JOURNAL OF VIROLOGY, May 1996, p. 2757–2763 Vol. 70, No. 5
0022-538X/96/$04.00+0

Identification of a Region in the Sindbis Virus Nucleocapsid Protein That Is Involved in Specificity of RNA Encapsidation

KATHERINE E. OWEN AND RICHARD J. KUHN*

Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

Received 6 November 1995/Accepted 15 January 1996

The specific encapsidation of genomic RNA by an alphavirus requires recognition of the viral RNA by the nucleocapsid protein. In an effort to identify individual residues of the Sindbis virus nucleocapsid protein which are essential for this recognition event, a molecular genetic analysis of a domain of the protein previously suggested to be involved in RNA binding in vitro was undertaken. The experiments presented describe the generation of a panel of viruses which contain mutations in residues 97 through 111 of the nucleocapsid protein. All of the viruses generated were viable, and the results suggest that, individually, the residues mutated do not play a critical role in encapsidation. However, one mutant which had lost the ability to specifically encapsidate the genomic RNA was identified. This mutant virus, which contained a deletion of residues 97 to 106, encapsidated both the genomic RNA and the subgenomic mRNA of the virus. It is proposed that the encapsidation of this second species of RNA, which is not present in wild-type virions, is the result of the loss of a domain of the nucleocapsid protein required for specific recognition of the genomic RNA packaging signal. The results suggest that this region of the protein is important in dictating specificity in the encapsidation reaction in vivo. The isolation and preliminary characterization of two independent second-site revertants to this deletion mutant are also described.

Sindbis virus (SIN) is an enveloped RNA virus, the prototype of the alphavirus genus. Its single-stranded positive-sense genome is enclosed within a spherical core made up of 240 symmetrically arranged, identical copies of the nucleocapsid protein (C) (4, 6, 26). This nucleocapsid core (NC), composed of the viral RNA and the nucleocapsid protein shell, is surrounded by a lipid bilayer membrane that is derived from the infected host cell. Distributed throughout the membrane and exposed on the surface of the virus particle are viral glycoproteins involved in recognition and attachment to a host cell and subsequent membrane fusion. These spikes are composed of two distinct viral glycoproteins, E1 and E2 (reviewed in reference 16). A domain of the E2 protein also interacts with the NC on the inside of the lipid bilayer (reviewed in reference 37).

The structural proteins of SIN are translated from a subgenomic mRNA as a polyprotein of C-E3-E2-6K-E1. The nucleocapsid protein has been shown by biochemical and genetic studies to possess a proteolytic activity that results in the autocatalytic cleavage of itself from the nascent structural polyprotein (1, 2, 14, 15). The high-resolution X-ray structure of the purified protein has determined that residues 114 to 264 have a structural fold similar to that of chymotrypsin (6, 39). Following its self-cleavage and release from the polyprotein, the capsid protein transiently associates with ribosomes, and within several minutes the protein binds to viral RNA and rapidly assembles into core particles (12, 36). No intermediates in NC assembly have been identified, and the role of ribosomes in this process is unclear. The NC eventually associates with the cytoplasmic domain of the E2 viral glycoprotein at the cell membrane, leading to budding and the formation of the mature virion (reviewed in reference 38).

Structural studies of alphaviruses have provided low-resolution information on the organization of the virion but only limited insight into the possible assembly mechanisms of the NC. Cryoelectron microscopy and image reconstruction have been recently carried out on Ross River virus (4). The data show the alphavirus NC exhibits T=4 icosahedral symmetry. The atomic structure previously obtained for the ordered amino acid residues 114 to 264 from the SIN nucleocapsid protein was modeled into the cryoelectron microscopy density. This modeling has oriented the nucleocapsid protein with respect to the inside and outside of the NC and shown that residues 114 to 264 of the SIN nucleocapsid protein project about 40 Å (1 Å = 0.1 nm) off the surface of the NC, which is found at a radius of 170 Å.

Residues 1 to 113 of the nucleocapsid protein, which are disordered in the X-ray crystal structure, are predicted to be on the interior of the NC in contact with the viral RNA. The N-terminal 96 amino acids consist of many charged residues, and this relatively nonconserved domain has been suggested to be involved in nonspecific ionic interactions with the viral RNA, as has been seen with other viral capsid proteins (9, 11, 29). The domain from residues 97 to 113 is remarkably conserved among alphaviruses, even though these residues have been shown to be dispensable for proteinase function both in vivo and in vitro (9, 24). This stretch of residues is highly charged, suggesting that it may be exposed on the surface of the protein and involved in ionic interactions. The region between residues 76 and 107 has been implicated by in vitro studies to be important in the specific interaction between the nucleocapsid protein and the binding domain on the viral RNA (11). On the basis of an independent set of in vitro studies, Wengler et al. (41) suggested that the ribosome competes for a ribosome binding site found between residues 99 and 113 of the SIN capsid protein and promotes the disassembly of the NC. Further results suggested that this interaction was mediated by rRNA (35). Together, these studies implicated a relatively small number of residues in the middle of the nucleocapsid protein as important for RNA recognition in vitro. We have consequently undertaken a molecular genetic study of this central domain of the SIN nucleocapsid protein, concentrating on the region between residues 97 and 111, to deter-

* Corresponding author. Phone: (317) 494-1164. Fax: (317) 496-1189. Electronic mail address: rjkuhn@bragg.bio.purdue.edu.

mine the importance of these residues in the assembly-disassembly pathway of the nucleocapsid core and to elucidate possible mechanisms of RNA binding.

MATERIALS AND METHODS

Viruses and cells. BHK-15 and Vero cells used for propagation of viruses were grown in Eagle minimal essential medium supplemented with 10% fetal bovine serum unless otherwise indicated. Secondary chicken embryo fibroblast cells were grown in the same medium supplemented with 5% tryptose phosphate broth and 3% fetal bovine serum.

Construction of DNA clones and mutagenesis. The full-length SIN cDNA clone pToto50 (27) was used as the parental cDNA for initial cloning steps. Two derivatives of pToto50 that contained two (pToto63) or three (pToto64) new unique restriction sites were constructed by site-directed mutagenesis. The substitutions required for these sites did not alter the amino acid sequence of the resulting clone. The virus that was recovered from pToto64 was used as the wild-type parental virus in all subsequent experiments.

Mutagenesis was performed in an M13mp18 subclone containing nucleotides 7334 to 9120 of the SIN genome, which includes the coding sequence for the entire nucleocapsid protein as well as flanking sequences. Unique *Bst*EII, *Xba*I, and *Mlu*I restriction sites were engineered by oligonucleotide-directed mutagenesis into the M13mp18 subclone, at positions corresponding to nucleotides 7472, 8529, and 8029, respectively. DNA containing these restriction sites was used to replace wild-type sequences in the full-length clone pToto50, using standard ligation methods (32). Mutagenesis of the putative RNA binding region (nucleotides 7935 to 7979) was carried out in the M13mp18 subclone containing all three unique sites (pJames811). In some cases, the mutation was engineered to create or destroy restriction sites to facilitate screening. In all cases, the mutation was confirmed by sequencing the single-stranded DNA. DNA fragments containing the desired mutations were excised from the M13mp18 subclones by using the unique *Bst*EII and *Xba*I restriction sites. These fragments were used to replace the corresponding fragment in pToto63 in a two-fragment ligation. Nomenclature for mutants is according to the suggestion of Kuhn et al. (19). For example, the mutant later discussed which contains a deletion of capsid protein residues 97 to 106 is labeled CΔ(97-106).

Virus isolation. Plasmid DNA containing the full-length SIN cDNA was digested with *Sac*I to linearize the template for transcription. RNA was transcribed by using SP6 RNA polymerase and was transfected into BHK cells either by DEAE-dextran (19) or by electroporation (23) essentially as described elsewhere. All virus isolations were done by using DEAE-dextran transfections followed by plaque purification. For electroporation, trypsin-EDTA was used to remove the subconfluent cells from culture flasks prior to transfection. After the cells (approximately 6×10^7) were washed twice with phosphate-buffered saline (PBS) and resuspended in 0.5 ml of PBS, they were placed into a 0.2-cm-gap cuvette and 10 μg of RNA was added. The cells were pulsed twice in a Bio-Rad Gene Pulser set at 1.5 kV, 25 μF, and 200 Ω and were replated onto 100-mm-diameter culture dishes in 10 ml of Eagle minimal essential medium containing 10% fetal bovine serum.

To determine the efficiency of each electroporation, a portion of the cells was assayed for the production of E2 by immunofluorescence. At 9 h posttransfection, cells were fixed onto coverslips by using methanol. Anti-E2 (SIN) polyclonal antibodies were added to the cells and incubated at 37°C for 45 min. Unbound antibodies were rinsed off with PBS, and goat anti-rabbit antibodies conjugated to fluorescein (Kirkegaard & Perry, Gaithersburg, Md.) were added. After incubation at 37°C for 45 min, excess antibodies were rinsed off and the coverslips were mounted onto slides. Electroporation efficiency was defined as the percentage of cells fluorescing. Multiple fields were counted to ensure accuracy.

Assay of NC. NC were analyzed as previously described (25). Briefly, transfected cells were incubated at 37°C for 6 h following electroporation, at which time actinomycin D was added to the cells to a final concentration of 1 μg/ml. At 7 h posttransfection, [5,6-^{3}H]uridine was added to a final concentration of 20 μCi/ml, and at 12 h posttransfection, the cells were washed and scraped from the plates into PBS. Cells were sedimented by low-speed centrifugation, washed, and resuspended in 1 ml of 10 mM Tris (pH 7.4)–10 mM NaCl–20 mM EDTA. Following a 15-min incubation on ice, 200 μl of 20% Triton X-100 was added and nuclei were removed by low-speed centrifugation. The resulting cytoplasmic extract was loaded onto a linear 10 to 40% sucrose gradient in 50 mM Tris (pH 7.6)–100 mM NaCl–1 mM EDTA–0.1% Triton X-100. The gradients were subjected to centrifugation for 2.5 h at 32,000 rpm at 4°C in an SW41 rotor and fractionated from the bottom into 600-μl fractions. The radioactivity in 10 μl from each fraction was quantitated by scintillation counting.

Western blot (immunoblot) analysis was performed on 10 μl from the peak fractions of the gradients. Samples from the seventh fraction were mixed with 10 μl of Laemmli sample buffer and electrophoresed in sodium dodecyl sulfate (SDS)–12% polyacrylamide gels. The proteins were transferred onto nitrocellulose, and Western blot detection was performed with a polyclonal anti-C antibody. An enhanced chemiluminescence detection kit (Amersham, Arlington Heights, Ill.) was used, and quantitation was performed by scanning the film and analyzing it with NIH Image software, version 1.52. Both the nucleocapsid accumulation assay and the Western blot analysis were completed twice and gave comparable results both times.

Viral growth analysis. Beginning at 3 h postelectroporation and every hour until 9 h postelectroporation, the medium over the transfected BHK cells was collected and replaced with fresh medium. Plaque-forming virus from each sample was assayed by titration on BHK cell monolayers at 37°C.

Virus replication was also examined following infection by the viruses as described previously (18). Cells were infected at a multiplicity of infection of 5. Following adsorption of the viruses to the cells for 1 h at room temperature, the cells were placed at 37°C, and the medium was replaced every 30 min for the first 2 h, and every hour after that until 12 h. Titers of released virus were determined by plaque assay on BHK cell monolayers. For both electroporation and infection, the entire assays were completed in duplicate. The results reported are the averages of the two counts obtained for each time point.

The nucleocapsid protein within virus released from the cells at the last time point following electroporation (9 h) and infection (12 h) was examined by Western blot analysis. Equal volume samples of the released viruses were denatured by boiling with Laemmli sample buffer and electrophoresed through SDS–12% polyacrylamide gels. The proteins were transferred to nitrocellulose, and Western blot detection was performed in duplicate as described above.

In vitro binding assay. Truncated nucleocapsid protein was expressed and purified from *Escherichia coli* by the pET vector system (Novagen, Madison, Wis.) (5). Briefly, the coding sequences for the SIN nucleocapsid protein amino acids 81 to 264 or 106 to 264 were inserted into the pET11a expression vector. Expressed proteins were purified to greater than 95% homogeneity as determined by silver staining (5). In addition to the truncated wild-type protein being expressed in this manner, the protein from CΔ(97-106) was expressed from amino acids 81 to 264 (maintaining the deletion of amino acids 97 to 106).

The truncated wild-type and CΔ(97-106) SIN nucleocapsid proteins expressed and purified from *E. coli* were examined for RNA binding activity in a nitrocellulose binding assay essentially as described previously (11). Eighty picomoles of *E. coli*-expressed capsid protein was immobilized on nitrocellulose, and the nonspecific binding sites were blocked by the addition of total *E. coli* RNA. An in vitro-transcribed [α-^{32}P]CTP-radiolabeled RNA which corresponds to the encapsidation sequence of SIN (nucleotides 945 to 1077) was bound to the immobilized capsid protein. Excess probe was removed by washing, and the bound probe was quantitated by scintillation counting. Results reported are the averages of triplicate trials.

Northern (RNA) blot and RNase protection analysis. Samples of RNA from purified virus were also used in Northern blot and RNase protection analyses as follows. Following a 48-h infection, supernatants from two 150-mm-diameter culture dishes were centrifuged in an SS34 rotor at 8,000 rpm for 20 min. One-quarter volume of 40% polyethylene glycol 8000–2.5 M NaCl was added, and the samples were incubated overnight on ice. The virus was then recovered by centrifugation in an HB-4 rotor at 8,000 rpm for 1 h and resuspended in 1 ml of 50 mM Tris (pH 7.6)–100 mM NaCl–1 mM EDTA. The entire sample was then loaded onto a 10 to 40% sucrose gradient and centrifuged for 1.5 h at 31,000 rpm in an SW41 rotor. Visible bands of virus were isolated and concentrated by high-speed centrifugation. The titer and plaque phenotype of each sample were determined by plaque assay. SDS was added to 1%, and phenol-chloroform extraction was used to remove protein from equal volumes of purified virus of known titer.

Following 9 h of infection with either Toto64 or CΔ(97-106), cytoplasmic extracts were also prepared as described (19). The samples were extracted with phenol-chloroform, and the cytoplasmic RNAs were precipitated with ethanol and resuspended in diethylpyrocarbonate-treated water. For Northern blot analysis, all of the RNA samples were prepared in formaldehyde-formamide loading buffer and electrophoresed through a 1% agarose denaturing formaldehyde gel. The RNA was transferred to a nylon membrane by capillary blotting for at least 22 h, and detection was accomplished by using a [α-^{32}P]dCTP-radiolabeled DNA probe from SIN nucleotides 8569 to 9804. Hybridization was carried out as described previously (32).

RNase A mapping was performed essentially as described previously (22). pGem-J1 was used to transcribe an [α-^{32}P]CTP-labeled, minus-sense RNA probe from SIN nucleotides 7716 to 7338, corresponding to the junction region between the 49S and 26S RNAs. The probe was hybridized to the purified RNA for 18 h at 55°C, and single-stranded RNA was digested with RNase A. The products were electrophoresed through a 6% acrylamide-urea sequencing gel. The gel was fixed, dried, and exposed to an imaging plate. A Fuji BAS2000 phosphorimaging system was used to analyze the resulting images. Both Northern and RNase analyses were completed in duplicate.

Molecular cloning of the revertants. Two mid-size-plaque variants of CΔ(97-106) were identified and isolated, one following infection of Vero cells and the other following infection of BHK cells. The viruses were plaque purified twice in BHK-15 cells and used to infect two 100-mm-diameter plates of BHK cells. At 9 h postinfection, cytoplasmic RNA was extracted from the cells (19). Avian myeloblastosis virus reverse transcriptase was used to generate cDNA, and PCR was used to amplify the sequences from nucleotides 7467 to 8541, including the entire nucleocapsid protein coding sequences and additional flanking sequences. The fragments were then digested with the unique restriction enzymes *Xba*I and *Bst*EII and inserted into both the full-length cDNA clone (pToto63) and the M13mp18 subclone (pJames811). Revertant viruses were rescued from the full-length clone as described above. Following insertion of the revertant cDNAs into pJames811, the complete nucleocapsid protein coding sequence was determined

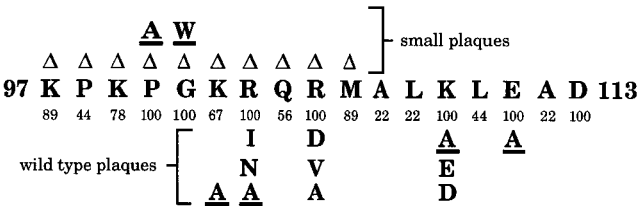

FIG. 1. Site-directed mutations in the SIN nucleocapsid protein from residues 97 to 113. Amino acids are shown in boldface; the number under each residue indicates the percent identity of the residue among nine sequenced alphaviruses (SIN [28], Aura virus [30], Semliki Forest virus [10], Ross River virus [8], eastern equine encephalitis virus [3], western equine encephalitis virus [13], Venezuelan equine encephalitis virus [17], O'nyong-nyong virus [21], and Mayaro virus [37]). Amino acid designations above and below the sequence indicate changes which were made to the residue. Underlined residues within a line indicate a double substitution. The mutant CΔ(97-106) is represented by the delta symbols above the 10 residues that were deleted. The nine mutants depicted below the sequence displayed wild-type plaques in BHK cells at 37°C; the two above displayed small plaques.

by dideoxy-chain termination sequencing using the M13 universal primer and SIN primers complementary to the positive strand.

RESULTS

Construction of pToto64. Residues 97 to 113 of the SIN nucleocapsid protein have been implicated as important in binding to both viral RNA and rRNA. There is considerable amino acid homology in this region among the sequenced alphaviruses, suggesting a conserved function for this domain of the protein (Fig. 1). We have therefore undertaken a molecular genetic approach to examine the importance of these residues in the assembly and disassembly processes of the alphavirus NC. Specifically targeting the conserved, charged amino acids 97 to 111 of the protein, we used mutagenesis to change or delete residues to identify essential interactions in which the nucleocapsid protein is involved. Since the plasmid containing the full-length cDNA clone of SIN is large (>13 kb), mutagenesis was performed in an M13mp18 subclone which contained the entire coding sequence for the nucleocapsid protein as well as flanking sequences. Unique *Bst*EII and *Xba*I restriction sites were introduced into both the M13mp18 subclone and subsequently the full-length SIN cDNA clone (pToto63) in regions which flank the coding sequence for the capsid protein. Although these substitutions altered the nucleotide sequence of the DNA, the addition of these sites did not change the amino acid sequence encoded by the nucleotides. Virus rescued following transcription and transfection of RNA from the clone containing the two new unique restriction sites (Toto63) produced plaques that were the same size as the parental virus. These unique sites were used to shuttle the DNA from the M13mp18 subclone following mutagenesis into the full-length SIN cDNA clone. To facilitate detection of full-length mutant clones, a unique *Mlu*I site was engineered into the middle of the coding sequence for the nucleocapsid protein (nucleotide 8029) in both the M13mp18 subclone and pToto63. Like the other restriction sites, this addition did not change the amino acids encoded by the RNA, and transcription and transfection of RNA from the resulting full-length clone (pToto64) resulted in virus displaying the wild-type plaque phenotype. All subsequent mutagenesis was carried out in the M13mp18 subclone that contained all three of the unique restriction sites (pJames811). By using the unique *Xba*I and *Bst*EII sites, the mutated sequences were ligated into a full-length cDNA which did not contain the *Mlu*I site (pToto63). The presence of the *Mlu*I site was therefore an indication of a full-length clone that contained the mutation of interest. Duplicate clones were generated, and virus was isolated for each mutation.

Single substitutions in the putative RNA binding domain do not have a drastic effect on virus survival. The initial strategy used was to carry out single amino acid substitutions at conserved, charged residues. Plaque phenotype was used as the primary indicator of a defective virus. Figure 1 shows that when single amino acids were substituted, the plaque phenotype in BHK cells at 37°C remained that of the wild-type virus. In addition, for two of the three cases in which double substitutions were made, the phenotype also remained wild type. In the mutant C(P100A,G101W), a slightly smaller plaque phenotype was observed. None of the viruses were temperature sensitive, as measured by plaque assay at 30 and 40°C in secondary chicken embryo fibroblast cells (data not shown). These results suggest that the single and double substitutions that were made do not dramatically affect the replication of the virus as measured by plaque phenotype.

Failing to identify specific amino acid residues critical for virus replication in this conserved region of the nucleocapsid protein, we used a different approach. Ten amino acids, from residues 97 through 106, were deleted, and a virus [CΔ(97-106)] was rescued from transfected cells. The viral plaques produced by CΔ(97-106) were approximately 20% the size of the wild-type virus plaques, indicating that although the deletion of 10 amino acids within this conserved, charged region was not lethal, it did exert a deleterious effect on virus replication. The presence of the deletion in the recovered virus was verified by sequencing the mutant cDNA.

Assembly of the nucleocapsid is affected in CΔ(97-106). Since CΔ(97-106) was the only virus to show a drastic reduction in plaque size, further characterization of this virus was done to determine if the interactions essential for forming the NC had been disrupted. An assay was performed to examine the relative amounts of NC found in cells transfected with the mutant or wild-type RNA. To avoid bias due to defects in entry or uncoating, electroporation was used to transfect infectious RNA into cells. Extensive analyses were performed to ensure that the electroporation efficiency between experiments remained constant. In all of the experiments performed, greater than 90% of the cells were positive for fluorescence in assays using an anti-E2 antibody.

Following transfection and labeling of RNA with [5,6-^{3}H] uridine, cytoplasmic extracts were prepared and fractionated by sucrose gradient centrifugation. The results of the nucleocapsid accumulation assays are shown in Fig. 2. The peak closest to the bottom of the gradient corresponds to RNA incorporated into intact NC. The second peak contains labeled viral RNA and nucleocapsid protein and has been suggested by others to be the result of interaction of these molecules with ribosomes (40). The area under the NC peak corresponding to labeled RNA is approximately threefold less for CΔ(97-106) than for wild-type NC. Western blot analysis of fractions following sucrose gradient centrifugation using anti-C polyclonal antibodies showed that the amount of nucleocapsid protein in the CΔ(97-106) NC peak was approximately equal to the amount of protein found in the wild-type NC peak (Fig. 2). These experiments did not address the nature of the RNA contained within the mutant NC; however, the data suggested that there was less RNA contained within equivalent numbers of NC. In addition, Western blot analysis of total cytoplasmic extracts suggested that approximately equivalent amounts of nucleocapsid protein were translated in cells transfected with either mutant or wild-type RNA (data not shown). Together,

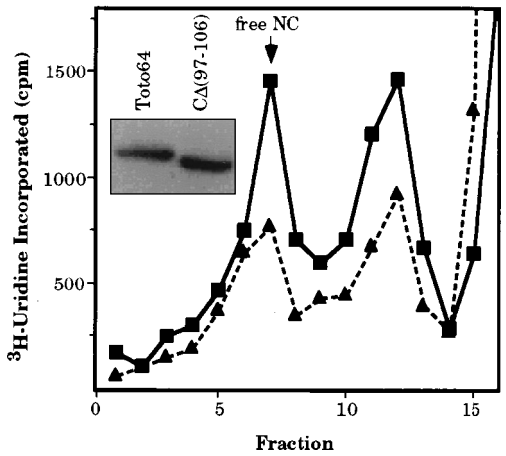

FIG. 2. Nucleocapsid core accumulation of CΔ(97-106). Cells were transfected with 10 μg of full-length wild-type or CΔ(97-106) RNA. Replicating RNA was labeled with [5,6-^{3}H]uridine from 7 to 12 h posttransfection, at which time the cells were lysed and the resulting cytoplasmic extracts were layered over 10 to 40% sucrose gradients. Gradients were centrifuged and fractionated from the bottom into 600-μl fractions, and radioactivity was measured by scintillation counting. The solid line is the profile of wild-type (Toto64) RNA in the gradient, while the dashed line represents CΔ(97-106) RNA. The peak closest to the bottom of the gradient corresponds to intact NC. The inset is a Western blot which shows the relative amounts of nucleocapsid protein found in equal-volume samples of the NC peaks for the wild type and CΔ(97-106). The image was scanned into Adobe Photoshop (version 3.0) and transferred to Canvas (version 3.01) for labeling.

these data indicate that the deletion mutant is affected in the proper assembly of the NC.

CΔ(97-106) is not deficient in releasing particles from infected cells. The virus released from transfected cells was also examined. Figure 3 shows the results of a time course assay of infectious virus released from BHK cells either electroporated with full-length RNA (Fig. 3A) or infected with virus (Fig. 3B). Throughout the two experiments, the wild-type virus consistently produced approximately 100-fold more PFU per milliliter than CΔ(97-106).

The inset Western blots in Fig. 3 show the amount of capsid protein present in an equal volume of supernatant from the last time point. The blots indicate that even though there was greater than a 100-fold decrease in infectious virus released for the deletion mutant, the mutant contained approximately equivalent amounts of nucleocapsid protein in the virus particle. This result suggests that the mutant was producing approximately equivalent numbers of virus particles as the wild type, but these particles were 100-fold less infectious than the wild-type virus particles. This assertion is supported by electron microscopy results which showed that at 10 h postinfection, cells infected with either wild-type or CΔ(97-106) virus appeared to be releasing roughly similar amounts of virus (data not shown). In addition, electron microscopy of purified virus gave no indication that CΔ(97-106) assembled into multicored particles.

We also performed an assay in which virus was double labeled with [5,6-^{3}H]uridine and [^{35}S]methionine-cysteine. Equal-volume samples of purified virus were quantitated by scintillation counting, and titers were determined on BHK cells. The results of this experiment showed that equivalent counts per minute corresponded to approximately 20-fold less PFU for CΔ(97-106) than for the wild-type virus (data not shown). This result suggested that CΔ(97-106) assembled many particles that were not infectious.

Purified mutant nucleocapsid protein is deficient in binding the encapsidation sequence in vitro. A binding assay was used to examine the interaction between the mutated capsid protein and the encapsidation signal for SIN in vitro. *E. coli*-expressed truncated nucleocapsid protein was bound to nitrocellulose filters, and the in vitro-transcribed radiolabeled encapsidation signal (nucleotides 945 to 1077) was incubated with the filters. Binding efficiency was quantitated by scintillation counting of labeled RNA which remained bound following multiple washes. The results of the binding experiments are displayed in Fig. 4. The data show that in this in vitro system, the truncated wild-type protein, corresponding to amino acids 81 to 264, bound the RNA at a level that was at least eight times higher than the background level. On the contrary, the wild-type truncated protein from residues 106 to 264 bound to the encapsidation signal only at background levels (i.e., the same as the level for bovine serum albumin [BSA]). Specificity in binding was established by performing a competition assay in which unlabeled RNA was shown to specifically compete with radiolabeled RNA in binding the nucleocapsid protein (data not shown). The data also show that the truncated protein (residues 81 to 264) from CΔ(97-106) bound the RNA at a very low level comparable to the background level. These results sug-

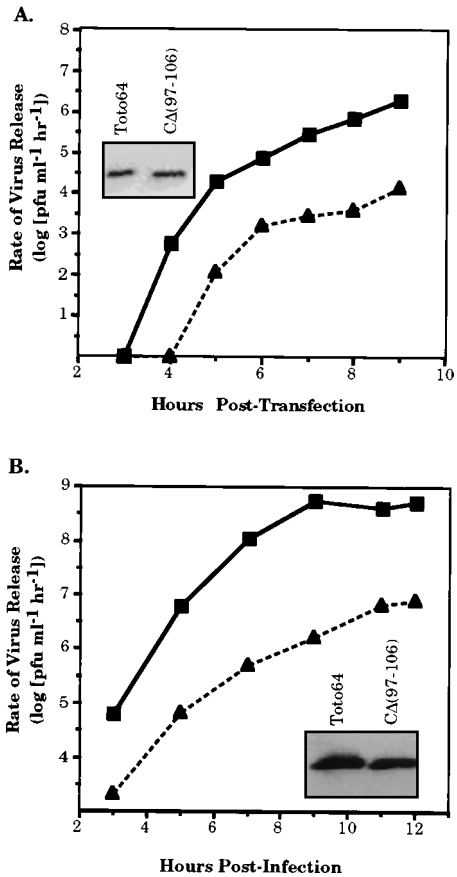

FIG. 3. Release of the wild type and CΔ(97-106) from transfected and infected cells. Cells were either transfected with 10 μg of in vitro-transcribed full-length wild-type (Toto64) or CΔ(97-106) RNA (A) or infected with one of the two viruses at a multiplicity of infection of 5 (B). Media were removed and replaced every hour, and titration of released virus was performed on BHK cell monolayers at 37°C. Wild-type virus is represented by the solid lines; the dashed lines represent CΔ(97-106). The inset Western blots show the amounts of nucleocapsid protein present in equal volumes of media from the last time point of each experiment.

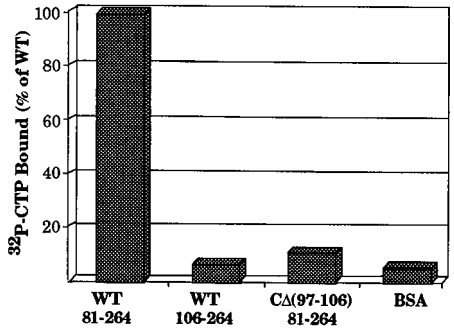

FIG. 4. In vitro RNA binding assay. Eighty picomoles of *E. coli*-expressed truncated protein was bound to a nitrocellulose membrane. Nonspecific binding sites were blocked with total *E. coli* RNA, and radiolabeled encapsidation signal was then added. Excess label was removed through successive washes, and bound labeled RNA was quantitated by scintillation counting. The first bar corresponds to wild-type (WT) nucleocapsid protein from amino acids 81 to 264; the second is wild-type nucleocapsid protein from amino acids 106 to 264; the third is CΔ(97-106) nucleocapsid protein from amino acids 81 to 264; the last is a BSA negative control.

gest that in vitro, residues 81 to 105 of the wild-type protein play a role in binding of the capsid protein to the encapsidation signal and that residues 97 to 106 are required for this binding. It should be noted that in this in vitro assay, the truncated wild-type protein did not bind the encapsidation signal as well as the full-length wild-type protein (33). Additional binding was obtained when the N terminus of the protein was present (11), but the interaction of the N terminus with the RNA has been suggested to be nonspecific (9).

CΔ(97-106) encapsidates both the 49S genomic and the 26S subgenomic RNAs. Since the in vitro results suggested the mutant capsid protein might be defective in binding the encapsidation signal, we examined the species of RNA found within virus particles. Virus was concentrated by polyethylene glycol precipitation and purified by sucrose gradient centrifugation. The amount of infectious virus present in these samples was determined by plaque assay on BHK cells. RNA was extracted from equal volumes of virus and used in Northern and RNase protection analyses. Northern blot analyses revealed there was an intense band present in the purified CΔ(97-106) virions which comigrated with the 26S RNA found in cytoplasmic extracts from cells infected with wild-type virus (data not shown). No visible 49S RNA was present in the deletion mutant. In addition, the Northern analysis suggested that the total amounts of both 49S and 26S RNAs present in cells infected with CΔ(97-106) were slightly decreased compared with the wild type (data not shown).

To confirm that the band comigrating with the 26S RNA did in fact correspond to the subgenomic mRNA, an RNase protection assay was performed (Fig. 5). A labeled negative-sense RNA transcript corresponding to the junction region of the 49S and 26S RNAs was hybridized to RNA isolated from purified virions. RNA extracted from the cytoplasm of wild-type-infected cells was used as a positive control for the two species of RNA. The data show that the smaller RNA species present in CΔ(97-106) virions was indeed 5′ coterminal with the 26S RNA and suggested that the deletion mutant encapsidated the subgenomic message. In addition, the data confirm that 26S RNA was not present in the wild-type virions. Also, the data showed that a small amount of 49S RNA was also present in the CΔ(97-106) virions. The molar ratio of 26S RNA to 49S RNA within the CΔ(97-106) virions was 5:1, as determined by quantitation of the radioactivity in the RNase protection assay. Together, these data confirm that specificity in encapsidation has been disrupted in CΔ(97-106).

Characterization of revertants to CΔ(97-106). Two independent variants that produced mid-size plaques were isolated from the small-plaque population of the deletion mutant and plaque purified twice in BHK cells. The coding sequences for the nucleocapsid protein and flanking sequences of both of these viruses were reverse transcribed into DNA and cloned into the full-length wild-type background. Full-length revertant cDNAs were transcribed into RNA and transfected into BHK cells. The ability to recover the mid-size plaque phenotype following transfection of the in vitro-transcribed RNA was an indication that the suppressor mutation(s) responsible for the reversions was contained within the coding sequences for the capsid protein or flanking sequences. DNA corresponding to the entire coding sequence for the nucleocapsid protein from both of the revertants was sequenced, and the results are displayed schematically in Fig. 6.

In addition to confirming the presence of the original deletion, sequencing determined the changes responsible for the revertant phenotype. In one of the revertants, the coding sequence for glutamic acid 176 was changed to code for lysine. The rest of the nucleotide sequence for the nucleocapsid protein was wild type. Surprisingly, in the second revertant, a duplication of residues 10 to 89 within the coding sequences for the capsid protein was found. It contained capsid protein amino acids in the following order: 1 to 89, 10 to 96, and 107 to 264 (Fig. 6). The finding that SIN can accommodate 80 extra amino acids in its nucleocapsid protein is unique and supports assertions that the N terminus of this protein is extremely flexible. A duplication of this highly charged yet mostly nonconserved region suggests the revertant is correcting a defect in

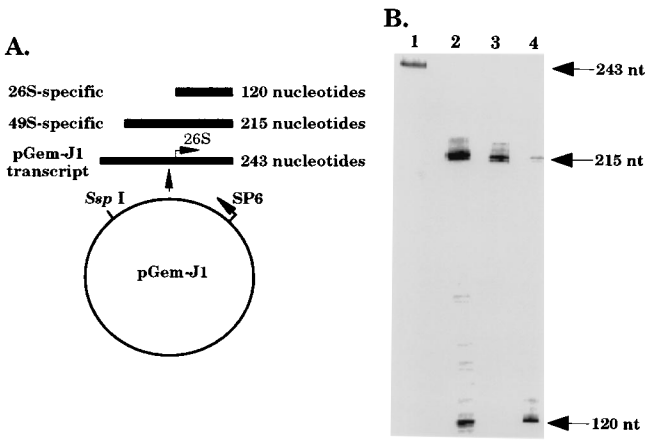

FIG. 5. RNase protection analysis of virion RNA. (A) Schematic of the transcript used in the RNase protection assay. Plasmid pGem-J1, which contains the coding sequences for the junction region between the 49S and 26S RNAs, was linearized at *Ssp*I and used to transcribe a negative-sense 243-nucleotide [α-^{32}P]dCTP-labeled RNA. The probe was complementary to 215 nucleotides of the 49S RNA and 120 nucleotides of the 5′ end of the 26S RNA. (B) RNA was extracted as described and hybridized with the 243-nucleotide (nt) probe. Single-stranded (unhybridized) RNA was digested with RNase A, and the resulting products were electrophoresed on a 6% acrylamide-urea sequencing gel. The gel was fixed and dried, and the exposed image was analyzed by the Fuji BAS2000 system. The image was transferred to Adobe Photoshop (Macintosh version 3.0); and subsequently to ClarisDraw (version 1.0) for labeling. Lane 1, unhybridized, undigested probe; lane 2, cytoplasmic RNA from Toto64-infected cells; lanes 3 and 4, RNA extracted from wild-type (Toto64) and CΔ(97-106) virus particles, respectively. The sizes of the RNAs were determined by mobility with a sequencing reaction.

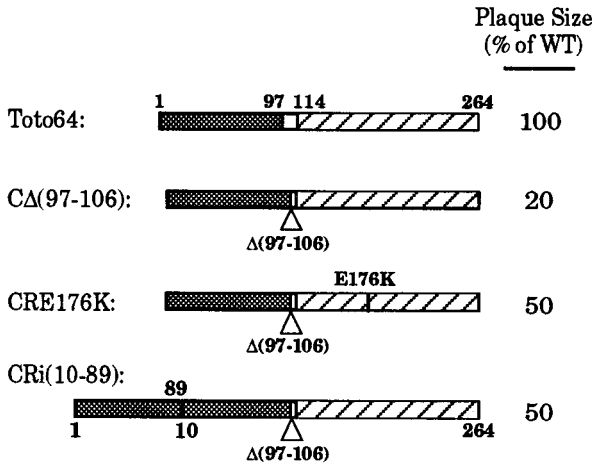

FIG. 6. Revertant viruses isolated from CΔ(97-106). Two independent mid-size-plaque variants were isolated from the small-plaque population of CΔ(97-106). The nucleotide changes responsible for each revertant were identified and confirmed as described in the text and are shown schematically. The three domains of the nucleocapsid protein represent functional delineations of the protein. Domain I (▩) has been suggested to be involved in nonspecific interactions with RNA. Domain II (▭) has been implicated in specifically recognizing RNA. Domain III (▨) is the proteinase domain of the protein and also has been suggested to interact with the cytoplasmic domain of the E2 glycoprotein. As shown, both of the revertant viruses maintained the original deletion in domain II of amino acids 97 to 106. One of the revertants (CRE176K) had a single additional change in domain III at residue 176 (E→K). The second revertant [CRi(10-89)] contained a 240-nucleotide duplication in domain I. This insertion created a nucleocapsid protein that has a direct duplication of amino acids 10 to 89. WT, wild type.

ionic interactions within the virion. The cDNA sequence from the original mutant CΔ(97-106) did not contain either of these suppressor mutations.

DISCUSSION

The nucleocapsid protein of SIN is a multifunctional protein which is involved in many macromolecular interactions required for the life cycle of the virus. Previous in vitro biochemical studies of the SIN nucleocapsid protein have implicated a relatively small region in the central part of this protein as being important in specific RNA recognition (11, 35, 41). However, the roles of individual residues in this protein-RNA interaction had not previously been examined. The results presented here show that single and double substitutions of charged residues within the domain do not dramatically affect viral replication as measured by plaque assay, despite considerable amino acid conservation. In addition, the deletion of residues 97 to 106 is tolerated, although the resulting virus, CΔ(97-106), is severely impaired in its replication capacity. Unlike the wild type, CΔ(97-106) encapsidates predominantly the subgenomic mRNA of the virus. Other groups have reported the encapsidation of defective interfering RNAs into virions (reviewed in reference 34), and the encapsidation of the subgenomic mRNA by an alphavirus was recently reported for Aura virus (31). However, this is the first example of a mutation in the alphavirus nucleocapsid protein which causes nonspecific encapsidation of viral RNA in vivo. It is unclear whether the mutant virus encapsidates a single copy or multiple copies of the smaller RNA. The data presented here suggest that an equivalent number of NC contain a threefold decrease in the total amount of RNA present. Since the subgenomic mRNA is approximately one-third of the size of the full-length genomic RNA, this finding probably indicates that many of the CΔ(97-106) cores contain a single copy of the 26S RNA.

The particle/PFU ratio is significantly increased for CΔ(97-106). At least part of the loss of infectivity is due to the misincorporation of RNA. The ratio of 26S RNA to 49S RNA within CΔ(97-106) virions is approximately 5:1, as determined by quantitation of the RNase protection assay. The data suggest that for CΔ(97-106), encapsidation of RNA occurs nonspecifically such that RNAs are packaged according to their relative abundance at the site of core assembly. Regardless of additional defects caused by the deletion of these residues, the data clearly show that this domain is involved in dictating specificity in the encapsidation reaction. It seems unlikely that the deletion of these residues creates an encapsidation signal on the 26S RNA, although this possibility cannot be ruled out by the data presented here. We propose instead that the deletion of these residues eliminates the ability of the nucleocapsid protein to distinguish between the RNAs found within the infected cell. However, our data do not preclude the possibility that the mutant nucleocapsid protein preferentially recognizes an endogenous, though inefficient, packaging signal on the 26S RNA.

Aura virus maintains all but one residue of the nucleocapsid protein sequence from residues 97 to 106 but still encapsidates the 26S RNA (30, 31). Therefore, even though the sequence on the capsid protein important for RNA recognition does not appear to be significantly altered in Aura virus, there is a disruption in encapsidation specificity. We suggest that specific recognition is not required to incorporate an RNA molecule into a virion. Thus, it is possible that instead of gaining an encapsidation sequence on its 26S RNA, Aura virus has lost the specific encapsidation sequence on its genomic RNA, leading to nonspecific RNA incorporation. In addition to encapsidating the abundant subgenomic RNA, CΔ(97-106) and Aura virus may also encapsidate cellular RNAs, although such RNAs are not detectable by the assays performed.

Preliminary characterization of the revertant viruses has suggested that they correct much of the particle/PFU ratio defect of CΔ(97-106), although they still encapsidate some 26S RNA (unpublished data). This finding suggests that another defect, in addition to nonspecific encapsidation, may contribute to the low infectivity of the virus. Transcription, translation, and budding did not appear to be dramatically affected in CΔ(97-106). Therefore, it is possible that entry and disassembly defects may also contribute to the high particle/PFU ratio for CΔ(97-106). Disassembly has been suggested to require a conformational change in the NC which primes the core for uncoating following entry (7, 20, 37). This change is likely to be triggered by the binding of the E2 glycoprotein to the assembled NC during virus budding (20). It has also been proposed that residues 99 to 113 of the nucleocapsid protein are important in binding ribosomes to facilitate NC uncoating (35, 40, 41). The priming event, which prepares the core for disassembly, may be required for efficient interaction of the NC with ribosomes. The deletion of residues 97 to 106 may disrupt the priming event, rendering the CΔ(97-106) cores inefficient in disassembly. The second-site mutations present in the revertants could obviate the need for priming of the NC during budding and thereby circumvent a defective interaction of the NC and the cytoplasmic domain of E2. Alternatively, deletion of residues 97 to 106 may remove residues required for the NC to directly interact with ribosomes during the disassembly process.

Collectively, the data suggest that RNA encapsidation and disassembly may be related processes, and residues involved in these events may be adjacent to or overlapping with one another in the primary sequence of the nucleocapsid protein.

Interestingly, double substitution of two hydrophobic amino acids just downstream of the region deleted in CΔ(97-106) has a significant effect on virus replication (20). The efficiency of core formation is decreased when mutations at Leu-108 and Leu-110 are simultaneously introduced into the virus. Although RNA binding has not yet been examined in these mutants, it seems unlikely that the nonpolar residues are involved in dictating specificity in encapsidation. Instead, it is more reasonable that these residues are involved in protein-protein interactions required for NC assembly.

The results presented here and elsewhere suggest that residues 97 to 113 of the SIN nucleocapsid protein are important in both RNA encapsidation and NC assembly and disassembly. Residues 97 to 106 are clearly involved in dictating specificity in encapsidation, and residues 108 and 110 appear to be involved in NC assembly. Although the relationship between these residues has not yet been defined, it is likely that mutations which affect core formation also affect core disassembly. Further examination of the revertant viruses as well as viruses containing double mutations at Leu-108 and Leu-110 may help elucidate the connection between RNA binding and assembly-disassembly in the life cycle of the alphavirus.

ACKNOWLEDGMENTS

We thank Zhu Zhu and Bonnie McKinney for excellent technical assistance and Jian-sheng Yao for help in establishing electroporation conditions. We also thank James Strauss and Ellen Strauss for critical reading of the manuscript and for the anti-E2 antibodies, and we thank Charles Rice for kindly providing plasmid pGEM-J1. Useful discussions with Sondra Schlesinger, Barbara Weiss, Jodi Muckelbauer, and Michael Rossmann are also gratefully acknowledged.

This work was supported by Public Health Service grant AI33982. Support for antibody production was through the Purdue University Cancer Center Antibody Production Facility. Support from a Lucille P. Markey Foundation grant for structural studies is also acknowledged. K.E.O. is a recipient of a Purdue Research Foundation fellowship.

REFERENCES

1. **Aliperti, G., and M. J. Schlesinger.** 1978. Evidence for an autoprotease activity of Sindbis virus capsid protein. Virology **90:**366–369.
2. **Boege, U., G. Wengler, G. Wengler, and B. Wittman-Liebold.** 1981. Primary structure of the core proteins of the alphaviruses Semliki Forest virus and Sindbis virus. Virology **113:**293–303.
3. **Chang, G.-J. J., and D. W. Trent.** 1987. Nucleotide sequence of the genome region encoding the 26S mRNA of Eastern equine encephalitis virus and the deduced amino acid sequence of the viral structural proteins. J. Gen. Virol. **68:**2129–2142.
4. **Cheng, R. H., R. J. Kuhn, N. H. Olson, M. G. Rossmann, H.-K. Choi, T. J. Smith, and T. S. Baker.** 1995. Nucleocapsid and glycoprotein organization in an enveloped virus. Cell **80:**621–630.
5. **Choi, H.-K., S. Lee, Y.-P. Zhang, B. R. McKinney, G. Wengler, M. G. Rossmann, and R. J. Kuhn.** Structural analysis of Sindbis virus capsid mutants involving assembly and catalysis. Submitted for publication.
6. **Choi, H.-K., L. Tong, W. Minor, P. Dumas, U. Boege, M. G. Rossmann, and G. Wengler.** 1991. Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion. Nature (London) **354:**37–43.
7. **Coombs, K., B. Brown, and D. T. Brown.** 1984. Evidence for a change in capsid morphology during Sindbis virus envelopment. Virus Res. **1:**297–302.
8. **Dalgarno, L., C. M. Rice, and J. H. Strauss.** 1983. Ross River virus 26S RNA: complete nucleotide sequence and deduced sequence of the encoded structural proteins. Virology **129:**170–187.
9. **Forsell, K., M. Suomalainen, and H. Garoff.** 1995. Structure-function relation of the NH_2-terminal domain of the Semliki Forest virus capsid protein. J. Virol. **69:**1556–1563.
10. **Garoff, H., A. M. Frischauf, K. Simons, H. Lehrach, and H. Delius.** 1980. The capsid protein of Semliki Forest virus has clusters of basic amino acids and prolines in its amino-terminal region. Proc. Natl. Acad. Sci. USA **77:** 6376–6380.
11. **Geigenmüller-Gnirke, U., H. Nitschko, and S. Schlesinger.** 1993. Deletion analysis of the capsid protein of Sindbis virus: identification of the RNA binding region. J. Virol. **7:**1620–1626.
12. **Glanville, N., and J. Ulmanen.** 1976. Biological activity of in vitro synthesized protein: binding of Semliki Forest virus capsid protein to the large ribosomal subunit. Biochem. Biophys. Commun. **71:**393.
13. **Hahn, C. S., S. Lustig, E. G. Strauss, and J. H. Strauss.** 1988. Western equine encephalitis virus is a recombinant virus. Proc. Natl. Acad Sci. USA **85:**5997–6001.
14. **Hahn, C. S., E. G. Strauss, and J. H. Strauss.** 1985. Sequence analysis of three Sindbis virus mutants temperature-sensitive in the capsid autoprotease. Proc. Natl. Acad. Sci. USA **82:**4648–4652.
15. **Hahn, C. S., and J. H. Strauss.** 1990. Site-directed mutagenesis of the proposed catalytic amino acids of the Sindbis virus capsid protein autoprotease. J. Virol. **64:**3069–3073.
16. **Harrison, S. C.** 1986. Alphavirus structure, p. 21–34. *In* S. Schlesinger and M. J. Schlesinger (ed.), The Togaviridae and Flaviviridae. Plenum Publishing Corp., New York.
17. **Kinney, R. M., R. J. B. Johnson, V. L. Brown, and D. W. Trent.** 1986. Nucleotide sequence of the 26S mRNA of the virulent Trinidad donkey strain of Venezuelan equine encephalitis virus and deduced sequence of the encoded structural proteins. Virology **152:**400–413.
18. **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.
19. **Kuhn, R. J., H. G. M. Niesters, Z. Hong, and J. H. Strauss.** 1991. Infectious RNA transcripts from Ross River virus cDNA clones and the construction and characterization of defined chimeras with Sindbis virus. Virology **182:** 430–441.
20. **Lee, S., K. E. Owen, H. K. Choi, H. Lee, G. Lu, G. Wengler, D. T. Brown, M. G. Rossmann, and R. J. Kuhn.** Identification of a protein binding site on the surface of the alphavirus nucleocapsid and its implications in virus assembly. Structure, in press.
21. **Levinson, R., J. H. Strauss, and E. G. Strauss.** 1990. Determination of the complete nucleotide sequence of the genomic RNA of O'Nyong-nyong virus and its use in the construction of phylogenetic trees. Virology **175:**110–123.
22. **Li, G., and C. M. Rice.** 1989. Mutagenesis of the in-frame opal termination codon preceding nsP4 of Sindbis virus: studies of translational readthrough and its effect on virus replication. J. Virol. **63:**1326–1337.
23. **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 6000-molecular-weight membrane protein modulates virus release. J. Virol. **65:**4107–4113.
24. **Little, C. N., and R. J. Kuhn.** Unpublished data.
25. **Lopez, S., J.-S. Yao, R. J. Kuhn, E. G. Strauss, and J. H. Strauss.** 1994. Nucleocapsid-glycoprotein interactions required for assembly of alphaviruses. J. Virol. **68:**1316–1323.
26. **Paredes, A. M., M. L. Simon, and D. T. Brown.** 1993. The mass of the Sindbis virus nucleocapsid suggests it has T=4 icosahedral symmetry. Virology **187:** 329–332.
27. **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.
28. **Rice, C. M., and J. H. Strauss.** 1981. Nucleotide sequence of the 26S mRNA of Sindbis virus and deduced sequence of the encoded virus structural proteins. Proc. Natl. Acad. Sci. USA **78:**2062–2066.
29. **Rossmann, M. G., and J. E. Johnson.** 1989. Icosahedral RNA virus structure. Annu. Rev. Biochem. **58:**533–573.
30. **Rümenapf, T., E. Strauss, and J. Strauss.** 1995. Aura virus is a new world representative of Sindbis-like viruses. Virology **208:**621–633.
31. **Rümenapf, T., E. G. Strauss, and J. H. Strauss.** 1994. Subgenomic mRNA of Aura alphavirus is packaged into virions. J. Virol. **68:**56–62.
32. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
33. **Schlesinger, S.** Personal communication.
34. **Schlesinger, S., and B. G. Weiss.** 1986. Defective RNAs of Alphaviruses, p. 149–166. *In* S. Schlesinger and M. J. Schlesinger (ed.), The Togaviridae and Flaviviridae. Plenum Press, New York.
35. **Singh, I., and A. Helenius.** 1992. Role of ribosomes in Semliki Forest virus nucleocapsid uncoating. J. Virol. **66:**7049–7058.
36. **Söderlund, H.** 1973. Kinetics of formation of Semliki Forest virus nucleocapsid. Intervirology **1:**354–361.
37. **Strauss, J. H., and E. G. Strauss.** 1994. The alphaviruses: gene expression, replication, and evolution. Microbiol. Biol. Rev. **58:**491–562.
38. **Strauss, J. H., E. G. Strauss, and R. J. Kuhn.** 1995. Budding of alphaviruses. Trends Microbiol. **3:**346–350.
39. **Tong, L., G. Wengler, and M. G. Rossmann.** 1993. Refined structure of the Sindbis virus core protein and comparison with other chymotrypsin-like serine proteinase structures. J. Mol. Biol. **230:**228–247.
40. **Wengler, G., and G. Wengler.** 1984. Identification of a transfer of viral core protein to cellular ribosomes during the early stages of alphavirus infection. Virology **134:**435.
41. **Wengler, G., D. Wurkner, and G. Wengler.** 1992. Identification of a sequence element in the alphavirus core protein which mediates interaction of cores with ribosomes and the disassembly of cores. Virology **191:**880–888.