JOURNAL OF VIROLOGY, Apr. 1990, p. 1639–1647
0022-538X/90/041639-09$02.00/0

Vol. 64, No. 4

Mutagenesis of the Conserved 51-Nucleotide Region of Sindbis Virus

HUBERT G. M. NIESTERS† AND JAMES H. STRAUSS*

Division of Biology 156-29, California Institute of Technology, Pasadena, California 91125

Received 27 October 1989/Accepted 27 December 1989

We have constructed 25 site-specific mutations in a domain of 51 nucleotides in Sindbis virus that is highly conserved among all alphaviruses sequenced to date. These 51 nucleotides are capable of forming two hairpin structures and are found from nucleotides 155 to 205 in Sindbis virus within the region encoding nsP1. Of the mutations, 21 were silent and did not lead to a change in the amino acid sequence encoded. These silent mutations changed not only the linear sequence but also the stability of the hairpins in most cases. Two double mutants that were constructed led to the replacement of one base pair by another so that the linear sequence was altered but the nature of the hairpins was not. All of the mutants with silent mutations were viable, but 19 of the 21 mutants were severely impaired for growth in both chicken and mosquito cells. Compared with the parental virus, they grew slowly and produced virus at rates of 10^{-1} to 10^{-4} times the parental rate. Surprisingly, however, the plaques produced by these mutants were indistinguishable from those produced by the parental virus. Two of the silent mutations, found within the first hairpin structure, produced virus at a faster rate than the parental virus. It is clear that the exact sequence of this region is important for some aspect of virus replication. We suggest that one or more proteins, either virus encoded or cellular, bind to the hairpin structures in a sequence-specific fashion in a step that promotes replication of the viral RNA. Of the mutations that resulted in a change of coding, only one of four was viable, suggesting that the amino acid sequence encoded in this domain is essential for virus replication.

Sindbis virus is a well-studied member of the alphavirus family (for a review, see reference 17). The genome of the virus consists of a single-stranded RNA of positive polarity 11,703 nucleotides in length that is capped at the 5′ end and polyadenylated at the 3′ end. The genomic RNA serves as a messenger for four nonstructural proteins, which are believed to be components of the viral replicase. A full-length minus-strand RNA complementary to the genomic sequence is synthesized during replication and serves as a template for the synthesis of a 26S subgenomic RNA as well as for the synthesis of new viral genomes. This subgenomic RNA is a messenger for a polyprotein that is cleaved posttranslationally to produce the viral structural proteins.

Comparative studies of the alphaviruses have identified four conserved sequences and/or structures that may serve as promoters for various functions during replication of the virus. The conserved 19-nucleotide sequence adjacent to the poly(A) tail at the 3′ terminus of the RNA is postulated to serve as a promoter for the initiation of minus-strand synthesis (14). There is a 21-nucleotide sequence in the junction region between the nonstructural and structural protein coding regions which contains the initiation site for the subgenomic 26S RNA and which acts as a promoter for subgenomic RNA synthesis (13; R. Levis, Ph.D. thesis, Washington University, St. Louis, Mo., 1988). Near the 5′ terminus of the genome, two conserved regions are found. At the very 5′ end is a conserved stem-loop structure that might serve as a promoter (in the minus-strand RNA) for the initiation of plus-strand RNA synthesis (13). Finally, there is a very highly conserved sequence and structure within the coding information for the nonstructural protein nsP1, the so-called 51-nucleotide region, whose function is not known (13). It has been postulated that this region might serve as an accessory promoter element for initiation of minus-strand RNA, because no minus-strand copy of the subgenomic 26S RNA is produced and thus non-26S RNA sequences are implicated in replication (17). Although no direct evidence for any function of this 51-nucleotide region has been found, removal of it reduces, but does not abolish, the replication efficiency of defective interfering particles (11).

The availability of a full-length cDNA clone of Sindbis virus (15) has facilitated the analysis of these conserved sequences of alphaviruses by using a molecular genetic approach (9). In this paper, we report the results of introducing a number of specific mutations in the 51-nucleotide region of Sindbis virus. The newly constructed plasmids were transcribed in vitro by SP6 RNA polymerase to produce full-length transcripts, and an attempt was made to recover virus. All of the mutations that did not lead to a change of coding resulted in viable viruses whose phenotypic properties were examined, whereas three of the four mutations resulting in a coding change were lethal.

MATERIALS AND METHODS

Cells and viruses. All viruses were propagated on secondary chicken embryo fibroblast (CEF) monolayers, as previously described (7). The cells were maintained in Eagle minimal essential medium supplemented with 3% fetal calf serum. The mosquito cell line C6/36 (*Aedes albopictus*) was obtained from the American Type Culture Collection, Rockville, Md. The wild-type parental virus used in all the experiments was derived from the full-length cDNA clone pToto50 of Sindbis virus (15), in which a unique *Xho*I restriction site following the 3′ terminal poly(A) tract and derived from the vector sequences of cDNA clone pToto1101 was used for runoff transcription. The resulting construct was called pToto51 and contained all the se-

* Corresponding author.
† Present address: Diagnostic Centre SSDZ, 2600 GA Delft, The Netherlands.

quences from the heat-resistant small plaque (HRSP) strain of Sindbis virus.

Site-directed mutagenesis and the construction of recombinant plasmids. All mutants were constructed in a similar way. Briefly, a 1,593-base-pair *Sac*I-*Pst*I fragment from the full-length clone pToto51 (nucleotides 13552 to 1507 [numbering in the clone begins with the first nucleotide of Sindbis virus cDNA]) was subcloned into bacteriophage M13mp18. After two rounds of plaque purification, the phage DNA was enriched in uracil by two rounds of infection in *Escherichia coli* BW313 (Ung^- Dut^-) (10). Phage single-stranded DNA (0.5 μg) was used as a template for in vitro mutagenesis, using 20 pmol of a phosphorylated synthetic oligonucleotide (15 or 16 nucleotides in length) (10, 21). Clones containing the desired mutation were identified by single-lane nucleotide sequencing (16) and digested with *Pst*I and *Sac*I to isolate the mutagenized fragment. The fragment was purified from low-melting-point agarose and recloned into the pToto51 background in a three-fragment ligation. The full-length cDNA clone was checked by double-stranded sequencing (1) to confirm the presence of the mutation. All virus mutants were named according to the nomenclature proposed by Kuhn et al. (9).

Transfection and plaque assays. Full-length mutagenized derivatives of pToto51 were prepared for transcription by linearization with the restriction enzyme *Xho*I for the production of runoff transcripts. RNA transcripts were synthesized in vitro with SP6 RNA polymerase essentially as described by Rice et al. (15). The resulting transcription mixture was diluted in phosphate-buffered saline (2) and used directly for transfection into secondary CEFs by using DEAE-dextran. After incubation for 1 h at room temperature, the cells were overlaid with 1% agarose (SeaKem ME, FMC Corp., Marine Colloids Div., Rockland, Maine) in minimal essential medium containing 3% fetal calf serum. Plaques were visualized by staining with neutral red after 48 to 72 h of incubation at 30 or 40°C. Virus from a single plaque was eluted overnight at 4°C in 0.5 ml of phosphate-buffered saline and used to infect secondary CEFs in a T25 flask for 10 to 14 h at 30°C. The resulting seed stock was often of low titer, and a working stock was prepared by infecting secondary CEFs with the seed stock for up to 20 h at 30°C.

One-step growth analysis. The ability of the mutants to replicate was examined in differential growth curves. Monolayers of secondary CEFs or C6/36 cells were infected with virus at a multiplicity of infection of 10 (as determined by a plaque assay of the infecting stock in CEFs) in 250 μl of phosphate-buffered saline containing 1% fetal calf serum. After absorption at room temperature for 1 h, the inoculum was removed and the monolayers were washed twice with phosphate-buffered saline. Finally, 1 ml of prewarmed medium was added and the cells were incubated at 30°C. The culture fluid was removed every hour, the cells were washed, and fresh medium was added. The samples were then assayed for plaque-forming virus by titration on monolayers of CEFs at 30°C.

RESULTS

Construction of nucleotide substitution clones and isolation of viral mutants. The 51-nucleotide region within the nonstructural protein nsP1 is highly conserved among the alphaviruses sequenced to date (Fig. 1A). This domain, which encompasses nucleotides 155 to 205 of Sindbis virus, is capable of forming two hairpin structures whose calculated free energy is such that these structures would be expected to exist stably in solution (Fig. 1B). The stem of the first hairpin structure is identical in all alphaviruses, in terms of both sequence and length, whereas there is considerable variability in the loop. The stem of the second structure is identical in all alphaviruses as to length but there is some variability in sequence. It is noteworthy that O'Nyong-nyong virus has a UA base pair replacing a CG in the second stem and that the substitution of A at nucleotide 206 (numbering for Sindbis virus) requires the substitution of C for the otherwise invariant U at nucleotide 182 to keep the stem length unchanged. At the boundaries of the stems of the structures, there is considerable variability in nucleotide sequence (Fig. 1A), suggesting that the structure rather than the nucleotide sequence per se is important. It is also clear that the amino acid sequence of this region is highly conserved, but as we have discussed previously, conservation of amino acid sequences in alphaviruses is not normally accompanied by nucleotide sequence conservation because of the degeneracy of the genetic code (17). Nucleotides that can be changed without a change in coding are indicated in Fig. 1B. The conservation in both sequence and structure of this domain suggests that it may have an important function in viral replication (13, 17) and that nucleotide changes might have a drastic effect on the growth of the virus. However, deletion of this sequence from defective interfering (DI) RNAs (11) was not lethal but led only to a reduction in the growth rate of the DI particles.

The availability of a full-length cDNA clone of Sindbis virus from which infectious RNA can be transcribed in vitro (15) makes it possible to analyze this sequence and structure with a molecular genetic approach. Furthermore, the effect of changes can be studied directly in the virus itself. The strategy we used was to substitute a large number of nucleotides within this conserved domain, most of which would disrupt the hairpin structure. Most substitutions were chosen to leave unaltered the amino acid sequence encoded by this region (the nucleotides that can be changed in this way are shown in Fig. 1B). After isolation of the mutant clone, the mutagenized fragment was used to replace the corresponding fragment in the full-length Sindbis virus cDNA clone. RNA transcripts were made in vitro with SP6 RNA polymerase and introduced into susceptible chicken cells by transfection with DEAE-dextran, and the viability of the mutant was tested by plaque assays at both 30 and 40°C to identify possible temperature-sensitive mutants. In every case, three serial dilutions of the RNA transcripts were tested and compared with the results in the same experiment with wild-type RNA transcribed from the parental clone pToto51, and for all viable mutants the specific infectivity of the RNA was within threefold that of wild-type RNA. Thus, there was no indication that revertants, same site or second site, were being selected, and we believe that all results obtained are the result of the introduction of the primary mutation (this analysis is complicated by the presence of two plaque sizes in two of the mutants, as described below). Stocks of viable mutants were produced, and most mutants were tested for their growth properties in one-step differential growth curves in both chicken cells and C6/36 mosquito cells in comparison with the wild-type virus recovered from the parental cDNA clone pToto51.

Nucleotide substitutions within the first hairpin structure. The 11 mutations made in the first hairpin structure are summarized in Table 1. All are located within the stem region, which is invariant in sequence among all alphaviruses sequenced to date, and all but three were silent mutations. The calculated free energy of the resulting stem-

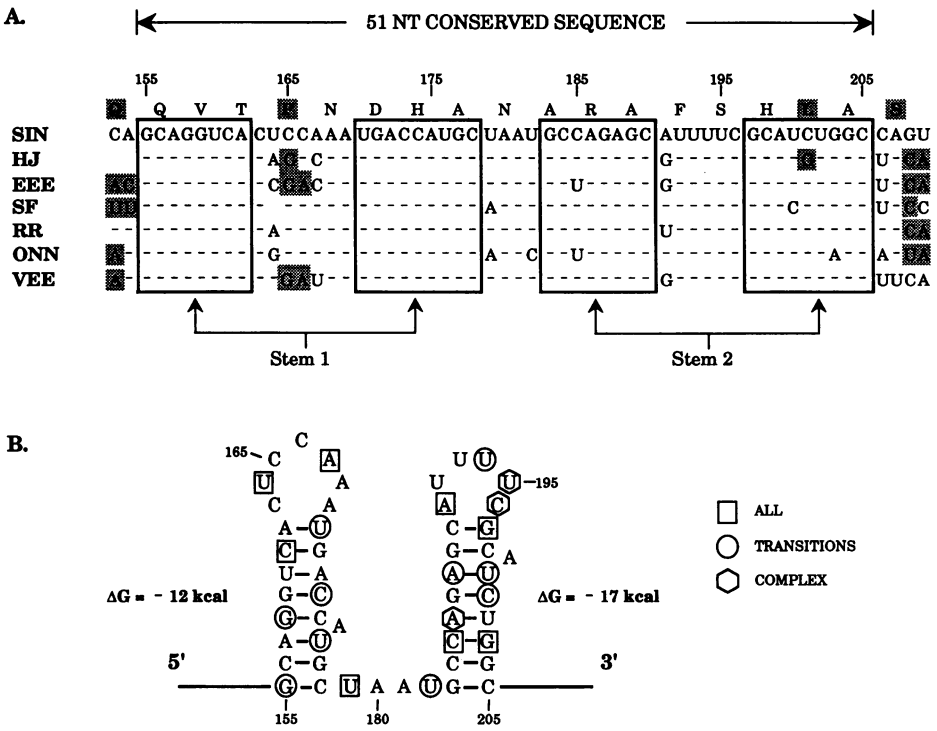

FIG. 1. The 51-nucleotide conserved sequence element of the alphaviruses. (A) The linear nucleotide sequences of the seven alphaviruses sequenced through this region to date are shown. Nucleotides represented with a dash are the same as in the Sindbis virus sequence in the first line; whenever a nucleotide differs from the Sindbis virus sequence, it is shown. Nucleotide changes that are silent are unshaded; shaded nucleotides result in a change in coding from the Sindbis virus sequence. Nucleotide numbers refer to Sindbis virus starting from the 5′ end of the RNA, and the amino acids encoded in the sequence of Sindbis virus are shown above the nucleotide sequence. Amino acids that are invariant in the seven viruses are unshaded; shaded amino acids differ in one or more of the other viruses. Two additional nucleotides are shown at the 5′ end of the conserved element, and four are shown at the 3′ end to illustrate that conservation terminates abruptly at the boundaries. The boxed nucleotides are those found in the base-paired stems of the stem-loop structure in panel B. SIN, Sindbis virus; HJ, Highlands J virus; EEE, Eastern equine encephalitis virus; SF, Semliki Forest virus; RR, Ross River virus; ONN, O'Nyong-nyong virus; VEE, Venezuelan equine encephalitis virus. The data for SIN, HJ, EEE, and SF are from Strauss and Strauss (17), the RR sequence is from Faragher et al. (3), VEE is from Kinney et al. (8), and the ONN sequence is from Levinson, Strauss, and Strauss (Virology, in press). (B) A possible secondary structure of the conserved 51-nucleotide region of Sindbis virus. The free energies were calculated by the method of Tinoco et al. (18). The nucleotides that can be substituted in the 51-nucleotide element without altering the encoded amino acid sequence are indicated. Symbols: □, nucleotides that can be changed to any of the other three nucleotides; ○, nucleotides for which a purine can be replaced by the other purine or a pyrimidine can be replaced by the other pyrimidine; ⬡, nucleotides that can be changed by other than a transition or that require that one or both of the other two nucleotides within the same codon also be changed. One calorie = 4.184 J.

loop structure is indicated. As shown, all of the mutants with silent mutations were viable, and none was temperature sensitive by plaque assays on chicken cells. The mutant P1-158U with the missense mutation which results in the substitution of Gln-33 by His was also viable. The two missense mutations at nucleotide 176, both of which result in the substitution of His-39 by Gln, were lethal, suggesting that this amino acid replacement is not tolerated and that this protein domain performs an essential function during virus replication.

The two single mutations which would most destabilize the first hairpin structure were P1-158A and P1-173U, which reduce the negative free energy to −3.8 and −4.8 kcal/mol, respectively (approximately −16 and −20 kJ/mol, respectively). Both constructs resulted in viable virus at both 30 and 40°C. However, a mixed plaque population was present in the primary transfection assay for both viruses. This mixed plaque population consisted of very large plaques (five to six times the size of plaques produced by wild-type virus) and plaques somewhat smaller than those of the wild type, with the small plaques in about a fivefold excess. Plaque purification led to homogeneous stocks of the large and small plaque variants, which were used to infect secondary chicken cells or C6/36 mosquito cells at a multiplicity of 10 in a differential one-step growth experiment in which these mutants were compared with the wild type for their ability to produce virus (Fig. 2). Note that the cumulative yield of virus was not determined. Instead, in these and all other growth curves, the media were changed every hour in order to determine the amount of virus released per hour at different times, and thus the curves show the rate of virus release as a function of time.

All four mutants grew well in both cell types, but for both P1-158A and P1-173U, the small plaque strains grew better than the large plaque strains in both cell types (thus the very large difference in plaque size is not correlated with the titer of virus produced). Intriguingly, P1-158A.SP grew even better than the wild type in both cell lines. In chicken cells, virus was produced at faster rates than for the wild type early in infection, but the rates of virus release for the mutant and wild type were the same after 8 h. In mosquito cells, the rate of virus production by the P1-158A.SP mutant was about threefold greater than that of the wild type throughout the infection cycle. Mutant P1-173U.SP also

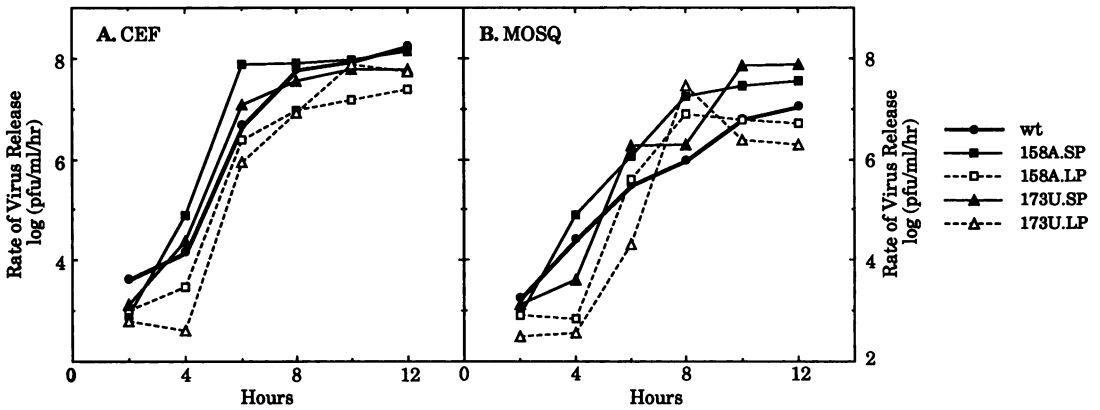

FIG. 2. Differential one-step growth curves of the small and large plaque variants of P1-158A and P1-173U in chicken (A) and mosquito (C6/36) (B) cells at 30°C. Monolayers of cells were infected with wild-type viruses or one of the four mutant viruses at a multiplicity of 10. After absorption for 1 h at room temperature, the inoculum was removed, the monolayers were washed, and the cells were covered with Eagle medium. The medium was replaced every hour with fresh medium, and titers of selected samples were determined by plaque assay on chicken monolayers at 30°C. The data are the average numbers of two independent experiments.

outgrew the wild type in mosquito cells but not in chicken cells. Sequence analysis of the genomic RNA of all four mutants between nucleotides 151 and 240 showed that the original mutation was still present and that no other change was present in this region. The source of the plaque size variation is thus obscure. The large and small plaque strains both grew well in chicken cells, and neither possessed a significant selective advantage over the other. Both were present in the original transfection mixture, and because the specific infectivity of the RNA was the same as that of wild-type RNA within the margin of experimental error, there is no reason to believe that revertants were being selected.

In order to investigate whether the mutations P1-158A and P1-173U together would have more effect on viral replication or plaque morphology than either one alone, the double mutant P1-158A173U, which led to further destabilization of the first hairpin, was constructed. This mutant was viable and grew fairly well in chicken cells, but there was a delay in initial virus production and virus were produced at a slightly reduced rate late in infection. In mosquito cells, however, the mutant produced virus at only about 10% of the wild-type rate throughout the infection cycle (Fig. 3). Thus, although either mutation alone has a little or even a slightly enhancing effect upon virus growth, the double mutant is significantly impaired for growth.

The effect of changing the C at position 161 to any of the other three nucleotides, which could be done without alteration of the amino acid sequence, was also tested. As shown in Table 1, a change to G or A destablizes the hairpin (ΔG = −5.6 kcal [~−23 kJ]/mol) more than a change to U (ΔG = −7.6 kcal [~−32 kJ]/mol) because of the ability of U to participate in a GU pair. All three mutants were viable. Growth curves of these three mutants in both chicken and mosquito cells are shown in Fig. 4A and B. All were significantly impaired in their ability to grow, showing a considerable lag in the release of virus and producing virus at a significantly lower rate. There is a clear difference in the three, and the ability to replicate did not correlate with the

TABLE 1. Mutations in the first hairpin structure (nucleotides 155 to 178)

Mutant[a]	Free energy (kcal/mol)[b]	Plaque phenotype[c]
P1-158A	−3.8	Large and small plaques[d]
P1-158U(Q33H)	−8.8	Wild type
P1-161A	−5.6	Wild type
P1-161G	−5.6	Wild type
P1-161U	−7.6	Wild type
P1-170C	−10.8	Wild type
P1-173U	−4.8	Large and small plaques
P1-158A173U	−3.6	Wild type
P1-176C	−8.6	Wild type
P1-176A(H39Q)	−8.6	Lethal
P1-176G(H39Q)	−8.6	Lethal

[a] Mutants are named for the change(s) made. Thus P1-158A has A as nucleotide 158 of the Sindbis virus genome in the region encoding nsP1 (the wild-type nucleotide is G, as shown in Fig. 1). All of the mutants have silent mutations except for three, for which the change in coding is indicated at the end of the name by a parenthesis which contains the single-letter for the original amino acid, the position in nsP1, and the single-letter code for the mutant amino acid. Thus, substitution 158U leads to the substitution of Gln-33 by His in nsP1.

[b] Calculated according to Tinoco et al. (18). One calorie = 4.184 J.

[c] Determined in CEF monolayers.

[d] Two sizes of plaques were present in the original transfection assays. Large plaques were five- to sixfold larger than wild-type plaques, whereas small plaques were slightly smaller than those of the wild type. Variants producing both plaque types were stable on plaque purification; the large plaque variant is indicated by .LP and the small plaque is indicated by .SP after the mutant name.

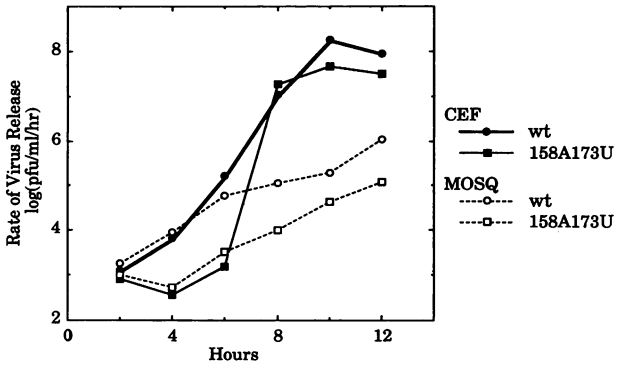

FIG. 3. Differential one-step growth curves of wild-type Sindbis virus and the mutant P1-158A173U in chicken and mosquito cells. Methods were the same as those described in the legend to Fig. 2.

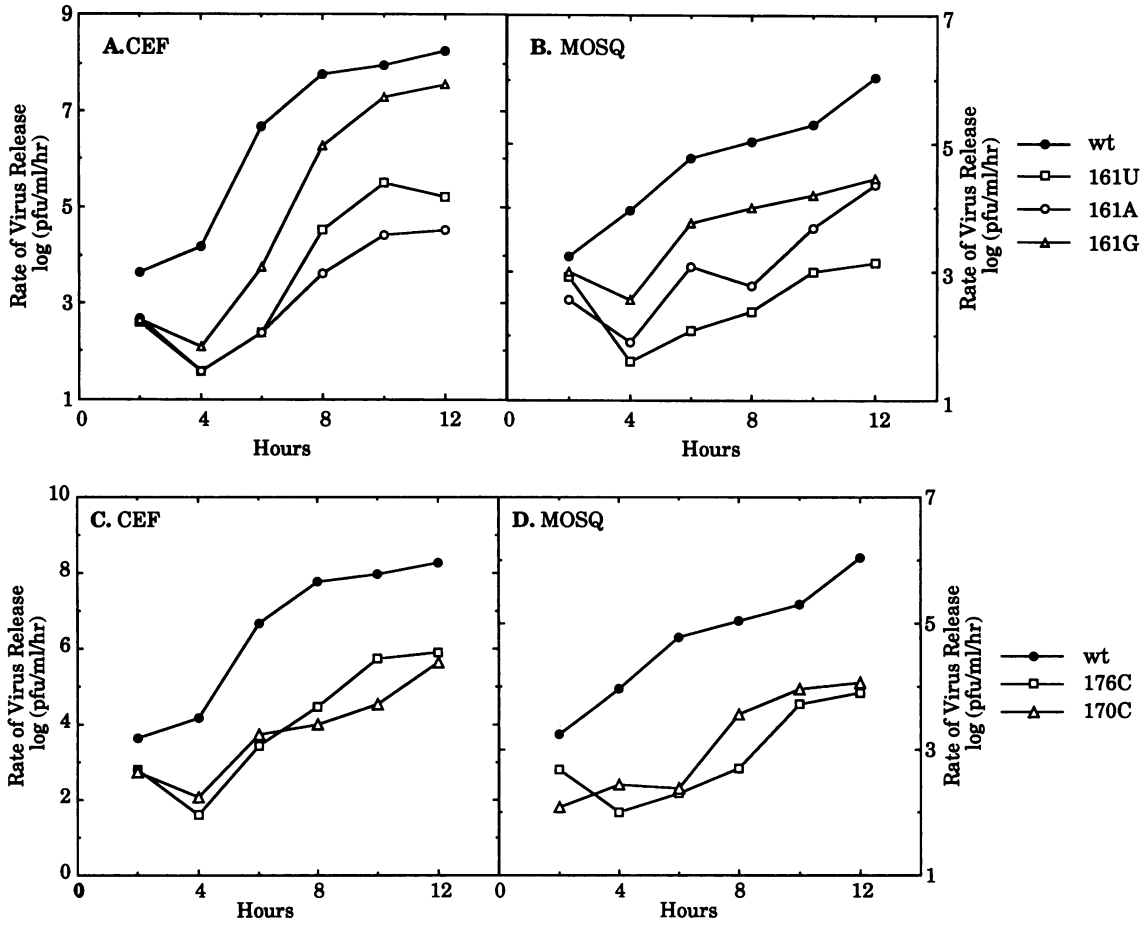

FIG. 4. Differential one-step growth curves for wild-type Sindbis virus and several mutants with mutations in the first stem structure. Growth curves are shown for the wild-type virus and for mutants P1-161A, P1-161G, and P1-161U grown in CEFs (A) and C6/36 cells (B) and for the wild type and mutants P1-170C and P1-176C grown in CEFs (C) and C6/36 cells (D). Methods were the same as those described in the legend to Fig. 2 except that for P1-170C, the growth curve represents the results of a single experiment rather than the average of two experiments.

stability of the hairpin structure. P1-161G grew best in both cell types, whereas 161A was worst in chicken cells and 161U was worst in mosquito cells. P1-161A produces virus at a rate four orders of magnitude less than the wild type in chicken cells (but, as noted in Table 1, its plaque-forming properties are the same as those of the wild type). P1-161U produces virus at a rate three orders of magnitude less than the wild type in both cell lines. Thus, it seems clear that the stability of the hairpin structure per se is not the determining factor, but either the structure of the stem of the hairpin or the linear sequence of this domain is crucial for replication.

We also tested the effects of nucleotide substitutions at positions U170 and U176, both of which could be changed into a C without altering the amino acid sequence (Table 1). These mutants were also viable, and growth curves in chicken and mosquito cells are shown in Fig. 4C and D. Both mutants were significantly impaired in their ability to grow, producing virus slowly and at rates three orders of magnitude less than the wild type in both cell types.

The effect of changing U176 to G or A, which results in the replacement of His-39 in nsP1 by Gln, was also tested. Both mutations were lethal, probably as a result of the amino acid change.

Finally, the effect of one other nucleotide substitution that resulted in an amino acid change was tested. Bulged As have been found to be important in the recognition of some RNAs by proteins (20), and bulged As are found in both stems in the 51-nucleotide sequence element. We wanted to test the effect of drawing the extrahelical A at nucleotide 175 into the stem. There is no simple way to do this without altering the amino acid sequence, but substitution of the G at position 158 with U, which results in the substitution of Gln-33 in nsP1 by His, will draw A175 into the stem and cause C174 to become extrahelical. Mutant P1-158U(Q33H) was viable, and growth curves for it are shown in Fig. 5. This mutant is significantly impaired in its ability to grow, but no more so than are many of the other (silent) mutants in this region. Because of the change in coding, it is impossible to determine at this time whether the effect is at the level of amino acid sequence or of nucleotide sequence, but it does seem clear that the bulged A is not essential for virus replication.

Nucleotide substitutions in the second hairpin structure. In contrast to the stem of the first hairpin structure, the stem of the second hairpin (nucleotides 183 to 205) is not absolutely conserved among alphaviruses (Fig. 1A). A number of differences are seen, some of which lead to a less stable hairpin structure. However, in each case, the second hairpin structure is more stable than is the first. It is interesting that in the Sindbis virus structure, nucleotides C185 and G203 form a CG base pair which was also present in the equivalent

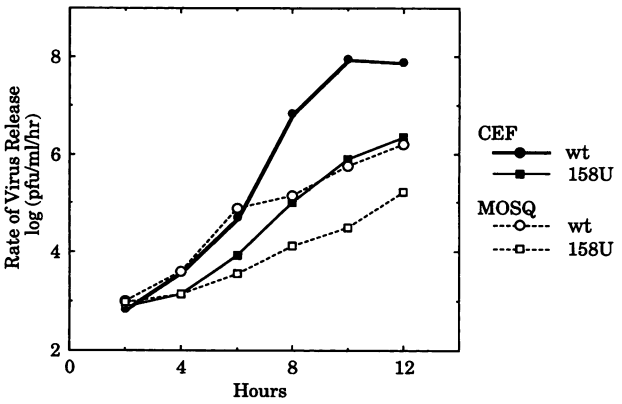

FIG. 5. One-step differential growth curves of the mutant P1-158U(Q33H) in chicken and mosquito cells. Methods were the same as those described in the legend to Fig. 2.

position in four of the other alphaviruses, but in Eastern equine encephalitis virus, this base pair is replaced with UG and in O'Nyong-nyong virus it is replaced with UA (Fig. 1). These substitutions have little effect on the structure that can be formed. The changes at C201 in Highlands J virus and U200 in Semliki Forest virus would be more disruptive.

We have tested the effects of a number of nucleotide substitutions in this structure, which are summarized in Table 2. All of the changes but one were silent, and almost all of the changes were chosen to disrupt the stem structure (but in all cases, the residual stability of the hairpin is significant). One double substitution, P1-185G203C, substituted a CG base pair with a GC base pair. A second double substitution, P1-188G200C, changed an AU base pair to GC. All of the silent mutants were viable, and all produced plaques with morphologies similar to that of the wild-type virus. No mutant was temperature sensitive by plaque assay in chicken cells.

Seven of the mutants were analyzed in differential one-step growth curves (Fig. 6). All of the mutants were significantly impaired in growth, both in chicken cells and in mosquito cells, producing virus at approximately 1% of the wild-type rate. Results for two pairs of mutants are shown in Fig. 6A and B. Mutant P1-185G disrupted a CG base pair, making the hairpin less stable ($\Delta G = -7.7$ kcal [~ -32 kJ]/mol), whereas the double mutant P1-185G203C replaced the CG base pair with a GC pair and restored the stem. Both mutants grew poorly. In chicken cells, these two mutants grew very similarly, with the double mutant growing just slightly better. In mosquito cells, the single mutant grew almost 10-fold better than the double mutant. Mutants P1-200C and P1-188G200C form a similar pair, but the AU base pair in the stem is replaced with a GC base pair in the double mutant. Again, both mutants grew poorly in both cell types. Clearly it is not the stability of the stem structure per se that determines the function, but the sequence of the stem is important as well.

TABLE 2. Mutations in the second hairpin structure (nucleotides 183 to 205)

Mutant[a]	Free energy (kcal/mol)[b]	Plaque phenotype[c]
P1-185G	−7.6	Wild type
P1-186C	−10.4	Wild type
P1-188U(R43S)	−17.0	Lethal
P1-197A	−10.4	Wild type
P1-197C	−10.4	Wild type
P1-197U	−10.4	Wild type
P1-195A196G197U	−10.4	Wild type
P1-200C	−13.4	Wild type
P1-188G200C	−20.8	Wild type
P1-201U	−12.4	Wild type
P1-203A	−7.6	Wild type
P1-203C	−7.6	Wild type
P1-203U	−7.6	Wild type
P1-185G203C	−17.0	Wild type

[a] Mutants are named as in Table 1. All are silent except for P1-188U(R43S).
[b] Calculated according to Tinoco et al. (18). One calorie = 4.184 J.
[c] Determined in CEF monolayers.

Growth curves for mutants P1-197A and P1-201U and for the triple mutant P1-195A196G197U, in which the serine codon UCG is replaced by AGU, are shown in Fig. 6C and D. This last mutant also tests the effect of changing a conserved UC in the second loop. All three mutants grew poorly in both cell lines, with the triple mutant growing as well as or better than many of the single mutations tested. Because the triple mutant has a change in the stem in addition to the two changes in the loop, the relative contribution of changes in the stem and in the loop to the phenotype of this mutant cannot be determined, but it is clear that changes in the loop structure do not have a more severe effect upon virus replication than do changes in the stem. In chicken cells, these last three mutants produced virus at about 1% of wild-type rate. In mosquito cells, P1-201U grew significantly less well than other mutants tested, producing only 0.1% of wild-type levels at all times tested.

Mutants P1-197C, P1-197U, P1-203A, P1-203C, and P1-203U were also constructed. Although all were viable and produced normal-size plaques, all grew very poorly such that the resulting stocks were low titer, and growth curves were not attempted because of the inability to produce an input multiplicity of 10. It seems clear, however, that these mutants are all significantly impaired in their ability to grow, at least in chicken cells.

We also tested one coding change mutant. Mutant P1-188U(R43S) was constructed in an attempt to test the importance of the bulged A in the second hairpin, as there was no simple way to draw this A into the hairpin without a coding change. This mutant was not viable, as shown by transfection attempts with three independent constructs. Even after blind passaging in liquid culture, no virus could be detected. It seems likely that the lethality is due to the amino acid substitution, although we cannot rule out the possibility that the extrahelical A is essential for virus replication.

DISCUSSION

A number of conserved nucleotide sequences have been found in alphavirus RNAs which are postulated to function as promoter elements in virus replication (for a review, see reference 17). Although there are probably other important sequence elements in the viral genomes (for example, a domain important for interaction with the viral capsid protein has recently been identified [19]), it seemed likely that investigation of the role of these short conserved sequence elements would be useful for our understanding of virus replication. We have begun with a mutational analysis of the 3′ nontranslated region (9) and the 51-nucleotide region reported here. Most of the 25 different mutations in the 51-nucleotide sequence element examined were silent so as

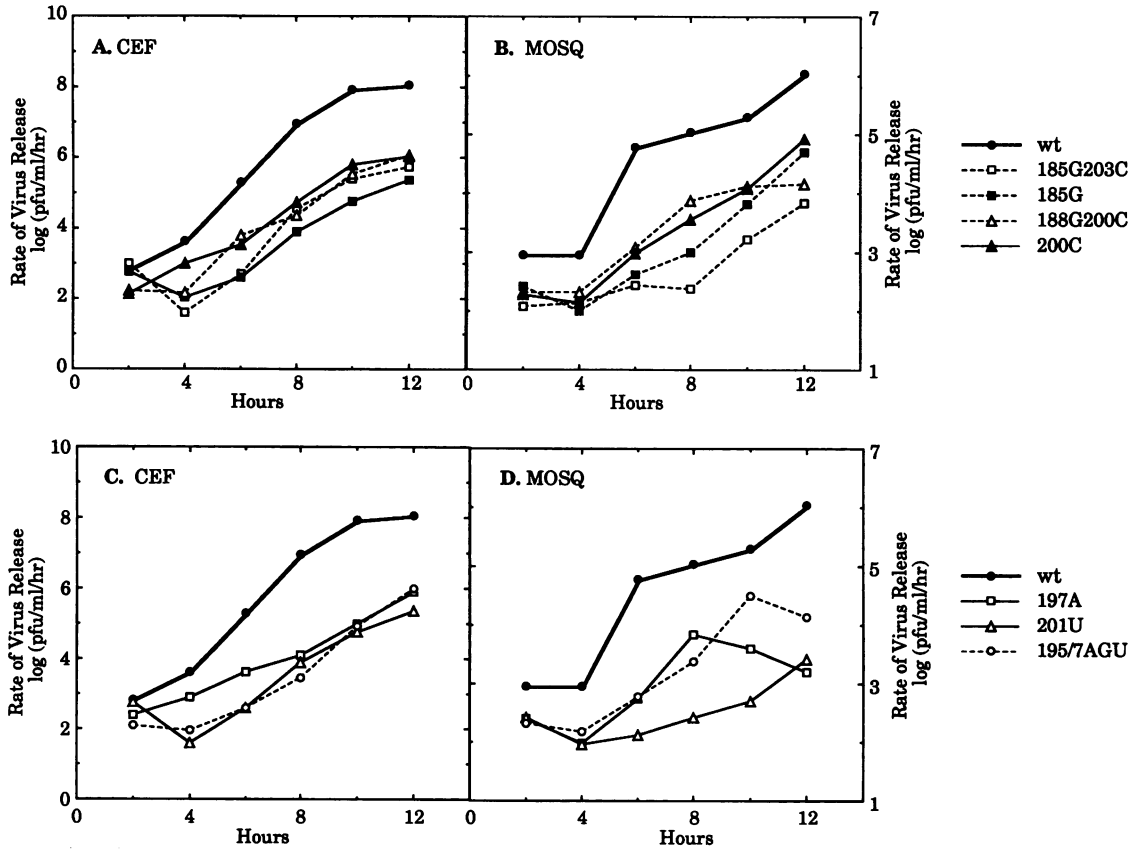

FIG. 6. Differential one-step growth curves of the mutants in the second hairpin structure. Methods were the same as those described in the legend to Fig. 2, except that the growth curves for P1-197A, P1-200C, P1-201U, and P1-195A196G197U (referred to as 195/7AGU in the figure) are from a single experiment rather than the average of two independent experiments.

not to complicate the analysis with changes in protein sequence, and most of these would disrupt one or the other of the hypothetical stem-loop structures, although two double mutants were constructed that replaced base pairs in the second stem with another base pair. The results make it clear that the sequence in this region is quite important in some way for virus replication, as predicted from its conservation.

The fact that two mutants, P1-158A and P1-173U, grow better than the wild type in tissue culture suggests that the wild-type viral sequence is a compromise. We presume that these two mutants are less well adapted for persistence in nature even though they replicate at faster rates in tissue culture cells. It is possible that they grow less well than the wild-type virus in the target tissues of their natural hosts. If these viruses replicated less efficiently in gut cells or salivary gland cells of the mosquito, for example, they would not be transmitted as efficiently. A second intriguing possibility, which can be explored in whole animals, is that strains that replicate too efficiently in mosquitoes might be selected against in nature because of lowered vector viability.

These two mutants that grow well are also strange in that both large and small plaque variants are present in the primary transfection plate. But the large and small plaque variants grow similarly, and thus neither has a significant selective advantage over the other. Furthermore, the specific infectivity of the RNA transcript is the same (within the margin of experimental error) as that of wild-type transcripts, so that there is no reason to suspect that revertants of some sort are being selected from the transfection mix. Both the large and small plaque variants have the predicted sequence within the 51-nucleotide region but clearly must differ somewhere else in the genome. We presume that the small plaque variant represents the parental mutant that was constructed because it is present in fivefold excess over the large plaque variant and that an RNA polymerase error leads to a second change elsewhere in the genome to produce large plaques. Because the small plaque variant is stable once isolated, it seems likely that the error is introduced by the SP6 polymerase and not by the viral replicase. It is unclear if such an SP6 polymerase error is one that is introduced often but not seen because its effect is seen only when the change in the 51-nucleotide region is present or if the mutation in this element induces the polymerase to make the error. It will be possible to map the second site change by using the same methods that have been used successfully for mapping temperature-sensitive mutations or other phenotypic markers (5, 6, 12).

The relationship of plaque size to virus growth rate is also unpredictable. In the case of the two mutants P1-158A and P1-173U, the small plaque variant grows at a faster rate than the large plaque variant, which produces plaques that are six times as large. But even more strangely, all of the other mutants produce plaques that are wild type in morphology, despite the fact that most produce virus at a rate only 1%, and in extreme cases 0.01%, that of the wild type rate and many exhibit a pronounced lag in the growth curve. Thus, primary plaque morphology cannot be used to predict the growth properties of the virus other than that it is viable. We

have obtained similar results with mutants in other regions of the genome (9).

Deletion mapping of DI RNAs shows that the entire 51-nucleotide domain is dispensible, although the DI particles do replicate somewhat more slowly (11). In our hands, changes in this region led to a reduction of virus growth rate of two orders of magnitude, which effectively makes the RNA noncompetitive and might be expected to inactivate DI RNA. It is unclear at present why DI RNA can dispense with the function of this region.

Although the function of this region is currently unknown, we presume that a protein, whether virus encoded or cellular, interacts with this domain in a specific fashion. It seems almost certain that the hairpin structures exist in solution. They have a large calculated free energy, and the sequence conservation ends precisely at the boundaries of the stems (Fig. 1). Furthermore, observations such as the fact that O'Nyong-nyong virus has compensating changes that prevent the second stem from being elongated by a base pair or that in some viruses one base pair is substituted for another (Fig. 1) lend credence to the hypothesis that this conserved domain functions as two hairpin structures. The mutational analysis here makes clear that the sequence within the stems of the hairpins is crucial, and thus these stems cannot exist simply to present the loop sequence to a protein. Furthermore, our results suggest that the presence of the extrahelical As in the stem structures is not the key feature (20), although we cannot rule this out now because of the difficulty of obtaining mutants with silent mutations to test this in a simple fashion. We suggest, however, that some protein interacts with the stem structures in the major or minor grooves in a fashion similar to that of a number of DNA binding proteins and that the distortions in the structure caused by nucleotide alterations (or the altered sequence when one base pair is replaced with another) lead to a lower affinity of binding and slower virus growth. In this model, two explanations are possible for the fact that the mutants are viable but grow poorly. One (the enhancer theory) is that the binding function can be dispensed with entirely such that the virus can replicate in the absence of binding of the putative binding factor but at a much slower rate. The second possibility is that binding is absolutely required for replication and the slower kinetics of growth result from a weakened affinity of the binding factor for its substrate. The DI particle results suggest the former, but it is unclear at what stage binding to this region is required and therefore if it is a function required for virus replication but not for DI particle replication in the presence of helper virus. Furthermore, the fact that the different mutants tested vary over three to four orders of magnitude in the rate of virus production might suggest that for many or all of these mutants there is some residual binding by the interacting protein.

Two hypotheses have been put forward for the possible function of this region (13, 17). One hypothesis suggests that this domain is involved in encapsidation, which seems unlikely now because of recent results that implicate a region between nucleotides 746 and 1226 in encapsidation (19). The second possible function is production of minus strands from a plus-strand template, because it is thought that non-3′-terminal sequences are required for production of minus strands. In this model, the replicase must bind simultaneously to both the 5′ and 3′ ends of the RNA to initiate minus-strand synthesis in an efficient fashion, a process which could be facilitated by the cyclization of alphavirus RNAs that is known to occur (4). Now that we have a series of mutants that have been characterized as to growth properties in two cell lines representing different phyla, we are in a position to explore the function of the domain in virus replication.

ACKNOWLEDGMENTS

We are grateful to E. G. Strauss for preparation of the figures and for help with the text, to R. J. Kuhn for many stimulating discussions, and to Z. Hong and E. Lenches for technical assistance.

This work was supported by grant DMB 86-17372 from the National Science Foundation. H.G.M.N. was supported in part by a fellowship from the Niels Stensen Foundation, a Gosney fellowship from the California Institute of Technology, and a travel grant from the Netherlands Organization for the Advancement of Pure Research, N.W.O.

LITERATURE CITED

1. **Chen, E., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4:** 165–170.
2. **Dulbecco, R., and M. Vogt.** 1954. One-step growth curves of western encephalomyelitis virus grown in vitro and analysis of the virus yields from single cells. J. Exp. Med. **99:**183–199.
3. **Faragher, S. G., A. D. J. Meek, C. M. Rice, and L. Dalgarno.** 1988. Genome sequences of a mouse-avirulent and a mouse-virulent strain of Ross River virus. Virology **163:**509–526.
4. **Frey, T. K., D. L. Gard, and J. H. Strauss.** 1979. Biophysical studies on circle formation by Sindbis virus 49S RNA. J. Mol. Biol. **132:**1–18.
5. **Hahn, Y. S., A. Grakoui, C. M. Rice, E. G. Strauss, and J. H. Strauss.** 1989. Mapping of RNA$^-$ temperature-sensitive mutants of Sindbis virus: complementation group F mutants have lesions in nsP4. J. Virol. **63:**1194–1202.
6. **Hahn, Y. S., E. G. Strauss, and J. H. Strauss.** 1989. Mapping of RNA$^-$ temperature-sensitive mutants of Sindbis virus: assignment of complementation groups A, B, and G to nonstructural proteins. J. Virol. **63:**3142–3150.
7. **Hardy, W. R., and J. H. Strauss.** 1988. Processing of the nonstructural polyproteins of Sindbis virus: study of the kinetics in vivo using monospecific antibodies. J. Virol. **62:**998–1007.
8. **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.
9. **Kuhn, R. J., Z. Hong, and J. H. Strauss.** 1990. Mutagenesis of the 3′ nontranslated region of Sindbis virus RNA. J. Virol. **64:**1465–1476.
10. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82:**488–492.
11. **Levis, R., B. G. Weiss, M. Tsiang, H. Huang, and S. Schlesinger.** 1986. Deletion mapping of Sindbis virus DI RNAs derived from cDNAs defines the sequences essential for replication and packaging. Cell **44:**137–145.
12. **Lustig, S., A. Jackson, C. S. Hahn, D. E. Griffin, E. G. Strauss, and J. H. Strauss.** 1988. Molecular basis of Sindbis virus neurovirulence in mice. J. Virol. **62:**2329–2336.
13. **Ou, J.-H., E. G. Strauss, and J. H. Strauss.** 1983. The 5′-terminal sequences of the genomic RNAs of several alphaviruses. J. Mol. Biol. **168:**1–15.
14. **Ou, J.-H., D. W. Trent, and J. H. Strauss.** 1982. The 3′-non-coding regions of alphavirus RNAs contain repeating sequences. J. Mol. Biol. **156:**719–730.
15. **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.
16. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.

17. **Strauss, E. G., and J. H. Strauss.** 1986. Structure and replication of the alphavirus genome, p. 35–90. *In* S. Schlesinger and M. J. Schlesinger (ed.), The Togaviridae and Flaviviridae. Plenum Publishing Corp., New York.
18. **Tinoco, I., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. Nature New Biol. **246**:40–41.
19. **Weiss, B., H. Nitschko, I. Ghattas, R. Wright, and S. Schlesinger.** 1989. Evidence for specificity in the encapsidation of Sindbis virus RNAs. J. Virol. **63**:5310–5318.
20. **Wu, H.-N., and O. C. Uhlenbeck.** 1987. Role of a bulged A residue in a specific RNA-protein interaction. Biochemistry **26**:8221–8227.
21. **Zoller, M. J., and M. Smith.** 1984. Laboratory methods: oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template. DNA **3**:477–488.