JOURNAL OF VIROLOGY, Feb. 1996, p. 1182–1190 Vol. 70, No. 2
0022-538X/96/$04.00+0

Translation of Sindbis Virus mRNA: Analysis of Sequences Downstream of the Initiating AUG Codon That Enhance Translation

ILYA FROLOV AND SONDRA SCHLESINGER*

Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri 63110-1093

Received 14 August 1995/Accepted 30 October 1995

Alphaviruses, particularly Sindbis virus and Semliki Forest virus, are proving to be useful vectors for the expression of heterologous genes. In infected cells, these self-replicating vectors (replicons) transcribe a subgenomic mRNA that codes for a heterologous protein instead of the structural proteins. We reported recently that translation of the reporter gene *lacZ* is enhanced 10-fold when the coding sequences of this gene are fused downstream of and in frame with the 5′ half of the capsid gene (I. Frolov and S. Schlesinger, J. Virol. 68:8111–8117, 1994). The enhancing sequences, located downstream of the AUG codon that initiates translation of the capsid protein, have a predicted hairpin structure. We have mutated this region by making changes in the codons which do not affect the protein sequence but should destabilize the putative hairpin structure. These changes caused a decrease in the accumulation of the capsid–β-galactosidase fusion protein. When these alterations were inserted into the capsid gene in the context of the intact Sindbis virus genome, they led to a decrease in the rate of virus formation but did not affect the final yield. We also altered the original sequence to one that has 12 contiguous G · C base pairs and should form a stable hairpin. The new sequence was essentially as effective as the original had been in enhancement of translation and in the rate of virus formation. The position of the predicted hairpin structure is important for its function; an insertion of 9 nucleotides or a deletion of 9 nucleotides decreased the level of translation. The insertion of a hairpin structure at a particular location downstream of the initiating AUG appears to be a way that alphaviruses have evolved to enhance translation of their mRNA, and, as a consequence, they produce high levels of the structural proteins which are needed for virus assembly. This high level of translation requires an intracellular environment in which host cell protein synthesis is inhibited.

There are multiple factors involved in regulating the translation of mRNAs in eucaryotic cells. One of the important elements is the sequence, and consequently the structure, of the mRNA itself. Effects of RNA sequence and structure on controlling translation have focused mainly on the 5′ and 3′ untranslated regions (7, 13, 24). There is increasing evidence that sequences within the coding region also affect translation (12, 28), and in a recent study, we showed that in cultured cells infected with Sindbis virus replicons, translation of the viral mRNA was enhanced more than 10-fold when the 5′ half of the capsid protein gene was inserted upstream of and in frame with the *lacZ* reporter gene (6).

Sindbis virus is classified in the *Togaviridae* family and is the prototype member of the *Alphavirus* genus. The genome, a positive-strand nonsegmented RNA of almost 12 kb, consists of two modules: the 5′ two-thirds, which codes for the nonstructural proteins (nsPs), and the 3′ one-third, which codes for the structural proteins (reviewed in reference 25). In infected cells, the genome functions directly as an mRNA but only for the nsPs, which are synthesized as a polyprotein that is cleaved to produce four polypeptides. These proteins are required for the replication of the viral genomic RNA and for the transcription of the subgenomic RNA which is identical in sequence to the 3′ one-third of the genome and codes for the viral structural proteins. Transcription of the subgenomic RNA is directed by an internal promoter located at the junction between the genes coding for the nsPs and structural proteins on the negative strand of the genome. The structural proteins are also synthesized as a polyprotein. The amino-terminal polypeptide, the viral capsid protein, is an autoprotease and cleaves itself from the remaining polypeptide chain. The remainder of the polypeptide undergoes several cleavages to produce two glycoproteins and a small hydrophobic protein, which become part of the envelope of the extracellular virion particle.

Sindbis virus has been engineered through its cDNA to serve as a vector for the expression of heterologous genes (3, 20, 29). These vectors are self-replicating RNAs (replicons) which contain the genes for the nsPs, but the structural protein genes are replaced by a heterologous gene (see Fig. 1A for a diagram of the replicon genome).

Infection of many cultured cells by Sindbis virus, as well as other related alphaviruses, leads to the inhibition of host cell protein synthesis (9, 25). This ability of a virus to inhibit the translation of host mRNAs under conditions in which the viral mRNAs can be translated has been observed for a variety of positive- and negative-strand RNA viruses, including poliovirus and influenza virus, and for DNA viruses such as adenovirus (4, 11, 21, 23). How this is accomplished remains an enigma, and different viruses may use different mechanisms to achieve this differential effect. The prevailing models suggest that factors required for initiation of protein synthesis are depleted in infected cells and that viral mRNAs are able to usurp these factors for their own translation.

* Corresponding author. Mailing address: Department of Molecular Microbiology, Box 8230, Washington University School of Medicine, 660 S. Euclid Ave., St. Louis, MO 63110-1093. Phone: (314) 362-2746. Fax: (314) 362-1232. Electronic mail address: sondra@borcim.wustl.edu.

Infection of baby hamster kidney (BHK) cells by Sindbis virus leads to a rapid inhibition of host cell protein synthesis, and beginning 4 to 6 h postinfection, the viral subgenomic RNA is the major mRNA being translated. Using Sindbis virus replicons, we showed that expression of only the nsP genes is required for inhibition of host cell protein synthesis (5). This inhibition occurs at the same rate in BHK cells in the absence of the viral structural proteins and under conditions in which the level of the viral subgenomic RNA is too low to be detected. The idea that limiting initiation factors are involved in the inhibition of host protein synthesis in alphavirus-infected cells comes from studies with Semliki Forest virus, an alphavirus related to Sindbis virus (26, 27). In vitro translation assays were used to show that the viral subgenomic RNA can be translated much more efficiently than either host RNA or viral genomic RNA. Experiments carried out with chimeric mRNAs containing only the 5′ untranslated sequences of either the subgenomic or genomic RNAs provided evidence that these sequences from the subgenomic RNA were the ones responsible for the stimulation of translation and that they had a reduced requirement for several initiation factors (1). The increased binding efficiency was proposed to be due to a lack of secondary structure at the 5′ end of the subgenomic mRNA. The importance of the structure of the 5′ untranslated sequences in controlling translation of mRNAs is well established.

In a previous study, we compared the translation of viral subgenomic RNAs produced in cells infected with a variety of Sindbis virus replicons (6). The subgenomic RNAs all had identical 5′ noncoding regions but differed in the number of nucleotides derived from the capsid gene upstream and fused in frame with the *lacZ* gene. RNAs which contained at least 226 nucleotides (nt) from the capsid gene were translated 10-fold more efficiently than were those lacking these sequences. The sequences responsible for enhancing translation have a predicted hairpin structure. These sequences enhance translation in infected cells but not in uninfected cells. Furthermore, a second viral RNA lacking these sequences (structure) is able to suppress translation of this RNA. We had suggested that a hairpin structure presents a barrier to the movement of the ribosomes during translation of the mRNA and had proposed that in infected cells, under conditions in which this subgenomic mRNA is essentially the only RNA being translated, a slowing down in the transit of the ribosomes could give factors present at low concentrations a chance to bind to the translation complex and permit a high level of functional complexes to be formed. In uninfected cells and in infected cells translating two different viral subgenomic mRNAs, a pause in the movement of the ribosomes along the RNA would no longer be an advantage, because required factors are now usurped by other translation complexes (6).

In the present study, we have examined the effects of making mutations that should destabilize the hairpin structure and have shown that these mutations lead to a decrease in translation. To support our contention that a structure is important for translation enhancement, we substituted the naturally occurring sequences with a different sequence, which should also form a stable hairpin. The latter sequence stimulated translation to almost the same extent as the original. We also determined that the position of these sequences influenced their ability to enhance translation.

MATERIALS AND METHODS

Sindbis virus replicons and Sindbis virus. The Sindbis virus replicons SINrep/LacZ, SINrep/447LacZ, SINrep/447Δ81–119LacZ, SINrep/447Δ118–181LacZ, and SINrep/447Δ178–271LacZ were described previously (6). In this nomenclature, 447 refers to the number of nucleotides from the 5′ end that are derived from the Sindbis virus subgenomic RNA. The other replicons used in this study were derived from the cDNA of SINrep/447LacZ (Fig. 1B and C). The replicons are identical from the 5′ terminus of the Sindbis virus genome to nt 7646. This region includes the 5′ untranslated sequences, the nsP genes, the promoter for transcription of the subgenomic RNA, and the 5′ noncoding region of this promoter (the first 49 nt of the subgenomic RNA). The downstream sequences (from nt 7647 to 8037 in SINrep/447LacZ) which code for the amino-terminal part of the capsid gene were mutated as illustrated in Fig. 1B and C. The modified fragments were fused in frame with the complete *lacZ* gene of *Escherichia coli* (3,270 nt). Point mutations and rearrangements in the amino-terminal part of the capsid coding sequence were done by standard PCR methods. All mutations were confirmed by sequencing with the Sequenase version 2 kit (Amersham Life Sciences Inc.). The mutationally altered DNA fragments were then inserted into cDNA plasmids to construct the appropriate replicons. The replicons all terminated with 310 nt of the 3′ noncoding region of the Sindbis virus genome and 37 A residues.

The mutationally altered DNA fragments (from *Bam*HI at nt 7334 in the nsP4 gene of the SINrep cDNAs to the *Aat*II site at nt 7995 in the capsid protein gene) were also used to construct complete copies of Sindbis virus cDNA genomes. The same nsP and envelope cDNAs were also used to construct a nonmutated intact Sindbis virus cDNA genome (designated Toto1101* [see Fig. 4]) to ensure that the only differences in the viral cDNAs were the ones we specifically inserted into the capsid gene. Subsequently, we used PCR to change either the second or the third methionine codon in the capsid protein gene of the Toto/mut1 cDNA from ATG to ATT. Detailed cloning information and the clones are available on request.

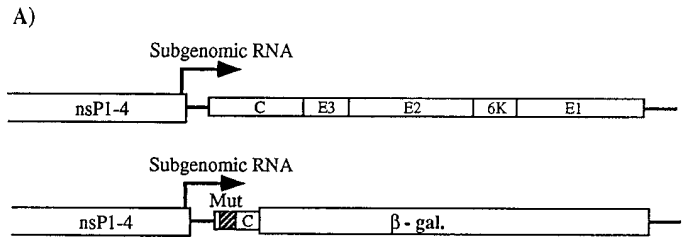
Transcriptions and transfections. The transcription and transfection procedures have been described elsewhere (2, 5). Transcriptions were carried out in the presence of the 5′ cap analog 7^{m}G5′ppp5′G (New England Biolabs). Replicons and helper RNAs were cotransfected, and Sindbis virus RNAs were transfected into BHK-21 cells by electroporation (18). The replicon particles released into the medium were harvested 24 to 30 h postelectroporation. Titers of replicon-containing particles were determined by the cytopathic effect assay on chicken embryo fibroblasts as described previously (5). Infectious virus particles were harvested 16 h postelectroporation. Virus titers were obtained by standard plaque assays.

Computer analysis of the RNA secondary structures. The optimal secondary structures of the RNAs were predicted by using the energy minimization program MulFold (8).

RESULTS

Packaged Sindbis virus replicons. Sindbis virus replicon RNAs contain the genes that code for the proteins required for replication of this RNA and for transcription of the subgenomic RNA. They do not contain the complete set of structural protein genes but can be packaged into extracellular particles if the structural-protein genes are expressed on a different viral genome (2). Replicons are packaged by cotransfection of cells with a defective helper RNA which contains the 5′ and 3′ *cis*-acting signals required for replication and the structural protein genes downstream of the subgenomic RNA promoter. Defective helper RNAs with large deletions in the nsP genes are not packaged to any significant extent. One advantage of using replicons packaged with the viral structural proteins for infection is that, just as with virus particles, essentially all of the cells can be synchronously infected. In contrast, the replicon RNA by itself must be transfected into cells, and many of the methods used for transfection are not as efficient as those used for infection. The method of electroporation which we use does lead to transfection of most of the cells, but it also requires several hours for the cells to recover from this treatment.

The replicons described here contain the *lacZ* gene in frame with and downstream of sequences of the capsid gene (Fig. 1A). The *lacZ* gene has been our choice for these experiments because it is not toxic for BHK cells even when the protein product accumulates to high levels and because β-galactosidase migrates in the polyacrylamide gels in a region that contains very few cellular proteins, so that its concentration in stained gels can be accurately measured. The replicons described here were prepared with the same defective helper



B)

```
                                                        M   N   R   G   F
ATAGTCAGCATAGTACATTTCATCTGACTAATACTACAACACCACCACC ATG AAT AGA GGA TTC -

SINrep/447LacZ
 F   N   M   L   G   R   R   P   F   P   A   P   T   A   M   W   R   P   R   R   R   R   Q   A   A
TTT AAC ATG CTC GGC CGC CGC CCC TTC CCG GCC CCC ACT GCC ATG TGG AGG CCG CGG AGA AGG AGG CAG GCG GCC
                ---------------------------------------------------------------------------------

SINrep/Cmut1/LacZ
 F   N   M   L   G   R   R   P   F   P   A   P   T   A   M   W   R   P   R   R   R   R   Q   A   A
TTT AAC ATG CTC GGa aGa aGg CCt TTt CCa GCa CCa ACT GCC ATG TGG AGG CCG CGG AGA AGG AGG CAG GCG GCC

SINrep/Cmut2/LacZ
 F   N   M   L   G   R   R   P   F   P   A   P   T   A   M   W   R   P   R   R   R   R   Q   A   A
TTT AAC ATG CTC GGC CGC CGC CCC TTC CCG GCC CCC ACT GCC ATG TGG AGG CCG CGG cGc cGc cGa CAa GCG GCC

SINrep/Cmut3/LacZ
 F   N   M   L   G   R   R   P   F   P   A   P   T   A   M   W   R   P   R   R   R   R   Q   A   A
TTT AAC ATG CTC GGa aGa aGg CCt TTt CCa GCa CCa ACT GCC ATG TGG AGG CCG CGG cGc cGc cGa CAa GCG GCC

SINrep/Cnew/LacZ
 F   N   M   L   G   R   R   A   S   R   G   A                                               A   A
TTT AAC ATG CTC GGC CGC CGC gCC tct aga ggc gcg ... ... ... ... ... ... ... ... ... ... ... GCG GCC
                ---------------------------------------------------------------------------------
```

C)

```
SINrep/447Δ81-119LacZ
 F   N   M   L   G                                                   R   R   R   R   Q   A   A
TTT AAC ATG CTC GGC C.. ... ... ... ... ... ... ... ... ... ... ... ... ... .GG AGA AGG AGG CAG GCG GCC

SINrep/447Δ118-181LacZ
 F   N   M   L   G   R   R   P   F   P   A   P   T   A   M   W   R   P
TTT AAC ATG CTC GGC CGC CGC CCC TTC CCG GCC CCC ACT GCC ATG TGG AGG CC. ... ... ... ... ... ... ...

SINrep/C+9/LacZ
          I   K   L
 F   N   M   L   G   R   R   P   F   P   A   P   T   A   M   W   R   P   R   R   R   R   Q   A   A
TTT AAC ATG  CTC GGC CGC CGC CCC TTC CCG GCC CCC ACT GCC ATG TGG AGG CCG CGG AGA AGG AGG CAG GCG GCC
                ---------------------------------------------------------------------------------
        att aaa tta

SINrep/CΔ9/LacZ
                F   G   R   R   P   F   P   A   P   T   A   M   W   R   P   R   R   R   R   Q   A   A
T.. ... ... ..TC GGC CGC CGC CCC TTC CCG GCC CCC ACT GCC ATG TGG AGG CCG CGG AGA AGG AGG CAG GCG GCC
                ---------------------------------------------------------------------------------
```

FIG. 1. (A) Diagrams illustrating the subgenomic RNA of Sindbis virus (top) and the Sindbis virus replicons (bottom) used in this study. The replicons contain the *lacZ* gene in place of the structural protein genes; most of them also contain some part of the capsid protein gene. The structural protein genes of Sindbis virus are designated C for the capsid protein gene and E3, E2, 6K, and E1 for the membrane protein genes. The horizontal arrows show the start of the subgenomic RNAs. Mut indicates the region which was mutationally altered and corresponds to the hairpin structure in the RNA. The nsP genes are indicated as nsP1 to nsP4; they are not drawn to scale. (B and C) Nucleotide and amino acid sequences at the beginning of the subgenomic RNAs in the different replicons. The boxed sequences include the 5′ noncoding region of the RNA and the first five codons of the capsid gene which are identical to those in the Sindbis virus genome and are the same in these replicons. The dashed lines underneath the sequences indicate the regions coding for the original hairpin structure and the new hairpin in the SINrep/Cnew/LacZ replicon. Underlined lowercase letters represent the point mutations and the insertion. The sequences of SINrep/Cmut1/LacZ show the second and third methionine codons in boldface type. These are the codons that were mutated to obtain the data shown in Fig. 4C. The dots indicate the in-frame deletions in the capsid coding sequences of the different replicons. SINrep/447Δ81–119/LacZ contains a deletion from nt 81 to 119 in the subgenomic mRNA, and SINrep/447Δ118–181/LacZ is deleted from nt 118 to 181. They were described previously (6).

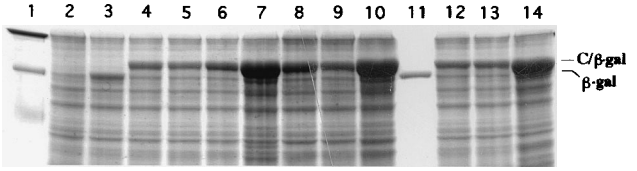


FIG. 2. Accumulation of β-galactosidase and the capsid-β-galactosidase fusion proteins in infected cells analyzed by Coomassie blue staining. BHK-21 cells (5×10^5 cells) were seeded in 35-mm-diameter dishes 4 h prior to infection. They were infected with packaged replicons (multiplicity of infection 25 infectious units per cell) as described previously (5, 6). At 18 h postinfection, cells were harvested, and the proteins were analyzed by gel electrophoresis and detected by Coomassie brilliant blue R staining. Samples equivalent to 5×10^4 cells were loaded onto the gel. Lanes: 1, molecular weight markers; 2, uninfected cells; 3, SINrep/LacZ; 4, SINrep/Cmut1/LacZ; 5, SINrep/Cmut3/LacZ; 6, SINrep/Cmut2/LacZ; 7, SINrep/Cnew/LacZ; 8, SINrep/C+9/LacZ; 9, SINrep/CΔ9/LacZ; 10, SINrep/447LacZ; 11, 1 μg of β-galactosidase; 12, SINrep/447Δ81–119/LacZ; 13, SINrep/447Δ118–181/LacZ; 14, SINrep/447Δ178–271/LacZ. The last of these contains a deletion from nt 178 to 271 (6).

RNA, one that was not (detectably) packaged itself but produced similar titers of replicon particles. Titers of replicon particles ranged between 4×10^8 and 1×10^9 infectious units per ml (5).

Point mutations in the capsid protein gene. We designed two series of point mutations in the hairpin region of the capsid-coding sequence of the SINrep/447LacZ replicon and then combined them (Fig. 1B, SINrep/Cmut1/LacZ, SINrep/Cmut2/LacZ, and SINrep/Cmut3/LacZ). These nucleotide modifications did not change the codon reading and would not affect the amino acid sequence of the capsid protein. They did, however, affect the accumulation of the capsid–β-galactosidase fusion protein (Fig. 2; Table 1). Cells infected with the replicon containing the first set of mutations (SINrep/Cmut1/LacZ) produced sevenfold less fusion protein than did cells infected with the original SINrep/447LacZ replicon. The second set of mutations also affected the accumulation of the fusion protein but not to the same extent (see below). Combining the two sets of mutations was essentially no worse than using the first series of mutations alone. The lower level of fusion protein was almost the same as that obtained in cells that had been infected with the SINrep/LacZ, SINrep/447Δ81–119LacZ, or SINrep/447Δ118–181LacZ replicon. The last two contain deletions in the predicted hairpin region, and we had reported previously that these deletions in the RNA led to a decrease in the translation of the chimeric proteins (6). Cells infected with a replicon that had a deletion downstream of this region (SINrep/447Δ178–271LacZ) produced a high level of fusion protein (Table 1) (6).

Cells infected with these replicons all synthesized similar levels of virus-specific genomic and subgenomic RNAs, so that differences in the chimeric proteins could be attributed to effects on translation and not to levels of RNA (data not shown). We had previously carried out pulse-chase experiments which showed that differences in the amounts of the chimeric proteins were due to differences in translation and not to differences in stability. The chimeric proteins translated in cells infected with the new replicons (SINrep/Cmut1/LacZ, SINrep/Cmut2/LacZ, and SINrep/Cmut3/LacZ) tested here had the same amino acid sequence as did the proteins whose stability we had tested previously.

The predicted RNA secondary structures of the nucleotide sequences from the 5′ end of the subgenomic mRNA to nt 170 are shown in Fig. 3. Only one optimal structure was predicted for this region for SINrep/447LacZ; the hairpin structure has a calculated free energy of −40 kcal/mol (−167 kJ/mol) (Table 2). The first 170 nt of both SINrep/Cmut1/LacZ and SINrep/Cmut2/LacZ could be folded into several different secondary structures with similar calculated free energies (Table 2). For the Cmut1/LacZ RNA, the only indication of a potentially stable stem-like structure was the one starting at nt 111, which was significantly downstream of the hairpin in SINrep/447LacZ, and, as described below, the location of the hairpin was important for enhancement of translation. The secondary structure shown for SINrep/Cmut2/LacZ (one of three) is the one that most closely resembled that predicted for SINrep/447LacZ. The potential of this sequence to fold into a stem-like structure may explain why these mutations did not cause the same decrease in the level of the fusion protein as did the first set of mutations (Table 1).

Substitution of the hairpin structure with another hairpin. Another way to determine if the enhanced accumulation of the fusion protein was due to sequence or to a hairpin structure (Fig. 3, nt 77 to 139) would be to change the sequence but maintain the hairpin. With this goal in mind, we designed a replicon (SINrep/Cnew/LacZ) in which the natural hairpin structure was replaced by a shorter one (Fig. 1B). These nucleotides translate into an amino acid sequence which differs from the naturally occurring capsid protein by 4 amino acids and has a deletion of 11 amino acids (Fig. 1B). The amount of fusion protein produced was almost identical to that in cells infected with SINrep/447LacZ (Fig. 2; Table 1). The new hairpin has a calculated free energy of −28 kcal/mol (−117 kJ/mol) (Table 2), which was lower than the −40 kcal/mol calculated for the original SINrep/447LacZ hairpin and was the same as the value calculated for the mut2 sequence shown in Fig. 3. The new hairpin secondary structure, however, has a predicted structure that includes 12 contiguous G · C base pairs. The original predicted hairpin (447/LacZ) has a total of 22 G · C base pairs; only the first 7 pairs represent a contiguous stretch, while the remainder are interspersed with unpaired bases and A · U base pairs. The hairpin structure shown for Cmut2/LacZ also has several unpaired bases that would disrupt the G · C base pairs and has an extensive loop (nt 86 to 93).

Effect of the position of the hairpin on the efficiency of translation. We designed two replicons in which the position of the hairpin structure was altered. In one (SINrep/Δ9LacZ),

TABLE 1. Accumulation of capsid–β-galactosidase fusion proteins

Replicon	Mutations in capsid-coding region of subgenomic RNA	Accumulation of fusion protein (% of maximum[a])
SINrep/447LacZ	None	100
SINrep/LacZ	No capsid-coding sequence	9, 7
SINrep/Cmut1/LacZ	Mutations destroying the hairpin	14, 16
SINrep/Cmut2/LacZ	Mutations destroying the hairpin	26, 27
SINrep/Cmut3/LacZ	Mutations destroying the hairpin	11, 14
SINrep/447Δ81–119 LacZ	Deletion destroying the hairpin	13, 18
SINrep/447Δ118–181LacZ	Deletion destroying the hairpin	12, 19
SINrep/447Δ178–271LacZ	Deletion downstream of the hairpin	62, 85
SINrep/Cnew/LacZ	New hairpin in capsid gene	93, 95
SINrep/C+9/LacZ	Insertion upstream of the hairpin	30, 32
SINrep/CΔ9/LacZ	Deletion upstream of the hairpin	13, 17

[a] The concentrations of the capsid–β-galactosidase fusion proteins were determined by laser densitometry of the Coomassie brilliant blue R-stained gels (Fig. 2). The data were normalized to the concentration of the cellular actin in each sample. The numbers represent values obtained in two independent experiments.

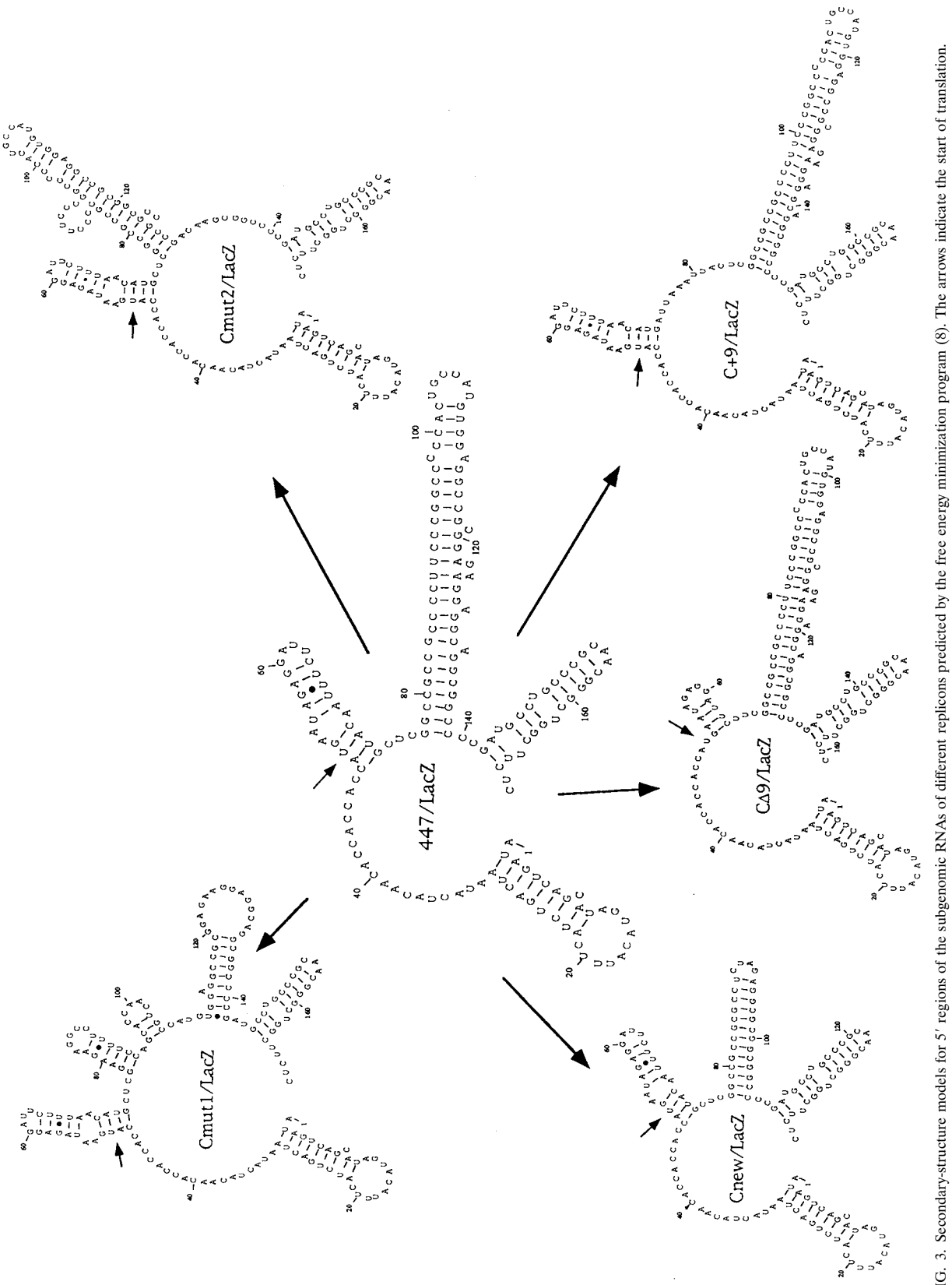


FIG. 3. Secondary-structure models for 5′ regions of the subgenomic RNAs of different replicons predicted by the free energy minimization program (8). The arrows indicate the start of translation.

TABLE 2. Predicted hairpin structures downstream of the initiating AUG codon in the Sindbis virus mRNA

Replicon	No. of secondary structures with similar calculated free energies (1–170 nt)	Calculated free energies (74–140 nt) (kcal/mol)[a]
SINrep/447LacZ	1	−40
SINrep/Cmut1/LacZ	4	−20 to −22
SINrep/Cmut2/LacZ	3	−26 to −28
SINrep/Cnew/LacZ	1	−28

[a] The energies were calculated with the program of Jaeger et al. (8). 1 kcal/mol = 4.184 kJ/mol.

the hairpin structure was moved 9 nt closer to the initiating AUG codon; in the other (SINrep/+9LacZ), the same structure was moved 9 nt downstream (Fig. 1C). These changes did not affect the predicted secondary structure of the sequences (Fig. 3), but they both led to a decrease in the accumulation of the fusion protein (Fig. 2; Table 1). The effect of moving the hairpin closer to the initiating AUG decreased translation to that seen with the mRNA (SINrep/LacZ) that did not contain any capsid nucleotides downstream of the initiating AUG. These results indicate that the location of the hairpin structure was important for its enhancing effect on translation.

Effect of mutations in the hairpin on the synthesis of the capsid protein and on the rate of virus formation. The mutations described above were inserted into the capsid protein gene of the intact Sindbis virus genome to determine how they affected the synthesis of the virus capsid protein and the formation of infectious virus. We analyzed the synthesis of the capsid protein in cells 8 h after infection with the different viruses by labeling cells with [^{35}S]methionine (Fig. 4A and B). Not only was the level of capsid protein decreased when the viral mRNA contained mutations or deletions in the hairpin sequence, but also several forms of the capsid protein were synthesized. The faster-migrating forms could be the result of translation being initiated at each of the two AUGs downstream of the first AUG (see the sequence of SINrep/Cmut1/LacZ in Fig. 1B). The differences in apparent molecular mass detected with the capsid protein, which has a molecular mass of about 30 kDa, would be too small to be seen with the larger LacZ fusion proteins.

To determine if the synthesis of the faster-migrating proteins was due to initiation at the two downstream AUGs, we mutated each of the AUGs in the capsid protein gene of the Sindbis virus containing Cmut1. The second ATG in the cDNA, beginning at nt 71 (amino acid 8), was changed to ATT (coding for isoleucine) in Cmut1/8, and the third ATG, beginning at nt 107 (amino acid 20), was changed to ATT in Cmut1/20. BHK cells were infected with Sindbis viruses containing these nucleotide changes, and the infected cells were labeled with [^{35}S]methionine as described above. When the second methionine codon was changed to isoleucine, the protein band migrating just below the capsid band disappeared (Fig. 4C, lane 3). The fastest-migrating capsid band was no longer seen when the third methionine codon was altered (Fig. 4C, lane 2). These results prove that the mRNA containing the Cmut1 nucleotides in the capsid gene can initiate translation at any one of the first three AUG codons. They suggest that the heterogeneity in the capsid protein bands found in cells infected with the other viruses containing mutations or deletions in the hairpin structure was also due to initiation at downstream AUGs.

The decrease in the synthesis of the capsid protein was correlated with a decrease in the rate of formation of infectious virions (Fig. 5). The rate of formation of Sindbis virus containing the mut1 capsid gene (which has a capsid protein identical in sequence to the original capsid protein) or the Δ81–119 capsid gene was detectably lower than that of the virus containing the normal capsid gene. The growth rates of Sindbis viruses containing either the mut2 capsid gene or the new hairpin in the capsid gene were almost the same as that of the normal virus. (Other viruses were not analyzed.) The decrease in the synthesis of the capsid protein containing mut2 changes did not affect the growth of the virus, and the amino acid changes resulting from the sequence change in the new hairpin did not interfere with the formation of viral nucleocapsids.

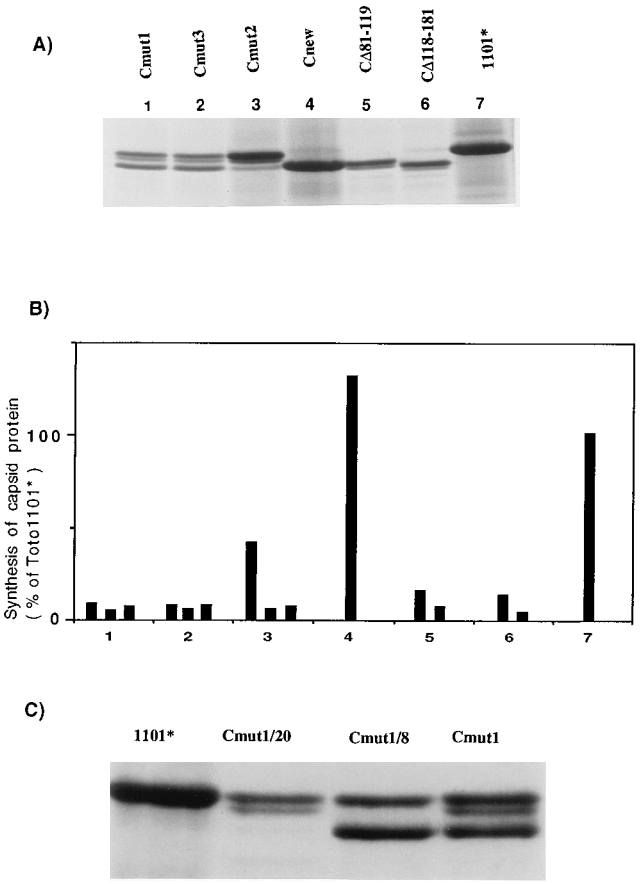


FIG. 4. Synthesis of radioactively labeled capsid protein in cells infected with Sindbis viruses. BHK-21 cells (5×10^5 cells) were seeded in 35-mm-diameter dishes. Four hours later, they were infected with Sindbis virus (multiplicity of infection, 20 PFU per cell). At 7.5 h postinfection, the cells were incubated for 30 min in medium without methionine and then for 30 min after the addition of [^{35}S]methionine (30 μCi/ml). Samples were harvested as described previously (6) and analyzed by polyacrylamide gel electrophoresis. (A) Autoradiogram showing the capsid protein bands from cells infected with Sindbis viruses containing mutations or deletions in the hairpin region of the capsid gene. The virus used for infection is indicated above the appropriate lanes (without the Toto prefix). The Toto/Cnew capsid protein is smaller than the original capsid protein and migrates faster. (B) Quantitation of the radioactive capsid bands with a molecular imager (Bio-Rad Laboratories). The sample numbers are the same as in panel A. Each bar corresponds to a band in the appropriate lane, with the highest-molecular-weight band on the left. (C) Autoradiogram showing the capsid protein bands from cells infected with Sindbis viruses in which either the second or third AUG codon was mutated. The virus used for infection is indicated above the appropriate lanes. In Toto/Cmut1/20, the third AUG (codon 20) was changed to AUU; in Toto/Cmut1/8, the second AUG (codon 8) was changed to AUU. Toto1101* is identical to Toto1101 but newly constructed (see Materials and Methods).

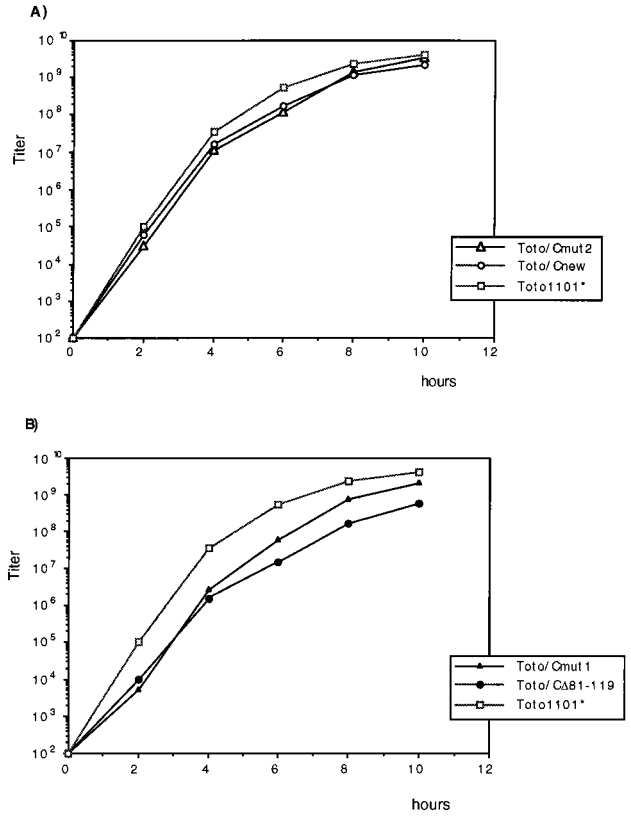


FIG. 5. Growth of Sindbis virus in BHK cells. BHK cells were infected with different Sindbis viruses at a multiplicity of infection of 5 PFU per cell. At the times indicated, the medium was removed from the dish and replaced with new medium. Virus titers were determined for the samples collected at the different times by standard plaque assays on secondary passages of chicken embryo fibroblasts. Data are shown for one of two experiments which gave essentially identical results.

DISCUSSION

The subgenomic RNA of many alphaviruses is translated to high levels in infected cells. This high level of expression was one of the attractions for developing Sindbis and Semliki Forest viruses as vectors for the expression of heterologous genes. In our studies with Sindbis virus as an expression vector, we found that when the heterologous genes were inserted directly downstream of the untranslated 49 nt of the subgenomic RNA, they were translated much less efficiently than the subgenomic RNA containing the virus capsid genes. Translation could be enhanced 10-fold by including the first 226 nt from the capsid gene inserted upstream and in frame with the heterologous gene (6). A deletion encompassing nt 81 to 119 (nt 32 to 70 in the capsid gene) obliterated the enhancement. The sequence between nt 77 and 139 in the subgenomic RNA could be folded into a stable hairpin structure (Fig. 3). The data presented here show the results obtained when we modified this sequence in ways that should affect the structure. We made changes in codons which should destabilize a hairpin structure but not affect the amino acid sequence. Two sets of mutations were made, both of which, separately and together, affected the predicted hairpin structure. They all caused a decrease in the accumulation of the capsid–β-galactosidase fusion protein. The structure predictions and energy calculations indicate that the hairpin had been destabilized by the mutations (Fig. 3; Table 2). The second modification that we made was to change the original sequence into one which should form a stable hairpin structure with 12 contiguous G · C pairs (Fig. 3, Cnew/LacZ). The new sequence was almost as effective in enhancing translation as the original had been. The same modifications were introduced into the intact capsid gene in the complete Sindbis virus genome. The effects on the synthesis of the viral capsid protein and on the rate of formation of infectious virus particles were similar to those that had been seen with the capsid–β-galactosidase fusion protein. These results, taken together, support our conclusion that a stable hairpin structure is the element required for the increased level of translation.

The data presented here also provide some indication that enhancement was occurring at the initiation of translation. Our finding that either an insertion or a deletion of 9 nt decreased the level of translation had first suggested that initiation was being affected, because position should be less critical at an elongation step. The observation that cells infected with Sindbis virus containing mutations or deletions in the hairpin region of the capsid gene synthesized several forms of the capsid protein is also consistent with an effect on initiation. The multiple bands detected on the gel (Fig. 4) were due to translation beginning not only at the first AUG in the sequence, the one normally recognized, but also at downstream AUG codons. Changing the first downstream AUG led to the loss of the protein band migrating just below the authentic capsid band, and altering the second downstream AUG caused the fastest-migrating band to disappear. We had proposed that the hairpin acts as a barrier to the movement of ribosomes, thereby permitting limiting (but undefined) factors to bind to the translation complex. The absence of the hairpin would prevent this pause and decrease the level of translation of the mRNA. A second consequence appeared to be initiation at downstream AUGs.

In the Sindbis virus mRNA, the distance from the first initiating AUG to the hairpin is 27 nt, a length that is spanned by a single ribosome. The translation enhancement that we observed with Sindbis virus replicons has also been reported for the replicons of Semliki Forest virus (22). These two viruses and other alphaviruses do not have significant homology in this region of the genome, but their sequences can form predicted hairpin-like structures located 27 to 29 nt from the initiating AUG (Fig. 6).

A measurement of the pausing of ribosomes during translation in vitro had shown that the centers of stacked ribosomes are about 27 to 29 nt apart and that a ribosome paused at the initiating AUG protects 12 to 13 nt 5′ to that codon (28). Previous RNase protection experiments had demonstrated that the distance between the initiating AUG and the leading edge of the ribosome is 12 to 15 nt (14–17). The ability of a hairpin downstream of the first AUG codon to influence which AUG is selected for the initiation of translation had been noted previously. In vitro translation studies showed that the insertion of a hairpin 14 nt downstream of an AUG in a suboptimal context led to increased recognition of that AUG; the same hairpin at distances of 8 or 32 nt from the AUG did not have the same effect (12). The hairpin was thought to slow scanning, and when it was at this position (14 nt) the 40S ribosomal subunit stalled over the AUG, permitting more time for its recognition (12). If the hairpin in the Sindbis virus mRNA were causing the ribosome to pause, it must not be pausing directly over the AUG, because a distance of 27 nt between the initiating AUG and the hairpin was more effective at enhancing translation than when the hairpin was moved closer to the initiating AUG. Instead, the ribosome may be pausing at a position which delays an oncoming 40S ribosomal subunit from forming a competent 80S initiation complex, al-

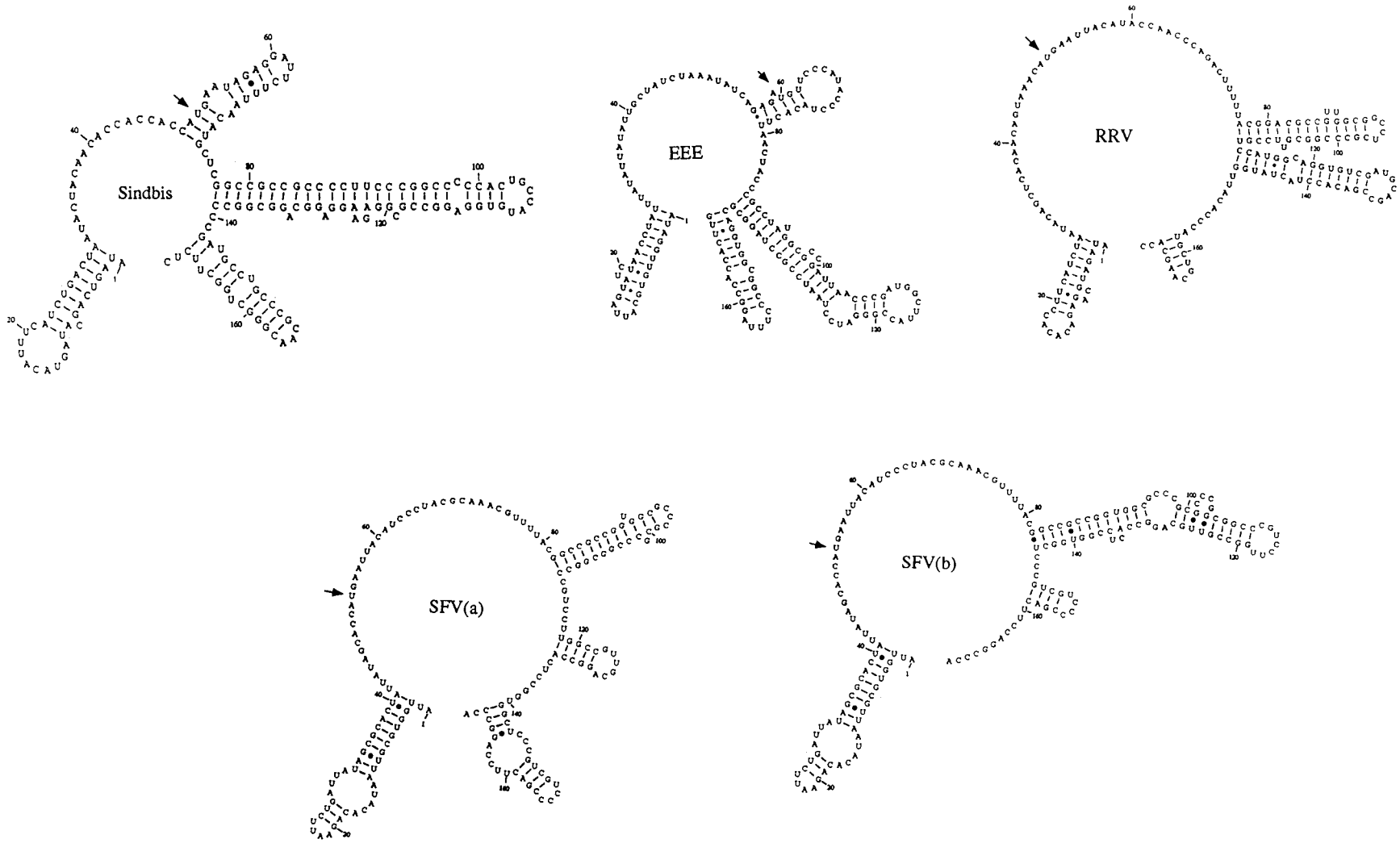


FIG. 6. Optimal secondary structures of the first 170 nt of the subgenomic RNAs of different alphaviruses predicted by the energy minimization program (8). EEE, eastern equine encephalitis virus; RRV, Ross River virus; SFV(a) and SFV(b), two alternative structures with the similar free energies predicted for Semliki Forest virus. The arrows indicate the start of translation.

lowing a necessary factor(s) to interact to form competent complexes. Alternatively, the hairpin structure at this position may be involved in some tertiary RNA structure which could either affect the binding of the 40S ribosomal subunit and initiation factors to the 5′ terminus of the mRNA or impede scanning.

Viruses have devised many strategies for translation of their mRNAs under conditions in which the translation of host cell mRNAs is inhibited (10, 19). The insertion of a hairpin structure at a particular location downstream of the initiating AUG appears to be a way that alphaviruses have evolved to enhance translation of their mRNA and as a consequence to produce high levels of the structural proteins which are needed for virus assembly. This high level of translation requires an intracellular environment in which host cell protein synthesis is inhibited (6). A more detailed understanding of this strategy may provide further insight into how the cellular translational machinery discriminates between host cell and viral protein synthesis. In addition, recognition of this structure as a translational enhancer will be an important factor in obtaining alphaviruses as vectors which express heterologous genes to high levels.

ACKNOWLEDGMENTS

This research was supported by grant AI11377 from the National Institutes of Allergy and Infectious Diseases.

We thank Sergey Dryga for help in sequencing of the recombinant clones.

REFERENCES

1. **Berben-Bloemheuvel, G., M. A. M. Kasperaitis, H. van Heugten, A. A. M. Thomas, H. van Steeg, and H. O. Voorma.** 1992. Interaction of initiation factors with the cap structure of chimaeric mRNA containing the 5′-untranslated regions of Semliki Forest virus RNA is related to translational efficiency. Eur. J. Biochem. **208:**581–587.
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. **Bredenbeek, P., and C. M. Rice.** 1992. Animal RNA virus expression systems. Semin. Virol. **3:**297–310.
4. **Ehrenfeld, E.** 1993. Translational regulation. Semin. Virol. **4:**199–268.
5. **Frolov, I., and S. Schlesinger.** 1994. Comparison of the effects of Sindbis virus and Sindbis virus replicons on host cell protein synthesis and cytopathogenicity in BHK cells. J. Virol. **68:**1721–1727.
6. **Frolov, I., and S. Schlesinger.** 1994. Translation of Sindbis virus mRNA: effects of sequences downstream of the initiating codon. J. Virol. **68:**8111–8117.
7. **Hershey, J. W. B.** 1991. Translational control in mammalian cells. Annu. Rev. Biochem. **60:**717–755.
8. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1990. Predicting optimal and suboptimal secondary structures for RNA. Methods Enzymol. **183:**281–306.
9. **Kääriäinen, L., and M. Ranki.** 1984. Inhibition of cell functions by RNA-virus infections. Annu. Rev. Microbiol. **38:**91–109.
10. **Katze, M. G.** 1993. Games viruses play: a strategic initiative against the interferon-induced dsRNA activated 68,000 Mr protein kinase. Semin. Virol. **4:**259–268.
11. **Kozak, M.** 1986. Regulation of protein synthesis in virus-infected animal cells. Adv. Virus Res. **31:**229–292.
12. **Kozak, M.** 1990. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. Proc. Natl. Acad. Sci. USA **87:**8301–8305.
13. **Kozak, M.** 1992. Regulation of translation in eukaryotic systems. Annu. Rev. Cell Biol. **8:**197–225.
14. **Kozak, M., and A. J. Shatkin.** 1977. Sequences and properties of two ribosome binding sites from the small size class of reovirus messenger RNA. J. Biol. Chem. **252:**6895–6908.
15. **Lazarowitz, S. G., and H. Robertson.** 1977. Initiator regions from the small size class of reovirus messenger RNA protected by rabbit reticulocyte ribosomes. J. Biol. Chem. **252:**7842–7849.
16. **Legon, S.** 1976. Characterization of the ribosome-protected regions of ^{125}I-labelled rabbit globin messenger RNA. J. Mol. Biol. **106:**37–53.
17. **Legon, S., H. D. Robertson, and W. Prensky.** 1976. The binding of ^{125}I-labelled rabbit globin messenger RNA to reticulocyte ribosomes. J. Mol. Biol. **106:**23–36.
18. **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.
19. **Mathews, M. B.** 1993. Viral evasion of cellular defense mechanisms: regulation of the protein kinase DAI by RNA effectors. Semin. Virol. **4:**247–257.
20. **Schlesinger, S.** 1993. Alphaviruses—vectors for the expression of heterologous genes. Trends Biotechnol. **11:**18–22.
21. **Schneider, R. J., and T. Shenk.** 1987. Impact of virus infection on host cell protein synthesis. Annu. Rev. Biochem. **56:**317–332.
22. **Sjöberg, E. M., M. Suomalainen, and H. Garoff.** 1994. A significantly improved Semliki Forest virus expression system based on translation enhancer segments from the viral capsid gene. Bio/Technology **12:**1127–1131.
23. **Sonenberg, N.** 1990. Measures and countermeasures in the modulation of initiation factor activities by viruses. New Biol. **2:**402–409.
24. **Sonenberg, N.** 1994. mRNA translation: influence of the 5′ and 3′ untranslated regions. Curr. Opin. Genet. Dev. **4:**310–315.
25. **Strauss, J. H., and E. G. Strauss.** 1994. The alphaviruses: gene expression, replication, and evolution. Microbiol. Rev. **58:**491–562.
26. **van Steeg, H., A. Thomas, S. Verbeek, M. Kasperaitis, H. O. Voorma, and R. Benne.** 1981. Shutoff of neuroblastoma cell protein synthesis by Semliki Forest virus: loss of ability of crude initiation factors to recognize early Semliki Forest virus and host mRNA. J. Virol. **38:**728–736.
27. **van Steeg, H., M. van Grinsven, F. van Mamsfeld, H. O. Voorma, and R. Benne.** 1981. Initiation of protein synthesis in neuroblastoma cells infected by Semliki Forest virus. FEBS Lett. **129:**62–66.
28. **Wolin, S. L., and P. Walter.** 1988. Ribosome pausing and stacking during translation of a eukaryotic mRNA. EMBO J. **7:**3559–3569.
29. **Xiong, C., R. Levis, P. Shen, S. Schlesinger, C. M. Rice, and H. V. Huang.** 1989. Sindbis virus: an efficient, broad host range vector for gene expression in animal cells. Science **243:**1188–1191.