JOURNAL OF VIROLOGY, Dec. 1989, p. 5310–5318
0022-538X/89/125310-09$02.00/0
Copyright © 1989, American Society for Microbiology
Vol. 63, No. 12

Evidence for Specificity in the Encapsidation of Sindbis Virus RNAs

BARBARA WEISS, HANS NITSCHKO, INGRID GHATTAS, REBECCA WRIGHT, AND SONDRA SCHLESINGER*

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

Received 20 July 1989/Accepted 30 August 1989

We investigated the interaction of the capsid protein of Sindbis virus with Sindbis viral RNAs and defined a region of the genome that is required for binding in vitro and for packaging in vivo. The binding studies were performed with purified capsid protein immobilized on nitrocellulose and ^{32}P-labeled RNAs transcribed in vitro from viral and nonspecific cDNAs. Genomic and defective interfering (DI) RNAs bound capsid protein significantly better than either the subgenomic (26S) RNA or nonspecific RNAs. Transcripts prepared from either truncated or deleted cDNAs were used to define the segment required for binding. This segment, which is represented twice in DI RNA, lies between nucleotides 746 and 1226 of the genomic RNA and is within the coding region of the nonstructural protein nsP1. Insertion of a domain covering these sequences into a nonviral RNA was able to convert it from a background level of binding to an activity that was 80% that of the Sindbis virus DI RNA. We analyzed DI RNA transcripts in detail because they could be studied not only for the ability to bind capsid protein in vitro but also for the ability to be replicated and packaged in vivo in the presence of helper virion RNA. The results obtained with three DI RNAs are reported. One (CTS14), which has one copy of the binding domain, bound efficiently to capsid protein in vitro and was packaged in vivo as measured by amplification on passaging. In contrast, a DI RNA (CTS1) which lacked this region did not bind to capsid protein and was not detected on passaging. By using lipofectin (P. L. Felgner, T. R. Gadek, M. Holm, R. Roman, H. W. Chan, M. Wenz, J. P. Northrop, G. M. Ringold, and M. Danielson, Proc. Natl. Acad. Sci. USA 84:7413–7417, 1987) to enhance RNA uptake, we were able to demonstrate that CTS1 RNA was replicated in the transfected cells. It was replicated to the same level as another DI RNA (CTS253) which has only the 3′ 279 nucleotides of the binding domain and these are located near the 3′ terminus of the RNA. CTS253 bound capsid protein to an intermediate level but was amplified on passaging. The binding studies and the in vivo packaging data, taken together, provide strong support for the conclusion that there is a specific capsid recognition domain in Sindbis virus RNA that plays a role in nucleocapsid assembly.

Sindbis virus is a positive-strand enveloped virus, the prototype member of the *Alphavirus* genus of the *Togaviridae* family. Both the enveloped particle and the internal nucleocapsid have icosahedral structures (8, 10). The genomic RNA (49S RNA) contains 11,703 nucleotides plus a poly(A) tail at the 3′ terminus. Strauss et al. (31) determined the complete sequence of the cDNA derived from the genomic RNA, and Rice et al. have constructed cDNAs that can be transcribed in vitro to produce infectious RNAs (24). In infected cells, the genomic RNA serves as a messenger for the nonstructural proteins and as a template for synthesis of the complementary negative strand (reviewed in reference 32). The latter is the template for synthesis of new genomic RNAs and also for synthesis of a subgenomic RNA (26S RNA) which is the mRNA for the viral structural proteins (the capsid protein and two membrane glycoproteins). These proteins are synthesized as a polyprotein with the capsid protein located at the amino-terminal end. The capsid protein contains protease activity responsible for cleaving the protein from the remaining nascent peptide, which is then transported into the rough endoplasmic reticulum. The peptide inserted in the membrane undergoes further proteolytic cleavages and co- and posttranslational modifications as the proteins are transported through the Golgi membranes to the plasma membrane of the cell (reviewed in reference 26). Assembly of virion particles requires at least two steps. The first is interaction of the 49S virion RNA with the capsid protein to form the nucleocapsid; the second is interaction of the nucleocapsid with the viral glycoproteins embedded in the cellular membranes. In animal cells, the latter interaction results in budding and release of the virus from the plasma membrane of the cell. Very little is known about the specificity of either of these steps, but there is specificity, since only the 49S virion RNA, and not the subgenomic 26S RNA, is encapsidated. Recent studies using idiotypic antibodies have provided evidence for a specific interaction between the cytoplasmic tail of the E2 glycoprotein and the nucleocapsid of the related Semliki Forest virus (37).

There are several examples of specificity in the encapsidation of viral nucleic acids. The best-defined system of virus assembly is that of tobacco mosaic virus, a helical plant virus particle in which assembly is initiated by specific interaction between the coat protein disk and an internal sequence in the RNA genome (reviewed in reference 4). Specificity of assembly of the icosahedral plant virus turnip crinkle virus was demonstrated by studies showing a preference for virion RNA over rRNA in the reassembly of the virus in vitro (29). There are domains in the genomes of retroviruses that are required for packaging of the RNA into virions (1, 20). Previous studies with Sindbis virus, however, have shown that in vitro, the viral capsid protein interacts with a variety of RNAs, as well as with other negatively charged macromolecules to form nucleocapsidlike structures (39–41). These results led Wengler to suggest that the specificity of encapsidation in vivo could be due to the greater stability of the nucleocapsid containing the viral genomic RNA rather than to any specific packaging signal (39).

Cloning of the cDNA of the Sindbis virus genome (24) and the smaller defective interfering (DI) RNA (19, 22) and the biological activity of their RNA transcripts has made it

* Corresponding author.

possible to analyze and define sequences in these genomes that are required for recognition in replication and to determine whether there are sequences required for encapsidation. In previous studies, we performed a deletion analysis of the cDNA of a DI RNA (19). In that work, the DI RNA transcripts were transfected into cells in the presence of helper Sindbis virus, and amplification of the DI RNA was analyzed after two or three passages on cells. We demonstrated that only deletions at the 5' and 3' termini of the DI genome prevented amplification. Furthermore, substitution of 75% of the internal viral sequences with foreign sequences did not destroy the ability of the DI RNA to be amplified (18). Because it was necessary to passage the virus to detect the DI genomes in infected cells, we could not distinguish between sequences required for replication and those that might be required for encapsidation. We have now been able to analyze the specificity of encapsidation by two independent means. (i) We identified sequences in the genomic RNA that interact specifically with the capsid protein in binding assays using purified capsid protein and RNA transcripts, and (ii) we were able to analyze the replication of DI RNAs in transfected cells by using lipofectin (7) to increase the sensitivity of our assays. We report here that the sequences of the genome required for binding to the capsid protein are located near the 5' terminus of the RNA, in the gene for nonstructural protein nsP1. These sequences were not essential for replication, but DI RNAs that lacked these sequences were not detected in passaged virus.

MATERIALS AND METHODS

Plasmids and in vitro transcription. The DI cDNA clones KDI25 and CTS253 have already been described in detail (18, 19, 22). For a diagram of their structures, see Fig. 5A. Rice et al. have reported the construction of the full-length cDNA clone of Sindbis virus 49S RNA (Toto1101) (24). Toto1102, a derivative of Toto1101, contains an *Xba*I linker insertion at an *Rsa*I site at position 7612. The Toto clones were provided by C. Rice and H. Huang, Washington University, St. Louis, Mo. Xiong et al. have described the self-replicating Sindbis virus vector TRCAT (43). The KDI25 cDNA derivative used to generate negative-strand DI RNA transcripts was obtained by insertion of a 20-nucleotide oligomer containing the T7 promoter downstream of the 3' terminus of the DI cDNA at an *Eco*RI site. An *Xho*I linker was inserted between the SP6 bacteriophage promoter and the start of DI RNA 5' sequences. Linearization of the DNA at the *Xho*I site permitted runoff transcription by use of the T7 polymerase to generate negative-strand RNAs. Plus-sense transcripts were synthesized from KDI25 by using cDNA linearized at a downstream *Pst*I site by using SP6 bacteriophage DNA-dependent RNA polymerase. CTS1, a derivative of CTS253, was constructed by ligation of a 2,520-nucleotide *Sca*I fragment from CTS253 with a 1,839-nucleotide *Sca*I-*Spl*I fragment from Toto1102 (see Fig. 5A). Plasmid pCAT was provided by H. Huang. It contains the chloramphenicol acetyltransferase (CAT) gene inserted downstream of the SP6 promoter in a pBR322-derived vector. Linearization of this plasmid at a unique *Sac*I site generated, upon runoff transcription, an RNA of 2,608 nucleotides, with 657 nucleotides representing the CAT gene sequence and the remainder derived from pBR322. The plasmid containing the gene for β-actin was provided by Dorothy Schafer (Washington University).

Transcriptions using SP6 bacteriophage polymerase were performed as described previously (19). ^{32}P-labeled RNA probes were synthesized by using [^{32}P]CTP in reactions in which the concentration of CTP was 200 μM. All RNA transcripts were analyzed by gel electrophoresis after denaturation with glyoxal (6) to verify their size and that they were intact.

Northwestern blot analysis. Sindbis virus purified by sucrose gradient centrifugation was disrupted in sodium dodecyl sulfate. The viral proteins were separated by electrophoresis in 7.5% polyacrylamide gels (15). The proteins were electrophoretically transferred to nitrocellulose as described by Towbin et al. (33). Identical membrane transfer strips were probed with ^{32}P-labeled RNAs by a modification of the procedure of Bowen et al. (5). The blots were washed in probe buffer (10 mM Tris [pH 7.5], 50 mM NaCl, 1 mM EDTA, 1× Denhardt solution) for 1 h at room temperature and incubated for another hour in this buffer in 250 μg of *Escherichia coli* total RNA per ml. Carrier RNA was required to block nonspecific binding. Commercially obtained *Saccharomyces cerevisiae* RNA did not give the same discrimination as *E. coli* RNA. RNA probes labeled with ^{32}P were then added to the nitrocellulose strips, followed by incubation for 1 h with continuous shaking. Unbound probe was removed by three consecutive 10-min washes with probe buffer.

Nitrocellulose dot blot binding assay. A dot blot manifold (Schleicher & Schuell, Inc.) was used to apply 600 to 1,200 ng of capsid protein in 1% acetic acid or water to nitrocellulose. Protein was allowed to bind for 30 min in a 20-μl volume, and then gentle suction was applied to the manifold. Each well was washed with 1% acetic acid or water, and the entire membrane was then washed twice for 30 min each time in probe solution. Individual filter squares containing capsid protein were exposed to carrier RNA as indicated above and then to radioactive probes. The filters were washed as described above. Each datum point was obtained by averaging the results from triplicate samples. The counts per minute (cpm) added to each filter ranged from 0.4×10^6 to 2.2×10^6 in the different experiments (see Table 1 for examples). Both positive (DI25) and negative (CAT) RNAs were included in the assays in all experiments, and the data were normalized to the bound fraction of DI25 by using the following equation: (bound cpm/input cpm)/(bound DI25 cpm/input DI25 cpm) × 100.

Purification of capsid protein from Sindbis virions. The procedure used to purify capsid protein was identical to that described by Wengler et al. (40). The purity and integrity of the final recovered capsid polypeptide were determined by polyacrylamide gel electrophoresis and Coomassie blue staining for protein.

RNA analysis. RNA was isolated from monolayers of chicken embryo fibroblasts (CEF) as previously described and analyzed by agarose gel electrophoresis following denaturation with glyoxal in dimethyl sulfoxide (6).

Lipofectin-mediated cotransfections, passaging of virus, and labeling of viral RNAs. The cationic lipid lipofectin (7) was used to facilitate uptake of Sindbis virus RNAs in secondary cultures of CEF. The efficiency of transfection with lipofectin was at least 10-fold better than that obtained with DEAE-dextran and permitted us to detect Sindbis virus-specific RNAs synthesized after transfection without passaging. For transfections, virion 49S RNA (0.4 μg) or TRCAT RNA (0.5 μg) and DI RNA (0.5 μg) were combined and diluted into 0.25 ml of serum-free minimal essential medium (MEM). The diluted RNAs were then mixed with 7.5 μg of lipofectin in 0.25 ml of MEM. The 0.5-ml lipid-RNA mixture was applied to confluent monolayers of CEF in 60-mm-

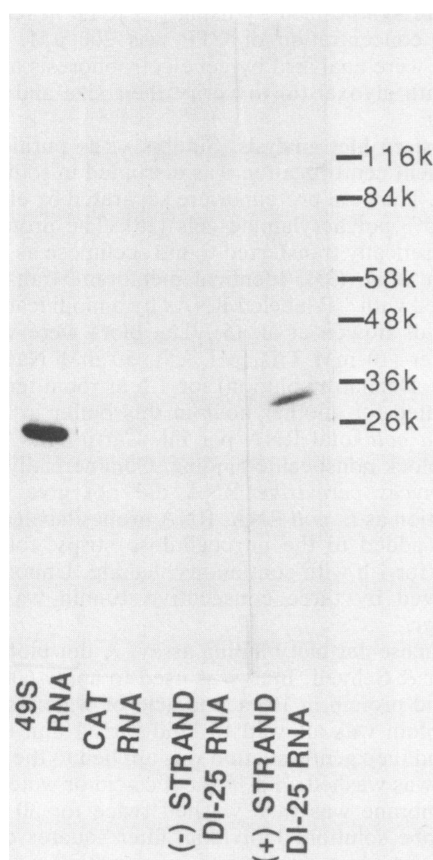

FIG. 1. Northwestern blots showing specific binding of Sindbis virus RNAs to viral capsid protein. Sindbis virus structural proteins were electrophoresed through a 7.5% polyacrylamide denaturing gel and electrotransferred to nitrocellulose. Following preincubation with *E. coli* RNA (250 μg/ml) in binding buffer, the nitrocellulose strips were probed with ^{32}P-labeled RNA transcripts. The RNAs used are indicated; negative-strand DI25 RNA is complementary to DI25 RNA and is described in Materials and Methods. The molecular weights (in thousands) of coelectrophoresed stained marker proteins are indicated in the margin.

diameter dishes that had been washed twice with serum-free MEM. The dishes were incubated at 30°C for 1 h. A 1.5-ml volume of serum-free MEM and dactinomycin (1 μg/ml) were added for an additional hour. This medium was then replaced with 2.0 ml of MEM containing 3% fetal bovine serum–1 μg of dactinomycin per ml–[^{3}H]uridine at 25 μCi/ml, and the cells were incubated overnight at 30°C. In some experiments, samples of medium from this first passage or from subsequent passages were used to infect CEF in dactinomycin and labeled as described above.

RESULTS

Binding of Sindbis virus RNAs to viral capsid protein. In cells infected with Sindbis virus, newly synthesized virion RNA interacts with the capsid protein to form viral nucleocapsids. This step is assumed to be one of self-assembly, because purified capsid protein and RNA can interact in vitro to form nucleocapsidlike structures (40). We investigated the possibility that there is a specific binding site(s) on the viral RNA by analyzing the specificity of the interaction between the RNAs and purified capsid protein immobilized on nitrocellulose. In the first experiments, the structural proteins of the virus were separated by polyacrylamide gel electrophoresis and transferred to nitrocellulose. Four different ^{32}P-labeled RNA transcripts were used as probes to test for their ability to bind to the proteins. The following three were Sindbis virus specific: plus-strand genomic 49S RNA, the plus strand of a DI RNA (DI25), and the negative strand of DI25 RNA. The fourth RNA was a transcript derived from plasmid pCAT and contained sequences from pBR322, as well as the complete CAT RNA sequences. Only two of these RNAs—the 49S genomic RNA and the plus strand of DI25—gave a strong signal in the capsid protein region (Fig. 1). None of the RNAs bound to the viral glycoproteins. Specificity was observed, however, only when carrier *E. coli* RNA was present during incubation.

TABLE 1. Dot blot assay of binding of RNA to capsid protein

RNA	Input cpm (10^6)[a]	cpm bound (± SD)	% Bound
Expt 1			
DI25	0.8	3,796 ± 542	0.47
DI25	4.3	22,765 ± 574	0.53
β-Actin	0.8	447 ± 277	0.06
Expt 2			
DI25	1.6	26,742 ± 824	1.7
CAT	2.0	4,350 ± 740	0.20

[a] In experiment 1, 10^6 cpm equalled 0.37 pmol; in experiment 2, 10^6 cpm equalled 0.2 pmol.

To analyze RNA binding more quantitatively, we purified the capsid protein by the method of Wengler et al. (40) so that it could be bound directly to nitrocellulose. Binding of ^{32}P-labeled transcripts to capsid protein was then determined under conditions in which the capsid protein was in at least 10-fold molar excess (Table 1). Binding of DI25 RNA was seven- to eightfold higher than that of two unrelated RNA transcripts (β-actin and CAT) of similar sizes and specific activities. The binding conditions described were optimized to achieve specificity (see Materials and Methods). Although reductions in pH (varied down to 5.5 in 0.05 M KPO_4 buffer) or addition of Mg^{2+} ions (5 to 10 mM $MgCl_2$) led to substantial increases in binding, they also caused an almost total loss of specificity, even in the presence of competitor RNA.

Since most of the sequences in DI25 are derived from the 5′-terminal region of genomic RNA (22), it seemed likely that if a specific binding domain were present, it would be in that region. We examined this possibility by comparing the binding of a truncated 49S RNA transcript containing the 5′-proximal 4,633 nucleotides of the genome with both DI25 and 26S RNA (Fig. 2). Two transcripts of DI25 were tested. One was transcribed from a cDNA that had been linearized at the *Pst*I site (the standard transcript of 2,358 nucleotides). The second was transcribed from a cDNA linearized at the downstream *Sac*I site so that the RNA, which contained both Sindbis virus and plasmid sequences, would be essentially the same size (4,241 nucleotides) as the truncated 49S RNA and the 26S RNA. The two DI RNAs and the truncated 49S transcript bound to the capsid protein to the same extent. In contrast, the 26S RNA transcript bound much less efficiently. This RNA, which represents the 3′ one-third of the genome, is not encapsidated in vivo.

Several other DI RNAs with alterations in the 5′ terminus were tested for binding. These included a DI RNA in which the 5′ end is identical to that of 49S RNA (19, 34) and a DI

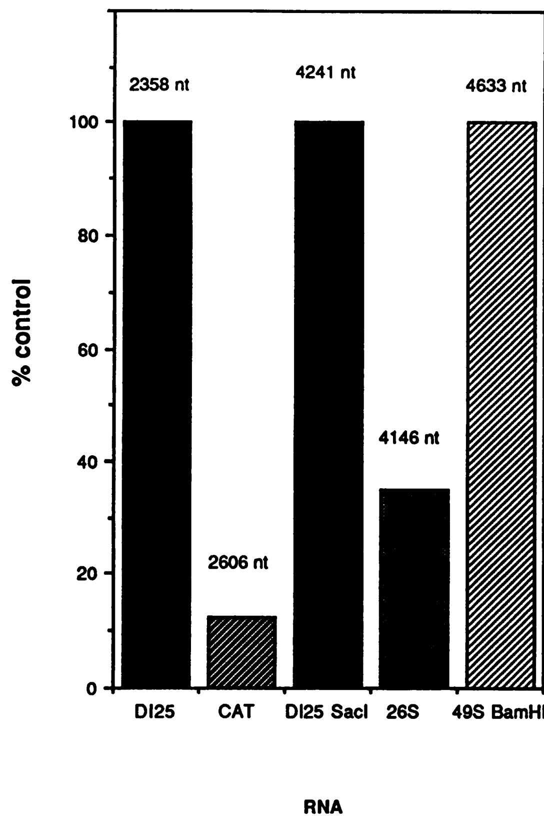

FIG. 2. Binding of Sindbis virus-specific RNAs to purified capsid protein. Capsid protein (600 ng per filter) was blotted onto nitrocellulose, followed by incubation with various ^{32}P-labeled RNA transcripts as described in Materials and Methods. The counts per minute added to each filter ranged from 0.6×10^6 to 1.8×10^6 for the different transcripts. Each sample was analyzed in triplicate. The data were normalized to the bound fraction of the 2,358-nucleotide (nt) DI25 RNA derived by transcription of the cDNA linearized at a *Pst*I site. A second DI25 transcript was obtained by using as a template cDNA linearized at the downstream *Sac*I site to generate an extended transcript of 4,241 nucleotides containing DI25 and pBR322 sequences. The latter transcript was essentially the same size as the 49S and 26S RNA transcripts. The 49S RNA used in this experiment was obtained by transcription of Toto 1102 cDNA linearized at a *Bam*HI site to generate a 4,633-nucleotide RNA.

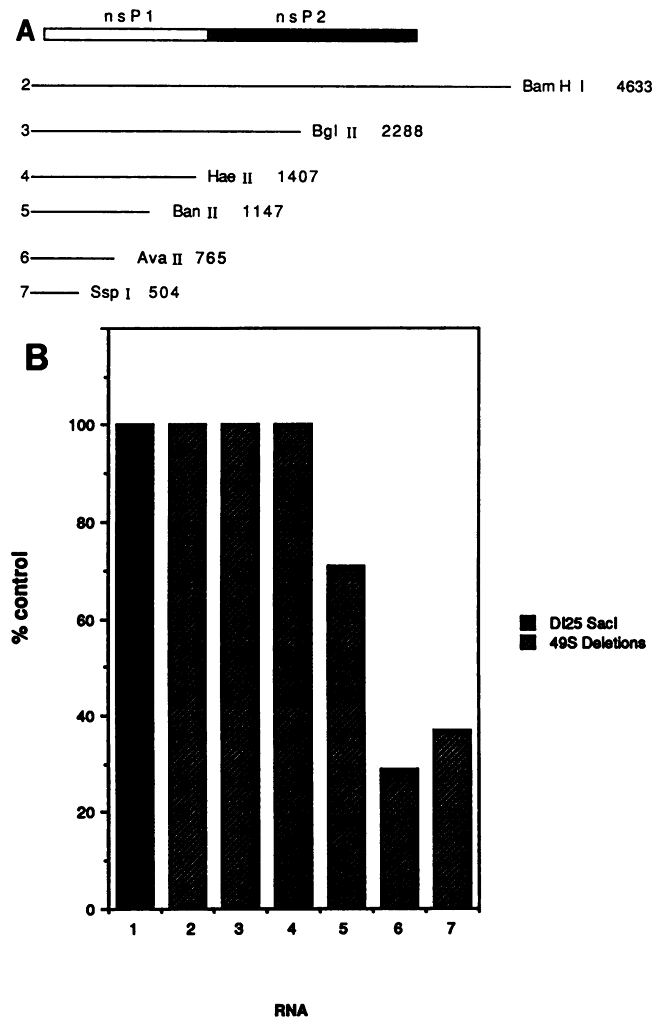

FIG. 3. Binding of 5' fragments of 49S RNA to capsid protein. (A) 49S RNA transcripts were derived by using as a template Toto 1102 cDNA linearized at the sites indicated. The length of each resulting transcript is given (in nucleotides) to the right of the enzyme used. The positions of the coding sequences for nsP1 and nsP2 are indicated by the open and closed bars, respectively. The nsP1 domain extends from nucleotides 60 to 1679 in 49S RNA (31). (B) Relative binding of RNA probes described in panel A to capsid protein using the binding assay described in Fig. 2 and in Materials and Methods. Each numbered bar in the histogram corresponds to the transcript generated by cleavage at the sites indicated in panel A.

RNA with a deletion from nucleotides 162 to 241 (19). This region includes a stretch of nucleotides that is highly conserved among alphaviruses and is thought to have some *cis*-acting function (23). These RNAs bound to capsid protein with the same efficiency as DI25 RNA (data not shown).

Location of the capsid-binding domain in 49S RNA in the nsP1 gene. A series of truncated RNAs was obtained by linearizing the 49S cDNA template at a variety of restriction enzyme sites upstream from the site at nucleotide 4633 (Fig. 3A). The transcripts consisted of a nested set of RNA probes differing in the amounts of the 3' sequence they retained; all had identical 5'-proximal sequences. No reduction in binding was observed until a transcript size of 1,147 nucleotides was reached, and binding was decreased to an even greater extent when the transcript was reduced in size to 765 nucleotides (Fig. 3B). These data suggested that the sequences that confer binding specificity on the RNA begin upstream from nucleotide 1407 of the genomic RNA and are in the coding region of nsP1.

A possible criticism of the data in Fig. 3B is that the RNAs analyzed were of different sizes. To address this potential problem, we analyzed RNA transcripts of the same size with internal deletions. One deletion encompassed 637 nucleotides from a *Sma*I site at position 767 to an *Eco*47III site at nucleotide 1404. The other deletion was only 177 nucleotides, extending from the *Sma*I site at nucleotide 767 to a second *Sma*I site at nucleotide 944. The deleted cDNAs and the parental 49S cDNA were linearized at restriction sites such that the transcripts were all of comparable sizes. The binding studies showed that the large deletion reduced binding by 75%. The smaller deletion also had a significant but smaller effect on binding (Table 2).

Enhanced binding to capsid protein by CAT RNA containing an insertion from nsP1. CAT RNA served as our negative control. If the region in nsP1 is actually a recognition site for

TABLE 2. Binding of internally deleted 49S transcripts to capsid protein

Transcript	Runoff site	Transcript size (nucleotides)	% Binding[a]
49S	*Bgl*II	2,288	100
SE2	*Sac*II	2,132	25
SE	*Bgl*II	2,111	60
CAT	*Sst*I	2,606	12

nsP1
49S 2288
SE2 767 1404 2771
S3 767 944 2288

[a] Binding values are normalized to binding of the 49S *Bgl*II fragment previously shown by comparison with DI25 RNA to be efficiently bound to capsid protein.

capsid binding, it should be possible to confer binding activity on CAT RNA by inserting this sequence into it. We inserted a 572-nucleotide *Hha*I fragment extending from nucleotides 683 to 1255 in 49S RNA into the pCAT plasmid immediately downstream from the start site for in vitro transcription. The fragment was inserted in pCAT DNA in two orientations such that it was represented in RNA transcripts in a plus or minus sense. When it was inserted in the plus-sense orientation, binding was enhanced fivefold over that of unmodified CAT RNA. This level of binding was about 80% of that observed with DI25 RNA (Fig.4). In contrast, when the insertion was in the opposite orientation, there was no increase in binding over the background value. These data support the conclusion that the region identified by deletion analysis contains a major capsid recognition sequence.

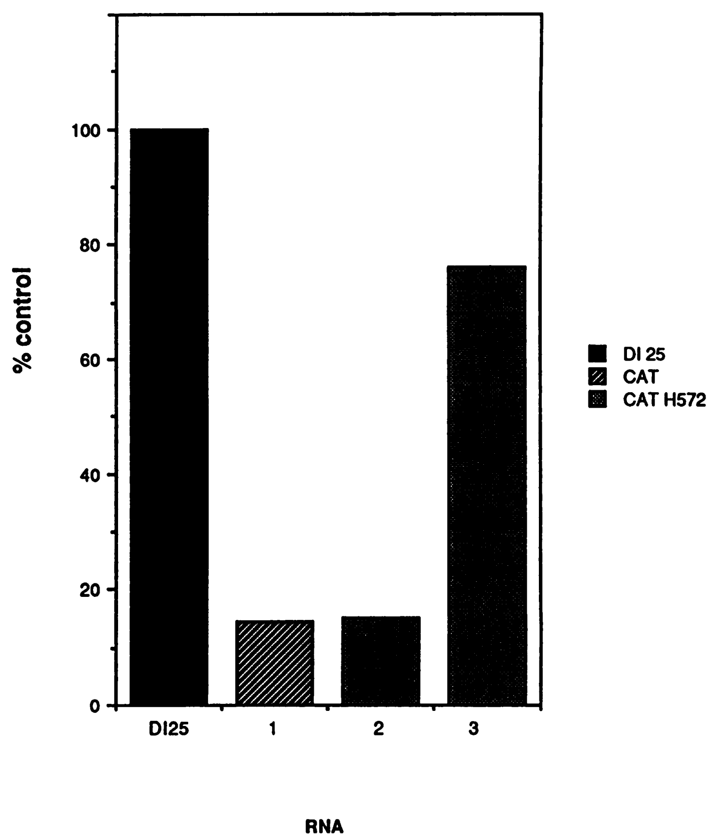

FIG. 4. The effect of insertion of 572 nucleotides from 49S virion RNA (the region from nucleotides 639 to 1255 in 49S virion RNA) into CAT RNAs. RNA-capsid binding was performed as described in Materials and Methods. Lanes: 1, unmodified CAT RNA transcript; 2, CAT RNA containing the 572-nucleotide insert in the inverse orientation; 3, CAT RNA containing the 572-nucleotide insertion in the correct orientation.

Capsid binding by DI RNAs containing CAT sequences. The domain in 49S RNA associated with specificity of binding to capsid protein is present in DI25 RNA in two copies (Fig. 5A). The first copy is located between nucleotides 823 and 1302; the second begins at nucleotide 1730 and terminates at nucleotide 2208 (19, 22). They correspond to nucleotides 746 to 1226 in the 49S virion RNA. A derivative of DI25, CTS253, in which foreign sequences were substituted for a large internal region of DI cDNA (Fig. 5A) had been constructed previously (18). The foreign sequences include CAT and simian virus 40 sequences. CTS253 retains the following two regions from DI25: the first 240 nucleotides at the 5′ end and the last 428 nucleotides at the 3′ end. This latter region contains 279 nucleotides from the 3′ region of the capsid-binding domain. CTS253 RNA is amplified on passaging and therefore is encapsidated (18). If in vivo assembly of this DI RNA requires sequences within the 279-nucleotide domain, DI RNAs which completely lack these sequences should not be packaged. To test this prediction, we constructed a cDNA derivative of CTS253 (CTS1) in which the 279 nucleotides of the binding domain were deleted (Fig. 5A). The first 938 nucleotides of CTS1 were derived from CTS253, and the remaining 1,359 were from the extreme 3′ terminus of 49S RNA. A similar cDNA clone, designated CTS14, which contained a 5′ insertion of the 572-nucleotide *Hha*I fragment (nucleotides 683 to 1255 of virion RNA) was also constructed (Fig. 5A).

DI RNAs transcribed from these modified cDNAs were examined for the ability to bind to capsid protein (Fig. 5B). CTS253 bound weakly to capsid protein. Binding of CTS1 was no better than that of CAT RNA, but CTS14 bound to capsid protein almost as well as did DI25.

Requirement of nucleotides in the capsid-binding domain for encapsidation of DI RNAs. The three DI RNA transcripts, CTS253, CTS1, and CTS14, that had been analyzed in the binding assay were tested in vivo for the ability to be packaged in cotransfection experiments. They were each cotransfected with 49S virion RNA into CEF. Cotransfected cells were labeled with [^{3}H]uridine in the presence of dactinomycin, and RNA was extracted and analyzed by agarose gel electrophoresis. All three DI RNAs were detected in passage 1 (Fig. 6). The medium harvested from the transfected cells (passage 1) was used for infection of new cells, and RNAs from this passage 2 infection were analyzed (Fig. 6). In passage 2, CTS253 and CTS14 RNAs were clearly evident, but CTS1 RNA was not. An additional passage (passage 3) led to amplification and enrichment of both CTS253 and CTS14 but not CTS1. These results demonstrated that CTS1 was lost on passaging and provide strong evidence that the binding domain we identified is required for packaging in vivo.

A quantitative analysis of the RNAs synthesized in passage 1 showed that the amounts of each species of RNA in lanes 1 and 2 were identical and that in each case the DI

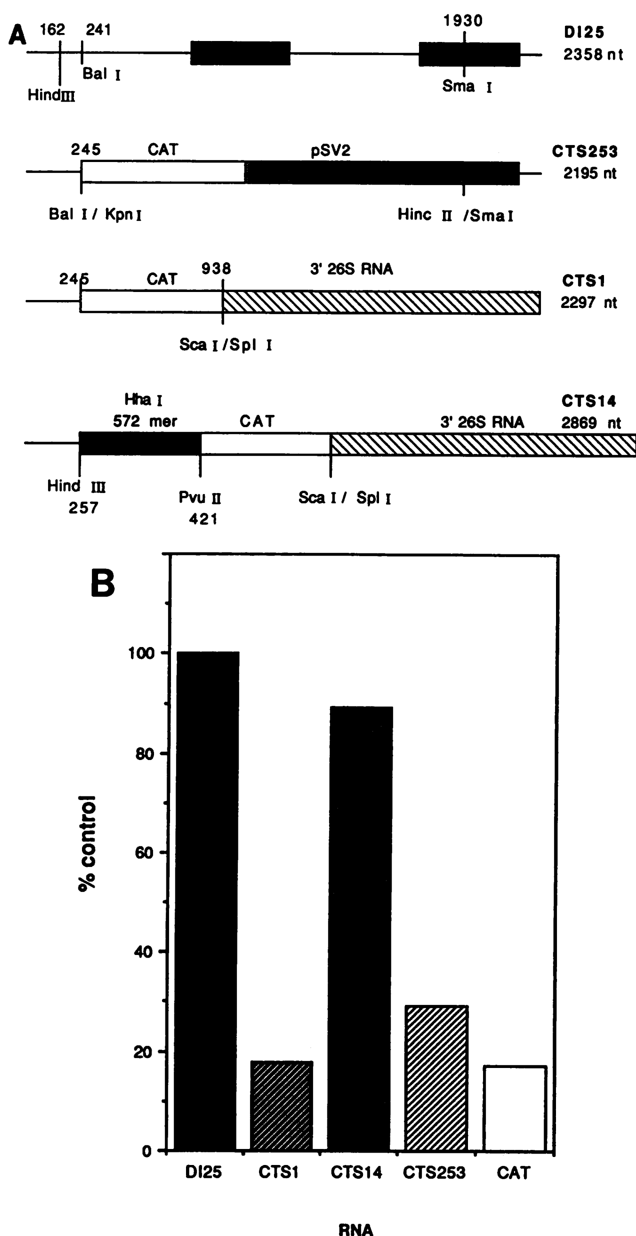

FIG. 5. Binding of CAT-containing DI RNAs to capsid protein. (A) Structures of the DI RNAs tested. The name and length of each transcript is given on the right. The black bars indicate regions defined as capsid-binding domains. The 3′ 26S RNA domain fused to CAT in CTS1 and CTS14 was derived from nucleotides (nt) 10381 through 11740 in Toto1102 cDNA. CTS14 has the 572-nucleotide insertion derived from nucleotides 639 to 1255 of 49S virion RNA. (B) Efficiency of binding of various ^{32}P-labeled DI RNAs containing CAT sequences. Binding was normalized to that of DI25 RNA (2,358 nucleotides), which was set at 100%.

RNA represented 10% of the total radioactivity. CTS14 RNA (lane 3) made up 40% of the total radioactivity. The molar ratios of DI RNA to 49S RNA were 1.7, 1.6, and 8.4 for CTS253, CTS1, and CTS14 RNAs, respectively. The higher level of CTS14 RNA in the transfected cells suggested that the 572-nucleotide insert affected replication as well as encapsidation. An alternative explanation is that CTS14 was packaged efficiently and, after being released from transfected cells, was able to infect new cells that would also be

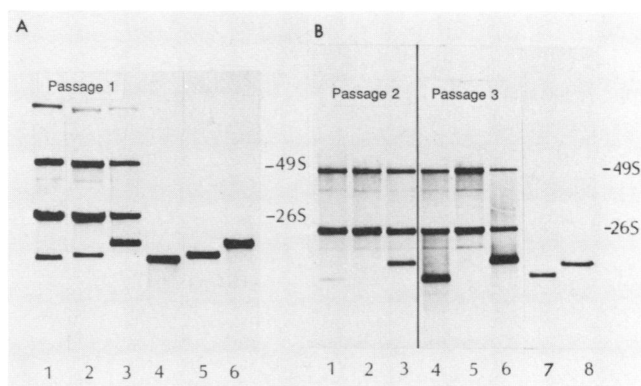

FIG. 6. Analysis of Sindbis virus RNAs after transfection or after passaging. (A) Electrophoretic analysis of Sindbis virus RNAs synthesized in cells transfected with 49S RNA and DI RNAs. Secondary CEF were cotransfected with 49S virion RNA and DI RNA transcripts and labeled with [^{3}H]uridine in the presence of dactinomycin for 18 h at 30°C. The isolated RNA was denatured with glyoxal and electrophoresed through a 1% agarose gel as previously described (19). Lanes: 1 to 3, RNAs from cells transfected with 49S RNA and CTS253 (lane 1), CTS1 (lane 2), or CTS14 (lane 3); 4 to 6, transcripts of CTS253 (lane 4), CTS1 (lane 5), and CTS14 (lane 6). (B) Electrophoretic analysis of viral RNAs synthesized in passages 2 and 3. One-fourth of the medium harvested either from transfected cells or from passage 2 was used to infect new monolayers. The labeling conditions were the same as those described for panel A for passage 1. Lanes: 1 and 4, CTS253; 2 and 5, CTS1; 3 and 6, CTS14; 7, CTS253 transcript; 8, CTS14 transcript. The viral RNA bands in lanes 1 and 3 in passage 2 are lower in intensity than the bands in these lanes in passage 1 (transfected cells) because the autoradiogram was exposed for less time than the one in panel A. Longer exposures of the passage 2 autoradiogram showed no CTS1 band in lane 2.

infected by helper virus. We were able to distinguish between these possibilities in two ways. First we repeated the cotransfection experiment with anti-Sindbis virus antibodies which prevented the spread of both infectious virus and DI particles. With these antibodies, the titer of newly synthesized virus was reduced to less than 200 PFU/ml. The ratios of DI RNAs to 49S and 26S RNAs were similar to those shown in Fig. 6, with CTS14 again representing a higher percentage of the total radioactivity than either of the other two DI RNAs (data not shown). A second and more definitive experiment was to use as the helper an RNA lacking the structural genes. This RNA (referred to as TRCAT) is a self-replicating vector in which the Sindbis virus structural protein genes were replaced with the gene for the bacterial enzyme CAT (43). Cells transfected with TRCAT synthesize a subgenomic RNA that is translated to produce CAT protein. No viral structural proteins are produced, and no viral particles are formed. Each of the DI RNAs was transfected into CEF with TRCAT RNA, and the virus-specific RNAs synthesized were analyzed. CTS14 RNA was present at a level significantly above that of the other two DI RNAs (Fig. 7). It represented 41% of the total radioactivity; CTS253 and CTS1, respectively, represented 14 and 25% of the total radioactivity. These results demonstrated that in transfected cells, CTS14 was replicated to a greater extent than either of the other DI RNAs.

DISCUSSION

The data presented in this report show that there is a specific interaction between Sindbis virus RNA and the viral

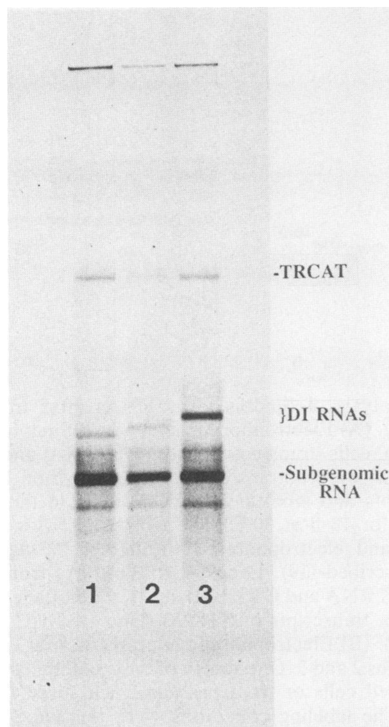

FIG. 7. Analysis of Sindbis virus DI RNAs after transfection with TRCAT RNA as the helper. CEF were cotransfected with TRCAT and CTS253 RNA (lane 1), CTS1 RNA (lane 2), or CTS14 RNA (lane 3). A 0.5-μg sample of each RNA was used. Transfections and labeling conditions were identical to those described in the legend to Fig. 6. The bands referred to as DI RNAs on the autoradiogram were identified by comparison with the appropriate transcripts run in adjacent lanes but not shown. The bands indicated as subgenomic RNAs have been identified as the subgenomic RNA produced when cells are transfected with TRCAT RNA (43).

capsid protein immobilized on nitrocellulose and that the specificity of binding is associated with a small region of the genome. This same region is required for a DI RNA to be packaged into particles released from infected cells. It lies between nucleotides 683 and 1255 of the virion RNA and is in the gene for the nonstructural protein nsP1. Deletion mapping of the virion cDNA demonstrated that almost 90% of the genome, those sequences downstream from nucleotide 1255, could be deleted without affecting the degree of binding of the RNA transcript. In contrast, deletions immediately upstream from that site inhibited binding of the RNA to capsid protein. These same sequences, whose absence in Sindbis virus RNAs prevented them from binding to the capsid, were able to impart binding activity to an RNA–the CAT RNA transcript—that had previously served to establish the nonspecific background value. There are almost 600 nucleotides in this region, but the actual sequence(s) required for recognition may be much smaller. Deletion of the 5′ 177 nucleotides decreased binding in vitro by 40% (Table 2), suggesting that at least some of these sequences are involved in binding recognition.

We had previously studied DI RNAs of Sindbis virus to identify the *cis*-acting sequences in their genomes that are essential for their amplification on passaging in the presence of helper virus (19). A major advantage of analyzing DI RNAs is that they are only one-fifth the size of the virion RNA and can be manipulated more easily than virion RNA. Most importantly, for these studies, it was possible to make deletions and insertions in the DI cDNAs and then analyze the RNA transcripts not only for binding to capsid protein but also for the ability to be replicated and packaged. We compared four DI RNAs. The first DI RNA transcript tested was DI25, the prototype of the DI RNAs and one that had been analyzed previously. This DI RNA, which has a 5′ terminus that is similar to that of $tRNA^{Asp}$ and distinct from that of virion RNA (21), bound capsid protein as well as virion RNA. The other three DI RNAs had some of the Sindbis virus sequences replaced by foreign sequences. Only one of these (CTS14) contained the entire capsid-binding domain. This DI RNA not only bound to the capsid protein with almost the same efficiency as DI25, but it also was packaged into extracellular particles as measured by its amplification on passaging. The DI RNA (CTS1) that lacked this domain did not bind the capsid and was not detected in infected cells in subsequent passages. The significant point is that CTS1 did replicate in transfected cells as well as did another DI RNA (CTS253) that is amplified on passaging (18; Fig. 6). In our previous study, we had demonstrated by S1 nuclease analysis that the CTS253 RNA detected in passage 3 is identical to the original transcript (18).

CTS14, the DI RNA into which the segment required for binding and for in vivo packaging was inserted, was present in cotransfected cells at higher levels than the DI RNAs that lacked all or part of this region, suggesting that a *cis*-acting replication-enhancing sequence also lies within this region. There are thus two regions near the 5′ terminus of the DI RNA that are not essential for but do stimulate RNA replication. The other region is a stretch of 51 nucleotides beginning about 150 nucleotides from the 5′ terminus of the virion genome that is highly conserved among alphaviruses (23). Our previous deletion mapping of DI25 RNA showed that DI RNAs are amplified without this conserved sequence, but not as well as when they are present (19, 27). The original analysis did not distinguish between effects on replication and those on encapsidation. Since DI RNAs that lack the 51 nucleotides showed no decrease in the ability to bind the capsid, we believe that they function in replication. The 572-nucleotide segment functioned both in encapsidation and replication, but further mapping will be necessary to determine whether the sequences that affect the two functions are identical.

CTS253 was something of an anomaly. It contains only 279 nucleotides at the 3′ end of the encapsidation domain and did not bind capsid protein as well as did DI25 or CTS14. The binding, however, was greater than that of either CTS1 or nonspecific CAT RNA. The inefficient binding could be due to absence of the complete recognition sequence or to its improper location at the 3′ terminus of the molecule. Binding of CTS253 was similar to that of 26S RNA, but only CTS253, and not 26S RNA, is packaged in vivo, even though 26S RNA is able to interact with capsid protein to form capsid-like structures in vitro (H. Nitschko et al., unpublished data). Other factors may be involved in the in vivo discrimination. 26S RNA is translated much more efficiently than CTS253 (18), and in infected cells, a greater percentage of it than of CTS253 RNA may be sequestered in polyribosomes and thus be unavailable for interaction with capsid protein. In experiments analyzing cells infected with Semliki Forest virus, Ulmanen et al. (36) could not find 26S RNA bound to capsid protein, although they were able to isolate a complex between capsid protein and the 60S ribosomal subunit. To account for the packaging specificity in vivo, Wengler et al. (39, 40) proposed that corelike particles containing heterol-

ogous RNAs are much less stable than particles containing virion RNA and that capsids containing 26S RNA may be unstable. Another possibility is that the interaction between 26S RNA and capsid protein has some function other than that involved in assembly. The capsid protein is not required for replication of Sindbis virus RNA, since viral RNA transcripts in which the structural genes have been replaced by a foreign gene are replication competent (43). Interactions between 26S RNA and capsid protein, however, could have some regulatory role in an infection.

The possibility that internal sequences of the alphavirus genome are involved in encapsidation had been suggested several years ago by Jalanko and Söderlund (13). They cloned a DI cDNA into a simian virus 40 vector and expressed the DI RNA in monkey kidney cells. The DI RNA was not replicated by superinfecting helper Semliki Forest virus nor was it able to interfere with replication of the virus. Packaging of the virus, however, was inhibited by about 60%. The DI RNA lacks the extreme 5′ terminus of the genomic RNA but contains repeated elements from the nonstructural region of the virion genome. The researchers suggest that the repeated regions might be involved in encapsidation. The general repeat unit contains approximately 200 nucleotides from nsP2 (starting at nucleotide 2728 of the Semliki Forest virus sequence), followed by 278 nucleotides from nucleotides 39 to 317 of the Semliki Forest virus genome (16, 17). Thus, that Semliki Forest virus DI RNA does not have a region from nsP1 comparable to the one we have defined in Sindbis virus DI RNAs. Semliki Forest virus DI RNAs are packaged inefficiently (14), and it will be interesting to know whether they contain a capsid-binding domain. There are several domains in the nsP1-, nsP2-, and nsP4-coding genes of Sindbis virus RNA that show strong amino acid homology, not only with Semliki Forest virus RNA but also with plus-stranded plant RNA viruses (2, 11). These homologies are thought to reflect evolutionary relatedness and conserved function of the proteins. The alphaviruses have also conserved sequence domains at the 5′ and 3′ termini and at the start of the subgenomic RNA (32). There is increasing evidence that the conserved sequences represent *cis*-acting regulatory or binding functions (19; R. Levis, S. Schlesinger, and H. Huang submitted for publication; R. Kuhn, E. Strauss, and J. Strauss, personal communication). It should now be possible to determine whether there is any conservation in the capsid-binding domain.

The coat proteins of a number of different positive-stranded RNA viruses have been demonstrated to bind to specific regions of their viral RNAs. For tobacco mosaic virus, a coat protein disk aggregate binds to an internal region of tobacco mosaic virus RNA (4). Encapsidation of heterologous RNA containing this specific recognition domain occurs both in vitro (9, 28) and in vivo (25). Studies of disassembly and reassembly of turnip crinkle virus suggest that dimers of the coat protein interact with the viral RNA and it is most likely this step that leads to the specificity of assembly (29). Other examples of high-affinity binding of proteins to viral RNAs may define functions that are not necessarily associated with encapsidation. A classic example is the coat protein of the RNA bacteriophages whose interaction with the viral RNA serves to repress translation (38). Alfalfa mosaic virus provides another important model. The genomic RNAs of this virus contain high-affinity binding sites for the viral coat protein, and binding of the coat protein to each of the genomic RNAs is an essential step in the abilities of these RNAs to be infectious. Binding sites were identified both at the 3′ terminus and at internal sites of the largest genomic segment and the subgenomic RNA, but the subgenomic RNA of alfalfa mosaic virus can also be encapsidated. These sites are thought to be involved in replicase recognition (12, 44), but one or more may also be important in assembly. Multiple binding sites may also exist on the RNAs of coronaviruses. Northwestern blot analysis of the interaction of the coronavirus nucleocapsid protein and viral RNAs has demonstrated that the viral nucleocapsid protein binds to the 3′ end of the viral leader RNA and therefore to all of the viral mRNAs, as well as to genomic RNA (3, 30). Stohlman et al. proposed that this interaction plays an important role in the transcription process (30). They also suggested that other sequences affect encapsidation.

These examples, and that of Sindbis virus, define a region(s) of the viral genome that imparts specificity to the binding of the viral capsid protein. In each case, the region defines a sequence but also a particular structure. Experiments with both tobacco mosaic virus (35) and Qβ bacteriophage (42) that have analyzed the binding characteristics in more detail conclude that structure plays a crucial part in determining specificity. One goal of our future studies will be to determine the role of structure in Sindbis virus RNA-capsid interactions.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AI11377 from the National Institute of Allergy and Infectious Diseases.

We thank our colleagues at Washington University for comments on the manuscript. We thank P. L. Felgner and Gordon M. Ringold for providing lipofectin (7) and P. L. Felgner for advice and suggestions on the use of lipofectin.

LITERATURE CITED

1. **Adam, M. A., and D. Miller.** 1988. Identification of a signal in a murine retrovirus that is sufficient for packaging of nonretroviral RNA into virions. J. Virol. **62:**3802–3806.
2. **Ahlquist, P., E. G. Strauss, C. M. Rice, J. H. Strauss, J. Haseloff, and D. Zimmern.** 1985. Sindbis virus proteins nsP1 and nsP2 contain homology to nonstructural proteins from several RNA plant viruses. J. Virol. **35:**536–542.
3. **Baric, R. S., G. W. Nelson, J. O. Fleming, R. J. Deans, J. G. Keck, N. Casteel, and S. A. Stohlman.** 1988. Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. J. Virol. **62:**4280–4287.
4. **Bloomer, A. C., and P. J. G. Butler.** 1986. Tobacco mosaic virus. Structure and self-assembly, p. 19–57. *In* M. H. V. Van Regenmortel and H. Fraenkel-Conrat (ed.), The plant viruses, vol 2. Plenum Publishing Corp., New York.
5. **Bowen, B., J. Steinberg, U. K. Laemmli, and H. Weintraub.** 1980. The detection of DNA-binding proteins by protein blotting. Nucleic Acids Res. **8:**1–20.
6. **Carmichael, G. G., and G. K. McMaster.** 1980. The analysis of nucleic acids in gels using glyoxal and acridine orange. Methods Enzymol. **65:**380–389.
7. **Felgner, P. L., T. R. Gadek, M. Holm, R. Roman, H. W. Chan, M. Wenz, J. P. Northrop, G. M. Ringold, and M. Danielson.** 1987. Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. Proc. Natl. Acad. Sci. USA **84:**7413–7417.
8. **Fuller, S. D.** 1987. The T=4 envelope of Sindbis virus is organized by interactions with a complementary T=3 capsid. Cell **48:**923–934.
9. **Gallie, D. R., D. E. Sleat, J. W. Watts, P. C. Turner, and T. M. A. Wilson.** 1987. In vivo uncoating and efficient expression of foreign mRNAs packaged in TMV-like particles. Science **236:**1122–1124.
10. **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.

11. **Haseloff, J., P. Goelet, D. Zimmern, P. Ahlquist, R. Dasgupta, and P. Kaesberg.** 1984. Striking similarities in amino acid sequence among nonstructural proteins encoded by RNA viruses that have dissimilar genomic organization. Proc. Natl. Acad. Sci. USA **81**:4358–4362.
12. **Houwing, C. J., and E. M. J. Jaspars.** 1982. Protein binding sites in nucleation complexes of alfalfa mosaic virus RNA 4. Biochemistry **21**:3408–3414.
13. **Jalanko, A., and H. Söderlund.** 1985. The repeated regions of Semliki Forest virus defective-interfering RNA interferes with the encapsidation process of the standard virus. Virology **141**: 257–266.
14. **Kääriäinen, L., R. F. Pettersson, S. Keränen, P. Lehtovaara, H. Söderlund, and P. Ukkonen.** 1981. Multiple structurally related defective-interfering RNAs formed during undiluted passages of Semliki Forest virus. Virology **113**:686–697.
15. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
16. **Lehtovaara, P., H. Söderlund, S. Keränen, R. F. Pettersson, and L. Kääriäinen.** 1981. 18S defective interfering RNA of Semliki Forest virus contains a triplicated linear repeat. Proc. Natl. Acad. Sci. USA **78**:5353–5357.
17. **Lehtovaara, P., H. Söderlund, S. Keränen, R. F. Pettersson, and L. Kääriäinen.** 1982. Extreme ends of the genome are conserved and rearranged in the defective interfering RNAs of Semliki Forest virus. J. Mol. Biol. **156**:731–748.
18. **Levis, R., H. Huang, and S. Schlesinger.** 1987. Engineered defective interfering RNAs of Sindbis virus express bacterial chloramphenicol acetyltransferase in avian cells. Proc. Natl. Acad. Sci. USA **84**:4811–4815.
19. **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.
20. **Mann, R., R. C. Mulligan, and D. Baltimore.** 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell **33**:153–159.
21. **Monroe, S. S., and S. Schlesinger.** 1983. RNAs from two independently isolated defective-interfering particles of Sindbis virus contain a cellular tRNA sequence at their 5′ ends. Proc. Natl. Acad. Sci. USA **80**:3279–3283.
22. **Monroe, S. S., and S. Schlesinger.** 1984. Common and distinct regions of defective-interfering RNAs of Sindbis virus. J. Virol. **49**:865–872.
23. **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.
24. **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.
25. **Sacher, R., R. French, and P. Ahlquist.** 1988. Hybrid brome mosaic virus RNAs express and are packaged in tobacco mosaic virus coat protein in vivo. Virology **167**:15–24.
26. **Schlesinger, M. J., and S. Schlesinger.** 1986. Formation and assembly of alphavirus glycoproteins, p. 121–148. *In* S. Schlesinger and M. J. Schlesinger (ed.), The Togaviridae and Flaviviridae. Plenum Press Publishing Corp., New York.
27. **Schlesinger, S., R. Levis, B. G. Weiss, M. Tsiang, and H. Huang.** 1986. Replication and packaging sequences in defective interfering RNAs of Sindbis virus, p. 241–250. *In* M. A. Brinton and R. R. Rueckert (ed.), Positive strand RNA viruses. Alan R. Liss, Inc., New York.
28. **Sleat, D. E., P. C. Turner, J. T. Finch, P. J. G. Butler, and T. M. A. Wilson.** 1986. Packaging of recombinant RNA molecules into pseudovirus particles directed by the origin-of-assembly sequence from tobacco mosaic virus RNA. Virology **155**:299–308.
29. **Sorger, P. K., P. G. Stockley, and S. C. Harrison.** 1986. Structure and assembly of turnip crinkle virus. II. Mechanism of reassembly in vitro. J. Mol. Biol. **191**:639–658.
30. **Stohlman, S. A., R. S. Baric, G. N. Nelson, L. J. Soe, L. M. Welter, and R. J. Deans.** 1988. Specific interaction between coronavirus leader RNA and nucleocapsid protein. J. Virol. **62**:4288–4295.
31. **Strauss, E. G., C. M. Rice, and J. H. Strauss.** 1984. Complete nucleotide sequence of the genomic RNA of Sindbis virus. Virology **133**:92–110.
32. **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.
33. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
34. **Tsiang, M., B. G. Weiss, and S. Schlesinger.** 1988. Effects of 5′-terminal modifications on the biological activity of defective interfering RNAs of Sindbis virus. J. Virol. **62**:47–53.
35. **Turner, D. R., L. E. Joyce, and P. J. G. Butler.** 1988. The tobacco mosaic virus assembly origin RNA. Functional characteristics defined by directed mutagenesis. J. Mol. Biol. **203**: 531–547.
36. **Ulmanen, I., H. Söderlund, and L. Kääriäinen.** 1976. Semliki Forest virus capsid protein associates with the 60S ribosomal subunit in infected cells. J. Virol. **20**:203–210.
37. **Vaux, J. T., A. Helenius, and I. Mellman.** 1988. Spike-nucleocapsid interaction in Semliki Forest virus reconstructed using network antibodies. Nature (London) **336**:36–42.
38. **Weber, K., and W. Konigsberg.** 1975. Proteins of the RNA phages, p. 51–84. *In* N. D. Zinder (ed.), RNA phages. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
39. **Wengler, G.** 1987. The mode of assembly of alphavirus cores implies a mechanism for the disassembly of the cores in the early stages of infection. Arch. Virol. **94**:1–14.
40. **Wengler, G., U. Boege, G. Wengler, H. Bischoff, and K. Wahn.** 1982. The core protein of the alphavirus Sindbis virus assembles into core-like nucleoproteins with the viral genome RNA and with other single-stranded nucleic acids in vitro. Virology **118**:401–410.
41. **Wengler, G., G. Wengler, U. Boege, and K. Wahn.** 1984. Establishment and analysis of a system which allows assembly and disassembly of alphavirus core-like particles under physiological conditions in vitro. Virology **132**:401–412.
42. **Witherell, G. W., and O. C. Uhlenbeck.** 1989. Specific RNA binding by Qβ coat protein. Biochemistry **28**:71–76.
43. **Xiong, C., R. Levis, P. Shen, S. Schlesinger, C. M. Rice, and H. V. Huang.** 1989. Sindbis virus: an efficient, broad host range vector for gene expression in animal cells. Science **243**:1188–1191.
44. **Zuidema, D., M. F. A. Bierhuizen, B. J. C. Cornelissen, J. F. Bol, and E. M. J. Jaspars.** 1983. Coat protein binding sites on RNA 1 of alfalfa mosaic virus. Virology **125**:361–369.