdRP Assay

DNA templates for *in vitro* transcription harboring the desired mutations in a 719-nt template containing the first 230 nt from the 5′-end and the last 489 nt of the genome (54) were prepared by an overlap PCR amplification strategy as follows. In the first PCR, the region spanning the 5′-UTR and the *NS5*-3′-UTR junction was amplified using the pSY2 plasmid as the template and the primers pSY2 EcoRI and NS5-3′-UTRr. In the second PCR, the full-length 3′-UTR was amplified from the replicon mutants using the start of 3′-UTR and 3′-end DEN2 primers (see supplemental Table 1 for primer sequences). The two fragments were joined together by an overlap PCR and purified from the agarose gel.

In Vitro Transcription of RNA Templates

Approximately 50 μg of WT and mutant replicons were linearized with EcoICRI and purified twice with phenol/chloroform/isoamyl alcohol and once with chloroform to remove excess phenol. DNA was ethanol-precipitated and resuspended in 20 μl of RNase-free water. Three microliters were used to make RNA with the Ampliscribe SP6 high yield transcription kit (Epicenter Biotechnologies, Madison, WI). The reaction mixtures (20 μl) contained 3.13 mM GTP, 8 mM $m^7G(5')ppp(5')G$ cap analog (New England Biolabs) and 80 units of RNase inhibitor (New England Biolabs) and incubated for 3–4 h at 37 °C. Following transcription, DNAse I (Epicenter) was directly added to the mixture and incubated for another 1 h at 37 °C. 0.5 μl of each reaction was visualized on a 0.5% agarose gel to verify the integrity of the RNA prior to storage or use. The plasmid encoding the rabbit α-globin 5′- and 3′-UTRs flanking the firefly luciferase gene (pGLGpA) (15) (kindly provided by Dr. Theo Dreher of Oregon State University) was linearized by Acc65I digestion and used for the *in vitro* transcription. The GLGpA RNA was used as an internal control for transfections of the WT and mutant replicons. GLGpA and (+) templates for the *in vitro* RdRP assays were made using the Ampliscribe T7 high yield transcription kit.

RNA Transfection

For each reaction, ~10^6 BHK-21 cells were resuspended in 100 μl of Ingenio solution (Mirus Bio, Madison, WI), and ~3 μg of replicon RNA and 0.1 μg of GLGpA RNA were added. Transfections were performed using a Nucleofector II electroporator (Amaxa Biosystems, Cologne, Germany) using program A031. Following pulsing, cells were carefully transferred in a tube containing 1 ml of prewarmed complete EMEM. Cells were allowed to recover at 37 °C for 5 min and then transferred to another tube with 3 ml of complete EMEM. Cells were plated in 8 wells of a 48-well plate, each containing 400 μl. The rest of the cells were plated in a 6-well plate containing 3 ml of medium. Each transfection experiment was done in quadruplicate and was repeated at least three times using at least two different batches of *in vitro* transcribed RNA.

Cap-independent Translation Assay

In vitro transcribed mutant RNAs were transfected into BHK21 cells as described above. After transfection, cells were allowed to recover in 1 ml of complete EMEM at 37 °C for 5 min. 100 μl was seeded in each of 10 wells of a 48-well plate containing 100 μl of medium. Half of the wells were used as no-treatment controls, whereas the other half contained LY294002 with a final concentration of 40 μM (Cayman Chemical, Ann Arbor, MI). Cells were lysed at 2 h post-transfection (hpt) and assayed as described below.

Luciferase Assays

At the desired time points, cells were lysed with 60 μl of 1× *Renilla* luciferase lysis buffer to determine the Rluc signal using a kit (Promega). Three additional wells were also lysed at 2 hpt with 1× Passive Lysis Buffer (Promega) to measure the GLGpA Fluc activity. To ensure complete lysis, plates were put on an orbital shaker for 30 min. Rluc activity was measured using a Centro LB 960 luminometer (Berthold Technologies) by injecting 100 μl of 1× *Renilla* luciferase substrate and reading for 10 s after a 2-s delay. Fluc readings were similarly measured using 1× luciferase assay reagent (Promega).

Immunofluorescence

Following transfection, cells were plated into wells of a 2-well chamber slides (BD Falcon). At 96 hpt, the culture medium was removed, and the cells were washed with PBS. The cells were fixed in 100% cold methanol and placed at −20 °C for 30 min. After fixing, the cells were blocked in 1× PBS with 1% nonfat dry milk for 1 h with rocking. Cells were washed three times with PBS and then incubated with rabbit anti-NS5 IgG (1:250) in PBS for 2 h with rocking. After washing, cells were incubated with goat anti-rabbit IgG conjugated to FITC (ICN, Solon, OH) (1:250) for 2 h with rocking. The secondary antibody was removed, and cells were washed six times in PBS, mounted using ProLong reagent (Molecular Probes), and visualized using an Olympus Fluoview FV300 laser confocal microscope.

Quantitative PCR

Early Time Points—To determine the RNA stabilities of replicon RNAs, we transfected 6 μg of each RNA by electroporation into 2×10^6 BHK-21 cells as described above, and the transfected cells were plated into 24-well plates (2×10^5 cells/well). At the desired time points (1, 2, 4, and 6 hpt), cells were harvested. Total RNA was extracted using 250 μl of TRIzol® reagent (Invitrogen) following the manufacturer's protocol, resuspended in 40 μl of diethylpyrocarbonate-treated water, and stored immediately at −80 °C until used for qPCR analysis. Each RNA sample was analyzed in duplicate wells of a 96-well plate, and the assay was performed on the Applied Biosystems 7900HT fast real-time PCR system. Primers and probe were targeted to amplify nucleotides 9725–9820 in the *NS5* gene. The qPCR reaction mixture (50 μl) contained 100 ng of extracted RNA, a 0.2 μM concentration of primers qPCR *NS5*F and qPCR *NS5*R (supplemental Table 1), 0.1 μM qPCR *NS5* probe (supplemental Table 1) (from Dr. Robin Levis, Food and Drug Administration), 1× rTth EZ buffer, 300 μM each dNTP, 3 mM $Mn(OAc)_2$, and 5 units of rTth DNA polymerase (Applied Biosystems). The reactions were subjected to 60 °C for 30 min for reverse transcription, followed by 94 °C for 1 min, and PCRs were amplified through 40 cycles of 94 °C for 15 s and 60 °C for 1 min. RNA was quantified by reference to RNA extracted from a virus stock with a known titer.

96 hpt Time Point—Transfected cells grown in 6-well plates for 96 hpt were washed with PBS and lysed with 1 ml of TRIzol® reagent, and total RNA was purified following the manufacturer's protocol. cDNA was synthesized from 1.5 μg of total RNA using the iScript cDNA synthesis kit (Bio-Rad). qPCR analysis was done on a iQ5 multicolor real-time PCR detection system (Bio-Rad) on 2 μl of the cDNA reaction together with a 0.25 μM concentration of qPCR NS1F and qPCR NS1R primers (supplemental Table 1; primer sequences from R. Takhampunya) in a 20-μl reaction using iQ SYBR Green Supermix (Bio-Rad). Each data point was done in triplicate. Thermocycling conditions were as follows: 95 °C for 3 min, 35 cycles of 95 °C for 20 s, and 30 s each at 55 °C and 72 °C. Measurements of SYBR Green signal were done at the annealing step. RNA copy numbers were determined using purified PCR products of the NS1 region of known concentrations. Values were further normalized from GAPDH qPCR results generated using the same protocol and cDNA batch but with primers qPCR GAPDHF and qPCR GAPDHR (supplemental Table 1).

Expression and Purification of Full-length NS5

E. coli Rosetta (DE3)pLysS cells transformed with pSUMO-DEN2*NS5*NHis (a gift from Dr. Craig Cameron of Pennsylvania State University) were cultured in 2 liters of Luria-Bertani (LB) growth medium supplemented with 25 μg/ml chloramphenicol, 30 μg/ml kanamycin (C25K30), and 0.1% glucose at 37 °C until mid-log phase (A_{600} 0.5–0.6). The medium was replaced with a fresh 2 liters of LB-C25K30 containing 1 mM isopropyl 1-thio-β-D-galactopyranoside. Cells were incubated for 48 h at 16 °C.

After induction, cells were harvested and lysed with 30 ml of lysis buffer (50 mM HEPES, 0.3 M NaCl, 10% glycerol, 1% Nonidet P-40, pH 7) supplemented with 500 μl of Protease Inhibitor Mixture VII (Calbiochem) and lysozyme (Pierce) on ice for 30 min. The lysates were sonicated on ice for 20 min total processing time (15 s on, 45 s off) and were centrifuged at 18,000 × *g* for 45 min. Supernatant was allowed to bind to TALON® resin (Clontech) in a 50-ml tube for 1 h in a cold room on a shaker.

The resin was washed five times with 10 ml of Buffer A (50 mM NaH_2PO_4, 0.3 M NaCl, 10% glycerol, pH 7). At the last wash, the slurry was transferred and packed in a 0.8 × 4-cm Poly-Prep chromatography column (Bio-Rad). The protein was eluted with 10 ml of buffer A containing 150 mM imidazole. Fractions of 500 μl were collected, and the presence of protein was monitored by a Bradford assay on a 96-well plate. The eluted fractions containing the most protein were combined in a single tube, and 20 units of SUMO protease I (Life Sensors, Malvern, PA) was added to cleave the N-terminal histidine tag to produce a full-length NS5 with an authentic N terminus. Digestion was allowed to proceed overnight on ice in the cold room. The purified protein was dialyzed using Spectra/Por 6 dialysis tubing (50,000 molecular weight cut-off; Spectrum Laboratories, Rancho Dominguez, CA) for 6 h in 1 liter of enzyme buffer (50 mM Tris-Cl, 50 mM NaCl, 5 mM $MgCl_2$, 40% glycerol, 1 mM fresh dithiothreitol, pH 7.5).

In Vitro RdRP Assay

Each 25-μl RdRP reaction contained 150 ng of RNA template, 4 μCi of [α-^{32}P]GTP, 0.5 mM ATP, CTP, UTP, 12.5 μM GTP, and 20 nM NS5 in 50 mM HEPES, 10 mM KCl, 5 mM $MgCl_2$, 2 mM $MnCl_2$, pH 8. The reaction was incubated at 37 °C for 30 min. The RdRP products were purified as described previously (10) and resuspended in 20 μl of diethylpyrocarbonate-treated water. A 5-μl aliquot was mixed with an equal amount of Gel Loading Buffer II (Ambion) and was heat-denatured at 70 °C for 10 min and then flash-cooled on ice for 5 min. Samples were run on a 8 M urea, 4% polyacrylamide gel in 1× TBE buffer at 125 V and visualized using a Storm 840 PhosphorImager (Amersham Biosciences). Band intensities were quantified using ImageJ version 1.45 software (55).

MPGAfold Analysis

For secondary structure prediction analysis, we used the MPGAfold (47–51, 56, 57) to fold a 719-nt subgenomic RNA (minigenome) from the DENV2 (New Guinea C strain) containing the 5′-terminal 226 nt, 42 nt from the C-terminal coding region of *NS5*, including the UAG termination codon, and the 451-nt 3′-UTR derived from pSY2 (10). Previous studies have shown that this RNA molecule contains the essential *cis*-acting elements required for efficient translation (15, 29, 30, 58) as well as negative strand RNA synthesis *in vitro* (10, 41, 54, 59). MPGAfold utilizes the genetic operators of mutation, recombination, and selection in parallel on a high performance computer. The program uses previously reported energy rules (60), including efn2 coaxial energy calculations at run time, to drive a population (*e.g.* 16,384) of RNA structures toward a set of structures that have high fitness (low free energy). MPGAfold is capable of predicting RNA secondary structure folding dynamics, focusing on the formation of significant final structures and folding intermediates (61–65). Thus, MPGAfold is especially suited for identifying and visualizing, in addition

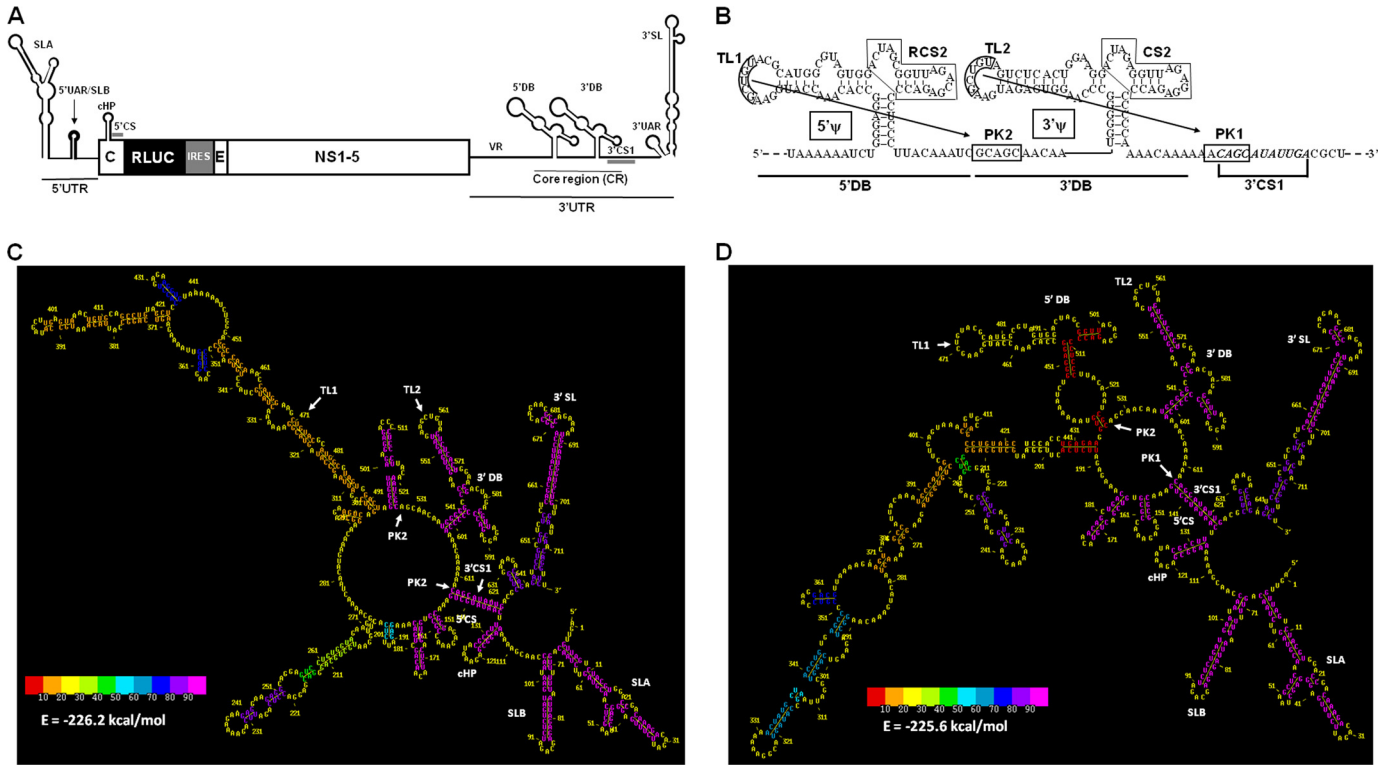


FIGURE 1. *A*, schematic of the DENV2 Rluc reporter replicon. Construction of WT and mutant replicons are as described under "Experimental Procedures." Briefly, using a DENV2 (New Guinea C strain infectious clone in yeast shuttle vector pRS424 (53), we have replaced the structural proteins with Rluc and EMCV IRES, retaining the first 75 nt of C containing the cHP and the 5′-CS, and the last 73 amino acids of E for proper endoplasmic reticulum translocation of NS1. Mutant replicons were made using this backbone for yeast recombination. *VR*, variable region. *B*, the secondary structure analysis of the core region of the 3′-UTR showed two DB structures (5′- and 3′-DB) (45). In the left arm of both DBs are identical pentanucleotide sequences, TL1 (nt 10,474–10,478) and TL2 (nt 10,562–10,566), which are predicted to base-pair with downstream sequences PK2 (nt 10530–10534) and PK1 (nt 10617–10621), respectively, to form two pseudoknot structures, 5′- and 3′-ψ (45). On the right arm of the DBs are duplicated conserved sequences RCS2 and CS2. *C*, a minigenome of DENV2 RNA sequence containing both 5′- and 3′-UTR sequences and other essential elements for RNA synthesis *in vitro* (10, 54) was used for analysis of secondary structures by MPGAfold as described under "Experimental Procedures." Based on this analysis, a stable best fit structure ($E = -226.2$ kcal/mol) has only the 3′-DB, whereas the appearance of the 5′-DB is visible in metastable structures such as in *D* (−225.6 kcal/mol). The frequency of occurrence of individual stems among the final predicted structures is *color-coded*, which increases from left to right (shown in the *bottom left*).

to highly stable RNA conformations with low free energy, metastable states that potentially occur in a folding pathway. Because the algorithm is nondeterministic, 100 independent runs at a 16,384 population level were performed for each WT and mutant RNA to determine the structure or structural components that form a consensus. StructureLab (56, 66) and, more specifically, the StemTrace (51) component of StructureLab were used to analyze the results and to obtain the frequencies of structural motifs (supplemental Table 2).

Statistical Methods

Rluc readings were normalized against the initial amount of RNA transfected and the corresponding Fluc signal. These were expressed as a percentage of WT. Because our objective is to determine the effect of each mutation on translation and replication, the values for each group were compared with WT using unpaired Student's *t* test ($\alpha = 0.01$), using GraphPad Prism version 5.03 (GraphPad Software, San Diego, CA).

RESULTS

The 3′-DB Forms a Stable Secondary Structure, Whereas the 5′-DB Occurs Mostly in RNA Folding Intermediates as Revealed by MPGAfold—A number of studies have indicated that secondary structures in both 5′- and 3′-terminal regions are required for efficient viral RNA translation (15, 29, 30, 58) and viral RNA synthesis (10–12, 16, 41). In this study, we focused on the role of 5′- and 3′-DBs and potential pseudoknot base pair interactions within the CR of DENV2 RNA in translation and replication, which is currently not understood. This knowledge is a prerequisite for a detailed analysis of the *trans*-acting factors that interact with these elements in mediating their effects in these processes.

Earlier RNA secondary structure prediction studies have revealed two almost identical DB-shaped structures (5′- and 3′-DBs) in the CR of DENV2 3′-UTR (45, 46, 67) (Fig. 1, *A* and *B*). These structural predictions and phylogenetic analyses were based on the 3′-UTR sequences alone from RNA folding algorithms. However, in previous studies using a subgenomic RNA of 719 nt and the viral replicase complex from DENV2-infected mammalian cells or purified NS5, we showed that long range interactions between 5′- and 3′-terminal regions of DENV2 RNA are important for RNA synthesis *in vitro*. For example, the 3′-UTR itself is not an active template for RNA synthesis *in vitro* by the viral polymerase unless the 5′-UTR is also added in *trans* or is present in the same RNA (10, 16, 41, 59). Therefore, in this study, we used the "minigenome," a subgenomic RNA sequence containing both 5′- and 3′-terminal sequences, for

secondary structure prediction analysis using the MPGAfold algorithm. The MPGAfold algorithm (47, 50, 56, 57), starting from a pool of randomly generated stem-loop structures from a given RNA sequence, recombines these parent structures over multiple generations until the population of RNA molecules evolves and reaches a consensus RNA structure. Because of the stochastic nature of MPGAfold, multiple independent runs must be done to determine the overall consensus structure for the given RNA.

Fig. 1*C* shows the best fit structure (−226.2 kcal/mol) from 100 different runs of the 719-nt DENV2 WT RNA sequence. The frequency of occurrence of each local structure is shown by a *color scheme*. The appearance of the already identified and functionally characterized DENV2 RNA secondary structures like the SLA, SLB (16, 27, 28), cHP (29, 30), 3′-SL (33, 35, 37, 68), and 5′-3′-CS1 base pairing (10–12, 14, 40, 41, 59) confirms the reliability of the MPGAfold-predicted RNA structures.

MPGAfold shows the formation of the 3′-DB with 97% frequency of occurrence (Fig. 1*C* and supplemental Table 2). However, the sequences involved in the proposed 5′-DB do not readily form the expected DB structure; instead, its upstream half, including TL1, is involved in a long stem-loop structure, whereas the lower half folds into another stem-loop with a frequency similar to that of the 3′-DB (>90%) (Fig. 1*C*).

In some runs, we noticed that the same RNA did form both 5′- and 3′-DBs as an energetically close suboptimal structure (Fig. 1*D*, −225.6 kcal/mol). The frequency of occurrence of the 5′-DB among the final structures reached only 6%. However, in an additional 62% of runs, it was found in intermediate structures.

This difference in the folding behavior of the two DBs led us to hypothesize that it is the 3′-DB that plays a significant role in the viral life cycle. MPGAfold analysis showed that TL2 is part of an unpaired loop in the 3′-DB in the best fit structure, whereas TL1 is alternatively base-paired with upstream sequences (Fig. 1*B*). Thus, despite their identical sequences of TL1 and TL2, it is TL2 that is available to bind PK1 in the 3′-DB to form the putative 3′-*ψ*. This notion is consistent with sequence analysis of different flavivirus 3′-UTRs showing that the exact sequence of TL2 is conserved in DENV, WNV, Murray Valley encephalitis virus, and JEV, whereas TL1 is only conserved within the DENVs (45).

Construction and Characterization of the DENV2 Rluc Replicon—To study the effect of different RNA motifs on viral translation and replication, we constructed a DENV2 New Guinea C strain replicon expressing Rluc (Fig. 1*A*). This was cloned into the yeast shuttle vector pRS424 directly following the SP6 promoter, which was used for *in vitro* transcription. The viral structural proteins were replaced by the Rluc gene fused with an EMCV IRES element, which directs cap-independent translation of the nonstructural proteins, NS1–NS5. Specific sequences from the N terminus of the capsid-coding region and the last 73 amino acids of the envelope protein were retained because these are required in *cis* for RNA replication. For specific details of the construction, see "Experimental Procedures." In addition, we constructed the replication-deficient GND replicon by substitution of the highly conserved GDD → GND in the polymerase domain (69, 70).

Kinetics of the Rluc activity following electroporation of the WT and GND replicon RNAs in BHK21 cells show a peak of Rluc activity at 2 hpt, indicating translation of the input WT and GND replicon RNAs (Fig. 2*A*). This activity declined over time, which was then followed by a second increase of luciferase activity in WT replicon up to 96 hpt due to translation of the newly replicated RNA; this second burst of Rluc activity was not observed in the GND mutant. This was confirmed by directly quantifying replicon RNA at 96 hpt through qPCR (Fig. 2*B*) and by immunofluorescent staining for detection of NS5 (Fig. 2*C*). These results indicated that the WT and GND replicons could be used in our analysis of the effects of *cis*-acting elements in translation and replication by quantifying the luciferase activities at 2 and 96 hpt, respectively.

The Roles of Conserved TL and PK Motifs in RNA Translation—To verify whether the RNA folding inequality and the frequency of occurrence between the 5′- and the 3′-DBs based on the MPGAfold analysis reflect differences in TL1 and TL2 functionalities (Fig. 1, *C* and *D*), we analyzed the effects of mutations engineered to disrupt the potential base pair interactions among TL1/PK2 in the 5′-DB and TL2/PK1 in the 3′-DB (Fig. 1*B*). We first deleted the conserved sequence of 5 nt in one or both TLs, creating mutants Δ^5TL1, Δ^5TL2, and Δ^5TL1Δ^5TL2 (Fig. 3*A*). The results at 2 hpt indicated that Δ^5TL2 modestly affected translation (72% of WT, $p < 0.0001$), whereas deleting TL1 did not have any effect (Fig. 3*B*). However, when both were deleted in tandem, translation further decreased to 40% of WT ($p < 0.0001$), suggesting that TL1 exhibits a cooperative synergy with TL2 in enhancement of translation of DENV2 RNA.

To further investigate the role of TL1 and TL2, we employed a second approach to disrupt the TL1/PK2 and TL2/PK1 base pair interactions. We introduced point mutations by "flipping" the TL sequences (5′-GCUGU-3′ → 5′-UGUCG-3′), generating $TL1_{Flip}$, $TL2_{Flip}$, and $TL1_{Flip}TL2_{Flip}$ (Fig. 3*A*). At 2 hpt, $TL1_{Flip}$ was efficiently translated at levels similar to WT, confirming the results obtained with Δ^5TL1 (Fig. 3*B*, *leftmost panel*). In the $TL2_{Flip}$ mutant, translation was reduced to 80% of WT ($p = 0.0006$) and 33% in $TL1_{Flip}TL2_{Flip}$ ($p < 0.0001$). The results obtained with Flip mutants thus confirm our results obtained with the deletion mutants described above. Next, we examined whether the reduction in translation of some of the mutant RNAs is due to differences in RNA stabilities. RNA stabilities of the WT and mutants were determined by qPCR. The results showed that there was no appreciable change in RNA stabilities of these mutants from that of the WT (Fig. 3*C*).

We sought to examine whether the putative partners of TL1 and TL2 would also have a similar effect on translation. Conserved sequences PK2 and PK1 were flipped to make $PK2_{Flip}$ and $PK1_{Flip}$. These mutations unexpectedly did not affect translation (Fig. 3*B*, *second panel*). Another set of PK2 and PK1 mutants, $PK2_{mut}$ and $PK1_{mut}$ (Fig. 3*A*), were constructed to verify this result. Both translated efficiently like WT, $PK2_{Flip}$, and $PK1_{Flip}$ (Fig. 3*B*, *rightmost panel*). To eliminate the possibility that a mutation in one PK might be rescued by the other non-mutated PK, we mutated both PK2 and PK1 in the same replicon RNA ($PK2_{mut}PK1_{mut}$). Fig. 3*B* shows that $PK2_{mut}PK1_{mut}$ still translated as efficiently as WT.

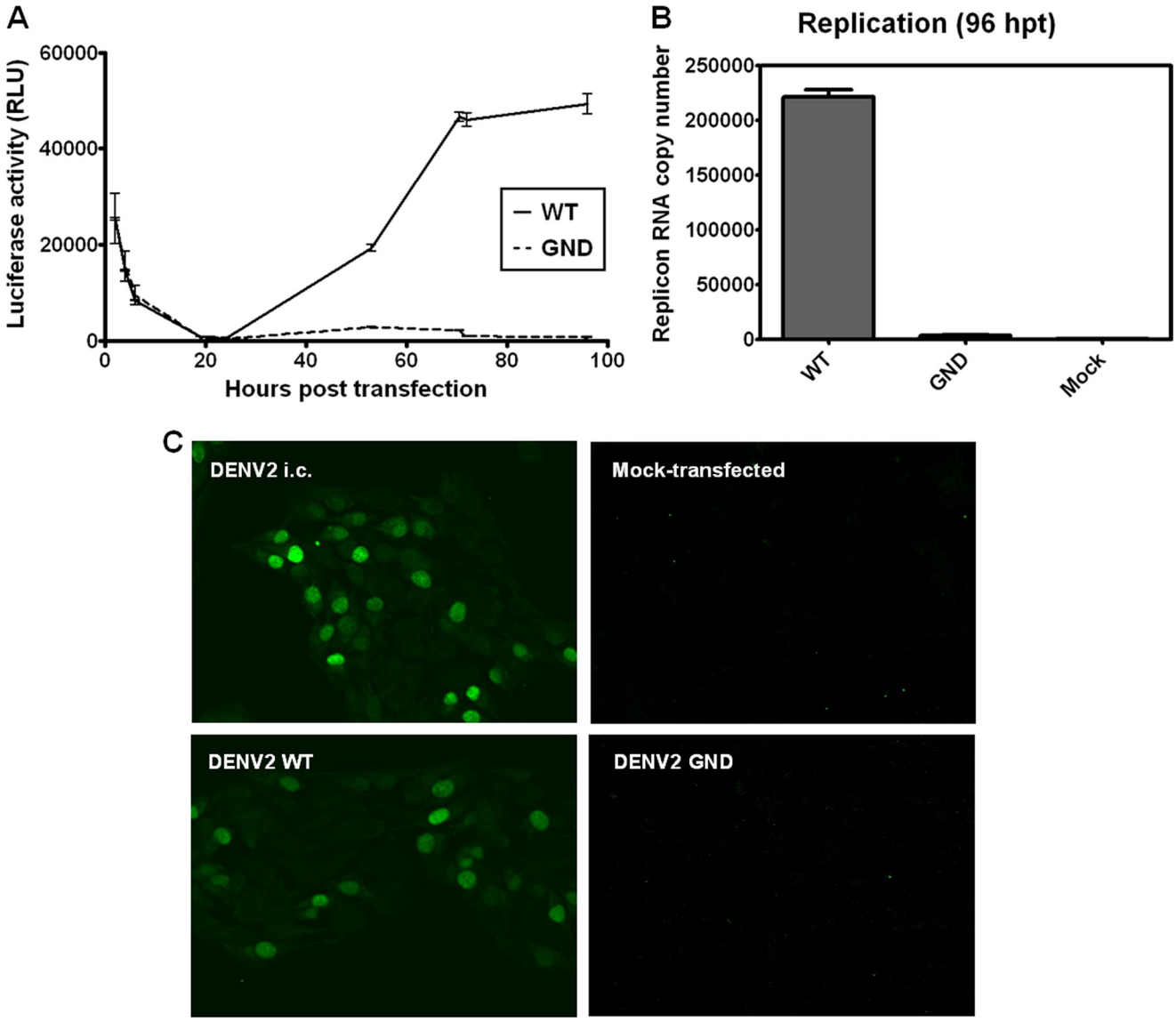


FIGURE 2. **Characterization of DENV2 replicon encoding Rluc reporter.** *A*, the kinetics of luciferase expression of WT and a replication-deficient mutant (GDD → GND) were followed at different time points post-transfection (hpt). 2 and 96 hpt were determined to be the measure of translation and replication, respectively, of input RNA electroporated into BHK21 cells. *RLU*, relative light units. *B*, direct quantification of replicated RNA by quantitative RT-PCR. Quantitative RT-PCR was carried out by RNA extraction from DENV2 replicon-transfected BHK21 cells at 96 hpt as described under "Experimental Procedures." *C*, detection of DENV2 NS5 by immunofluorescence staining for NS5. Expression of viral proteins was confirmed by immunofluorescence for detection of DENV2 NS5 in BHK21 cells transfected with DENV2 replicon as described under "Experimental Procedures." Cells transfected with infectious RNA of DENV2 were used for a positive control (*DENV i.c.*). *Error bars*, S.E.

Interestingly, the restoration of the base pairing between the TLs and PKs in $TL2_{Flip}PK1_{Flip}$ and $TL1_{Flip}PK2_{Flip}TL2_{Flip}PK1_{Flip}$ having the 20 and 67% defect in translation, respectively, with non-viral sequences (Fig. 3*A*) restored translation back to WT levels (Fig. 3*B*, *second panel*). Similar restoration of translation to WT levels was observed in another set of mutants that swapped the positions of TL2 and PK1 sequences (*TL2 → PK1/PK1 → TL2* in Fig. 3*B*, *third panel*).

Taken together, our evidence suggests that TL1 and TL2 play an important role in translation with unequal contributions and that their specific sequences are not required. Substitutions with non-viral sequences can still maintain efficient translation at WT levels provided that they retain base pairing with PK2 and PK1.

The TLs and PKs Are Critical for Viral Replication—Replication of WT and mutant replicons was assayed by measuring the luciferase signal at 96 hpt. The results indicated that replication of all of the mutants was attenuated (Fig. 3, *D* and *E*). TL1 and TL2 have different contributions to replication. A deletion or substitution mutation of the TL1 motif resulted in a 70–80% decrease in Rluc signal (Fig. 3*D*) and a 50% drop in RNA levels (Fig. 3*E*). Mutation of TL2, on the other hand, had a much greater effect, with replication at ~10% of WT. Mutating both sequences in one RNA generated a more severe phenotype. In contrast to their phenotypes in translation, PK2 and PK1 mutants are defective in replication (Fig. 3, *D* and *E*).

We also sought to examine whether restoring the base pairing in TL and PK mutants with non-viral sequences could rescue the

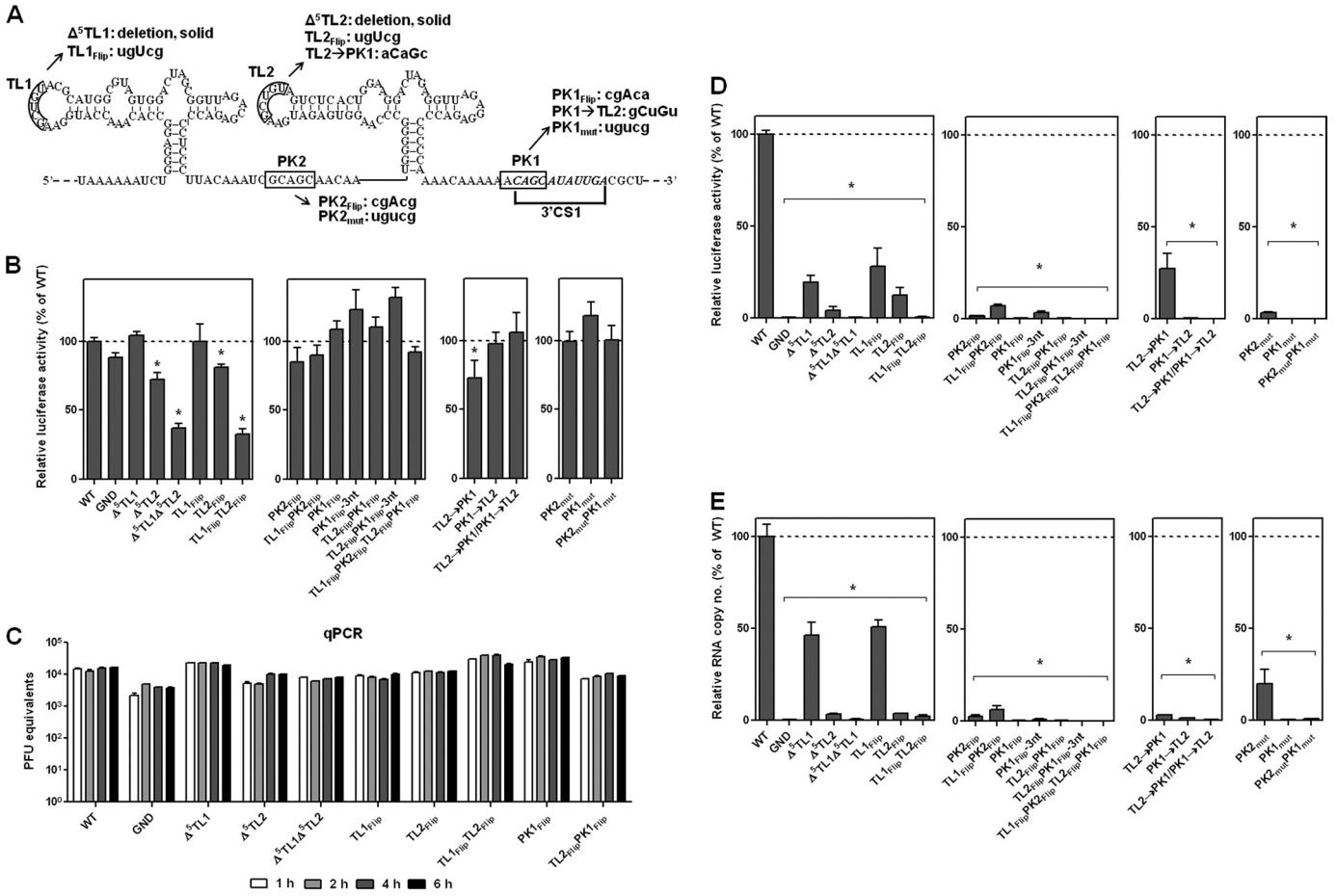


FIGURE 3. **Mutational analysis of conserved putative pseudoknot elements in translation and replication.** *A*, WT and mutant RNA elements in the 5′- and 3′-DBs are shown. *B*, the effects of various mutations in translation were examined by measuring the luciferase activities in BHK21 cells transfected with replicon RNAs at 2 hpt. *C*, the RNA stabilities of selected mutants were examined directly by qPCR. *PFU*, plaque-forming units. *D* and *E*, the effects of various mutations in replication were determined. Transfected replicon RNAs in BHK21 cells were incubated for 96 hpt, and both Rluc activities (*D*) and RNA levels by qPCR were measured (*E*). *Error bars*, S.E. *, $p < 0.001$ compared with WT.

defect in replication. Unlike in translation, the base pairing between mutant TL and PK motifs did not restore RNA synthesis ($TL1_{Flip}PK2_{Flip}$, $TL2_{Flip}PK1_{Flip}$, $TL1_{Flip}PK2_{Flip}TL2_{Flip}PK1_{Flip}$, and $TL2 \rightarrow PK1/PK1 \rightarrow TL2$ in Fig. 3, *D* and *E*). Because PK1 overlaps with the 3′-CS1 (Fig. 1*B*), we sought to determine whether the loss of replication was due to the decrease in base pairing between the cyclization sequences. We introduced compensatory mutations in the 5′-CS to restore the loss of base pairing between the 5′-3′-CS1 in $PK1_{Flip}$ and $TL2_{Flip}PK1_{Flip}$ (indicated by "-3nt"). MPGAfold analysis showed that the frequency of occurrence of the 5′-3′-CS1 base pairing is reduced as expected in the $PK1_{Flip}$ mutant because of a reduction in the number of base pairs between the cyclization sequences. This frequency was restored in the $PK1_{Flip}$-3nt and the $TL2_{Flip}PK1_{Flip}$-3nt mutants (supplemental Table 2). These compensatory mutations for restoration of the 5′-3′-CS1 base pairing in these replicons did not improve their replication (Fig. 3, *D* and *E*), suggesting that specific sequences within the 5′-3′-CS1 are important for replication.

The Conserved Nucleotides Surrounding TL2 Motif Also Play a Role in Translation and Replication—Our data in Fig. 3*B* indicate the possibility that the TL2/PK1 interaction to form the putative 3′-ψ might be playing a role in stabilizing a structure suitable for translation. We hypothesized that if we increase the strength of the 3′-ψ by increasing the number of base pairs, translation efficiency could be enhanced. We chose to mutate the nucleotides adjacent to TL2 that are also conserved and yet are not predicted to participate in the 3′-ψ base pairing (45, 46) and have not been functionally studied previously.

We generated various mutants that increased the number of base pairs by 1–4 bp in the region surrounding TL2 or PK1 (Fig. 4*A*). Mutating the A immediately downstream or upstream of TL2 (+1 down or +1 up) while increasing the base pairing to 6 nucleotides has a dramatic effect on translation (45 or 60% of WT, respectively; Fig. 4*B*). Mutations of other residues resulting in 7–9 bp overall had negative effects on translation except for one mutant containing 7 continuous bp, which did not show an appreciable effect (+2 up in Fig. 4, *A* and *B*) ($p = 0.1118$). The data collectively suggest that the conserved nucleotides adjacent to TL2 are also important for efficient translation.

In contrast to their phenotypes in translation, these mutants exhibit a more pronounced defect in replication (Fig. 4*C*). It is interesting to note that even one base change (+1 down) can result in a 90% reduction in RNA synthesis. The critical roles of these nucleotides in translation and replication thus explain why these sequences are conserved (45, 46).

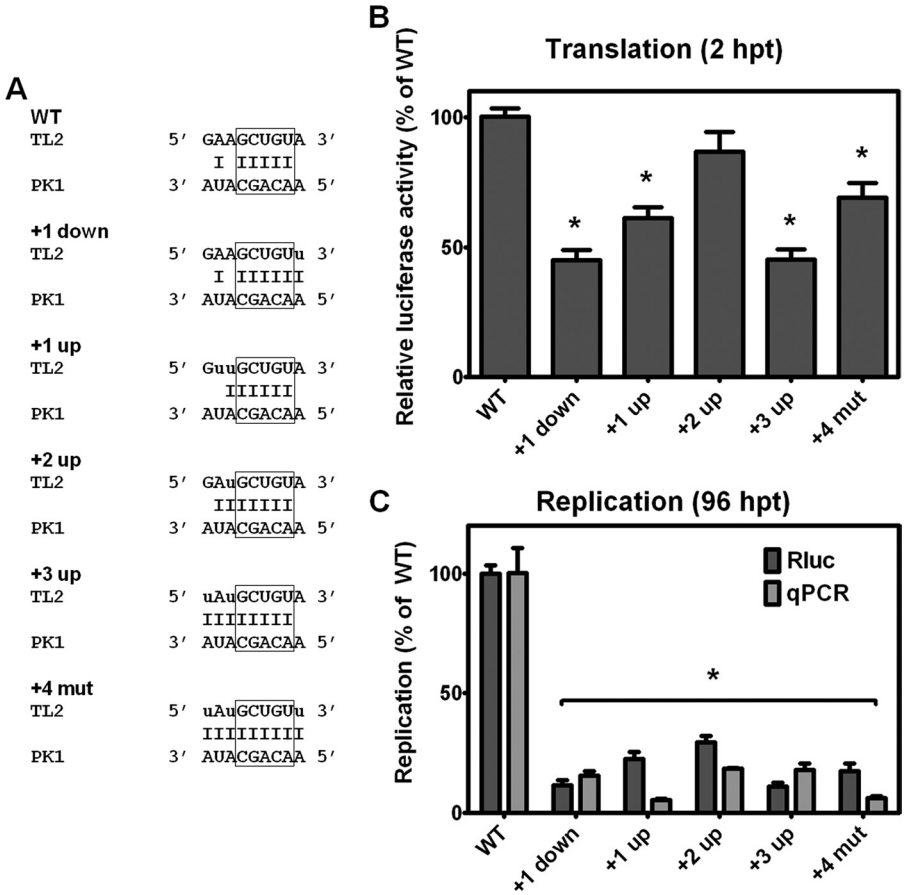


FIGURE 4. **Effect of mutations of the conserved sequences in the vicinity of TL2.** *A*, the conserved nucleotides surrounding the TL2 motif were mutated to strengthen the TL2/PK1 base pairing. Bases in *lowercase type* denote mutation. *B*, translation of WT and mutant replicons was measured by Rluc activities at 2 hpt. *C*, replication of WT and mutant replicons was measured by Rluc activities at 96 hpt (*dark gray*) and by qPCR (*light gray*). *Error bars*, S.E. *, $p < 0.001$ compared with WT.

The CS2 Sequence, but Not Secondary Structure, Determines Its Role for Its Functionality—As a part of our overall objective to study the role of 3′-DB in translation and replication, we examined the contribution of the CS2 region within the 3′-DB in these two processes. One previous study done using WNV replicon reported that the entire deletion of CS2 resulted in a significant impairment of replication (12). In this study, we sought to determine the role of CS2 by creating substitution mutations without altering its secondary structure ($CS2_{mut}$) (Fig. 5*A*) as predicted by MPGAfold (data not shown). To determine which part of the CS2 structure or sequence is important, we made three additional replicons that introduced mutations in the 3′-DB junction (Loop A_{mut}), stem ($Stem_{mut}$), and loop (Loop B_{mut}) (Fig. 5*A*). All three subdomains affected both translation and replication to a similar extent (Fig. 5, *B* and *C*). The combination of all three mutations did not have an additive effect, suggesting that all subdomains most likely functioned as one structural motif that was essential for CS2 function.

Cis-Acting Elements in the Core Region Are Involved in a Noncanonical Cap-independent Translation—It was reported that DENV2 RNA required both 5′- and 3′-UTR to initiate translation under conditions of inhibition of translation initiation factor 4E by the addition of LY294002 (58). We sought to determine whether the RNA elements within the 3′-DB are required for the viral RNA to be translated by a noncanonical mechanism. This compound inhibits the phosphotidylinositol-3 kinase, which leads to the hypophosphorylation of 4E-BP1 (eIF4E-binding protein 1) (71). This form of 4E-BP1 is able to sequester eIF4E, thus suppressing cap-dependent translation of the majority of capped mRNAs in the compound-treated host cells. We selected mutant replicons that would give a good representation of all of the mutants we have done so far: $\Delta^5TL1\Delta^5TL2$, $TL1_{Flip}TL2_{Flip}$, $TL1_{Flip}PK2_{Flip}TL2_{Flip}PK1_{Flip}$, $PK2_{mut}PK1_{mut}$, +4 mut, and $CS2_{mut}$. The mRNA containing the 5′- and 3′-UTR of the α-globin gene flanking a firefly luciferase coding sequence with an engineered poly(A) tail (15) was used as an internal control for the transfections of WT and mutant Rluc replicons. The time course of this control RNA was followed in the presence and absence of LY294002. The addition of LY294002 to the transfected cells inhibited the cap-dependent translation of this control mRNA (Fig. 6*A*). However, WT replicon RNA was translated efficiently in the presence of LY294002 (Fig. 6*B*, *WT*) (58). We consistently observed an increase in the Rluc signal obtained from the translation of WT replicon RNA in the presence of the LY294002 compared with the signal obtained in the absence of the drug. Moreover, the 5′- and 3′-DB mutants that showed deficient translation in the absence of the drug LY294002 were still defective in the presence of the drug (Fig. 6*B*). However, the cap-dependent translation of the control RNA GLGpA (15) included in all transfec-

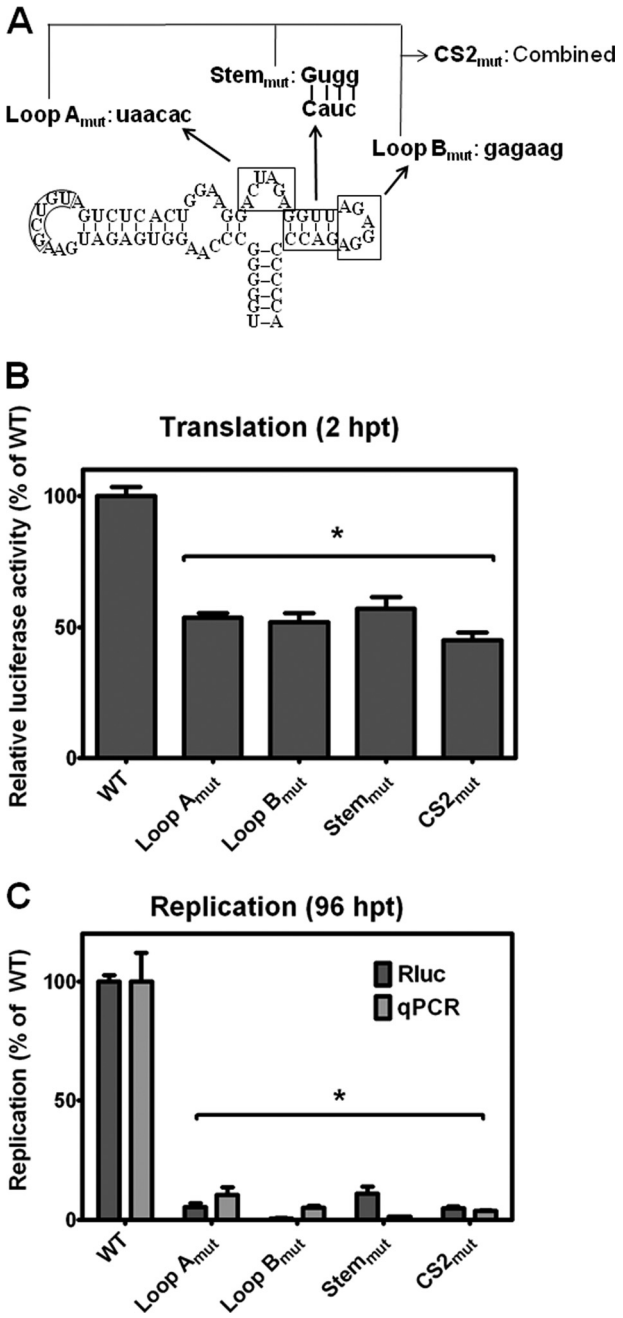


FIGURE 5. **Effect of mutations in the conserved CS2 region of 3′-DB.** *A*, subsections of the CS2 region were mutated such that the predicted secondary structure is conserved. Bases in *lowercase type* denote mutation. Translation (*B*) and replication (*C*) of WT and mutant replicons transfected into BHK21 cells were measured at 2 and 96 hpt, respectively. Replication measured by Rluc assay (*dark gray*) and directly by qPCR (*light gray*) is shown. *Error bars*, S.E. *, $p < 0.001$ compared with WT.

tion experiments was inhibited in the presence of LY294002 (Fig. 6*C*), indicating that the drug was active in these assays. These results indicated that the motifs in the CR are required for the noncanonical cap-independent translation of the viral RNA when the cap-dependent translation is inhibited. Interestingly, restoration of base pairing between TK1/PK2 and TL2/PK1 in the mutant, $TL1_{Flip}PK2_{Flip}TL2_{Flip}PK1_{Flip}$, brought translation back to the WT level (Fig. 6*B*), suggesting that there is a functional interaction between these motifs.

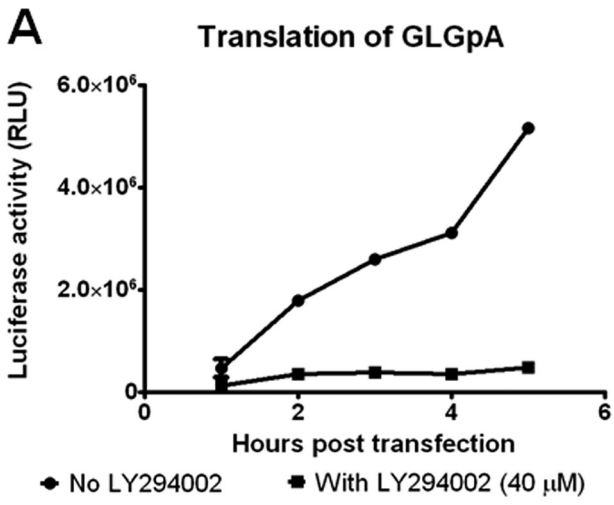
The CR Elements Do Not Affect (−)-strand Synthesis—To explore the step at which the conserved sequences in the CR act during replication, we tested the efficiencies of RNA synthesis *in vitro* of (+)-strand RNA templates containing selected mutations using a full-length NS5 with authentic N and C termini. The RdRP assay showed that all mutants produced (−)-strand efficiently with the exception of $PK1_{Flip}$ and $TL2_{Flip}PK1_{Flip}$ (Fig. 7). This defect in replication was rescued by restoring the 5′-3′-CS1 complementarity ($PK1_{Flip}$-3nt and $TL2_{Flip}PK1_{Flip}$-3nt). These are in contrast to our replicon data that showed that the 5′-3′-CS1 and CR sequences were required for replication (Fig. 3, *D* and *E*). This suggests that these nucleotides are required at subsequent steps of replication.

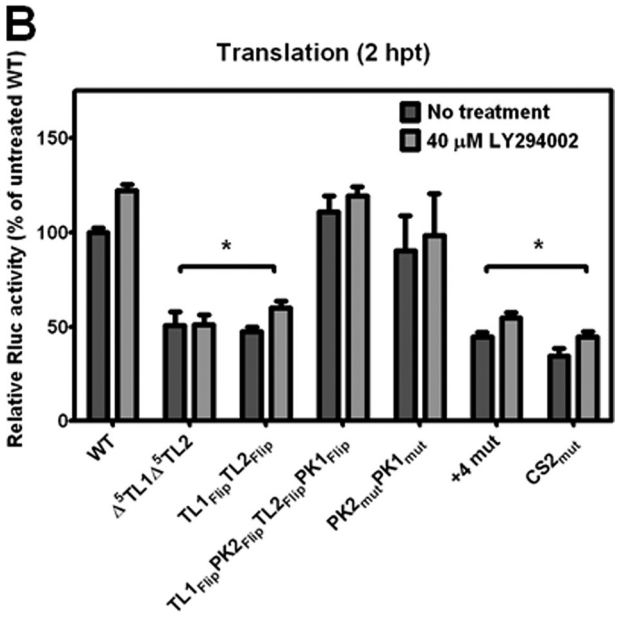
DISCUSSION

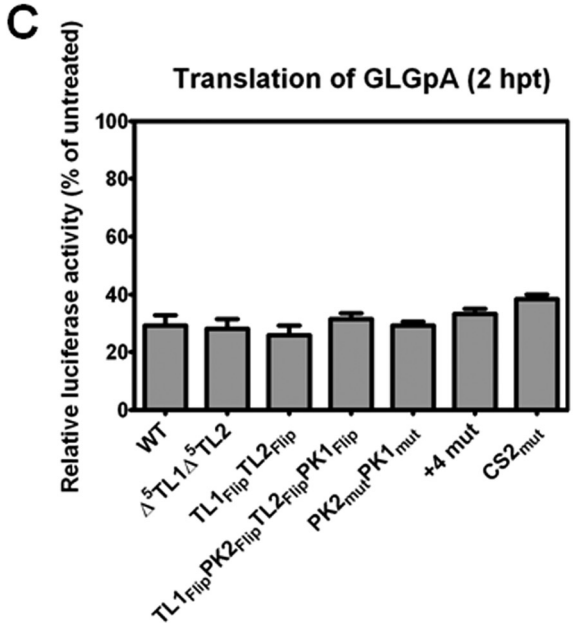
Previous *in silico* analyses of various DENV2 3′-UTR sequences have predicted that the CR folds into two similar DB-like structures, 5′- and 3′-DB (45, 46, 67). Our study is unique in that we take into consideration the influence of 5′-terminal sequences and structures on the folding of 3′-UTR by using the MPGAfold for predicting RNA secondary structures.

Moreover, MPGAfold is able to capture metastable intermediates in an RNA folding pathway. These metastable intermediates as well as stable and best fit structures predicted by MPGAfold have been shown to have biological relevance in a number of previous studies (61–65, 72). Another advantage of MPGAfold is that besides computing the free energy of the RNA, it also provides us with information on the relative frequencies of occurrence of local secondary structural motifs.

Although MPGAfold reveals the formation of RNA secondary structural elements, such as SLA, SLB, cHP, 5′-CS/3′-CS1 cyclization, 3′-DB, and 3′-SL readily (Fig. 1*C*), the 5′-DB and 5′-3′-UAR interaction is not seen in this stable structure. Surprisingly, the base pairing between 5′- and 3′-UAR elements occurs only 8% of the time in the final structure with an additional 27% in the intermediate structures (supplemental Fig. 1 and supplemental Table 2). This is consistent with the observation that the 5′-3′-UAR binding occurs only after the 5′-3′-CS1 circularization has been established (73), suggesting that the latter is the primary driving force in genomic end-to-end communication. This conclusion is also supported by a stem trace plot of the minigenome RNA maturation in an MPGAfold run (data not shown). However, there is strong experimental evidence that 5′-3′-UAR base pairing is essential for relieving the repression of RNA synthesis by RdRP *in vitro* by 3′-SL (74). The 3′-SL includes a part of 3′-UAR element, which upon a conformational change in the presence of the viral polymerase is likely to form the duplex 5′-3′-UAR (74). Conformational changes occurring within the 3′-SL due to 5′-3′-terminal RNA-RNA interaction as well as the NS5/RNA interactions by structure probing and footprinting methods have also been reported for WNV (75). Thus, these results in light of MPGAfold analysis suggest that the 5′-UAR forms part of a long stem loop and that the 3′-UAR is a part of the 3′-SL in the best fit structure (Fig. 1*C*); the 5′-3′-UAR duplex, which occurs only in a metastable structure (supplemental Fig. 1 and supplemental Table 2),







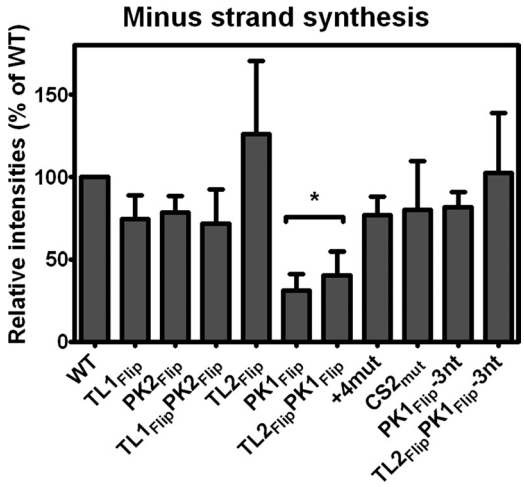


FIGURE 7. **RdRP assay of (+)-copies of 719-nt minigenomes containing CR mutations.** WT or mutant (+)-RNAs (150 ng) were used as templates for (−)-strand synthesis by 20 nM purified NS5 in the presence of [α-^{32}P]GTP. RNA products were run on 4% 8 M urea-polyacrylamide gel, detected by a PhosphorImager. Band intensities were measured by ImageJ software and expressed as a percentage of WT. The experiment was repeated five times. *Error bars*, S.E. *, $p < 0.05$ compared with WT.

could possibly be formed upon polymerase binding prior to RNA synthesis.

The previously predicted RNA structures of the CR using 3′-UTR sequences showed the formation of both the 5′- and 3′-DBs (45, 46, 67). However, the best fit structure produced by MPGAfold after 100 independent runs contained only the 3′-DB with a high frequency (Fig. 1*C*). The simultaneous appearance of both DBs occurs mostly in suboptimal folding intermediates in which the population of RNA structures contains 62% 5′-DB (Fig. 1*D*), although among the final structures, the 5′-DB was predicted at a very low frequency (6%). Thus, the long range interaction of the 5′- and 3′-termini is an important determinant in RNA structure prediction of flaviviruses as revealed in this study. In addition to the folding influence of the 5′-terminal sequence that includes 5′-CS, the rarity of formation of the 5′-DB may also be explained by the nature of the bases in its long arm (Fig. 1*B*). The base pairing in this arm is interrupted midway by an internal loop, reducing the number of stacked pairs to five. On the other hand, the 3′-DB has eight uninterrupted base pairs (Fig. 1*B*). The position of the break in the 5′-DB arm increases the potential for that structure to breathe open and allow its sequences to interact with other nucleotides.

Because most of the final structures show the 5′-DB sequences in an alternative stem motif, TL1 remains largely inaccessible to the proposed PK2 interaction to form the 5′-ψ.

FIGURE 6. **Noncanonical cap-independent translation of selected mutants of the CR.** *A*, the translation of the control reporter containing 5′- and 3′-UTRs and the 3′ poly(A) of the α-globin gene flanking the Fluc gene (GLGpA as described in Ref. 15) in the presence and absence of 40 μM LY294002 is shown. *B*, the translation of selected mutant replicons was measured in the presence or absence of 40 μM LY294002, an inhibitor of cap-dependent translation. Rluc activity of WT replicon measured in the absence of LY294002 was set as 100%. *C*, the control RNA GLGpA was co-transfected with WT and mutant DENV2 replicon RNAs in each transfection experiment. The translation of control RNAs measured as Fluc activity was plotted as a percentage of Rluc activity of WT replicon RNA-transfected cells in the absence of LY294002. *Error bars*, S.E. *, $p < 0.001$ compared with WT.

The stability of the 3′-DB structure, however, enables TL2 to be maintained as a single-stranded loop that can readily bind to the PK1 and form the 3′-ψ. In addition, the number of base pairs at the stalk of the 3′-DB corresponds to roughly a half helix turn, positioning TL2 nearer to PK1, suggesting that the 3′-ψ formation is sterically feasible. The differences in the frequency of formation of the DBs and the accessibility of TL2 for interaction with PK1, respectively, support the notion that the 3′-DB elements are more important for the virus. This prediction agrees with the sequence alignment data showing that the 3′-DB elements are present in the entire flavivirus genus, whereas the 5′-DB is only found in the DENV and JEV serological groups (40, 76). Moreover, the TL2 and PK1 are conserved within the flaviviruses, whereas TL1 and PK2 sequences are only conserved within the DENV serotypes (45). Our results show that the contribution of TL1 to translation can only be seen when TL2 is also mutated (Fig. 3*B*), suggesting that TL1 cooperates with TL2 in translation.

MPGAfold analyses for RNA folding were based on the minigenome, whereas the effects of mutations in translation and replication were analyzed in the context of the luciferase-based replicon RNAs (10,210 nt). Therefore, we sought to assess the potential impact of the luciferase coding region on the structures in the 5′- and 3′-terminal regions that contain the RNA elements of interest. Due to the computational constraints of the MPGAfold algorithm, we surveyed the full replicon secondary structures with the aid of two fast dynamic programming algorithms: Mfold (60, 77) and RNAfold (78). A common self-contained subdomain (1184 nt in length) bordered by conserved interactions between the 5′-end of the Rluc gene and the 3′-end of the *NS5* gene was identified by both algorithms (data not shown). Independent folds of this subdomain by the two dynamic programming algorithms and MPGAfold indicated a potential for interactions between the 5′-end of the Rluc gene and the 5′-DB sequences upstream of its head SL (data not shown). The key structural motifs discussed here (SLA, UAR, CS, 5′-DB head, 3′-DB, and 3′-SL) were present in the structures predicted by the dynamic programming algorithms, whereas the full 5′-DB was still found in energetically close suboptimal structures, suggesting that the interference of the Rluc gene in the RNA folding pathway is minimal.

Our results of replication assays suggest that mutations of these conserved motifs all affect replication. The replication-deficient $PK1_{Flip}$ and $TL1_{Flip}PK2_{Flip}TL2_{Flip}PK1_{Flip}$ mutations impact mainly the frequency of occurrence of the structural motifs rather than inducing major full secondary structure alterations as seen by MPGAfold and yet function differently (supplemental Fig. 3, *A* and *B*). Therefore, there seems to be a strong sequence-specific requirement for their function in replication because these mutations result in attenuation (Figs. 3, *D* and *E*, 4*C*, and 5*C*). This sequence specificity for replication is most evident in the $TL2_{Flip}PK1_{Flip}$-3nt mutant because the restoration of the TL2-PK1 and 5′-3′-CS1 base pairs was not sufficient to rescue the lethal phenotype (Figs. 3, *D* and *E*). A similar conclusion was reported in which the 3′-CS1 involving the same nucleotides that constitute PK1 was mutated in the WNV replicon (43). Their results showed that specific bases in the 5′- and 3′-CS1 are important for replication, regardless of complementarity. This sequence-specific requirement of PK1 in the 3′-CS1 and its complementary sequence in the 5′-CS in replication is not yet understood. Moreover, the *in vitro* RdRP assays using (+)-strand templates containing WT and 5′- and 3′-DB mutant sequences and purified full-length DENV2 NS5 polymerase (54, 59) revealed that the minus strand synthesis was not affected by these mutations (Fig. 7). These results suggest that the RNA elements in the 5′- and 3′-DBs are involved at a subsequent step in viral replication.

Our results, taken together, suggest that TL1 and TL2 have variable modulatory effects in translation. It is noteworthy that the translational defects in $TL2_{Flip}$, $TL1_{Flip}TL2_{Flip}$, and TL2 → PK1 could be rescued by a compensatory mutation in PK2 and PK1, although mutations in PK1 and PK2 motifs *per se* do not show any effect. One possible explanation is that there are alternative base pair interactions in which TL1 or TL2 could participate when PK1 and PK2 are mutated (supplemental Fig. 2). For example, TL2 could participate in an alternative base pair interaction with a motif within the top loop structure of the 3′-SL when PK1 is mutated.

More importantly, the ability of the viral RNA to switch to a noncanonical cap-independent pathway in translation was also restored in the mutant, $TL1_{Flip}PK2_{Flip}TL2_{Flip}PK1_{Flip}$ (Fig. 6*B*), supporting the model that there is a functional interaction between these motifs that is important in translation. A previous study examined the folding of the 5′- and 3′-DB regions of DENV4 3′-UTR RNA by RNA probing analysis, which revealed that PK2 and PK1 were susceptible to double strand-specific RNase V1 digestion. However, their binding partners as well as the effect of the 5′-terminal region on the folding of the CR were not investigated (79). A more thorough structural analysis is warranted to determine the solution structure of the CR and identify short and long range tertiary interactions between this region and different elements in the genome. Recent advances in structure probing have made it possible to analyze RNA structures in the context of the full viral genome. The high throughput selective 2′-hydroxyl acylation analyzed by primer extension has been used to determine the complete structure of the HIV-1 genome (80). A similar method can be applied to DENV to solve the secondary structure of the flaviviral genome and identify the binding partners of the elements in the CR.

Mutations of the surrounding nucleotides of the TL2 loop by increasing the number of base pairs affected translation (Fig. 4, *A* and *B*), suggesting the possibility that increasing the putative base pairing between TL2/PK1 destabilizes the 3′DB structure. Another possibility is that these four nucleotides also exhibit a sequence-specific requirement for binding of a protein. CS2 has a similar sequence-specific requirement for translation. It seems that mutation of any subregion results in a similar 50% decrease in translation (Fig. 5, *A* and *B*). Because the CS2 sequence is duplicated in the 5′-DB as RCS2 (Fig. 1*B*), it is likely that RCS2 may cooperate with CS2 in translation.

Previous studies revealed that some host proteins bind within the 3′-UTR, although the precise roles of these RNA-protein interactions in the virus life cycle are not fully understood at present. For example, EF-1α binds to the 3′-SL of WNV (81, 82) and DENV (83). The La protein and polypyrimidine tract-binding protein bind to the 3′-UTR of DENV4 (83).

The polypyrimidine tract-binding protein also interacts with DENV NS4A protein, and this interaction is required for viral replication (84). The host proteins containing the RNA recognition motif, such as TIAR and TIA-1, were shown to interact with WNV 3′-SL of (−)-strand RNA and increase the efficiency of WNV growth (85). On the other hand, the Y-box-binding protein-1 binds to the 3′-SL of DENV RNA and represses translation (86). Only the poly(A)-binding protein, which plays an important role in host mRNA translation, was reported to bind within the DENV CR (87).

There are a number of examples in which the RNA elements within the 3′-UTRs of positive strand RNA viruses play an important role in translation. In the Turnip crinkle virus 3′-UTR, a tRNA-like structure binds to 80 S ribosome and the 60 S ribosomal subunit (64, 65). In addition, in many plant virus RNAs, there are different classes of cap-independent translation elements within the 3′-UTRs that take over the function of the 5′-cap (for a review, see Ref. 88). Although DENV has a 5′-cap for translation of its genome, when the cap-dependent translation is limiting, a similar mechanism may take over under conditions of inhibition of translation initiation factor 4E required for cap-dependent translation, as reported previously (58). The identification of the RNA elements within the CR that seem to be required for both cap-dependent and a noncanonical cap-independent translation warrants further investigation to identify host factors that mediate in this pathway.

Acknowledgments—We thank Alan Gee for help with early MPGA-fold analysis, Dr. Sofia Alcaraz for initial assistance with transfections, and Dr. Masako Nomaguchi for cloning some DNA fragments used in replicon construction. We also thank Dr. Theo Dreher (Oregon State University) for the pGLGpA plasmid, Dr. Craig Cameron (Pennsylvania State University) for the pSUMO-DEN2NS5NHis and E. coli Rosetta (DE3)pLysS cells, and Dr. Maja Maric (Georgetown University) for the use of the Amaxa Nucleofector.

REFERENCES

1. Westaway, E. G., Mackenzie, J. M., and Khromykh, A. A. (2003) *Adv. Virus Res.* **59,** 99–140
2. Heinz, F. X., and Allison, S. L. (2000) *Adv. Virus Res.* **55,** 231–269
3. Gould, E. A., de Lamballerie, X., Zanotto, P. M., and Holmes, E. C. (2001) *Adv. Virus Res.* **57,** 71–103
4. Gubler, D. J. (1998) *Clin. Microbiol. Rev.* **11,** 480–496
5. Halstead, S. B., Lan, N. T., Myint, T. T., Shwe, T. N., Nisalak, A., Kalyanarooj, S., Nimmannitya, S., Soegijanto, S., Vaughn, D. W., and Endy, T. P. (2002) *Emerg. Infect. Dis.* **8,** 1474–1479
6. Monath, T. P. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91,** 2395–2400
7. Irie, K., Mohan, P. M., Sasaguri, Y., Putnak, R., and Padmanabhan, R. (1989) *Gene* **75,** 197–211
8. Lindenbach, B. D., and Rice, C. M. (2003) *Adv. Virus Res.* **59,** 23–61
9. Men, R., Bray, M., Clark, D., Chanock, R. M., and Lai, C. J. (1996) *J. Virol.* **70,** 3930–3937
10. You, S., and Padmanabhan, R. (1999) *J. Biol. Chem.* **274,** 33714–33722
11. Khromykh, A. A., Meka, H., Guyatt, K. J., and Westaway, E. G. (2001) *J. Virol.* **75,** 6719–6728
12. Lo, M. K., Tilgner, M., Bernard, K. A., and Shi, P. Y. (2003) *J. Virol.* **77,** 10004–10014
13. Alvarez, D. E., Lodeiro, M. F., Ludueña, S. J., Pietrasanta, L. I., and Gamarnik, A. V. (2005) *J. Virol.* **79,** 6631–6643
14. Alvarez, D. E., De Lella Ezcurra, A. L., Fucito, S., and Gamarnik, A. V. (2005) *Virology* **339,** 200–212
15. Chiu, W. W., Kinney, R. M., and Dreher, T. W. (2005) *J. Virol.* **79,** 8303–8315
16. Filomatori, C. V., Lodeiro, M. F., Alvarez, D. E., Samsa, M. M., Pietrasanta, L., and Gamarnik, A. V. (2006) *Genes Dev.* **20,** 2238–2249
17. Alvarez, D. E., Filomatori, C. V., and Gamarnik, A. V. (2008) *Virology* **375,** 223–235
18. Holden, K. L., and Harris, E. (2004) *Virology* **329,** 119–133
19. Mukhopadhyay, S., Kuhn, R. J., and Rossmann, M. G. (2005) *Nat. Rev. Microbiol.* **3,** 13–22
20. Perera, R., and Kuhn, R. J. (2008) *Curr. Opin. Microbiol.* **11,** 369–377
21. Kapoor, M., Zhang, L., Ramachandra, M., Kusukawa, J., Ebner, K. E., and Padmanabhan, R. (1995) *J. Biol. Chem.* **270,** 19100–19106
22. Cui, T., Sugrue, R. J., Xu, Q., Lee, A. K., Chan, Y. C., and Fu, J. (1998) *Virology* **246,** 409–417
23. Yon, C., Teramoto, T., Mueller, N., Phelan, J., Ganesh, V. K., Murthy, K. H., and Padmanabhan, R. (2005) *J. Biol. Chem.* **280,** 27412–27419
24. Westaway, E. G., Mackenzie, J. M., and Khromykh, A. A. (2002) *Curr. Top. Microbiol. Immunol.* **267,** 323–351
25. Bartenschlager, R., and Miller, S. (2008) *Future Microbiol.* **3,** 155–165
26. Brinton, M. A., and Dispoto, J. H. (1988) *Virology* **162,** 290–299
27. Cahour, A., Pletnev, A., Vazielle-Falcoz, M., Rosen, L., and Lai, C. J. (1995) *Virology* **207,** 68–76
28. Lodeiro, M. F., Filomatori, C. V., and Gamarnik, A. V. (2009) *J. Virol.* **83,** 993–1008
29. Clyde, K., and Harris, E. (2006) *J. Virol.* **80,** 2170–2182
30. Clyde, K., Barrera, J., and Harris, E. (2008) *Virology* **379,** 314–323
31. Friebe, P., Shi, P. Y., and Harris, E. (2011) *J. Virol.* **85,** 1900–1905
32. Friebe, P., and Harris, E. (2010) *J. Virol.* **84,** 6103–6118
33. Zeng, L., Falgout, B., and Markoff, L. (1998) *J. Virol.* **72,** 7510–7522
34. Brinton, M. A., Fernandez, A. V., and Dispoto, J. H. (1986) *Virology* **153,** 113–121
35. Mohan, P. M., and Padmanabhan, R. (1991) *Gene* **108,** 185–191
36. Elghonemy, S., Davis, W. G., and Brinton, M. A. (2005) *Virology* **331,** 238–246
37. Holden, K. L., Stein, D. A., Pierson, T. C., Ahmed, A. A., Clyde, K., Iversen, P. L., and Harris, E. (2006) *Virology* **344,** 439–452
38. Markoff, L. (2003) *Adv. Virus Res.* **59,** 177–228
39. Yu, L., and Markoff, L. (2005) *J. Virol.* **79,** 2309–2324
40. Hahn, C. S., Hahn, Y. S., Rice, C. M., Lee, E., Dalgarno, L., Strauss, E. G., and Strauss, J. H. (1987) *J. Mol. Biol.* **198,** 33–41
41. You, S., Falgout, B., Markoff, L., and Padmanabhan, R. (2001) *J. Biol. Chem.* **276,** 15581–15591
42. Corver, J., Lenches, E., Smith, K., Robison, R. A., Sando, T., Strauss, E. G., and Strauss, J. H. (2003) *J. Virol.* **77,** 2265–2270
43. Suzuki, R., Fayzulin, R., Frolov, I., and Mason, P. W. (2008) *J. Virol.* **82,** 6942–6951
44. Zhang, B., Dong, H., Stein, D. A., Iversen, P. L., and Shi, P. Y. (2008) *Virology* **373,** 1–13
45. Olsthoorn, R. C., and Bol, J. F. (2001) *RNA* **7,** 1370–1377
46. Proutski, V., Gould, E. A., and Holmes, E. C. (1997) *Nucleic Acids Res.* **25,** 1194–1202
47. Shapiro, B. A., Bengali, D., Kasprzak, W., and Wu, J. C. (2001) *J. Mol. Biol.* **312,** 27–44
48. Shapiro, B. A., and Navetta, J. (1994) *J. Supercomput.* **8,** 195–207
49. Shapiro, B. A., and Wu, J. C. (1996) *Comput. Appl. Biosci.* **12,** 171–180
50. Wu, J. C., and Shapiro, B. A. (1999) *J. Biomol. Struct. Dyn.* **17,** 581–595
51. Kasprzak, W., and Shapiro, B. (1999) *Bioinformatics* **15,** 16–31
52. Agis-Juárez, R. A., Galván, I., Medina, F., Daikoku, T., Padmanabhan, R., Ludert, J. E., and del Angel, R. M. (2009) *J. Gen. Virol.* **90,** 2893–2901
53. Polo, S., Ketner, G., Levis, R., and Falgout, B. (1997) *J. Virol.* **71,** 5366–5374
54. Nomaguchi, M., Ackermann, M., Yon, C., You, S., and Padmanabhan, R. (2003) *J. Virol.* **77,** 8831–8842
55. Abramoff, M. D., Magelhaes, P., and Ram, S. J. (2004) *Biophotonics Int.* **11,** 36–42
56. Shapiro, B. A., and Kasprzak, W. (1996) *J. Mol. Graph.* **14,** 194–205
57. Shapiro, B. A., Wu, J. C., Bengali, D., and Potts, M. J. (2001) *Bioinformatics* **17,** 137–148
58. Edgil, D., Polacek, C., and Harris, E. (2006) *J. Virol.* **80,** 2976–2986

59. Ackermann, M., and Padmanabhan, R. (2001) *J. Biol. Chem.* **276,** 39926–39937
60. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) *J. Mol. Biol.* **288,** 911–940
61. Gee, A. H., Kasprzak, W., and Shapiro, B. A. (2006) *J. Biomol. Struct. Dyn.* **23,** 417–428
62. Kasprzak, W., Bindewald, E., and Shapiro, B. A. (2005) *Nucleic Acids Res.* **33,** 7151–7163
63. Linnstaedt, S. D., Kasprzak, W. K., Shapiro, B. A., and Casey, J. L. (2006) *RNA* **12,** 1521–1533
64. Zhang, J., Zhang, G., Guo, R., Shapiro, B. A., and Simon, A. E. (2006) *J. Virol.* **80,** 9181–9191
65. Stupina, V. A., Meskauskas, A., McCormack, J. C., Yingling, Y. G., Shapiro, B. A., Dinman, J. D., and Simon, A. E. (2008) *RNA* **14,** 2379–2393
66. Shapiro, B. A., Kasprzak, W., Grunewald, C., and Aman, J. (2006) *J. Mol. Graph. Model.* **25,** 514–531
67. Shurtleff, A. C., Beasley, D. W., Chen, J. J., Ni, H., Suderman, M. T., Wang, H., Xu, R., Wang, E., Weaver, S. C., Watts, D. M., Russell, K. L., and Barrett, A. D. (2001) *Virology* **281,** 75–87
68. Yu, L., Nomaguchi, M., Padmanabhan, R., and Markoff, L. (2008) *Virology* **374,** 170–185
69. Khromykh, A. A., Kenney, M. T., and Westaway, E. G. (1998) *J. Virol.* **72,** 7270–7279
70. Ribas, J. C., and Wickner, R. B. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89,** 2185–2189
71. Haghighat, A., Mader, S., Pause, A., and Sonenberg, N. (1995) *EMBO J.* **14,** 5701–5709
72. Tortorici, M. A., Shapiro, B. A., and Patton, J. T. (2006) *RNA* **12,** 133–146
73. Polacek, C., Foley, J. E., and Harris, E. (2009) *J. Virol.* **83,** 1161–1166
74. Filomatori, C. V., Iglesias, N. G., Villordo, S. M., Alvarez, D. E., and Gamarnik, A. V. (2011) *J. Biol. Chem.* **286,** 6929–6939
75. Dong, H., Zhang, B., and Shi, P. Y. (2008) *Virology* **381,** 123–135
76. Gritsun, T. S., and Gould, E. A. (2006) *J. Gen. Virol.* **87,** 3297–3305
77. Zuker, M. (2003) *Nucleic Acids Res.* **31,** 3406–3415
78. Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R., and Hofacker, I. L. (2008) *Nucleic Acids Res.* **36,** W70–W74
79. Romero, T. A., Tumban, E., Jun, J., Lott, W. B., and Hanley, K. A. (2006) *J. Gen. Virol.* **87,** 3291–3296
80. Watts, J. M., Dang, K. K., Gorelick, R. J., Leonard, C. W., Bess, J. W., Jr., Swanstrom, R., Burch, C. L., and Weeks, K. M. (2009) *Nature* **460,** 711–716
81. Blackwell, J. L., and Brinton, M. A. (1995) *J. Virol.* **69,** 5650–5658
82. Blackwell, J. L., and Brinton, M. A. (1997) *J. Virol.* **71,** 6433–6444
83. De Nova-Ocampo, M., Villegas-Sepúlveda, N., and del Angel, R. M. (2002) *Virology* **295,** 337–347
84. Anwar, A., Leong, K. M., Ng, M. L., Chu, J. J., and Garcia-Blanco, M. A. (2009) *J. Biol. Chem.* **284,** 17021–17029
85. Li, W., Li, Y., Kedersha, N., Anderson, P., Emara, M., Swiderek, K. M., Moreno, G. T., and Brinton, M. A. (2002) *J. Virol.* **76,** 11989–12000
86. Paranjape, S. M., and Harris, E. (2007) *J. Biol. Chem.* **282,** 30497–30508
87. Polacek, C., Friebe, P., and Harris, E. (2009) *J. Gen. Virol.* **90,** 687–692
88. Miller, W. A., Wang, Z., and Treder, K. (2007) *Biochem. Soc. Trans.* **35,** 1629–1633