JOURNAL OF VIROLOGY, Mar. 1993, p. 1620–1626
0022-538X/93/031620-07$02.00/0

Vol. 67, No. 3

Deletion Analysis of the Capsid Protein of Sindbis Virus: Identification of the RNA Binding Region

UTE GEIGENMÜLLER-GNIRKE, HANS NITSCHKO,† AND SONDRA SCHLESINGER*

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

Received 9 October 1992/Accepted 6 December 1992

The capsid protein of Sindbis virus has multiple functions in the life cycle of the virus. One essential function is to interact with the genomic RNA of the virus to form the nucleocapsid. The experiments described in this article define a region of the protein that is required for binding to Sindbis virus RNA. The assay we used measured the binding of in vitro-translated proteins to RNA on the basis of their migration with the RNA during electrophoresis in an agarose gel. Binding to RNA showed specificity; more protein bound to an RNA containing the previously defined packaging signal in Sindbis virus RNAs than to a similar RNA lacking this sequence. We were able to produce a variety of deleted forms of the capsid protein by constructing cDNAs with in-frame deletions throughout the coding region of the capsid protein gene. These cDNAs were then transcribed into mRNAs and translated in vitro. C-terminal deletions in the capsid protein were obtained by preparing transcripts from cDNAs linearized at sites within the coding region. Our studies identified a 32-amino-acid region that is essential for the specificity in RNA binding, and they defined a 68-amino-acid minimal sequence which displays almost the complete specific RNA binding activity of the intact Sindbis virus capsid protein containing 264 amino acids.

Sindbis virus is a positive-strand RNA-enveloped virus and the prototypical member of the *Alphavirus* genus of the *Togaviridae* family (reviewed in references 21 and 23). The genomic 49S RNA serves as the mRNA for translation of a polyprotein precursor for the nonstructural proteins encoded in the 5′ two-thirds of the viral genome. These proteins are required for the replication and transcription of viral RNA. The viral structural proteins are also translated as a polyprotein, but from a subgenomic 26S mRNA identical in sequence to the 3′-terminal one-third of the 49S RNA. The viral capsid protein is translated first and cleaves itself autoproteolytically from the nascent polyprotein, leaving the precursor for the envelope proteins (1). These envelope proteins are synthesized and processed in the rough endoplasmic reticulum and are transported through the Golgi network to the plasma membrane of the infected cell.

The viral capsid protein plays several important roles in the life cycle of the virus. In addition to its proteolytic activity, it interacts with the genomic RNA to form the icosahedral nucleocapsid. This interaction is specific; only genomic RNA, not the 26S subgenomic RNA, is encapsidated. In an analysis of the binding of Sindbis virus-derived RNAs to the capsid protein, Weiss et al. showed that there is a region located between nucleotides 746 and 1226 in the genomic RNA that is required for the binding of RNA to the capsid protein (29). They also showed that this region is important for the encapsidation of viral RNAs in infected cells. Defective interfering RNAs of Sindbis virus containing these sequences (the packaging signal) are packaged into extracellular virions to a much greater extent than similar RNAs which lack these sequences. The capsid polypeptide also undergoes protein-protein interactions in forming the nucleocapsid. In the context of the nucleocapsid, the capsid protein must interact with the cytoplasmic tail of the viral glycoprotein E2 in the final step in assembly: the budding and release of mature virions from the plasma membrane of the infected cell.

These activities of the capsid protein are essential for virus assembly, but the alphavirus capsid may also play a role in the initial steps of infection. The capsid binds to ribosomes, both in vitro and in vivo, and it has been proposed previously that this interaction is involved in the disassembly of the nucleocapsid (24, 25, 31, 32). The capsid protein can interfere with the binding of host mRNA to translation initiation complexes, suggesting that it is a factor in the shutoff of host cell protein synthesis (26). The observations that the capsid protein of Semliki Forest virus can have diverse effects on host cell protein synthesis and that it is transported to the nuclei of cells suggest that the capsid protein may have some additional function in the disruption of host cell functions (4, 14).

The alphavirus capsid polypeptide can be divided into two domains on the basis of its amino acid sequence (6, 17). The N-terminal domain of about 113 residues is very basic, and it contains many proline residues. The C-terminal domain has a more conventional amino acid composition which is conserved among alphaviruses. Previous analysis of temperature-sensitive mutants of Sindbis virus and site-directed mutagenesis of the capsid protein gene identified His-141, Asp-147, and Asp-163 as the amino acids essential for autoprotease activity (7, 8). The sequence information and mutagenesis laid the foundation for the proposal that the capsid protein is a chymotrypsin-like serine protease, a proposal which has now been confirmed by X-ray diffraction data (3). The capsid exists in the crystal structure as a dimer with the monomer-monomer contacts made up of residues 185 to 190 in a β-strand and residue 222. The structure of the C-terminal domain (amino acids 114 to 264) was determined, and it is similar to the structure of mammalian serine proteases of the chymotrypsin family. In contrast, the N-terminal 113 residues remain largely unstructured in the crys-

* Corresponding author.
† Present address: Max von Pettenkofer Institute, D-8000 Munich, Germany.

tals. The basic nature of the N-terminal arm and its apparent flexibility are reminiscent of the coat proteins of several plant RNA viruses (9, 10, 18) and suggest that it is this region that interacts with the genomic RNA in the nucleocapsid.

In previous studies, we identified sequences in the Sindbis virus genomic RNA required for the RNA to bind to the capsid protein (29). In the studies described here, we examined the amino acids involved in this binding. We were able to produce a variety of deleted forms of the capsid protein by constructing cDNAs with in-frame deletions throughout the coding region of the capsid protein gene. These cDNAs were then transcribed into mRNAs and translated in vitro. We also obtained C-terminal deletions in the capsid protein by preparing transcripts from cDNAs linearized at sites within the coding region. The binding of the in vitro-translated proteins to RNA was detected by their migration with the Sindbis virus RNA during electrophoresis in an agarose gel. Our studies identified a 32-amino-acid region that is essential for the specificity in RNA binding, and they defined a 68-amino-acid minimal sequence which displayed almost the complete specific RNA binding activity of the Sindbis virus capsid protein.

MATERIALS AND METHODS

Plasmid constructions. (i) π(cap.stop). The first 974 bp of the Sindbis virus 26S cDNA, containing the nontranslated 5′ end, the capsid protein gene, and part of the E3 gene, were put under the control of the SP6 DNA-dependent RNA polymerase promoter and inserted into the πAN7 vector (13). The πAN7 vector contains relatively few restriction sites, so many of the sites within the capsid protein coding sequence were unique. In the original construction, sufficient sequences downstream of the C terminus of the capsid protein were included to retain its autoproteolytic activity. Since deletions in the capsid protein would probably interfere with the protease activity, we inserted two stop codons downstream of the capsid protein cDNA by means of polymerase chain reaction (PCR) (AmpliTaq; Perkin-Elmer Cetus, Norwalk, Conn.). The upstream primer was within the capsid protein coding region, and the downstream primer contained the sequences complementary to the 3′ 20 nucleotides of the capsid protein gene, two opal stop codons, and a polylinker site. The PCR product was cut with *Nru*I within the capsid protein coding region and with *Hin*dIII at the 3′ end of the polylinker and ligated into the *Nru*I–*Hin*dIII-cut π(capsid) vector. The PCR-derived sequence in the resulting π(cap.stop) DNA was confirmed by dideoxy nucleotide sequencing (20) with the Sequenase version 2.0 kit from U.S. Biochemical Corp., Cleveland, Ohio, and the upstream PCR primer.

(ii) Capsid deletion mutants. In-frame deletions within the capsid protein were obtained by cutting π(cap.stop) DNA at the restriction sites indicated in Fig. 2A, blunting the termini with T4 DNA polymerase, and religating the DNA. For the construction of Δ2, the reading frame was restored by insertion of a 10-bp *Sal*I linker into the deletion site. Δ13 and Δ18 contain out-of-frame deletions from the blunt-ended *Dra*III to the *Nru*I sites. In these mutants, a stop codon was moved into the reading frame 1 amino acid downstream of the deletion. All deletion mutants were sequenced across the deletion junctions.

(iii) cDNAs with Sindbis virus capsid protein-BMV coat protein fusions. cDNAs containing the sequences of the Sindbis virus capsid protein-brome mosaic virus (BMV) coat protein fusion proteins (Fig. 2D) were constructed from π(cap.stop) DNA that was cut with *Nar*I or *Nco*I, blunt ended, and then cut with *Hin*dIII. An intermediate containing a 10-bp *Sal*I linker in the filled-in *Afl*II site of π(cap.stop) DNA was cut with *Sal*I, filled in, and cut with *Hin*dIII. A 589-bp *Ava*II-*Hin*dIII fragment of pB3RS25 (18) containing the coding sequence of amino acids 27 to 190 of the coat protein of BMV was then inserted into the π(cap.stop) DNA vector fragments described above. All three fusion protein cDNA clones were sequenced across the fusion junction. pB3RS25 was provided by Paul Ahlquist, University of Wisconsin, Madison.

(iv) CTS-1 and CTS-14. The two cDNAs, CTS-1 and CTS-14, have been described previously; CTS-14, but not CTS-1, contains the packaging signal for Sindbis virus RNA (29). The original cDNAs were downstream of an SP6 DNA-dependent RNA polymerase promoter which was replaced by the T7 polymerase promoter.

(v) KS-G1. The packaging signal for the Sindbis virus genome is contained within a 572-bp region of Sindbis virus cDNA (29). KS-G1 was constructed by inserting a *Hin*dIII-cut 582-bp fragment generated by PCR corresponding to positions 721 to 1302 of Toto1102 (16) into a *Sma*I–*Hin*dIII-cut Bluescript vector (Stratagene, La Jolla, Calif.).

In vitro transcription. T7 RNA polymerase-dependent in vitro runoff transcription of *Ssp*I-cut T7CTS-1, *Dra*I-cut T7CTS-14, or *Hin*dIII-cut KS-G1 was performed under standard conditions (15). These RNAs were not capped. T7 RNA polymerase was a gift from Birgit Lewicki, Max-Planck-Institut fur Molekulare Genetik, Berlin, Germany.

The original π(cap.stop) DNA was linearized with *Hin*dIII downstream of the capsid protein coding sequence. Capsid protein-BMV coat protein fusion constructs were cut with *Eco*RI in the vector sequence. Truncated capsid protein mRNAs for the synthesis of C-terminal deletion mutants (Fig. 2B and C) were generated by cutting the cDNA within the coding region of the capsid protein at an *Afl*II, *Aat*II, *Nco*I, or *Hin*cII site. Runoff transcriptions from a DNA cut with *Aat*II or *Dra*III, which leave a 3′ overhang, resulted in multiple RNA species, probably because of nonspecific initiation at the 3′ overhang. Removal of the 3′ overhang by treatment with T4 DNA polymerase prior to transcription eliminated these multiple RNA bands. For *Dra*III runoff transcriptions, the problem was circumvented by using the out-of-frame deletion mutants Δ13 and Δ18 cut at the *Hin*dIII site. The linearized fragments of those DNAs cut within the coding region of the capsid protein were gel purified. The transcription reactions with SP6 DNA-dependent RNA polymerase were performed by following the protocol described by the manufacturer. All transcripts were analyzed for size and intactness by gel electrophoresis under denaturing conditions. RNAs larger than 300 nucleotides were denatured by treatment with glyoxal (2) and run on a 1% agarose gel in 10 mM sodium phosphate buffer. Transcripts smaller than 300 nucleotides were denatured in 40% formamide–8 mM EDTA and separated on a 6% polyacrylamide–9 M urea gel in 50 mM Tris-borate-EDTA (TBE).

In vitro translation. In vitro-transcribed (and -capped) mRNAs coding for intact capsid protein and the various deletion and fusion mutants were translated in vitro in a wheat germ translation system (Promega, Madison, Wis.) according to the manufacturer's protocol, except that 20 μCi of [^{35}S]methionine and 1 μg of in vitro-transcribed RNA were used per 25 μl of reaction mixture. The incorporation of labeled amino acid, determined by acid precipitation after alkaline hydrolysis of the charged tRNA, was about 20% for the full-length (wild-type [wt]) capsid protein. The values for

percent incorporation differed reproducibly for the various capsid protein mutants. Translation products were analyzed on sodium dodecyl sulfate (SDS)-polyacrylamide gels (10, 12, or 15% acrylamide) (12) and generally showed one defined band (see Fig. 3). When the in vitro translation was performed with a rabbit reticulocyte system, the capsid protein mRNAs gave rise to several bands on the acrylamide gel which may be due to ubiquitination of the proteins.

Assay to detect binding of the capsid protein to RNA. The capsid protein or mutated forms of the protein were incubated with CTS-1 or CTS-14 RNA in a 10-μl solution containing 10 mM Tris-HCl (pH 7.0), 100 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, 5.3 μg of HB101 RNA, 10 U of RNasin, 0.4 μg of CTS-1 or CTS-14 RNA, and 0.5 to 2 μl of the total in vitro translation reaction mixture. The exact amount of translation mix added was determined on the basis of the extent of incorporation of [^{35}S]methionine into protein during the in vitro translation reaction. The percent acid-insoluble counts per minute was corrected for the number of Met molecules per capsid protein, and equal molar amounts of capsid protein molecules were used. After the addition of the translation reaction mixture containing the capsid protein, the binding reaction mixture was not vortexed. Samples were incubated for 20 min at room temperature, 1 μl of dye (20% Ficoll–0.05% DEPC (diethylpyrocarbonate)–0.05% bromophenol blue in 25 mM TBE) was added, and the entire sample (11 μl) was loaded onto a 1% agarose gel in 25 mM TBE which had been prerun in this buffer for at least 20 min at 40 mA. The gel was then run for 3 to 4 h at 40 mA, soaked overnight in methanol, and dried. The binding to RNA was analyzed with a Betascope 603 blot analyzer (Betagen, Waltham, Mass.). Comparisons between the wt and deleted capsid proteins were usually made at more than one concentration of protein. Varying the protein concentration had little or no effect on the relative binding values of the different capsid proteins.

The extent of binding to RNA of the various altered forms of the capsid protein in different experiments was normalized to that of the binding obtained for the wt capsid protein analyzed in the same experiment, after correction for the number of Met residues in the different samples. The reported binding values were calculated as the means of at least 3 and up to 12 individual binding assays. The range of binding values is reflected in the standard error. For most of the altered capsid proteins (all except Δ11, Δ15, Δ16, and Δ18), products from at least two different translation reactions were used for binding experiments.

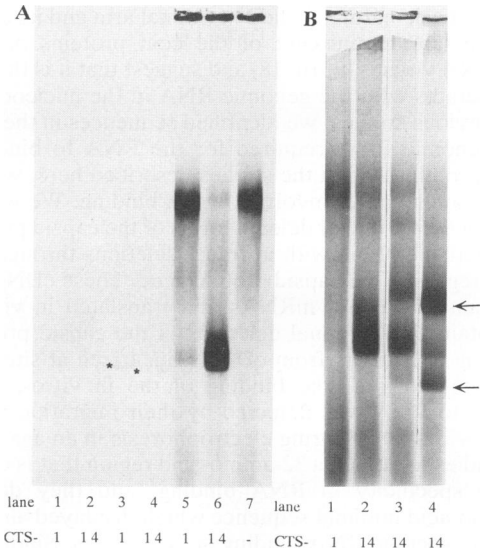

FIG. 1. Binding of capsid protein to RNA. (A) Comparison of the binding of the capsid protein to CTS-14 and CTS-1 RNAs. Binding conditions were as described in Materials and Methods except that CTS-1 and CTS-14 RNAs were labeled with [^{3}H]UTP at a specific activity 10-fold higher than that in the other experiments. One microliter of mock translation reaction mixture was first added to CTS-1 (lane 1) or CTS-14 (lane 2) RNA. CTS-1 (lane 3) and CTS-14 (lane 4) RNA in the absence of the in vitro translation extract are also shown. The bands in lanes 3 and 4 are indicated by asterisks. Although it may be difficult to see the bands in this figure, they were clearly visible on the autoradiogram and were equal in intensity. In all of the other lanes, the bands represent ^{35}S-labeled protein. The level of ^{3}H-labeled RNA was too low to be detected. In other lanes, 1 μl of translation reaction mixture with full-length capsid protein was added to CTS-1 RNA (lane 5), CTS-14 RNA (lane 6), or a binding assay mixture containing only HB101 RNA (lane 7). (B) Competition between the packaging signal from Sindbis virus RNA and CTS-14 RNA. Binding conditions were as described in Materials and Methods except that the amounts of CTS-1 and CTS-14 RNA were reduced to 0.2 μg per reaction mixture. Lanes: 1 and 2, CTS-1 RNA and CTS-14 RNA, respectively, with no competitor RNA; 3, an equimolar amount of competitor added to CTS-14 RNA; 4, a 10-fold molar excess of competitor added to CTS-14 RNA. The competitor RNA was KS-G1–*Hin*dIII RNA. It contains only the sequences from nucleotides 721 to 1302 in Sindbis virus 49S RNA and 55 nucleotides from the Bluescript vector. This RNA migrates as two bands (indicated by arrows) in agarose gels unless it is first denatured.

RESULTS

Binding of the Sindbis virus capsid protein to Sindbis virus RNA. We made a series of deletions in the capsid protein to test their effect on the binding of the protein to viral RNA. The deletions were made in cDNAs containing the gene for the capsid protein. RNAs transcribed in vitro from these DNAs were translated in vitro. This approach made it possible to prepare many different deleted forms of the capsid protein and to generate C-terminal deletions of the protein by runoff transcription of the RNA followed by translation. The specific binding of the capsid protein was detected by incubating the [^{35}S]methionine-labeled protein with RNA and then analyzing the amount of ^{35}S label that migrated with the RNA in an agarose gel. The binding of protein did not lead to a shift in the migration of the RNA in the gel, but this assay does not require a gel shift of the RNA. Two different RNAs were used in the assay. Both were derived from a defective interfering RNA of Sindbis virus. CTS-14 contains the 572 nucleotides defined as the packaging signal required for encapsidation of the RNA, and CTS-1 lacks these sequences (29). CTS-14, but not CTS-1, binds to the capsid protein and is packaged into extracellular particles in infected cells. For the experiments described here, the cDNAs were linearized upstream of the 3′ terminus so that the size of the CTS-14 RNA was 849 nucleotides and that of the CTS-1 RNA was 831 nucleotides. The data in Fig. 1A show the difference in binding of the capsid protein to the two RNAs (lanes 5 and 6). The in vitro-translated capsid protein migrated as a high-molecular-weight aggregate (Fig. 1A, lane 7), which may be because of nonspecific binding to RNAs present in the in vitro translation mix. This high-molecular-weight band decreased when the intact capsid bound to CTS-14 RNA (lane 6).

To demonstrate that the difference in binding between

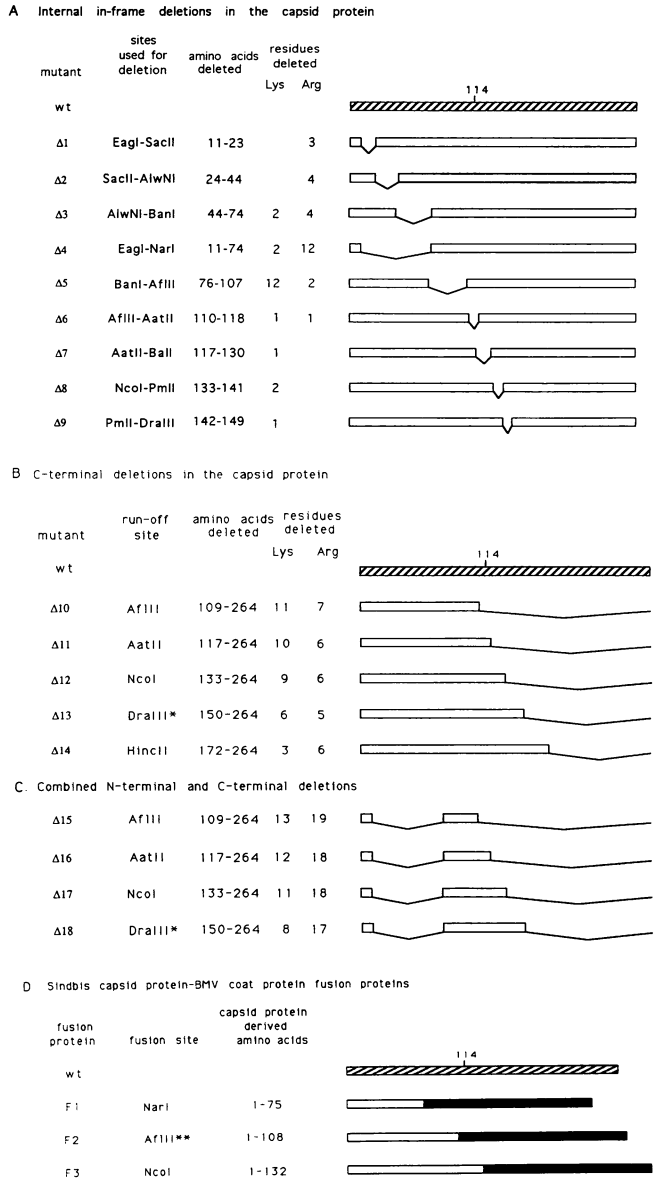

FIG. 2. Diagrams of capsid protein showing locations of the deletions. Open boxes signify the amino acid regions retained in the capsid protein, solid lines show the deleted regions, and the wt is the full-length capsid protein. Amino acid 114 of the wt capsid protein is highlighted in panels A, B, and D to indicate the junction between the unstructured N-terminal and the structured C-terminal parts of the protein. (A) Internal in-frame deletions in the capsid protein. (B) Deletions in the C terminus of the capsid protein. The *Dra*III runoff mutant, indicated by asterisks, was made from a *Hin*dIII-cut *Dra*III-*Nru*I internal out-of-frame deletion mutant which moved a stop codon into the reading frame 1 amino acid (arginine) downstream of the deletion. (C) Combined N- and C-terminal deletions. Mutants Δ15 to Δ18 contain the same internal deletion as Δ4 (*Eag*I-*Nar*I) in addition to various C-terminal deletions. (D) Sindbis virus capsid protein-BMV coat fusion proteins. Solid boxes represent BMV coat protein-derived sequences. The *Afl*II fusion protein, indicated by double asterisks, was constructed by insertion of a 10-bp *Sal*I linker into the filled-in *Afl*II site followed by fusion of the *Sal*I-cut capsid protein gene to the deleted BMV coat protein gene. Three amino acids (Asp, Gly, and Arg) were added between the Sindbis virus and the BMV-derived amino acids.

TABLE 1. Quantitative results from Betascope analysis[a]

Capsid mutant	Total binding	Nonspecific binding	Specific binding		% of wt[b]
			Initial	Corrected[b]	
wt	1.47	0.34	1.13	1.13	100
Δ1	1.29	0.40	0.89	0.99	88
Δ2	2.21	0.71	1.50	1.66	147
Δ3	2.18	0.80	1.38	1.38	122
Δ4	1.15	0.47	0.68	0.86	76
Δ5	0.12	0.18	−0.6	−0.7	0

[a] Data analyzed were obtained from the experiment shown in Fig. 4. Data in the table, except as indicated, are the ^{35}S counts (10^4) collected in a 30-min scan with the Betascope blot analyzer. Total binding is the binding to CTS-14 RNA; nonspecific binding is the binding to CTS-1 RNA; specific binding is the difference between the total (CTS-14-bound) and nonspecific (CTS-1-bound) counts.

[b] The specific binding was corrected for the number of Met residues per capsid mutant polypeptide (10 Met residues per full-length capsid protein). Corrected specific binding values obtained for capsid mutants were then expressed as the percentage of corrected specific binding obtained with the full-length (wt) capsid protein.

CTS-14 and CTS-1 was due to the 572-nucleotide sequence, we carried out the binding assay in the presence of a competitor RNA that contained only the 572-nucleotide sequence derived from Sindbis virus RNA (and an additional 55 nucleotides derived from the vector). The competitor RNA was added at concentrations equal to or in 10-fold excess of those of CTS-14 RNA. Specific binding to CTS-14 was reduced 2- and 10-fold, respectively (Fig. 1B, lanes 3

TABLE 2. Specific RNA binding activities of capsid protein mutants

Characteristics and mutant	Amino acids deleted	No. of expts	Binding of mutant as % of wt binding ($\bar{x} \pm \alpha\bar{x}$)[a]
Internal deletions			
Δ1	11–23	4	80 ± 20
Δ2	24–44	6	154 ± 53
Δ3	44–74	6	170 ± 59
Δ4	11–74	12	101 ± 42
Δ5	76–107	6	1.5 ± 2
Δ6	110–118	8	14 ± 8
Δ7	117–130	8	17 ± 10
Δ8	133–141	7	11 ± 5
Δ9	142–149	7	13 ± 5
C-terminal deletions			
Δ10	109–264	6	11 ± 8
Δ11	117–264	5	37 ± 21
Δ12	133–264	6	50 ± 8
Δ13	150–264	10	48 ± 17
Δ14	172–264	6	12 ± 9
Combined N- and C-terminal deletions			
Δ15	11–74, 109–264	4	4 ± 3
Δ16	11–74, 117–264	7	38 ± 21
Δ17	11–74, 133–264	10	70 ± 25
Δ18	11–74, 150–264	3	38 ± 14
Fusion proteins			
F1	76–264	6	3 ± 3
F2	109–264	9	17 ± 10
F3	133–264	9	31 ± 16

[a] $\bar{x}$, mean of n experiments (given in previous column); $\alpha\bar{x}$, standard deviation.

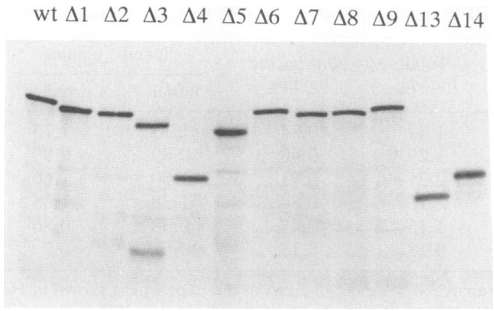

FIG. 3. Analysis of in vitro translation products. mRNAs that code for the capsid protein and various deleted forms of this protein were translated in vitro by using wheat germ extracts. Samples of the proteins were analyzed on an SDS–10% polyacrylamide gel. Approximately 5×10^5 acid-precipitable cpm was loaded in each lane.

and 4). Without denaturation, the 572-nucleotide RNA transcript migrates as two bands in an agarose gel (arrows in Fig. 1B, lane 4). This complicated the analysis of the agarose gel and made it difficult to use this smaller RNA in binding assays.

We had originally planned to use the [^{3}H]uridine-labeled RNA as an internal standard, but incubation of the RNA with a mock translation reaction mixture containing everything but ^{35}S-labeled capsid protein led to degradation of the RNA (Fig. 1A; compare lanes 1 and 2 with lanes 3 and 4). The intact capsid protein seemed to protect both RNAs from degradation. Because the protection could be variable between experiments, we did not normalize the data to the amount of [^{3}H]uridine-labeled RNA recovered. The lack of internal standards may have contributed to the considerable variation between experiments, as seen in the standard deviations (see Table 2). The experiments were all repeated multiple times, with the intact capsid protein included as a control in every experiment.

The effect of deletions in the N-terminal half of the capsid protein on binding to RNA. The internal in-frame deletions that we constructed in the amino-terminal half of the capsid protein are depicted in Fig. 2A. The [^{35}S]methionine-labeled proteins produced by translation in vitro were analyzed on acrylamide gels (Fig. 3), and the results of a gel-binding assay using these proteins are shown in Fig. 4. Only the ^{35}S-labeled protein bands were detected in this autoradiogram. Binding was quantified by analysis of the dried gel on a Betascope blot analyzer (Tables 1 and 2).

Deletion of amino acids 11 to 23 (Δ1) reduced only marginally the specific binding of the capsid protein to RNA, and deletion of amino acids 24 to 44 (Δ2) or 44 to 74 (Δ3) may have enhanced binding (Table 2). The latter two peptides showed an increase both in the nonspecific binding to CTS-1 and the specific binding to CTS-14 (Table 1). A deletion covering the entire region from amino acids 11 to 74 (Δ4) had no significant effect on the binding activity, even though this deletion eliminated 14 positive charges (12 arginines and 2 lysines) (Tables 1 and 2). In contrast, deletion of amino acids 76 to 107, which removed an identical number of positive charges (12 lysines and 2 arginines), decreased nonspecific binding to CTS-1 (Table 1) and almost completely abolished specific binding (residual binding of 1.5%) (Table 2).

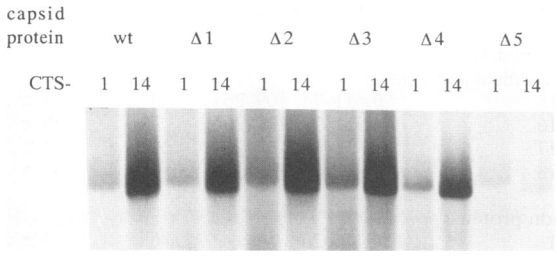

FIG. 4. Agarose gel analysis of the binding of capsid proteins with N-terminal in-frame deletions to Sindbis virus RNAs. Binding conditions were as described in Materials and Methods. Numbers obtained from the Betascope blot analysis are given in Table 1.

The effect of deletions in the C-terminal part of the capsid protein on binding to RNA. Two types of deletions were introduced into this region: internal in-frame deletions (Fig. 2A) and C-terminal truncations (Fig. 2B). The mutant (Δ10) lacking the C-terminal 156 amino acids retained only 11% of the specific binding activity (Table 2). A truncation of only 148 of these amino acids (Δ11) restored the activity to 37%. Two mutants, one (Δ12) lacking half of the C-terminal amino acids and the other (Δ13) lacking 115 of the C-terminal amino acids, retained about 50% of the specific binding activity of the full-length capsid protein (Table 2). Surprisingly, removal of the C-terminal 92 amino acids (Δ14) reduced the binding activity to only 12% that of the complete protein. Similarly, small internal deletions (Δ8 and Δ9) within this region also led to reductions in binding activity more severe than those that remove the entire region.

The effect of deletions in both the N-terminal and the C-terminal domains of the capsid protein on binding to RNA. The capsid protein with a large deletion in the amino-terminal half (Δ4) retained all of the binding activity of the intact capsid protein. C-terminal deletions of this protein were made by in vitro translation of truncated RNA transcripts (Fig. 2C). One of these mutants, Δ17, which contains amino acids 1 to 10 and 75 to 132 of the capsid protein, was almost fully active (70% residual binding activity) (Table 2). Δ16, which retained amino acids 1 to 10 and 75 to 116, still showed 38% of the original binding activity, but only 4% of the activity was retained by Δ15 (amino acids 1 to 10 and 75 to 108).

These results taken together with those obtained from the internal and the C-terminal deletions show that the amino acids between residues 75 and 117 were the most important for binding to the packaging signal in Sindbis virus RNA. They also reveal that in some instances, small internal deletions (Δ8 and Δ9) or partial deletions (Δ14) were more detrimental than more extensive deletions (Δ11 and Δ12 or Δ16 and Δ17).

Binding of Sindbis virus capsid protein-BMV coat protein fusion proteins to Sindbis virus RNA. Another approach to examining the abilities of fragments of the capsid protein to bind to Sindbis virus RNA which we took was to examine fusion proteins in which different amounts of the N-terminal domain of the capsid protein had been fused to the BMV coat protein. We chose the coat protein of BMV because it is similar in size to the Sindbis capsid protein and it also encapsidates a genomic RNA into an icosahedral capsid. The region important for specific RNA binding in the BMV coat protein has been assigned to the N terminus (19). Several different fragments of the Sindbis virus capsid protein cDNA were fused to nucleotide 79 of the BMV coat protein gene, eliminating the proposed RNA binding region of the latter. The Sindbis virus capsid-derived fragments in the fusion proteins contained the N-terminal 75 (F1), 108

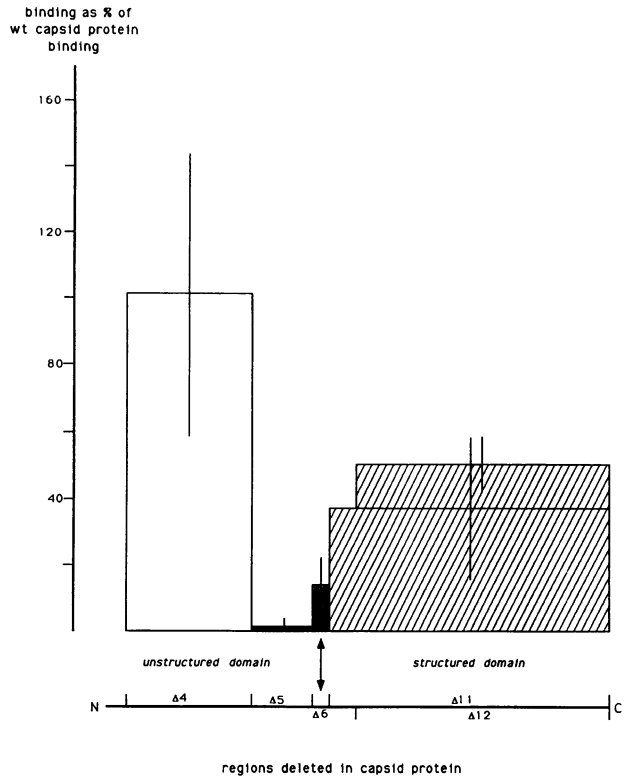

FIG. 5. Specific binding of deleted forms of the capsid protein to Sindbis virus RNAs. Each rectangle symbolizes one of the deleted forms of the capsid protein. Width reflects the amino acid region deleted, and height represents the specific RNA binding normalized to that of the full-length capsid protein. Binding values are the means of at least five experiments (Table 2). Bars at the top of each rectangle reflect standard deviations. Mutants are grouped into four classes according to their binding activities compared with that of the full-length capsid protein: (i) no reduction, (ii) 100-fold reduction, (iii) 5 to 10× reduction, and (iv) 2 to 3× reduction.

(F2), or 132 (F3) amino acids (Fig. 2D). F1 showed very little RNA binding activity (3% that of the full-length Sindbis virus capsid protein) (Table 2), but the activities of F2 (17%) and F3 (31%) were significant. These results again point to the importance of amino acids 76 to 108 in particular but also to that of amino acids 109 to 132 for the specific RNA binding activity of the capsid protein. In addition, the low-level binding activity of F1 argues against a major role of amino acids 1 to 10 in RNA binding (this region was not covered in the deletion analysis).

DISCUSSION

We have described an assay for analyzing the binding of an in vitro-translated protein to RNA on the basis of its migration with the RNA during electrophoresis in an agarose gel. This method of detection is a variation on the theme of the gel shift, in which it is the protein that shifts when it binds to the nucleic acid. Using this assay, we analyzed a series of deletions in the Sindbis virus capsid protein and identified a region in the protein that binds to the previously defined packaging signal in Sindbis virus RNA. Figure 5 illustrates four different classes of deletion mutants that do not overlap, in spite of the large standard errors. The first class, consisting of mutants located between amino acids 11 and 74, had little or no effect on the ability of the protein to bind RNA. In contrast, the second class, composed of the deletion from amino acids 76 to 107, defined a region of the protein that is essential for binding. The third class, exemplified in Fig. 5 by the deletion between amino acids 110 and 118, had only 10% the binding activity of the wt capsid, but the activity was 10-fold higher than that obtained with the class ii mutant. The fourth class lacked almost all of the C-terminal half of the protein, and it had intermediate activity.

The determination of the structure of the Sindbis virus capsid protein by Choi et al. (3) provides further insight into our results. The residues from Arg-114 to the C-terminal tryptophan are well ordered and are folded into two similar β-barrel domains with the connecting loops from approximately residue 170 to residue 187. One C-terminal deletion (Δ14) from amino acid 172 to the C terminus removed the second domain, causing a 90% reduction in the binding activity of the protein. A significant fraction of the activity was restored, however, with more extensive deletions in the C terminus (Δ11, Δ12, and Δ13). Deletion of only the C-terminal 93 residues may lead to a distortion in the structure of the protein, which is corrected by the removal of additional amino acid residues.

Deletions between amino acids 117 and 149 (Δ7, Δ8, and Δ9) were similar to the deletion from amino acid 172 to the C terminus (Δ14) in their abilities to decrease the binding activity about 90%. Deletion of the entire C terminus, including these residues, reduced the activity only about 50%. This contradiction suggests that deletions in this region of the protein also affect structure and can be compensated for by more extensive deletions.

Our data support the conclusion that the actual binding region lies between amino acids 76 and 116. The presence of amino acids 75 to 116 in the combined N-terminal- and C-terminal-deleted protein (Δ16) stimulated binding to 40% that of the complete capsid. Deletion of amino acids 76 to 107 almost completely eliminated binding, and the presence of this stretch of amino acids in the protein fused to the BMV coat protein increased the binding activity of that protein from 3 to 17%. The capsid protein exists as a dimer in the crystal structure. The region containing the amino acids involved in the monomer-monomer contacts (residues 185 to 190 and residue 222) can be deleted from the capsid protein without affecting binding to RNA, indicating that the oligomeric state of the protein is not involved in its ability to bind RNA.

Several different types of studies have shown previously that the capsid protein of one alphavirus can encapsidate the RNA of a different alphavirus (11, 30). This suggests that the amino acids involved in binding to genomic RNA would be similar in the different capsid proteins. Thus, a potential binding motif is the stretch of 10 amino acids in this part of the protein that is highly conserved among alphaviruses. The amino acid sequence in the Sindbis virus capsid protein from residues 97 to 106 is Lys Pro Lys Pro Gly Lys Arg Gln Arg Met. The underlined residues are invariant among eight different alphaviruses.

The determination of the structure of several RNA plant viruses provides a description of the domains of their coat proteins (9, 10, 18). The N-terminal residues of many of these proteins contain a large number of positively charged amino acids that are presumed to be important in neutralizing the negatively charged RNA. Removal of the N terminus from the coat protein of several of these viruses by proteolytic cleavage destroyed the ability of the protein to interact with RNA and led to the assembly of RNA-free particles (5,

22, 27, 28). More recent evidence for the importance of the N-terminal domain in assembly comes from the studies of Sacher and Ahlquist with BMV (19). They constructed a coat protein mutant lacking the first 25 N-terminal amino acids and showed that the protein was unable to interact with RNA to form virus either in plant protoplasts or in a systemic infection of barley plants. The interaction of the Sindbis virus capsid protein with the packaging signal in Sindbis virus RNA cannot be solely due to the high density of basic residues in the N-terminal domain of the protein. Our results show that the highly basic region from residues 11 to 74 plays no role in the specific binding of the protein to Sindbis virus RNA. We suggest that the interaction between the capsid protein and Sindbis virus RNA described here represents the nucleation event that confers specificity on the assembly process. Other regions of the protein may then function in the subsequent steps in the process that lead to the formation of the nucleocapsids.

ACKNOWLEDGMENTS

We thank Birgit Lewicki for providing the T7 polymerase and Paul Ahlquist for sending us the pB3RS25 DNA used for preparing the fusion clones. We also thank Barbara Weiss for comments and discussions throughout the course of this work.

This research was supported by grant AI11377 from the National Institute of Allergy and Infectious Diseases. U.G.-G. received fellowships from the scientific committee of NATO and the Keck Foundation.

REFERENCES

1. **Aliperti, G., and M. J. Schlesinger.** 1978. Evidence for an autoprotease activity of Sindbis virus capsid protein. Virology **90:**366–369.
2. **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.
3. **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 protease and the organization of the virion. Nature (London) **354:**37–43.
4. **Elgizoli, M., Y. Dai, C. Kempf, H. Koblet, and M. R. Michel.** 1989. Semliki Forest virus capsid protein acts as a pleiotropic regulator of host cellular protein synthesis. J. Virol. **63:**2921–2928.
5. **Erickson, J. W., and M. G. Rossmann.** 1982. Assembly and crystallization of a T=1 icosahedral particle from trypsinized southern bean mosaic virus coat protein. Virology **116:**128–136.
6. **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.
7. **Hahn, C. S., E. G. Strauss, and J. H. Strauss.** 1985. Sequence analysis of three Sindbis virus mutants temperature-sensitive in the capsid protein autoprotease. Proc. Natl. Acad. Sci. USA **82:**4648–4652.
8. **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.
9. **Harrison, S. C.** 1983. Virus structure: high resolution perspectives. Adv. Virus Res. **28:**175–240.
10. **Harrison, S. C.** 1990. Common features in the structures of some icosahedral viruses: a partly historical overview. Semin. Virol **1:**387–403.
11. **Kuhn, R. J., S. Lopez, and J. H. Strauss.** 1992. Personal communication.
12. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
13. **Lutz, C. T., W. C. Hollifield, B. Seed, J. M. Davie, and H. V. Huang.** 1987. Syrinx 2A: an improved λ phage vector designed for screening DNA libraries by recombination *in vivo*. Proc. Natl. Acad. Sci. USA **84:**4379.
14. **Michel, M. R., M. Elgizoli, Y. Dai, R. Jakob, H. Koblet, and A.-P. Arrigo.** 1990. Karyophilic properties of Semliki Forest virus nucleocapsid protein. J. Virol. **64:**5123–5131.
15. **Milligan, J. F., D. R. Groebe, G. W. Witherell, and O. C. Uhlenbeck.** 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res. **15:**8783–8798.
16. **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.
17. **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.
18. **Rossman, M. G., and J. Johnson.** 1989. Icosahedral RNA virus structure. Annu. Rev. Biochem. **58:**533–573.
19. **Sacher, R., and P. Ahlquist.** 1989. Effect of deletions in the N-terminal basic arm of brome mosaic virus coat protein on RNA packaging and systemic infection. J. Virol. **63:**4545–4552.
20. **Sanger, F. S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
21. **Schlesinger, S., and M. Schlesinger.** 1990. Replication of Togaviridae and Flaviviridae, p. 697–711. *In* B. N. Fields and D. M. Knipe (ed.), Virology. Raven Press, New York.
22. **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.
23. **Strauss, E. G., and J. H. Strauss.** 1986. Structure and replication of the alphavirus genome, p. 35–82. *In* S. Schlesinger and M. J. Schlesinger (ed.), The Togaviridae and Flaviviridae. Plenum Press, New York.
24. **Ulmanen, I., H. Soderlund, and L. Kaariainen.** 1976. Semliki Forest virus capsid protein associates with the 60S ribosomal subunit in infected cells. J. Virol. **20:**203–210.
25. **Ulmanen, I., H. Soderlund, and L. Kaariainen.** 1979. Role of protein synthesis in the assembly of Semliki Forest virus nucleocapsid. Virology **99:**265.
26. **Van Steeg, H., M. Kasperaitis, H. O. Voorma, and R. Benne.** 1984. Infection of neuroblastoma cells by Semliki Forest virus. The interference of viral capsid protein with the binding of host messenger RNAs into initiation complexes is the cause of the shut-off of host protein synthesis. Eur. J. Biochem. **138:**473–478.
27. **Vriend, G., M. A. Hemminga, B. J. M. Verduin, J. L. DeWit, and T. J. Schaafsma.** 1981. Segmental mobility involved in protein-RNA interaction in cowpea chlorotic mottle virus. FEBS Lett. **134:**167–171.
28. **Vriend, G., B. J. M. Verduin, and M. A. Hemminga.** 1986. Role of the N-terminal part of the coat protein in the assembly of cowpea chlorotic mottle virus. A 500 MHz proton nuclear magnetic resonance study and structural calculations. J. Mol. Biol. **19:**453–460.
29. **Weiss, B., H. Nitschko, I. Ghattas, R. Wright, and S. Schlesinger.** 1989. Evidence for specificity in the encapsidation of Sindbis virus RNAs. J. Virol. **63:**5310–5318.
30. **Weiss, B., and S. Schlesinger.** 1981. Defective interfering particles of Sindbis virus do not interfere with the homologous virus obtained from persistently infected BHK cells but do interfere with Semliki Forest virus. J. Virol. **37:**840–844.
31. **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.
32. **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:**435–442.