JOURNAL OF VIROLOGY, Sept. 2009, p. 8327–8339 Vol. 83, No. 17
0022-538X/09/$08.00+0 doi:10.1128/JVI.00586-09

Structural and Functional Elements of the Promoter Encoded by the 5′ Untranslated Region of the Venezuelan Equine Encephalitis Virus Genome▽

Raghavendran Kulasegaran-Shylini,[2] Svetlana Atasheva,[3] David G. Gorenstein,[2] and Ilya Frolov[1]*

Department of Microbiology and Immunology, University of Texas Medical Branch, Galveston, Texas 77555-1019[1]; Department of Biochemistry and Molecular Biology and the Sealy Center for Structural Biology and Molecular Biophysics, University of Texas Medical Branch, Galveston, Texas 77555-1157[2]; and Department of Microbiology, University of Alabama, Birmingham, Alabama 35294-2170[3]

Received 22 March 2009/Accepted 3 June 2009

Venezuelan equine encephalitis virus (VEEV) is one of the most pathogenic members of the *Alphavirus* genus in the *Togaviridae* family. The pathogenesis of this virus depends strongly on the sequences of the structural proteins and on the mutations in the RNA promoter encoded by the 5′ untranslated region (5′UTR) of the viral genome. In this study, we performed a detailed investigation of the structural and functional elements of the 5′-terminal promoter and analyzed the effect of multiple mutations introduced into the VEEV 5′UTR on virus and RNA replication. The results of this study demonstrate that RNA replication is determined by two synergistically functioning RNA elements. One of them is a very 5′-terminal AU dinucleotide, which is not involved in the stable RNA secondary structure, and the second is a short, G-C-rich RNA stem. An increase or decrease in the stem's stability has deleterious effects on virus and RNA replication. In response to mutations in these RNA elements, VEEV replicative machinery was capable of developing new, compensatory sequences in the 5′UTR either containing 5′-terminal AUG or AU repeats or leading to the formation of new, heterologous stem-loops. Analysis of the numerous compensatory mutations suggested that at least two different mechanisms are involved in their generation. Some of the modifications introduced into the 5′ terminus of the viral genome led to an accumulation of the mutations in the VEEV nsPs, which suggested to us that there is a direct involvement of these proteins in promoter recognition. Furthermore, our data provide new evidence that the 3′ terminus of the negative-strand viral genome in the double-stranded RNA replicative intermediate is represented by a single-stranded RNA. Both the overall folding and the sequence determine its efficient function as a promoter for VEEV positive-strand RNA genome synthesis.

Alphaviruses are a group of important human and animal pathogens. They are widely distributed both in the New and the Old Worlds and circulate between mosquito vectors and vertebrate hosts (45). In mosquitoes, they cause a persistent, life-long infection characterized by virus accumulation in salivary glands, which is required for infecting vertebrate hosts during a blood meal (50). In vertebrates, alphaviruses develop high-titer viremia, and their replication induces a variety of diseases with symptoms depending on both the host and the causative virus (11). Venezuelan equine encephalitis virus (VEEV), the New World alphavirus, is one of the most pathogenic members of the genus (16, 45). Representatives of the VEEV serocomplex circulate in Central, South, and North America and cause severe, and sometimes fatal, encephalitis in humans and horses (3, 11, 16, 24). Accordingly, VEEV represents a serious public health threat in the United States (39, 48, 51, 53), and during VEEV epizootics, equine mortality can reach 83%, and in humans, neurological diseases can be detected in up to 14% of all infected individuals, especially children (15). The overall mortality rate for humans is below 1%, but it is usually higher among children, the elderly, and, most likely, immunocompromised individuals (49). In spite of the continuous threat of VEEV epidemics, the biology of this virus, its pathogenesis, and the mechanism of replication are insufficiently understood. To date, no safe and efficient vaccine and therapeutic means have been developed for this pathogen.

The VEEV genome is represented by a single-stranded, almost 11.5-kb-long RNA molecule of positive polarity. This RNA mimics the structure of cellular mRNAs by containing a cap at the 5′ ends and a poly(A) tail at the 3′ ends of the genome (18). The genomic RNA encodes two polyproteins: the 5′-terminal open reading frame (ORF) is translated into viral nonstructural proteins (nsP1 to nsP4), forming the replication enzyme complex (RC). The second ORF corresponds to the 3′-terminal one-third of the genome and encodes all of the viral structural proteins, C, E2, and E1. The latter proteins are translated from the subgenomic RNA synthesized during virus replication (45).

The replication of the alphavirus genome is a highly regulated, multistep process, which includes the synthesis of three different RNA species (45). The regulation of their synthesis is achieved by differential processing of viral nsPs (22, 23, 43). First, the initially synthesized nonstructural polyprotein is partially processed by the nsP2-associated protease into P123 and nsP4, and this complex is active in negative-strand RNA synthesis (22). The latter RNA is present in the double-stranded RNA (dsRNA) replicative intermediate and is associated with

* Corresponding author. Present address: Department of Microbiology, BBRB 373/Box 3, University of Alabama, 1530 Third Avenue South, Birmingham, AL 35294-2170. Phone: (205) 996-8957. Fax: (205) 996-4008. E-mail: ivfrolov@UAB.edu.

▽ Published ahead of print on 10 June 2009.

plasma membrane and endosome-like vesicular organelles (8). Further processing of the polyproteins into individual nsP1 to nsP4 makes the RC capable of the synthesis of the positive-strand genome and subgenomic RNA but not of negative-strand RNA (23, 41, 42). Thus, the completely processed nsPs utilize only the promoters located on the negative strand of the viral genome.

The defined promoters in the alphavirus genomes include (i) a 3′-terminal 19-nucleotide (nt)-long, conserved sequence element (CSE) adjacent to the poly(A) tail (12, 13, 19); (ii) the subgenomic promoter in the negative-strand copy of the viral genome (25); and (iii) the promoter for the synthesis of the positive-strand viral genome (45). The latter promoter is located at the 3′ end of the negative strand of the viral genome and has a complex structure. The two identified elements include the sequence, encoded by the 5′ untranslated region (5′UTR) (a core promoter) (5, 9, 32), and a 51-nt CSE, found ~150 nt downstream of the genome's 5′ terminus in the nsP1-encoding sequence. Our previous results and those of other research groups demonstrated that the 51-nt CSE functions as a replication enhancer in a virus- and cell-dependent mode (4, 33). Clustered mutations in the VEEV 51-nt CSE or its complete deletion either had deleterious effects on RNA replication or completely abolished RNA synthesis (30). However, RNA replication was ultimately recovered due to an accumulation of compensatory, adaptive mutations in either VEEV nsP2 or nsP3 (30). Thus, the 51-nt CSE in the VEEV genome is not absolutely essential for virus replication, but its presence is highly beneficial for achieving the most efficient growth rates in cells of both vertebrate and invertebrate origins. Alphavirus core promoters demonstrate a very low level of sequence conservation and also function in cell- and virus-specific modes (9). Previous studies suggested that the sequence and/or secondary structure of the VEEV core promoter plays a critical role in virus pathogenesis, and the G3→A (A3) mutation, found in an attenuated strain of VEEV TC-83, is one of the determinants of its less pathogenic phenotype (17, 55). However, information about functional elements of the VEEV core promoter remains incomplete, and its structural and functional elements have not yet been dissected.

In this study, we applied a combination of molecular approaches to further define the functional components of the VEEV 5′UTR-specific core promoter, which mediates positive-strand genome synthesis. Our results demonstrate the presence of three structural RNA elements, two of which synergistically determine promoter activity. The first element of the promoter is a very short, 5′-terminal sequence, which is not involved in a stable secondary structure. Point mutations in the very 5′-terminal nucleotides have a deleterious effect on genome RNA replication. The second element is the short RNA stem, located in close proximity to the 5′ end of the genome. Mutations changing either the stability or sequence of the stem strongly affect virus replication and cause its rapid evolution, leading to the appearance of heterologous repeating elements in the unpaired 5′ terminus or the generation of other sequences that might potentially fold into stem structures. Surprisingly, the third structural RNA element, the loop, appears to play no important role in RNA replication and can be replaced either by a shorter loop or by the loop having a heterologous sequence without a detectable effect on virus and RNA replication.

MATERIALS AND METHODS

Cell cultures. BHK-21 cells were kindly provided by Sondra Schlesinger (Washington University, St. Louis, MO) and were propagated in alpha-minimal essential medium (MEM) supplemented with 10% fetal bovine serum (FBS) and vitamins.

Plasmid constructs. p(G3)/VEE/SINV and p(A3)/VEE/SINV, encoding the genomes of (G3)VEE/SINV and (A3)VEE/SINV chimeric virus genomes, were described elsewhere (20). Briefly, p(G3)/VEE/SINV encoded the 5′UTR of the VEEV Trinidad donkey (TRD) genome (18), the 3′UTR, the subgenomic promoter, and the nonstructural polyprotein-encoding sequence from the genome of VEEV TC-83 and the structural polyprotein-encoding sequence, derived from Sindbis virus (SINV) Toto1101 (38). p(A3)/VEE/SINV differed by one nucleotide (A3) only, which was specific for the VEEV TC-83 5′UTR. In both cDNA clones, the poly(A) sequence was followed by a MluI restriction site. Other 5′UTR mutants described in the text had essentially the same design but differed in the 5′-terminal sequences of the genome. Sequences of the 5′ termini of their genomes are presented in the corresponding figures.

pUbi/Luc and its derivatives encoded the promoter for the SP6 DNA-dependent RNA polymerase, followed by tested 5′UTRs, a VEEV genome fragment encoding 49 amino acids of nsP1 fused with the entire ubiquitin gene, and a firefly luciferase-encoding sequence. This ORF was followed by the VEEV TC-83-specific 3′UTR, poly(A), and an MluI restriction site, required for the linearization of the plasmid before an in vitro transcription reaction.

Plasmids pDI/Luc and pA3/DI/Luc encoded defective viral genomes under the control of the SP6 promoter. These genomes contained tested 5′UTRs followed by nt 45 to 519 derived from the VEEV TC-83 genome, nt 7291 to 7564 (encoding the VEEV subgenomic promoter and the 5′UTR of the subgenomic RNA), a firefly luciferase gene, 3′-terminal nt 11202 to 11446 of the VEEV genome, a poly(A) tail, and an MluI restriction site. The p(G3)/VEErep/Pac-encoded VEEV replicon, in which viral structural genes were replaced by a puromycin acetyltransferase (Pac) sequence, was described elsewhere previously (37).

All of the plasmids encoding modified VEEV genomes and luciferase were constructed by standard PCR-based mutagenesis and cloning methods. All of the cloned PCR fragments were sequenced to exclude the possibility of spontaneous mutations. The details of the cloning procedures and sequences will be provided upon request.

RNA transcripts. Plasmids were purified by centrifugation in CsCl gradients. Prior to transcription, the VEE/SINV genome-coding plasmids were linearized by MluI, and RNAs were synthesized by SP6 RNA polymerase in the presence of a cap analog (38) under the conditions recommended by the manufacturer (Invitrogen). The yield and integrity of the transcripts were monitored by gel electrophoresis under nondenaturing conditions, followed by analysis of the RNA concentration with a FluorChem imager (Alpha Innotech). For virus rescue and analysis of defective interfering (DI) RNA replication, the appropriate volumes of reaction mixtures were used directly for electroporation. For comparative studies of the translation efficiencies of the luciferase-encoding RNAs, the transcripts were additionally purified using RNeasy columns (Qiagen), and the RNA concentration was measured as described above.

RNA transfections. BHK-21 cells were electroporated with 2 μg of in vitro-synthesized RNA transcripts under previously described conditions (29). For analysis of DI RNA replication, 2 μg of replicon and DI RNAs were coelectroporated under the same conditions. Equal amounts of cells were seeded into 35-mm dishes, and luciferase activity was measured at different times posttransfection by using a luciferase assay kit according to the manufacturer's instructions (Promega). To assess the translation efficiency of the templates, equal amounts of Ubi/Luc RNAs were electroporated into BHK-21 cells, and equal aliquots of the cells were seeded into 35-mm dishes. At the indicated time points, cells were lysed, and luciferase activity was measured by using a Dual-Luciferase assay kit according to the manufacturer's instructions (Promega).

Infectious center assay. In standard experiments, 2 μg of in vitro-synthesized viral genome RNA was transfected into BHK-21 cells using previously described conditions (29). Tenfold dilutions of electroporated cells were seeded into six-well Costar plates containing naïve BHK-21 cells. After a 1-h incubation at 37°C, the cells were overlaid with 2 ml of 0.5% agarose (Invitrogen) containing MEM supplemented with 3% FBS. Plaques were stained with crystal violet after 2 days of incubation at 37°C. Before staining, some of the plaques were randomly isolated for analysis of the pseudorevertants. The remaining electroporated cells were seeded into 100-mm tissue culture dishes for generating viral stocks or into

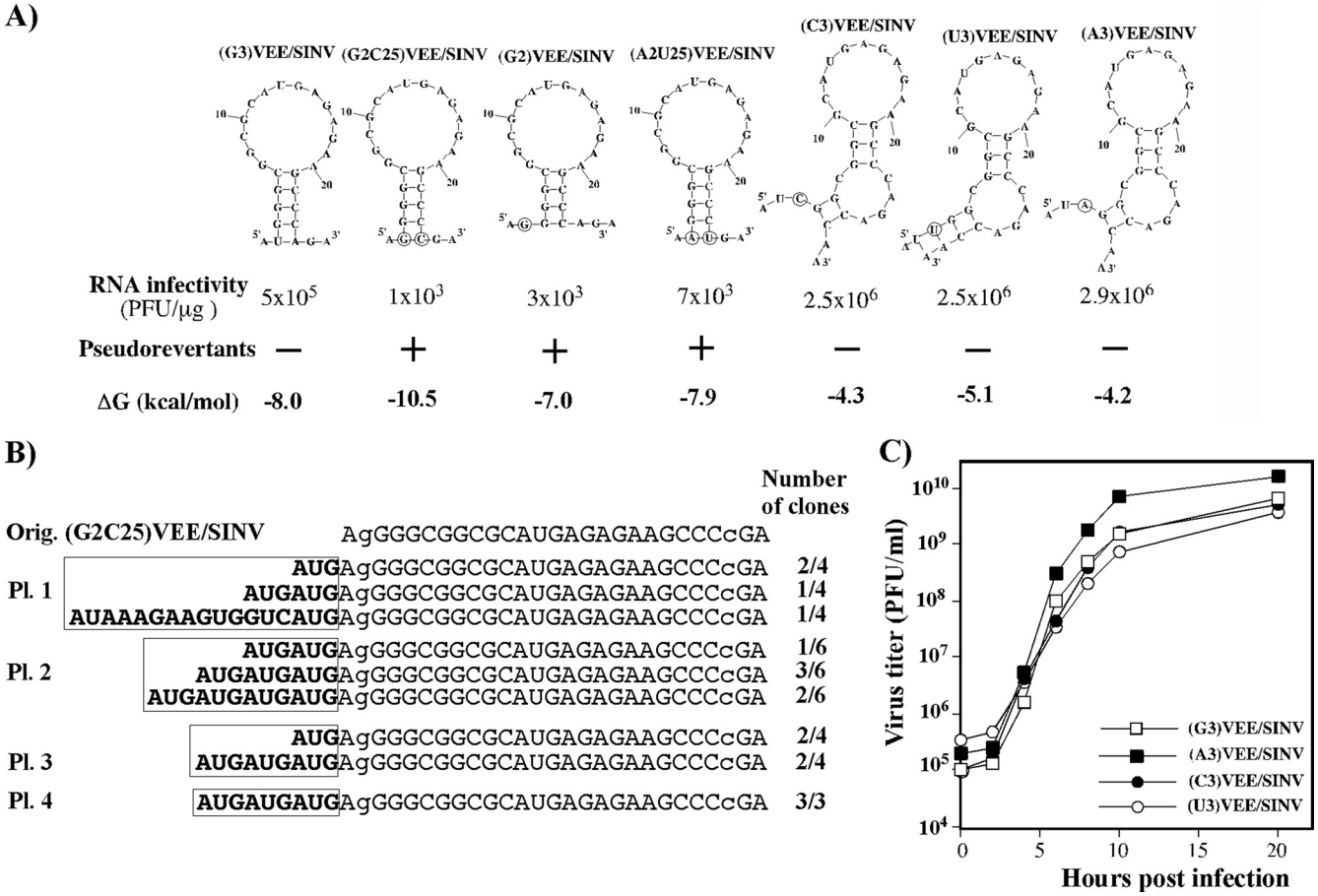

FIG. 1. Analysis of the effects of nt-2- and nt-3-specific mutations on VEE/SINV viability and replication. (A) Computer-predicted (M-fold) (14) secondary structures of the 5′ termini in the genomes of VEE/SINV variants having mutations in nt 2 and nt 3, RNA infectivity in the infectious center assay, and calculated overall free energies of the 5′ termini. Open circles indicate positions of the mutations. (B) Nucleotide sequences of the heterologous 5′ ends in the genomes of (G2C25)VEE/SINV pseudorevertants isolated from the plaques, which were randomly selected in the infectious center assay. Heterologous sequences are indicated by boldface type. Lowercase type indicates mutations introduced into (G3)VEE/SINV to generate (G2C25)VEE/SINV. The number of cDNA clones encoding a particular sequence is indicated. Orig., origin. (C) BHK-21 cells were infected with the indicated (G3)VEE/SINV, (C3)VEE/SINV, (A3)VEE/SINV, and (U3)VEE/SINV variants at an MOI of 10 PFU/cell. The media were harvested and replaced at the indicated time points. Released virus titers were measured by plaque assay on BHK-21 cells.

35-mm dishes to evaluate the rates of virus replication. At the time points indicated in the corresponding figures, media were replaced, and virus titers in the corresponding samples were assessed by a plaque assay on BHK-21 cells (21).

Analysis of virus replication. BHK-21 cells were seeded at a concentration of 5×10^5 cells per 35-mm dish, respectively. Monolayers were infected at a multiplicity of infection of 10 PFU/cell for 1 h, washed with phosphate-buffered saline, and overlaid with 1 ml of complete medium. At the indicated times, media were replaced by fresh media, and virus titers in the harvested samples were determined by a plaque assay on BHK-21 cells as described elsewhere previously (21).

Selection of pseudorevertants and sequencing of viral genomes. In order to identify adaptive mutations accumulating in viral genomes in response to mutations in the 5′UTR, plaques were isolated from an agarose overlay directly in the infectious center assay. Viruses were then eluted into 1 ml of alpha-MEM supplemented with 10% FBS, and 0.5 ml of this medium was used to infect naïve BHK-21 cells in 35-mm dishes. After 16 to 24 h of incubation at 37°C in 5% CO_2, medium was harvested, and RNA was isolated from the cells by using TRIzol reagent according to the manufacturer's instructions (Invitrogen). The isolated RNAs were used to generate cDNA fragments for the 5′ terminus by utilizing the commercially available FirstChoice RLM-Race kit (Ambion). The amplified DNA fragments were purified by agarose gel electrophoresis and cloned into plasmid pRS2. Plasmids isolated from multiple colonies were used for sequencing. For some of the viruses, PCR fragments covering the entire nonstructural polyprotein-encoding region were synthesized, purified by agarose gel electrophoresis, and also sequenced.

Analysis of viral RNA synthesis. BHK-21 cells were infected with chimeric viruses at a multiplicity of infection (MOI) of 10 PFU/cell. At the times indicated in the figure legends, viral RNAs were labeled with [^{3}H]uridine (20 mCi/ml) in the presence of 1 μg of dactinomycin/ml for 4 h at 37°C in 5% CO_2. RNAs were isolated from the cells by using TRIzol reagent according to the manufacturer's instructions (Invitrogen). The RNAs were denatured with glyoxal in dimethyl sulfoxide and analyzed by agarose gel electrophoresis under previously described conditions (2). For quantitative analysis, the RNA bands were excised from the 2,5-diphenyloxazole-impregnated gels, and radioactivity was measured by liquid scintillation counting.

RESULTS

Critical role of the 5′-terminal nucleotides in VEEV replication. In our previous studies (20, 30), the 5′-terminal sequences of the VEEV TRD and TC-83 genomes were predicted to fold into different stem-loop structures [(G3)VEE/SINV and (A3)VEE/SINV] (Fig. 1A). Data from enzymatic

and nuclear magnetic resonance analyses were in agreement with computer models (20). The computer-predicted folding of the 3′ terminus of the (G3)VEE/SINV negative-strand RNA intermediate mirrors the stem-loop structure of the 5′UTR on the positive strand (20). However, the TC-83-specific 3′ terminus of the negative-strand RNA was predicted to have multiple conformations, suggesting that this is likely to be a dynamic structure (data not shown). To further dissect functional components of the core promoter in the 5′UTR of the VEEV TRD genome, we provisionally divided this fragment into three elements: (i) a very short 5′-terminal sequence that is not involved in the stable base pairing, (ii) a stable stem containing four G-C base pairs, and (iii) a loop (nt 7 to 20 of the viral genome) (Fig. 1A). In this study, we performed extensive mutational analyses aimed at identifying the function of each element in virus and RNA replication.

Previously published data unambiguously demonstrated that alphavirus structural proteins are dispensable for the replication of the viral genome and the transcription of subgenomic RNA (10, 28, 37). This study was based on the construction, selection, and analysis of numerous virus variants having modified 5′UTRs, and such modifications could make VEEV, even its attenuated variant TC-83, more pathogenic (17). Therefore, the applied experimental system was based on using the chimeric virus (G3)VEE/SINV, whose genome encoded all of the VEEV-specific nsPs and *cis*-acting elements and structural protein genes derived from SINV (7). This virus was capable of replicating to high titers but was poorly cytopathic (7, 20) and highly attenuated in mice (data not shown). As a result, the experiments did not require high-level biocontainment conditions, and this strongly promoted the research.

The geographically isolated alphaviruses that belong to different antigenic subtypes demonstrate a very low level of identity in the 5′UTRs of their genomes (45), but all of them start from the same AU dinucleotide, which is suggestive of its critical role in RNA replication. The third nucleotide in the VEEV genome is either G (in wild-type [wt] VEEV TRD) or A (in vaccine strain VEEV TC-83), which suggests to us that it might be of less importance. To test these assumptions, in the initial experiments, we assessed the functioning of the very 5′-terminal 3 nt in RNA and virus replication (Fig. 1A).

First, we evaluated the importance of nt 2 by developing three different mutants: (G2)VEE/SINV, (G2C25)VEE/SINV, and (A2U25)VEE/SINV. (G2)VEE/SINV had a single U2→G mutation in the 5′UTR (Fig. 1A). To preserve the 5′-terminal RNA secondary structure and mimic the VEEV TRD-specific folding of the 5′UTR, the (G2C25)VEE/SINV variant contained both U2→G and the compensatory A25→C mutations (Fig. 1A). The (A2U25)VEE/SINV mutant also mimicked the VEEV TRD-specific folding of the 5′ terminus, but the U2-A25 base pair was replaced by A2-U25. The infectivity of these in vitro-synthesized RNAs were almost 3 orders of magnitude lower than that of the original (G3)VEE/SINV RNA. This was an indication that plaques detected in the infectious center assay (see Materials and Methods for details) were developed by true revertants or pseudorevertants of the originally designed construct. To identify the adaptive mutations, we randomly selected four plaques of (G2C25)VEE/SINV in the infectious center assay and sequenced the 5′UTRs of the isolated variants. Direct sequencing of the PCR product suggested the presence of multiple variants in each plaque (data not shown). Therefore, the 5′UTR-containing fragments were cloned into the plasmid, and multiple clones were used for identifying the 5′ termini. The results presented in Fig. 1B demonstrate that the (G2C25)VEE/SINV genome acquired a spectrum of new sequences at the 5′ terminus, and their characteristic feature was the presence of an additional one to four AUG repeats at the very terminus. (Note that the natural VEEV TRD genome also starts from an AUG.) Only one of the variants had a heterologous sequence (plaque 1, variant 3), and its genomic RNA started from the VEEV TC-83-specific AUA. Taken together, the results strongly suggested that the U2 nucleotide plays a critical role in the replication of VEEV-specific RNA and cannot be replaced by either A or G even if the later nucleotides are involved in secondary-structure formation, like U2 in the VEEV TRD genome.

Next, we evaluated the role of following nt 3 in virus replication by replacing TRD-specific G3 by U, C, and TC-83-specific A. All of the mutated nucleotides were predicted not to be involved in the formation of stable stems (Fig. 1A) and increased the calculated overall free energy of the 5′ RNA terminus. The designed (U3)VEE/SINV and (C3)VEE/SINV were viable, and their RNAs demonstrated the same efficiency of plaque formation in the infectious center assay as did the RNAs of (G3)VEE/SINV and (A3)VEE/SINV. Moreover, sequencing of the 5′ ends of the genomes from the variants in the randomly selected plaques showed no reversions or other mutations. Thus, the mutations of nt 3 do not have a deleterious effect on virus replication. However, in contrast to A3 (20), both the U3 and C3 mutations were not beneficial for both virus (Fig. 1C) and RNA (data not shown) replication. The (A3)VEE/SINV mutant was intensively characterized in our previous study (20), where we demonstrated that an A3 mutation leads to a strong increase in levels of replication of the viral genome and downregulates the transcription of subgenomic RNA.

Thus, the initial data suggested that the 5′UTR-specific promoter in the VEEV genome might be very simple and represented by an AU. However, they certainly did not rule out a possibility that other RNA elements are critically involved in promoter function. The following nucleotide could be any one albeit with A3 functioning most efficiently in RNA replication.

Modifications of the loop do not affect virus replication. In another set of experiments, we tested the effect(s) that a modification of the loop sequence may have on VEE/SINV replication. The loop in the 5′ terminus of the (A3U24)VEE/SINV genome was replaced by the same sequence cloned in the opposite orientation (Fig. 2A). The originally designed (A3U24) VEE/SINV variant demonstrated more efficient genomic RNA synthesis and a low subgenomic-to-genomic (SG-to-G) RNA synthesis ratio compared to those found for (G3)VEE/SINV (Fig. 2B). Therefore, we expected it to be a more sensitive model for detecting a decrease in the 5′UTR-specific RNA promoter activity in response to manipulations with the loop.

The designed extensive modification of the loop did not affect the overall, computer-predicted secondary structures of both the 5′UTR (Fig. 2A) and its complement at the 3′ end of the negative-strand intermediate (data not shown). The in vitro-synthesized (LOOP)VEE/SINV RNA was as efficient in plaque formation in the infectious center assay as the control

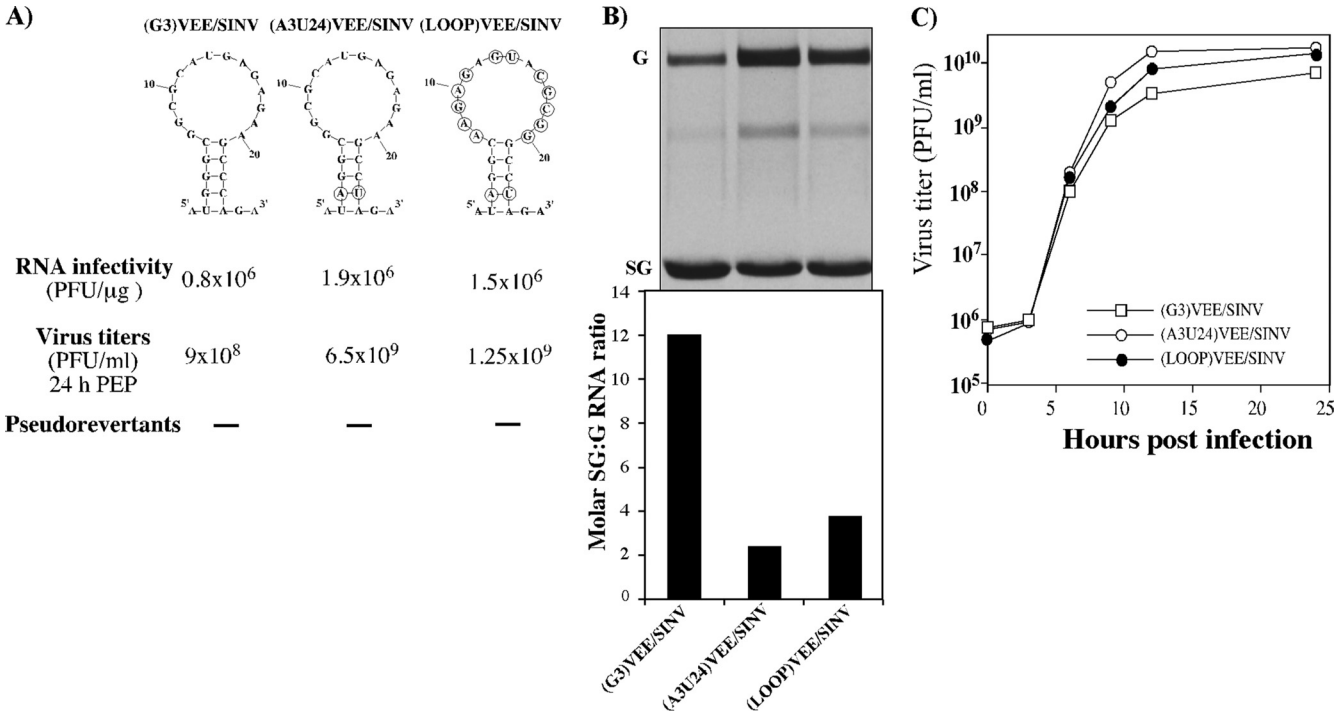

FIG. 2. Extensive modifications of the loop sequence do not affect virus replication. (A) Computer-predicted secondary structures of the 5′ termini in the genomes of VEE/SINV variants used in the experiments, infectivities of the in vitro-synthesized RNAs in the infectious center assay, and virus titers at 24 h post-RNA transfection. Open circles indicate positions of the mutations. PEP, postelectroporation. (B) Autoradiograph of the dried gels with the RNAs metabolically labeled with [^{3}H]uridine. Equal numbers of BHK-21 cells were infected with the indicated VEEV variants at an MOI of 10 PFU/cell. The RNAs were metabolically labeled with [^{3}H]uridine (20 μCi/ml) in the presence of 1 μg of dactinomycin/ml between 4 and 8 h postinfection and then isolated and analyzed by agarose gel electrophoresis under denaturing conditions as described in Materials and Methods. Bands corresponding to genomic and subgenomic RNAs were visualized by autoradiography. G and SG indicate positions of the viral genomic and subgenomic RNAs, respectively. RNA bands were excised, radioactivity in the genomic and subgenomic RNAs was measured by liquid scintillation counting, and the molar SG-to-G ratios are presented. (C) BHK-21 cells were infected with the indicated variants at an MOI of 10 PFU/cell. Virus replication was assessed as described in Materials and Methods.

RNA of (A3U24)VEE/SINV. Growth rates of the harvested viruses were also merely identical. We did not detect any effect of the loop replacement on RNA replication (Fig. 2B) and the SG-to-G ratio of RNA synthesis compared to (A3U24)VEE/SINV. Thus, taken together, the results indicated that a radical change in the loop sequence had no significant effect on virus and viral RNA replication. Moreover, the experiments with other efficiently replicating mutants, which had smaller loops (see Fig. 4), suggested that loop size is not likely to be critical for promoter activity. Therefore, further experiments on the modification of loop sequences were discontinued.

Higher stability of the 5′-terminal stem has a deleterious effect on virus replication. The VEEV TRD genome has a short, G-C-rich RNA stem located almost at the very 5′ terminus of the viral genome (Fig. 3A). Such 5′-terminal stem-loop structures are generally believed to be disadvantageous for RNA translation; therefore, it was reasonable to expect that its presence could have a negative effect on the synthesis of viral nsPs. On the other hand, the same structure could be essential for promoter function and promote RNA replication. To test the role of the indicated stem in VEEV replication, we designed a set of mutants having a higher or lower (see below) stability of the predicted, 5′UTR-specific stems. To achieve the higher stability, the loop sequence in the TRD-specific 5′UTR of the (G3)VEE/SINV genome was modified to promote base pairing (Fig. 3A). The M-fold-calculated minimum free energies of the 5′ ends in the (C19C20)VEE/SINV and (C17C19C20)VEE/SINV genomes were −18.4 and −23.5 kcal/mol, respectively, and, thus, lower than that calculated for the (G3)VEE/SINV-specific stem-loop (−8 kcal/mol). These modifications had a deleterious effect on virus replication. Compared to RNAs having either TC-83- or TRD-specific 5′UTRs, the in vitro-synthesized mutant RNAs demonstrated dramatically lower infectivities in the infectious center assay (Fig. 3A), and the magnitude of reduction correlated with the increase in predicted stem stability. An infectivity that was a few orders of magnitude lower was indicative of an accumulation of adaptive mutations leading to a more efficient replication in tissue culture and ultimately making them capable of causing cytopathic effects and plaque formation. The genomic 5′ ends of the variants, isolated from the randomly selected plaques, demonstrated that they had acquired three to six additional AUG repeats at the very 5′ termini. Of note, as in the above-described experiments with the (G2C25)VEE/SINV mutant, some heterogeneity in the 5′ termini was detected even in the plaque-purified variants. This finding was an indication of the continuing evolution and presence of multiple quasispecies in plaque isolates. One of the plaque-purified variants [see plaque 1 in (C19C20)VEE/SINV] differed by having a shorter version of the 5′UTR in which a 5′-terminal part of the stem

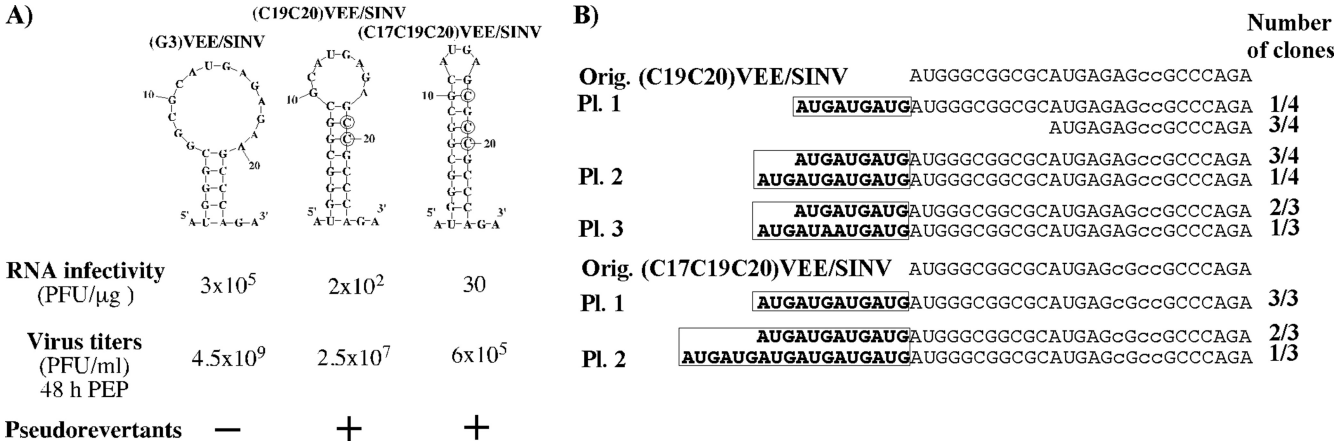

FIG. 3. Analysis of the effects of the mutations stabilizing the 5′-terminal stem of the VEE/SINV genome on virus viability and replication. (A) Computer-predicted secondary structures of the 5′ termini in the genomes of VEE/SINV variants having the mutations that increased the stability of the RNA stem. Open circles indicate positions of the mutations. RNA transfections were performed as described in Materials and Methods. RNA infectivities in the infectious center assay and titers of the viruses at 48 h posttransfection are indicated. PEP, postelectroporation. (B) Nucleotide sequences of the heterologous 5′ ends in the genomes of (C19C20)VEE/SINV and (C17C19C20)VEE/SINV pseudorevertants present in the plaques, which were randomly selected in the infectious center assay. Heterologous sequences are indicated by boldface type. Lowercase type indicates mutations introduced into (G3)VEE/SINV to generate (C19C20)VEE/SINV and (C17C19C20)VEE/SINV. The number of cDNA clones encoding a particular sequence is indicated. Orig., origin.

was deleted. Nevertheless, its genome still started from the AUG, signifying the importance of this nucleotide combination in RNA replication.

Our finding of the same AUG repeats in response to both stem-loop modification and the mutation of nt 2 led us to suggest that this might be a universal path in VEEV evolution in response to the 5′UTR modification. However, we did not sequence the entire genomes of the plaque-purified viruses, and thus, we could not rule out the possibility that other, additional mutations were also required for their viability. Therefore, in order to additionally test the importance of the identified repeats, we modified the (C19C20)VEE/SINV genome by adding one, two, or three AUGs (Fig. 4A) and analyzed their effects on virus replication (Fig. 4B). In the infectious center assay, the in vitro-synthesized RNAs of the new variants demonstrated infectivities that were more than 3 or-

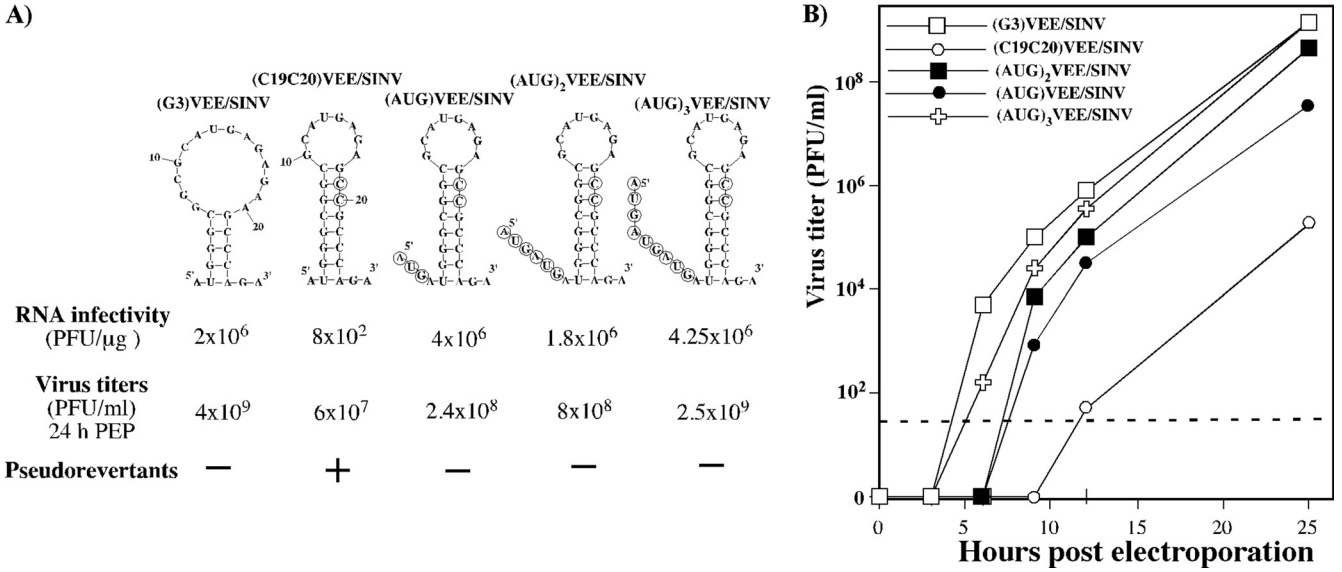

FIG. 4. Replication of (C19C20)VEE/SINV variants having additional AUG repeats at the 5′ termini of their genomes. (A) Computer-predicted secondary structures of the 5′ termini of the designed genomes, RNA infectivities in the infectious center assay, and virus titers at 24 h post-RNA transfection. PEP, postelectroporation. (B) Analysis of replication of the designed mutants. Two micrograms of the in vitro-synthesized RNAs was transfected into BHK-21 cells by electroporation, and one-fifth of the cells were seeded into 35-mm dishes. At the indicated times posttransfection, media were replaced, and titers of virus in the harvested samples were measured by a plaque assay of BHK-21 cells. Electroporation was used instead of a more traditional infection, because the designed variants most likely continued to evolve and accumulate quasispecies with higher numbers of repeats by late times posttransfection. The dashed line indicates the limit of detection.

ders of magnitude higher than those of the original (C19C20) VEE/SINV construct (Fig. 4A), and this infectivity was essentially the same as that of (G3)VEE/SINV RNA. This was an indication that the addition of at least one more AUG to the 5′ terminus of the (C19C20)VEE/SINV genome was sufficient for virus viability, and no other mutations in the nonstructural genes and/or noncoding regions were required. Furthermore, an increase in the number of AUG repeats correlated with earlier virus release from the electroporated cells and its faster replication rates (Fig. 4B). The three-AUG-containing variant was capable of replicating almost as efficiently as wt (G3)VEE/SINV. However, note that in plaque isolates of (C19C20)VEE/SINV and (C17C19C20)VEE/SINV pseudorevertants, we detected variants having four and six AUG repeats, suggesting that replication rates could likely achieve even higher levels. A more detailed characterization of the variants containing different numbers of AUG repeats was complicated by the likelihood of their further evolution, leading to the appearance of quasispecies with higher numbers of AUGs, as we described above for the plaque-purified variants.

We also engineered a variant with the above-described deletion in the 5′ terminus (Fig. 5A). The (Del)VEE/SINV mutant was viable, and the in vitro-synthesized RNA demonstrated an infectivity of 2×10^5 PFU/μg, which was close to the level demonstrated by (G3)VEE/SINV RNA. The designed construct initially replicated very inefficiently and developed pinpoint plaques, which were difficult to count. However, during the passaging that ensued, it evolved into the better-replicating virus, forming larger plaques. Sequencing of the 5′ ends of the genomes of the variants recovered from the randomly selected plaques revealed no additional changes in the 5′UTRs, but the adaptive mutations were found in the carboxy terminus of nsP1, the amino terminus of nsP2, or nsP4 (Fig. 5B). This suggested that the inefficient replication resulting from strong modifications in the promoter sequence could be compensated for not only by the evolution of the 5′UTR but also by mutations in different virus-specific nsPs. The effects of nsP1- and nsP4-specific mutations on (Del)VEE/SINV replication were additionally examined. (Del)VEE/SINV(nsP1) and (Del)VEE/SINV(nsP4) viruses were capable of forming larger, clear plaques and demonstrated growth rates that were higher than that of the original (Del)VEE/SINV construct (Fig. 5C). The effects of these mutations on wt virus replication are now under further investigation.

Lower stability of the 5′-terminal stem strongly affects replication of the chimeric viruses. The results of the experiments described above demonstrated that an increase in the stability of the 5′-terminal RNA stem had a strong negative effect on virus viability and/or rates of replication. Moreover, the experiments with the (A3)VEE/SINV and (A3U24)VEE/SINV mutants, having less stable stems, demonstrated higher rates of RNA replication (Fig. 2) (20). Therefore, one might assume that even the original, VEEV-specific stem-loop could be disadvantageous for RNA replication, and the conservation of this structure could be supported by other mechanisms required for virus replication in vivo. Therefore, to further assess the effect of stem stability on RNA and virus replication, we designed another set of mutants. The first one, termed (AUstem) VEE/SINV, contained a less stable stem consisting mainly of A-U pairs. In the second variant, (NOstem)VEE/SINV, the original stem was completely destroyed by a set of point mutations. The minimum free energies of the 5′ ends were −1.8 and −0.4 kcal/mol, respectively. Both variants were viable, and in the infectious center assay, their in vitro-synthesized RNAs demonstrated infectivities that were comparable to that of the original (G3)VEE/SINV (Fig. 6A). However, upon RNA transfection into the cells, replicating viruses were not only

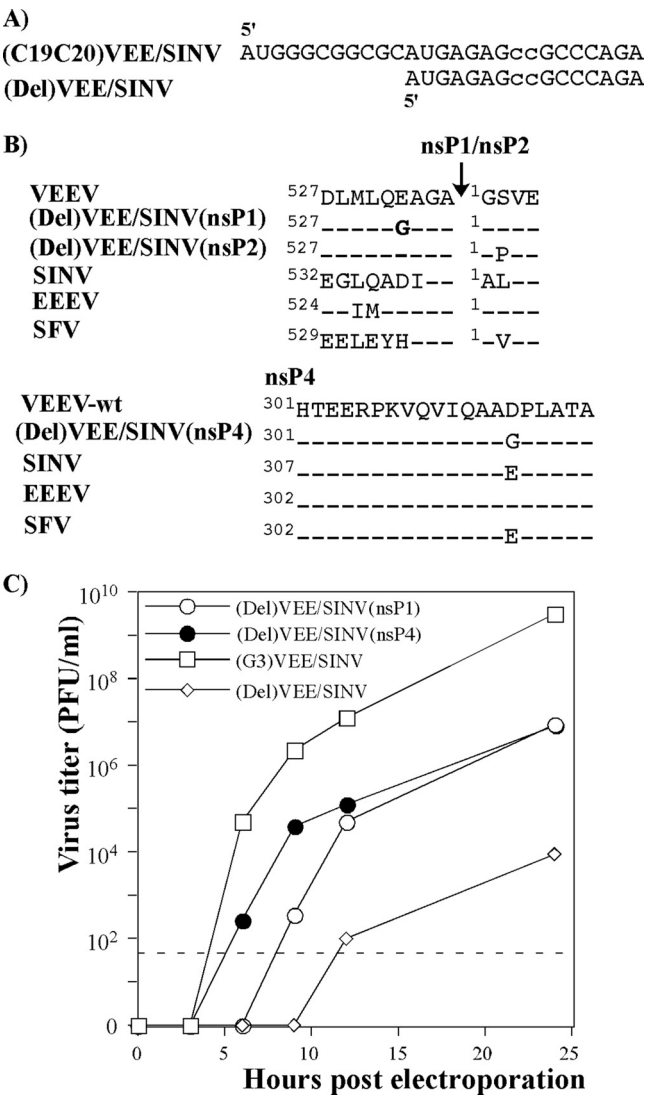

FIG. 5. Analysis of the adaptive mutations in viral nonstructural genes generated in response to a deletion in the 5′UTR. (A) Nucleotide sequences of the 5′UTR in (C19C20)VEE/SINV and (Del)VEE/SINV viral genomes. (B) Sequence alignments and mutations found in (Del)VEE/SINV plaque isolates adapted for growth in BHK-21 cells. Alignments of the nsP1/nsP2 junction and a fragment of nsP4 for several alphaviruses, including VEEV (18), eastern equine encephalitis virus (EEEV) (52), Semliki Forest virus (SFV) (46), and SINV (44). Residues identical to those in the VEEV sequence are indicated by dashes. Orig., origin. (C) Analysis of replication of the designed mutants. Two micrograms of the in vitro-synthesized RNAs were transfected into BHK-21 cells by electroporation, and one-fifth of the cells were seeded into 35-mm dishes. At the indicated times posttransfection, media were replaced, and titers of virus in the harvested samples were measured by a plaque assay using BHK-21 cells. The dashed line indicates the limit of detection.

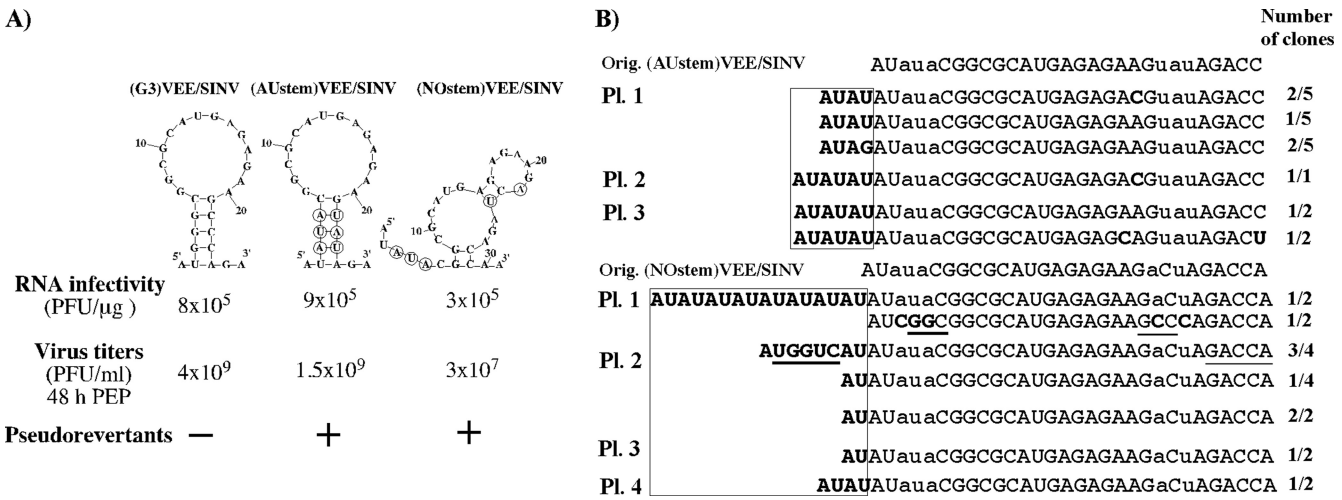

FIG. 6. Analysis of effects of mutations destabilizing the 5′-terminal stem of the VEE/SINV genome on virus viability and replication. (A) Computer-predicted secondary structures of the 5′ termini in the genomes of VEE/SINV variants having mutations that strongly decreased the stability of the stem. Open circles indicate positions of the introduced mutations. RNA transfections were performed as described in Materials and Methods. RNA infectivities in the infectious center assay and titers of the viruses at 48 h posttransfection are presented. PEP, postelectroporation. (B) Nucleotide sequences of the heterologous 5′ ends of the genomes of (AUstem)VEE/SINV and (NOstem)VEE/SINV variants, isolated from the randomly selected plaques in the infectious center assay. Heterologous 5′-terminal sequences and other mutations are indicated by boldface type. Lowercase type indicates mutations introduced into (G3)VEE/SINV to generate (AUstem)VEE/SINV and (NOstem)VEE/SINV. Sequences potentially capable of forming stems are underlined. The numbers of cDNA clones encoding particular sequences are indicated.

incapable of causing profound cytopathic effects but also unable to develop clear plaques under agarose cover in the infectious center assay. This phenotype was greatly unstable, and by 48 h posttransfection, media already contained variants that were capable of forming clear plaques, and their titers approached 1.5×10^9 and 3×10^7 PFU/ml [for (AUstem)VEE/SINV and (NOstem)VEE/SINV, respectively]. Viruses forming small, opaque plaques were no longer detectable in the samples harvested at late times post-RNA transfection. The more efficiently replicating variants, isolated from the randomly selected plaques, demonstrated modifications in the originally designed 5′UTRs. Sequencing detected two types of adaptive mutations. First, a majority of variants contained repeating, 5′-terminal AU sequences. One of the plaques isolated for (AUstem)VEE/SINV (Fig. 6B) contained an additional AUAG. The presence of G either was a result of the mutation in the AUAU repeat or, most likely, originated from the transfected, in vitro-synthesized RNA. Guanosine in the +1 position of cDNA is required for efficient functioning of the SP6 promoter during the in vitro transcription. Second, viruses recovered from two plaques of the (NOstem)VEE/SINV variant demonstrated an evolution different from the standard $(AU)_n$ generation (Fig. 6B). Their 5′ termini contained heterologous sequences, which could potentially form stem-loops, positioned similarly to that in the (G3)VEE/SINV genome.

Taken together, the results suggested that an RNA stem with particular stability in the 5′ terminus of the VEEV genome plays an important role in viral RNA replication. Modification of the stem or its complete destabilization by multiple point mutations has a strong negative effect on virus replication and leads to the modification of the 5′ terminus by either adding new repeating elements or, less frequently, developing new sequences that may potentially form stem-loop structures with the downstream fragments of the 5′UTR.

Stem sequence plays a critical role in virus replication. The results of the above-described experiments underlined the importance of the 5′-terminal stem and its stability for virus replication. However, they did not answer another key question as to whether the stem's stability is the only critical parameter or if its sequence is an important characteristic as well. To test the significance of the nucleotide sequence in the stem, we designed new 5′UTRs in which one or a few base pairs were positioned in the opposite orientation (Fig. 7A). The introduced changes did not affect the free energy of the stems, and it was reasonable to expect that overall 5′UTR RNA folding remained similar to that determined previously for (G3)VEE/SINV (36). The replacement of even one G-C base pair with a C-G base pair in (Rev1BP)VEE/SINV altered the rates of virus replication (Fig. 7A). More extensive mutagenesis of 2 or 4 bp in the stem [(Rev2BP)VEE/SINV and (Rev4BP)VEE/SINV] (Fig. 7A) had deleterious effects on virus growth. The in vitro-synthesized RNAs demonstrated dramatically lower infectivities (Fig. 7A) in the infectious center assay, and viruses were replicating to very low titers (Fig. 7B) and formed pinpoint-sized plaques. In these experiments, we did not detect a formation of the true revertants or pseudorevertants capable of replicating with an efficiency similar to that of (G3)VEE/SINV. Their appearance likely required multiple passages, and thus, further analysis was not performed. Taken together, the data strongly suggested that the 5′UTR-encoded stem is a critical element of the promoter for RNA replication, and the nucleotide sequence plays a role as important as the stem's stability.

Modifications of the 5′-terminal RNA secondary structure affect RNA replication rather than translation of the encoded polyprotein. The negative effects of the 5′-terminal RNA mod-

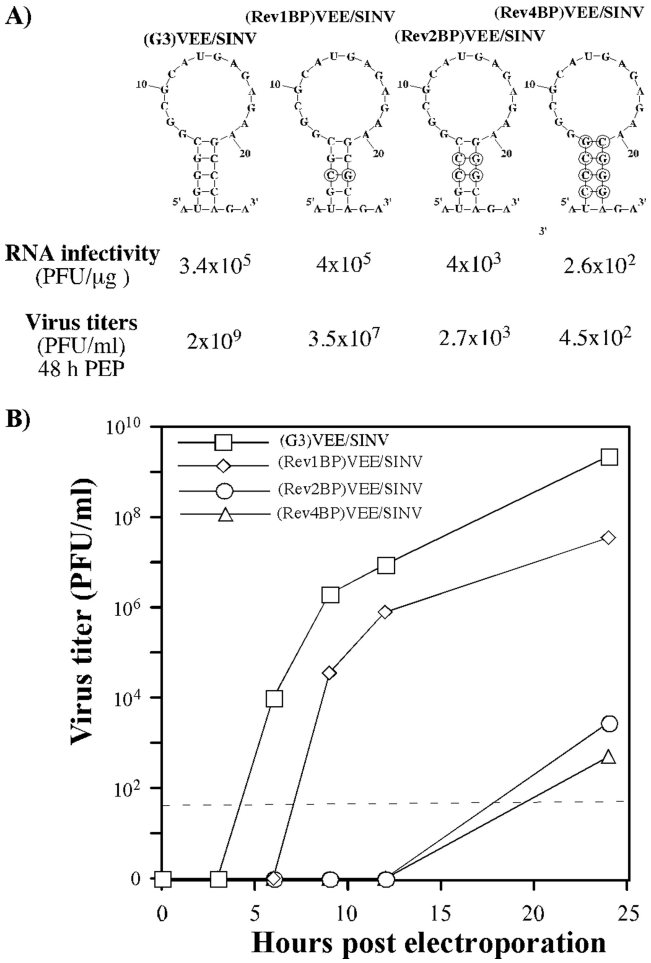

FIG. 7. Effect of stem sequence modification on VEE/SINV viability and replication rates. (A) Computer-predicted secondary structures of the 5′-terminal sequences of the designed mutant genomes, infectivities of the in vitro-synthesized RNAs in the infectious center assay, and titers of the viruses at 24 h postelectroporation. PEP, postelectroporation. (B) Analysis of virus replication. BHK-21 cells were transfected with 2 μg of the in vitro-synthesized RNAs, media were replaced at the indicated time points, and virus titers were assessed by plaque assay on BHK-21 cells. The dashed line indicates the limit of virus detection.

ifications on virus replication do not point directly to a lower activity of the RNA promoter elements. A downregulation of replication can potentially result from lowering the template activity of the RNA in the translation of the nonstructural polyprotein. In particular, this can be attributed to mutations increasing the secondary structure of the 5′UTR. To distinguish between the effects of the mutation on RNA replication and/or translation, we initially designed a number of cassettes that were incapable of self-replication and contained different 5′UTRs. The tested UTRs controlled the translation of the firefly luciferase gene, fused in frame with the Ubi sequence and the 147-nt-long fragment of the VEEV nsP1 gene (Fig. 8A and see Materials and Methods for details). The latter sequence preserved the natural secondary structure of the 5′ termini, and the Ubi gene was required for promoting the production of luciferase in a more active, free form. Equal amounts of in vitro-synthesized capped RNAs were transfected into BHK-21 cells by electroporation. Another *Renilla* luciferase-encoding RNA was cotransfected into the cells as a control for transfection efficiency. The activities of both luciferases were evaluated at different times posttransfection (see Materials and Methods for details). The results demonstrated that all of the modifications of the 5′ terminus downregulated the translation less than twofold. Only the (C17C19C20)VEE/SINV-specific 5′UTR, having the most stable stem, drove translation threefold less efficiently but did not abrogate it. Moreover, the luciferase expression by the (AUG)VEE/SINV-derived 5′UTR was higher than that determined for similar constructs with two or three AUGs [$(AUG)_2$VEE/SINV and $(AUG)_3$VEE/SINV 5′UTRs, respectively]. Thus, the applied modifications of the VEEV 5′UTR had no strong effect on RNA translation.

To dissect the effect of the mutation on RNA replication, we designed defective viral genomes encoding different 5′UTRs, 475 nt of the nsP1 gene, and a firefly luciferase gene under the control of the subgenomic promoter (Fig. 8B). The DI RNAs did not encode any full-length nonstructural proteins and were capable of replication only in the presence of nsPs supplied in *trans* by the helper VEEV replicon VEErep/Pac. Upon delivery into the same cell, replicon genomes were translated to produce VEEV nsPs, which could then be used for its own and DI RNA replication and transcription of the subgenomic, luciferase-expressing RNA. Thus, the efficiency of DI RNA replication and luciferase expression was independent of the 5′UTR function in translation initiation but was determined by the promoter sequences at the DI RNAs' 5′ ends. The VEEV replicon and the DI genomes were cotransfected into BHK-21 cells (see Materials and Methods for details), and luciferase activity was evaluated at different times postelectroporation. The DI RNAs having mutated, more stable 5′-terminal stems [(C19C20)DI/Luc and (C17C19C20)DI/Luc] demonstrated a profound decrease in luciferase expression, but the addition of one or more AUGs to the C19C20-containing 5′UTR made (AUG)DI/Luc, $(AUG)_2$DI/Luc, and $(AUG)_3$DI/Luc capable of replication and efficient luciferase expression (Fig. 8B). As we expected, the (G2C25)DI/Luc RNA did not demonstrate any detectable replication, and the NOstem 5′UTR-containing RNA was surprisingly efficient in luciferase expression. On the other hand, its ability to express luciferase correlated with a high level of RNA infectivity of the corresponding viral genome in the infectious center assay and very rapid evolution to the large-plaque phenotype.

Thus, the results of these experiments demonstrated that the modifications of the very 5′ terminus of the genome, described above, affected viral RNA replication rather than the translation of the VEEV nsPs. Additional 5′-terminal AUG repeats, acquired by mutants with more stable 5′-terminal RNA secondary structures, did not enhance the translation of the encoded proteins but increased the efficiency of the RNA replication that resulted in the generation of viable viruses.

DISCUSSION

The synthesis of the positive strand of the alphavirus genome depends directly on the promoter located at the 3′ termini of the RNA negative strand in the dsRNA replicative

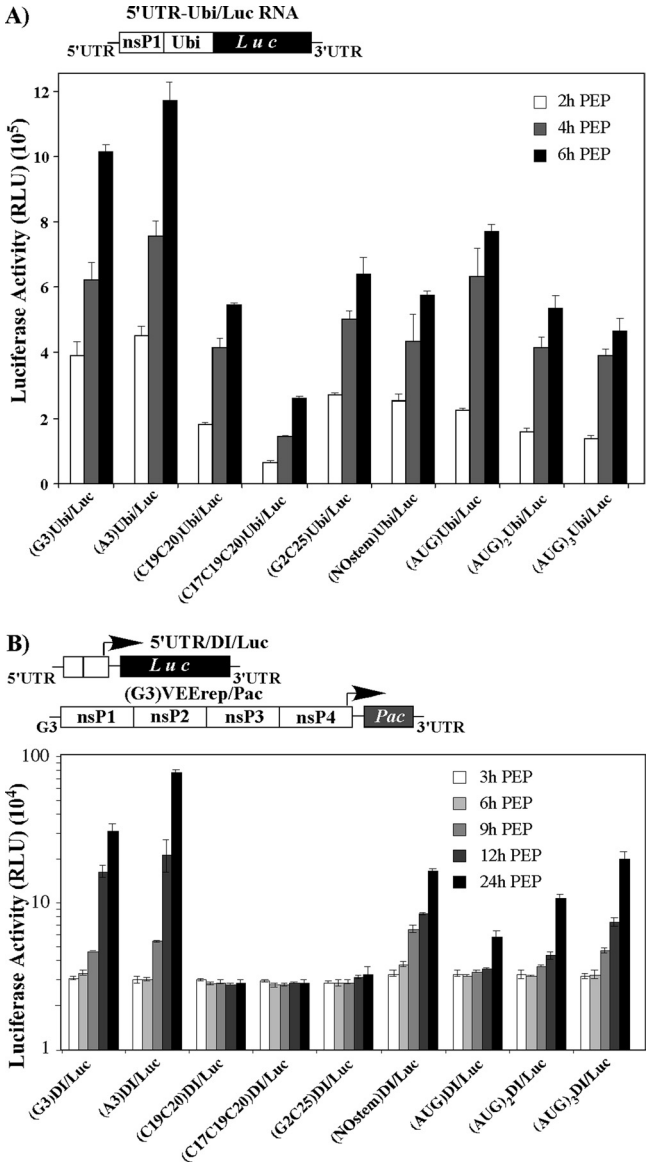

FIG. 8. Analysis of effects of mutations in the VEEV 5′UTR on the translation of the encoded proteins and RNA replication. (A) Schematic representation of the firefly luciferase-encoding constructs used for evaluating the effect of the 5′UTR-specific mutations on template translation. A detailed description of the constructs is presented in Materials and Methods. Shown are values for firefly luciferase expression in BHK-21 cells transfected with the in vitro-synthesized, capped RNAs. One microgram of each template was mixed with 0.1 μg of *Renilla* luciferase-encoding RNA and electroporated into BHK-21 cells as described in Materials and Methods. Equal aliquots of transfected cells were seeded into 35-mm dishes, and luciferase activities were determined at the indicated time points using a dual-luciferase system (Promega). (B) Schematic representation of DI RNAs having mutations in the 5′UTR and helper VEEV replicon used in the study. BHK-21 cells were cotransfected with 2 μg of replicons and 2 μg of the indicated DI RNAs. Equal numbers of electroporated cells were seeded into 35-mm dishes and incubated at 37°C in 5% CO_2. At the indicated time points, cells were lysed, and luciferase activity was measured as described in Materials and Methods. PEP, postelectroporation; RLU, relative light units.

intermediates. The structure and functions of this promoter encoded by the 5′ terminus of the viral genome are not well understood. This RNA fragment not only determines the promoter activity in a virus- and cell-specific mode (9, 31) but also has other functions in virus replication. The 5′UTR directs the translation of viral nsPs, and thus, its structure ultimately determines the accumulation of the replication- and other virus-specific protein complexes. The SINV-specific 5′UTR also plays a key role in negative-strand synthesis (5, 9) and can be considered a part of the promoter for the synthesis of the replicative intermediate. The same function might be attributed to the 5′UTRs of other alphaviruses.

The VEEV TRD genome-specific 5′UTR was predicted to form a stem-loop structure, and its presence was confirmed by nuclear magnetic resonance and enzymatic analysis (20). The 3′-terminal sequence at the negative strand was also predicted to form a stem-loop that mirrors that of the 5′ terminus of the viral genome. In the present study, the loop sequence was shown to play no direct role in driving virus replication in vitro. It could be replaced by a different sequence without any noticeable effect on virus replication. Moreover, because of the A3 mutation, the predicted loop was different in the VEEV TC-83 genome (20), but the replication of the virus was not affected. One might argue that in vitro-synthesized (C19C20) VEE/SINV and (C17C19C20)VEE/SINV RNAs, having extended RNA stems and smaller loops, exhibited very low infectivities in the infectious center assay, suggesting a possibility that the loop might also contribute to viral genome replication. However, in all of the selected pseudorevertants, the latter smaller loops remained intact, and the detected extensions of the 5′ termini in the recovered variants were similar to those developed in other pseudorevertants in response to stem modifications. Thus, it is highly unlikely that the loop sequence is directly involved in the promoter function.

Two other 5′-terminal RNA elements, the very short unpaired fragment and the stem, appear to represent important components of the VEEV-specific promoter. The two nucleotides (AU) at the very 5′ terminus are critical and are absolutely conserved in the genomes of other alphaviruses. In this study, we demonstrated that the G2 and A2 mutations have deleterious effects on virus replication regardless of the ability to form base pairs. Moreover, the AU was found to be the first 2 nt in the genomes of all of the pseudorevertants generated in this study. Mutants with a destabilized stem generated multiple, compensatory AU repeats. The importance of the five very 5′-terminal nucleotides in viral RNA replication was previously suggested by experiments with SINV (32), but the pseudorevertants were neither detected nor characterized. The critical role of the 5′-terminal AU suggested that the additional repeats might have a positive effect on the replication of the original (G3)VEE/SINV. However, this was not the case. Viruses with an additional AU or two AUs were viable but demonstrated growth and RNA replication rates that were very similar to those of the parental (G3)VEE/SINV (data not shown). This suggested that in VEEV TRD (having the wt 5′UTR), the 5′ sequence of the genome had already been optimized by the previous evolution of VEEV, and further improvement is an unlikely event.

The VEEV TRD-specific third nucleotide, G3, plays a less critical role in RNA replication, as the virus can tolerate the

presence of any other nucleotide in this position. The C3 and U3 mutations affected the secondary RNA structure but did not increase levels of either virus (Fig. 1) or RNA (data not shown) replication. However, the role of nt 3 should not be underestimated, because the VEEV TC-83-specific A3 mutation led to a significant attenuation of the virus in vivo (17, 55), a detectable increase in RNA replication rates in vitro, and lower rates of subgenomic RNA transcription (20).

The stem, following the unpaired RNA fragment, is another promoter element that plays a critical role in viral RNA synthesis. Mutations that caused either an increase or decrease in its thermodynamic stability or the complete elimination of the stem structure had a deleterious effect on virus replication (Fig. 3 and 6). All of these mutants reverted to an efficiently replicating phenotype mostly by generating additional AUG or AU repeats at the very 5′ terminus. This was an unexpected means of neutralizing the negative effect of stem modifications. Even a single additional AUG was capable of making the (C19C20)VEE/SINV virus viable. Further increases in the number of repeats correlated with the higher replication efficiencies of the mutants. Other experiments demonstrated that the 5′-terminal AUGs had a very minor positive effect on the translation of the following coded sequence, and their stimulatory effect on virus replication resulted rather from a more efficient RNA promoter function. Interestingly, the majority of arising variants with more stable stems developed AUG repeats (AUG is present in the VEEV TRD genome), and the repeats, developed in response to stem destabilization, consisted mainly of repeating AUs. The latter sequences correlated with the presence of AUA in the VEEV TC-83 5′UTR, in which the stem-loop structure also had a high level of free energy and a low level of stability. Importantly, the sequence of the stem was proven to be as critical as its stability, and the replacement of the G-C base pairs by similar C-G ones had a very strong negative effect on RNA and virus replication. The magnitude of the effect correlated with the number of replaced base pairs.

We previously detected the ability of SINVs to acquire AU-rich sequences at their 5′ termini in response to the replacement of their natural 5′UTRs by that derived from the distantly related Semliki Forest virus (9). These new sequences were very heterogeneous and strongly promoted virus replication in a sequence-specific mode. VEEV-specific replicative machinery is likely to be more conservative and capable of utilizing a very limited number of sequences in which AU and AUG repeats strongly dominate. The use of repeats was determined by particular stem modifications. Based on our previously reported SINV-related data and the results of the present study, we propose that even after strong modifications of the 5′-terminal fragments, alphaviruses sustain some ability for negative-strand RNA synthesis and are capable of generating heterologous sequences at the 3′ terminus of the negative-strand RNA intermediate. The particular sequences capable of functioning as a promoter for positive-strand RNA synthesis are then selected during the subsequent rounds of RNA replication. Recent studies have demonstrated that the core catalytic domain of SINV nsP4 possesses terminal adenylyltransferase activity (47). Experiments using full-length nsP4 further affirmed that finding (40). It is likely that full-length nsP4 by itself or in complex with other nsPs could possess terminal nucleotidyltransferase activity, by which it could add combinations of nucleotides to the 3′ end of the template negative strands. The efficiency of this process is most likely not very high, because we were never able to detect heterologous 5′ ends in the genomes of plaque-purified (G3)VEE/SINV and (A3)VEE/SINV viruses (data not shown), indicating that such mutant RNAs are present in virus populations at very low concentrations, if at all.

The high frequency of perfect repeats suggests another possible mechanism for 5′-terminal insertions. RNA replication in alphaviruses follows de novo, primer-independent initiation for RNA synthesis from the template strand. Modifications of the stem might cause the initiation of positive-strand synthesis to be inefficient. The replication complex likely dissociates from the template after synthesizing the first 2 or 3 nt and then reinitiates RNA synthesis with the previously synthesized oligonucleotide still in place, resulting in a nontemplated sequence being added to the 5′ terminus of the positive-strand RNA. This putative mechanism could explain the appearance of many, but not all, of the heterologous sequences, because some of them are represented by imperfect repeats. Thus, the template reinitiation mechanism and the terminal nucleotidyltransferase mechanism might function synergistically in 5′UTR modification.

Taken together, the data suggest that during RNA replication, the AU (or AUG) sequence functions as a critical element of the promoter, and along with the stem structure, it is likely the site for RC binding during the initiation of positive-strand RNA synthesis. A negative effect of the mutations in this stem can be compensated for by addition of AU or AUG repeats to the 5′UTR of the VEEV genome, and the number of repeats positively correlates with replication efficiency. We previously demonstrated that the alphavirus RNA replicative intermediate is a dsRNA (8). However, the results of this study strongly support the possibility that the 5′ terminus of the positive strand and the 3′ terminus of the negative strand in the VEEV replicative intermediate do not form the dsRNA duplex, and the RC interacts with the stem structures in single-stranded RNA (ssRNA). Figure 9 presents two models of the end dsRNA folding and promoter recognition by the VEEV RC. Model A represents the case where the promoter for RNA synthesis in terminal ends of the plus-strand RNA and the minus strand exists as dsRNA. Model B represents the scenario where the 5′ and the 3′ ends of the positive and negative strands, respectively, in this replicative intermediate do not form a dsRNA duplex but fold into individual stem-loop structures. Results from our studies strongly support the hypothesis presented in model B, as model A does not explain the effect of the point mutations increasing the stability of the stem on viral RNA replication (Fig. 9). The model that the promoter elements are represented by ssRNA fragments explains the results of our experiments and is supported by the data of other research groups. First, the single-stranded, negative-strand RNA is recognized by SINV nsPs and is utilized as a template for positive-strand RNA synthesis in vitro (26, 27). Second, the 3′ end of the negative strand of SINV RNA contains binding sites for a number of cellular proteins (34, 35), which is indicative of their possible function in RNA replication. Third, the data from SINV-related studies also suggest that the entire 5′ terminus in the SINV genome can be replaced by heterologous tRNA- or subgenomic RNA-derived

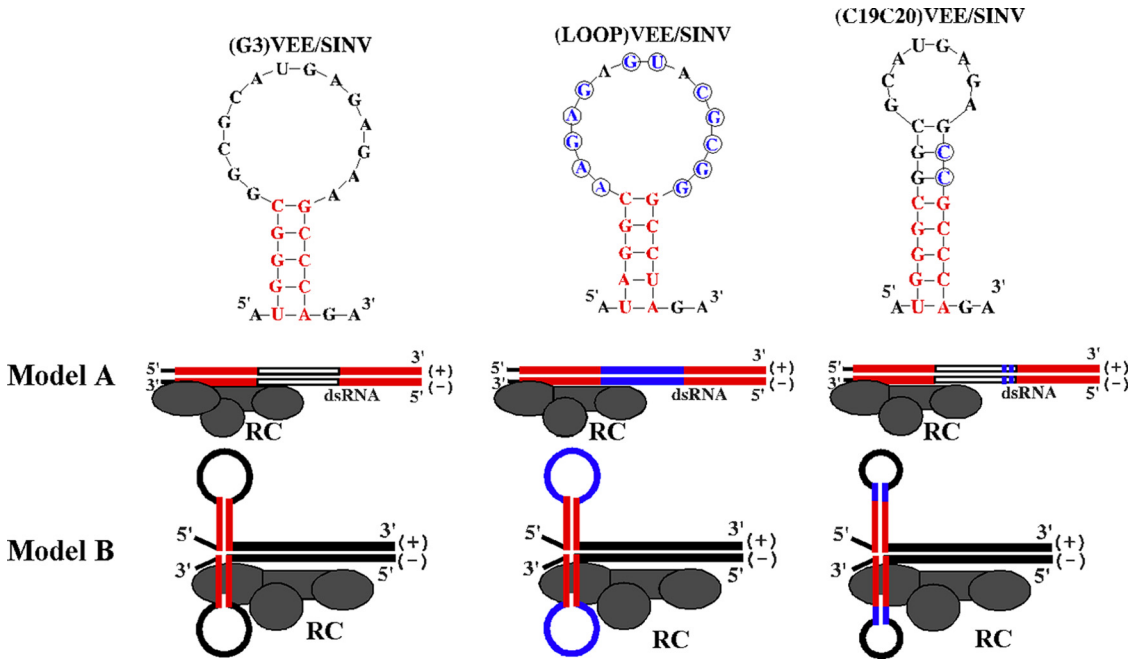

FIG. 9. Two models of the secondary structure of the dsRNA terminus and VEEV RC binding. Model A is based on the assumption that in the replicative intermediate, the 5′ end of the positive strand and the 3′ end of the negative strand exist as a complete dsRNA duplex. According to this model, the VEEV RC recognizes the very terminus of the (G3)VEE/SINV-specific dsRNA duplex and the following sequence. Model A cannot explain the lack of any effect of loop replacement in (LOOP)VEE/SINV and, on the other hand, the strong negative effect of two or three point mutations in (C19C20)VEE/SINV and (C17C19C20)VEE/SINV on RNA replication. Model B depicts the dsRNA terminus as two ssRNA fragments, which fold into the stem-loop structures. According to this model, the VEEV RC uses both the very 3′-terminal, unpaired fragment and the stem (but not the loop) of the negative-strand RNA as two elements synergistically functioning in RC binding and the initiation of positive-strand RNA synthesis. Thus, (LOOP)VEE/SINV is expected to have no defects in RNA replication. However, the point mutations in (C19C20)VEE/SINV and (C17C19C20)VEE/SINV as well as other mutations that change the stem's stability and/or sequence strongly affect the initiation of replication. This leads, in turn, to the selection of the variants having compensatory, repeating sequences in another, unpaired, promoter element. Model B is representative of the results of this study.

sequences (4, 54). It is highly unlikely for these totally different sequences to function as a promoter in the dsRNA duplex. However, their ability to fold into short stem-loops corroborates their activity as promoters according to model B (Fig. 9).

All of the experiments presented have been performed using cells of vertebrate origin, and thus, the heterologous sequences identified in the 5′UTR of VEE/SINV genomes might function with different efficiencies in the cells of insect origin. This was the case in our previously reported studies of the evolution of SINV with a heterologous 5′UTR and a modified 51-nt CSE (4, 9). The host cell-specific mode of functioning of the SINV RC and the coisolation of cellular proteins with viral nsPs (1, 6, 8) strongly suggest a direct involvement of host factors in the initiation of RNA synthesis on the replicative intermediate. However, this possibility remains to be experimentally tested for replicating VEEV RNA.

In summary, the results of the study demonstrated that (i) in the VEEV genome, both sequence folding into the RNA stem structure and the very 5′-terminal AU dinucleotide play critical roles in promoter's function; (ii) both the increase and decrease in the stem's stability have deleterious effects on virus and RNA replication; (iii) the stem sequence plays as important role, as does its stability, in the promoter's function; (iv) the VEEV replicative machinery is capable of developing compensatory sequences, either containing AUG or AU repeats or leading to new, heterologous stem-loop formation; (v) some of the 5′-terminal extended mutations lead to an accumulation of the compensatory mutations in VEEV nsPs; and (vi) our data provide new evidence that the 3′ terminus of the negative-strand viral genome in the dsRNA intermediate is represented by a ssRNA, whose overall folding, and a particular sequence, determines its function as a promoter for positive-strand RNA genome synthesis.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AI070207 to I.F. and S.A., NHLBI grant N01HV28184, and Welch Foundation grant H1296 to D.G.G. R.K.-S. was supported by the McLaughlin Endowment predoctoral fellowship.

REFERENCES

1. **Atasheva, S., R. Gorchakov, R. English, I. Frolov, and E. Frolova.** 2007. Development of Sindbis viruses encoding nsP2/GFP chimeric proteins and their application for studying nsP2 functioning. J. Virol. **81:**5046–5057.
2. **Bredenbeek, P. J., I. Frolov, C. M. Rice, and S. Schlesinger.** 1993. Sindbis virus expression vectors: packaging of RNA replicons by using defective helper RNAs. J. Virol. **67:**6439–6446.
3. **Dal Canto, M. C., and S. G. Rabinowitz.** 1981. Central nervous system demyelination in Venezuelan equine encephalomyelitis infection. J. Neurol. Sci. **49:**397–418.
4. **Fayzulin, R., and I. Frolov.** 2004. Changes of the secondary structure of the 5′ end of the Sindbis virus genome inhibit virus growth in mosquito cells and lead to accumulation of adaptive mutations. J. Virol. **78:**4953–4964.
5. **Frolov, I., R. Hardy, and C. M. Rice.** 2001. Cis-acting RNA elements at the 5′ end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. RNA **7:**1638–1651.
6. **Frolova, E., R. Gorchakov, N. Garmashova, S. Atasheva, L. A. Vergara, and**

I. Frolov. 2006. Formation of nsP3-specific protein complexes during Sindbis virus replication. J. Virol. **80:**4122–4134.
7. **Garmashova, N., R. Gorchakov, E. Volkova, S. Paessler, E. Frolova, and I. Frolov.** 2007. The Old World and New World alphaviruses use different virus-specific proteins for induction of transcriptional shutoff. J. Virol. **81:** 2472–2484.
8. **Gorchakov, R., N. Garmashova, E. Frolova, and I. Frolov.** 2008. Different types of nsP3-containing protein complexes in Sindbis virus-infected cells. J. Virol. **82:**10088–10101.
9. **Gorchakov, R., R. Hardy, C. M. Rice, and I. Frolov.** 2004. Selection of functional 5′ *cis*-acting elements promoting efficient Sindbis virus genome replication. J. Virol. **78:**61–75.
10. **Gorchakov, R., E. Volkova, N. Yun, O. Petrakova, N. S. Linde, S. Paessler, E. Frolova, and I. Frolov.** 2007. Comparative analysis of the alphavirus-based vectors expressing Rift Valley fever virus glycoproteins. Virology **366:**212–225.
11. **Griffin, D. E.** 2001. Alphaviruses, p. 917–962. *In* D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus (ed.), Fields virology, 4th ed. Lippincott Williams & Wilkins, Philadelphia, PA.
12. **Hardy, R. W.** 2006. The role of the 3′ terminus of the Sindbis virus genome in minus-strand initiation site selection. Virology **345:**520–531.
13. **Hardy, R. W., and C. M. Rice.** 2005. Requirements at the 3′ end of the Sindbis virus genome for efficient synthesis of minus-strand RNA. J. Virol. **79:**4630–4639.
14. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1989. Improved predictions of secondary structures for RNA. Proc. Natl. Acad. Sci. USA **86:**7706–7710.
15. **Johnson, K. M., and D. H. Martin.** 1974. Venezuelan equine encephalitis. Adv. Vet. Sci. Comp. Med. **18:**79–116.
16. **Johnston, R. E., and C. J. Peters.** 1996. Alphaviruses, p. 843–898. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), Fields virology, 3rd ed. Lippincott-Raven Publishers, Philadelphia, PA.
17. **Kinney, R. M., G. J. Chang, K. R. Tsuchiya, J. M. Sneider, J. T. Roehrig, T. M. Woodward, and D. W. Trent.** 1993. Attenuation of Venezuelan equine encephalitis virus strain TC-83 is encoded by the 5′-noncoding region and the E2 envelope glycoprotein. J. Virol. **67:**1269–1277.
18. **Kinney, R. M., B. J. B. Johnson, J. B. Welch, K. R. Tsuchiya, and D. W. Trent.** 1989. The full-length nucleotide sequences of the virulent Trinidad donkey strain of Venezuelan equine encephalitis virus and its attenuated vaccine derivative, strain TC-83. Virology **170:**19–30.
19. **Kuhn, R. J., Z. Hong, and J. H. Strauss.** 1990. Mutagenesis of 3′ nontranslated region of Sindbis virus RNA. J. Virol. **64:**1465–1476.
20. **Kulasegaran-Shylini, R., V. Thiviyanathan, D. G. Gorenstein, and I. Frolov.** 2009. The 5′UTR-specific mutation in VEEV TC-83 genome has a strong effect on RNA replication and subgenomic RNA synthesis, but not on translation of the encoded proteins. Virology **387:**211–221.
21. **Lemm, J. A., R. K. Durbin, V. Stollar, and C. M. Rice.** 1990. Mutations which alter the level or structure of nsP4 can affect the efficiency of Sindbis virus replication in a host-dependent manner. J. Virol. **64:**3001–3011.
22. **Lemm, J. A., and C. M. Rice.** 1993. Roles of nonstructural polyproteins and cleavage products in regulating Sindbis virus RNA replication and transcription. J. Virol. **67:**1916–1926.
23. **Lemm, J. A., T. Rümenapf, E. G. Strauss, J. H. Strauss, and C. M. Rice.** 1994. Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus and plus-strand RNA synthesis. EMBO J. **13:**2925–2934.
24. **Leon, C. A.** 1975. Sequelae of Venezuelan equine encephalitis in humans: a four year follow-up. Int. J. Epidemiol. **4:**131–140.
25. **Levis, R., S. Schlesinger, and H. V. Huang.** 1990. Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription. J. Virol. **64:**1726–1733.
26. **Li, M. L., and V. Stollar.** 2007. Distinct sites on the Sindbis virus RNA-dependent RNA polymerase for binding to the promoters for the synthesis of genomic and subgenomic RNA. J. Virol. **81:**4371–4373.
27. **Li, M. L., and V. Stollar.** 2004. Identification of the amino acid sequence in Sindbis virus nsP4 that binds to the promoter for the synthesis of the subgenomic RNA. Proc. Natl. Acad. Sci. USA **101:**9429–9434.
28. **Liljeström, P., and H. Garoff.** 1991. A new generation of animal cell expression vectors based on the Semliki Forest virus replicon. Biotechnology **9:**1356–1361.
29. **Liljeström, P., S. Lusa, D. Huylebroeck, and H. Garoff.** 1991. In vitro mutagenesis of a full-length cDNA clone of Semliki Forest virus: the small 6,000-molecular-weight membrane protein modulates virus release. J. Virol. **65:**4107–4113.
30. **Michel, G., O. Petrakova, S. Atasheva, and I. Frolov.** 2007. Adaptation of Venezuelan equine encephalitis virus lacking 51-nt conserved sequence element to replication in mammalian and mosquito cells. Virology **362:**475–487.
31. **Nickens, D. G., and R. W. Hardy.** 2008. Structural and functional analyses of stem-loop 1 of the Sindbis virus genome. Virology **370:**158–172.
32. **Niesters, H. G. M., and J. H. Strauss.** 1990. Defined mutations in the 5′ nontranslated sequence of Sindbis virus RNA. J. Virol. **64:**4162–4168.
33. **Niesters, H. G. M., and J. H. Strauss.** 1990. Mutagenesis of the conserved 51-nucleotide region of Sindbis virus. J. Virol. **64:**1639–1647.
34. **Pardigon, N., E. Lenches, and J. H. Strauss.** 1993. Multiple binding sites for cellular proteins in the 3′ end of Sindbis alphavirus minus-sense RNA. J. Virol. **67:**5003–5011.
35. **Pardigon, N., and J. H. Strauss.** 1992. Cellular proteins bind to the 3′ end of Sindbis virus minus-strand RNA. J. Virol. **66:**1007–1015.
36. **Peshkova, E. A., and E. K. Bakhtin.** 1978. Enzymatic and ultrastructural changes in isolated liver mitochondria from mice infected with several group A arboviruses. Vopr. Virusol. **1978:**350–357. (In Russian.)
37. **Petrakova, O., E. Volkova, R. Gorchakov, S. Paessler, R. M. Kinney, and I. Frolov.** 2005. Noncytopathic replication of Venezuelan equine encephalitis virus and eastern equine encephalitis virus replicons in mammalian cells. J. Virol. **79:**7597–7608.
38. **Rice, C. M., R. Levis, J. H. Strauss, and H. V. Huang.** 1987. Production of infectious RNA transcripts from Sindbis virus cDNA clones: mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. J. Virol. **61:**3809–3819.
39. **Rico-Hesse, R., S. C. Weaver, J. de Siger, G. Medina, and R. A. Salas.** 1995. Emergence of a new epidemic/epizootic Venezuelan equine encephalitis virus in South America. Proc. Natl. Acad. Sci. USA **92:**5278–5281.
40. **Rubach, J. K., B. R. Wasik, J. C. Rupp, R. J. Kuhn, R. W. Hardy, and J. L. Smith.** 2009. Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity in vitro. Virology **384:**201–208.
41. **Sawicki, D. L., S. Perri, J. M. Polo, and S. G. Sawicki.** 2006. Role for nsP2 proteins in the cessation of alphavirus minus-strand synthesis by host cells. J. Virol. **80:**360–371.
42. **Sawicki, D. L., and S. G. Sawicki.** 1987. Alphavirus plus and minus strand RNA synthesis, p. 251–259. *In* M. Brinton and R. Ruckert (ed.), Positive-strand RNA viruses. Alan R. Liss, Inc., New York, NY.
43. **Shirako, Y., and J. H. Strauss.** 1994. Regulation of Sindbis virus RNA replication: uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. J. Virol. **68:**1874–1885.
44. **Strauss, E. G., C. M. Rice, and J. H. Strauss.** 1984. Complete nucleotide sequence of the genomic RNA of Sindbis virus. Virology **133:**92–110.
45. **Strauss, J. H., and E. G. Strauss.** 1994. The alphaviruses: gene expression, replication, evolution. Microbiol. Rev. **58:**491–562.
46. **Takkinen, K.** 1986. Complete nucleotide sequence of the non-structural protein genes of Semliki Forest virus. Nucleic Acids Res. **14:**5667–5682.
47. **Tomar, S., R. W. Hardy, J. L. Smith, and R. J. Kuhn.** 2006. Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity. J. Virol. **80:**9962–9969.
48. **Weaver, S. C., and A. D. Barrett.** 2004. Transmission cycles, host range, evolution and emergence of arboviral disease. Nat. Rev. Microbiol. **2:**789–801.
49. **Weaver, S. C., C. Ferro, R. Barrera, J. Boshell, and J. C. Navarro.** 2004. Venezuelan equine encephalitis. Annu. Rev. Entomol. **49:**141–174.
50. **Weaver, S. C., and I. Frolov.** 2005. Togaviruses, p. 1010–1024. *In* B. W. J. Mahy and V. ter Meulen (ed.), Virology, vol. 2, Topley and Wilson's principles of bacteriology, virology and immunity. Edward Arnold, Salisbury, United Kingdom.
51. **Weaver, S. C., A. Hagenbaugh, L. A. Bellew, L. Gousset, V. Mallampalli, J. J. Holland, and T. W. Scott.** 1994. Evolution of alphaviruses in the eastern equine encephalomyelitis complex. J. Virol. **68:**158–169.
52. **Weaver, S. C., A. Hagenbaugh, L. A. Bellew, S. V. Netesov, V. E. Volchkov, G. J. Chang, D. K. Clarke, L. Gousset, T. W. Scott, D. W. Trent, et al.** 1993. A comparison of the nucleotide sequences of eastern and western equine encephalomyelitis viruses with those of other alphaviruses and related RNA viruses. Virology **197:**375–390.
53. **Weaver, S. C., R. Salas, R. Rico-Hesse, G. V. Ludwig, M. S. Oberste, J. Boshell, R. B. Tesh, et al.** 1996. Re-emergence of epidemic Venezuelan equine encephalomyelitis in South America. Lancet **348:**436–440.
54. **Weiss, B. G., and S. Schlesinger.** 1991. Recombination between Sindbis virus RNAs. J. Virol. **65:**4017–4025.
55. **White, L. J., J. G. Wang, N. L. Davis, and R. E. Johnston.** 2001. Role of alpha/beta interferon in Venezuelan equine encephalitis virus pathogenesis: effect of an attenuating mutation in the 5′ untranslated region. J. Virol. **75:**3706–3718.