

Structure-Function Relation of the NH₂-Terminal Domain of the Semliki Forest Virus Capsid Protein

KERSTIN FORSELL, MAARIT SUOMALAINEN, AND HENRIK GAROFF*

Center for Biotechnology, Karolinska Institute, Novum, S-141 57 Huddinge, Sweden

Received 13 July 1994/Accepted 18 November 1994

The capsid (C) protein of alphaviruses consists of two protein domains: a serine protease at the COOH terminus and an NH₂-terminal domain which is thought to interact with RNA in the virus nucleocapsid (NC). The latter domain is very rich in positively charged amino acid residues. In this work, we have introduced large deletions into the corresponding region of a full-length cDNA clone of Semliki Forest virus, expressed the transcribed RNA in BHK-21 cells, and monitored the autoprotease activity of C, the formation of intracellular NCs, and the release of infectious virus. Our results show that if the gene region encoding the whole NH₂-terminal domain is removed, the expressed C protein fragment cannot assemble into NCs and virus particles but it is still able to function as an autoprotease. Thus, these results underline the general importance of the NH₂-terminal domain in the virus assembly process and furthermore show that the serine protease domain can function independently of the NH₂ terminus. Surprisingly, analysis of additional C protein deletion variants showed that not all of the NH₂-terminal domain is required for virus assembly, but large deletions involving up to one-third of its positively charged residues are still compatible with NC and virus formation. The fact that so much flexibility is allowed in the structure of the NH₂-terminal domain of C suggests that most of this region is involved in nonspecific interactions with the encapsidated RNA, probably through its positively charged amino acid residues.

The alphavirus genome is represented by an approximately 11-kb-long single-stranded RNA molecule which is of plus polarity (10, 28, 33). In the infected cell, this RNA associates with 240 copies of a capsid (C) protein (33 kDa) into a nucleocapsid (NC) structure (5). The NC matures into a virus particle by budding at the cell surface (25). Through this process, the NC becomes enveloped into a lipid bilayer with transmembrane spike proteins.

A central question in the replication of the virus is how it is able to selectively encapsidate its own genome into an NC structure that can be used for virus particle formation. Most models include both specific and nonspecific interactions between virus proteins and the genome. It is thought that the former interaction is responsible for the selective incorporation of the genome and the latter is responsible for the packaging of the negatively charged nucleic acid into a protein coat.

An examination of the C protein structure of alphaviruses shows two distinct domains: a COOH-terminal one which is folded into a serine protease and an NH₂-terminal one which appears unstructured when analyzed by X-ray crystallography (5). The latter domain has been considered the RNA binding domain of the C protein because it is exceptionally rich in positively charged amino acid residues (Arg and Lys), which are thought to be necessary for making nonspecific interactions with RNA (9, 23). The serine protease domain is necessary for release of the C protein from the structural polyprotein C-p62-E1, where p62 and E1 represent the two membrane protein subunits of the virus spike (1, 14, 20). Processing of this polyprotein occurs on the nascent chain (11).

Weiss et al. have identified a stretch of nucleotides between positions 746 and 1226 of the Sindbis virus genome that functions as an encapsidation signal (40). Most recently, the signal

has been limited to a 132-nucleotide fragment between positions 945 and 1076 (39). The same group has also found that a stretch of amino acid residues from 76 to 116 in the NH₂-terminal domain of Sindbis virus C protein (264 residues long) is responsible for binding of C to the RNA encapsidation signal in an *in vitro* binding assay (12). Within this stretch of amino acid residues, there is a very conserved sequence, K(97)-P-K-P-G-K-R-Q-R-M(106), which is proposed to represent an RNA binding motif. Thus, this peptide could mediate the specific C-RNA interaction that is thought to initiate the RNA encapsidation process in infected cells. Accordingly, the rest of the NH₂-terminal domain (i.e., almost all of it) is expected to be involved in nonspecific interactions with the RNA, for instance via the extensive clusters of positively charged residues. In this work, we have introduced large deletions into that region of the full-length cDNA clone of Semliki Forest virus (SFV) that encodes the NH₂-terminal domain of C and expressed the corresponding RNA in BHK-21 cells in order to monitor the autoprotease activity of C, the formation of intracellular NCs, and the release of infectious virus (18). We find that if the gene region encoding the whole amino-terminal domain is removed, the expressed C protein fragment still functions as an autoprotease but cannot assemble into NCs and virus particles. These results underline the general importance of the amino-terminal domain in the virus assembly process. However, to our surprise, not all of the NH₂-terminal protein domain is required for virus assembly; large deletions involving up to one-third of its positively charged residues are still compatible with NC and virus formation.

MATERIALS AND METHODS

Preparation of stock virus. The SFV CΔ66-78 stock virus was prepared essentially as described previously for wild-type (wt) virus (38). As an inoculum, we used plaque-titrated media from SFV CΔ66-78 RNA-transfected cells (see below). About 10⁷ BHK-21 cells in a bottle were infected with mutant virus (multiplicity of infection of 0.1) and incubated for 18 h before harvesting of progeny virus. This virus preparation still carried the SFV CΔ66-78 RNA gene since it

* Corresponding author. Mailing address: Center for Biotechnology, Karolinska Institute, Novum, S-141 57 Huddinge, Sweden. Phone: 46-8-608 91 25. Fax: 46-8-774 55 38.

directed the synthesis of the deleted C protein in infected cells and produced small plaques in plaque assay (see Results).

Plasmids. Plasmid pSP6-SFV4 contains a full-length cDNA clone of the SFV genome under the control of the SP6 polymerase promoter (18). Plasmids pSFV-CΔ11-29, pSFV-CΔ11-63, pSFV-CΔ66-78, and pSFV-CΔ1-112 are derivatives of pSP6-SFV4 with deletions in the C gene region. The numbers indicate the deleted codons in the C gene. The deletion variants were constructed via sub-cloning steps according to standard techniques (19). A brief description of the constructions of the latter plasmids is given below.

Plasmids pSFV-CΔ11-29 and pSFV-CΔ11-63 were constructed by deleting the 59-bp *Xma*III-*Bst*XI and 159-bp *Xma*III-*Stu*I fragments, respectively, from the C gene region. The ends of the DNA molecules were blunted with mung bean nuclease prior to ligation.

For the construction of pSFV-CΔ66-78, we synthesized in vitro a 133-bp *Stu*I-*Csp*45I C gene fragment in which the C gene sequence from nucleotides 247 to 285 had been deleted (numbering is from the beginning of the SFV 26S mRNA sequence [10]). This was used to replace the wt 172-bp *Stu*I-*Csp*45I fragment of the C gene in pSP6-SFV4. The deleted fragment was synthesized by PCR using Vent DNA polymerase (New England Biolabs). The primers used were 5'TAGTGTCTAATAAGGCTCC*ACGCAGCCCAAGAAGATCAA CCG3' (the 5'-end primer) and 5'CCTTTGTCGCCACCAGG3' (the 3'-end primer). The 5'-end primer contains the deletion (*). Conditions for PCR were as recommended by manufacturer.

To generate pSFV-CΔ1-112, the following synthetic adaptor was used:

```
5'GATCCGCACCATGAAGATTGAAAATGACTGTATC
3'GCGTGGTACTTCTAACTTTTACTGACATAG
TTCGAAGCCG3'
AAGCTTCGGCCTAG5'
```

This molecule has a cohesive *Bam*HI site at both ends. The first ATG corresponds to Met-113 of the wt C gene. This is preceded by the same nucleotide sequence as is the ATG corresponding to the initiator Met of the SFV structural polyprotein. The adaptor nucleotide was first inserted into *Bam*HI-linearized pSFV-1 (17). Clones containing the adaptor in the correct orientation were identified by sequencing. A plasmid with the adaptor in the correct orientation was then digested with *Csp*45I and *Spe*I, and the 10,279-bp fragment was ligated to the 3,732-bp *Csp*45I-*Spe*I fragment of pSP6-SFV4 to give pSFV-CΔ1-112. All constructs were verified by DNA sequencing.

RNA transcription, electroporation, and metabolic labelling of transfected cells. In vitro transcription of viral RNA was done as previously described (18). The in vitro-made RNA was transfected into BHK-21 cells (American Type Culture Collection) by electroporation (31). The suspension of electroporated cells was diluted 1:20 in complete BHK-21 medium (BHK-21 medium [GIBCO] supplemented with 5% fetal calf serum, 20 mM HEPES [N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid; pH 7.3], and 20% tryptose phosphate broth), and cells were plated on several 3-cm-diameter dishes. At 7 to 8 h postelectroporation, culture media were replaced with a methionine-free minimum essential medium (GIBCO) supplemented with 10 mM HEPES. After 30 min at 37°C, the medium was replaced with identical medium into which 100 μCi of [³⁵S]methionine (Amersham) was added per ml. The cells were incubated (pulsed) at 37°C for 5 or 30 min. After the pulse, cells were washed once with BHK-21 medium containing 10 mM HEPES and a 10-fold excess of unlabeled methionine and then incubated (chased) in the same medium for 15 to 300 min. After the chase, culture media were collected and clarified by centrifugation in an Eppendorf centrifuge (5,000 rpm for 5 min at 4°C). Cell monolayers were placed on ice, washed twice with ice-cold phosphate-buffered saline, and solubilized with 1% Nonidet P-40 (NP-40) lysis buffer containing the following cocktail of proteolytic inhibitors: 10 mM iodoacetamide, 2.5 mM phenylmethylsulfonyl fluoride, 2 μg of aprotinin per ml, 5 μg of antipain per ml, 0.5 μg of leupeptin per ml, and 0.7 μg of pepstatin per ml. Cell nuclei were removed from cell lysates by centrifugation in an Eppendorf centrifuge (5,000 rpm for 5 min at 4°C).

Analysis of virus structural proteins and particles. Viral structural proteins in cell lysates were analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) after being extracted by immunoprecipitation with a monoclonal anti-C antibody (36-1-9 [13]), polyclonal anti-C antiserum (8), or monoclonal anti-E1 antibody (UM 8.139 [4]). Virus particles in media were recovered either by immunoprecipitation with the monoclonal anti-E2 antibody (UM 5.1 [4]) or by pelleting through a 10% sucrose cushion (Beckman JA 18.1 rotor at 17,000 rpm for 2 h at 4°C). Equal volumes (100 μl) of cell lysates and corresponding culture supernatants were analyzed. The immunocomplexes were precipitated with 40 μl of a 1:1 (vol/vol) slurry of protein A-Sepharose (Pharmacia Biotech, Uppsala, Sweden) in 10 mM Tris-HCl (pH 7.5). When necessary, rabbit anti-mouse immunoglobulins (Dakopatts a/s, Glostrup, Denmark) were used as linking antibodies. Immunoprecipitates were washed and processed for SDS-PAGE (10% gel) as described previously (30). Radioactivity in protein bands was quantitated as described earlier (38) or by using a PhosphorImager (26). All immunoprecipitations were compared with trichloroacetic acid (TCA) precipitations of corresponding samples. These studies showed that the immunoreactions were capturing between 90 and 60% of the corresponding antigen.

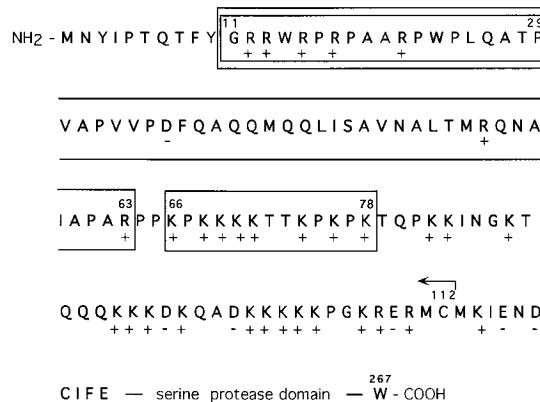


FIG. 1. C protein deletion variants studied in this work. The amino acid sequence of the NH₂-terminal protein domain of SFV C protein is shown. Amino acid residues are indicated by the one-letter code. The COOH-terminal part of the C protein sequence forms a serine protease, and this is indicated at the end of the sequence. The deletions introduced into the C protein sequence are either boxed (Δ11-29, Δ11-63, and Δ66-78) or indicated with an arrow (Δ1-112). All positively (+) and negatively (-) charged amino acid residues are also marked.

Analysis of intracellular NCs. Intracellular NCs were analyzed by sedimentation in sucrose gradients essentially as described before (34). Briefly, lysates (150 μl) of transfected and pulse-labelled cells were incubated with EDTA (final concentration, 25 mM) on ice for 15 min before application onto 15 to 30% (wt/wt) sucrose gradients in TNE buffer (50 mM Tris-HCl [pH 7.4], 100 mM NaCl, 1 mM EDTA) containing 0.1% NP-40. The EDTA treatment results in the dissociation of ribosomes into separate subunits. In some cases, lysates were, in addition, treated with 100 μg of RNase A per ml for 10 min on ice. Centrifugation was performed in a Beckman SW41 rotor at 40,000 rpm for 2 h at 4°C. Sucrose gradients were fractionated from the bottom, and fractions were analyzed by SDS-PAGE. In most cases, protein and protein complexes in the fractions were concentrated either by extraction with monoclonal anti-C antibody or by TCA precipitation. The latter precipitation was done by adding to an aliquot of the fraction an equal volume of 20% TCA. The precipitated immunocomplexes and the TCA pellets were resuspended in sample buffer for SDS-PAGE.

Determination of virus titers. To measure the infectivity of particles produced from transfected cells, serial dilutions of media were used for infection of BHK-21 cells. After 1 h of infection, cells were overlaid with 1.8% low-melting-point agarose (FMC Bioproducts, Rockland, Maine) in complete BHK-21 medium. After incubation at 37°C for 48 h, plaques were visualized by staining with neutral red. Alternatively, the infectivity of particles was determined by infecting BHK-21 cells grown on coverslips with serial dilutions of 3-h chase media. At 6 h postinfection, cells were processed for indirect immunofluorescence (35) and stained with a polyclonal anti-E2 antiserum or, alternatively, with the monoclonal anti-E1 antibody. Fluorescein-conjugated sheep anti-mouse and rhodamine-conjugated goat anti-rabbit antibodies (BioSys, Compiègne, France) were used as second antibodies. Virus titers of 3-h chase media (infectious units per milliliter) were determined by counting the cells positive for anti-E2 or anti-E1 staining.

RESULTS

Figure 1 shows the amino acid sequence of the NH₂-terminal domain of SFV C protein. The border between the NH₂-terminal and COOH-terminal domains is not possible to define exactly, but two hallmarks are that (i) the Sindbis virus C protein is resistant to trypsin digestion after residue 110 and (ii) the folding of the serine protease domain can be monitored in the crystal structure of Sindbis virus C from residue 114, which corresponds to residue 119 in the SFV C sequence (5, 29, 42). The sequence of the amino acid residues in the NH₂-terminal domain contains a large number of positively charged residues which can be grouped approximately into three clusters: an NH₂-terminal one around residue 15, a middle one around residue 69, and a COOH-terminal one around residue 97. In the sequence shown in Fig. 1, we have indicated those C protein deletion variants which are expressed by the four different SFV RNA mutants used in this study. There are three

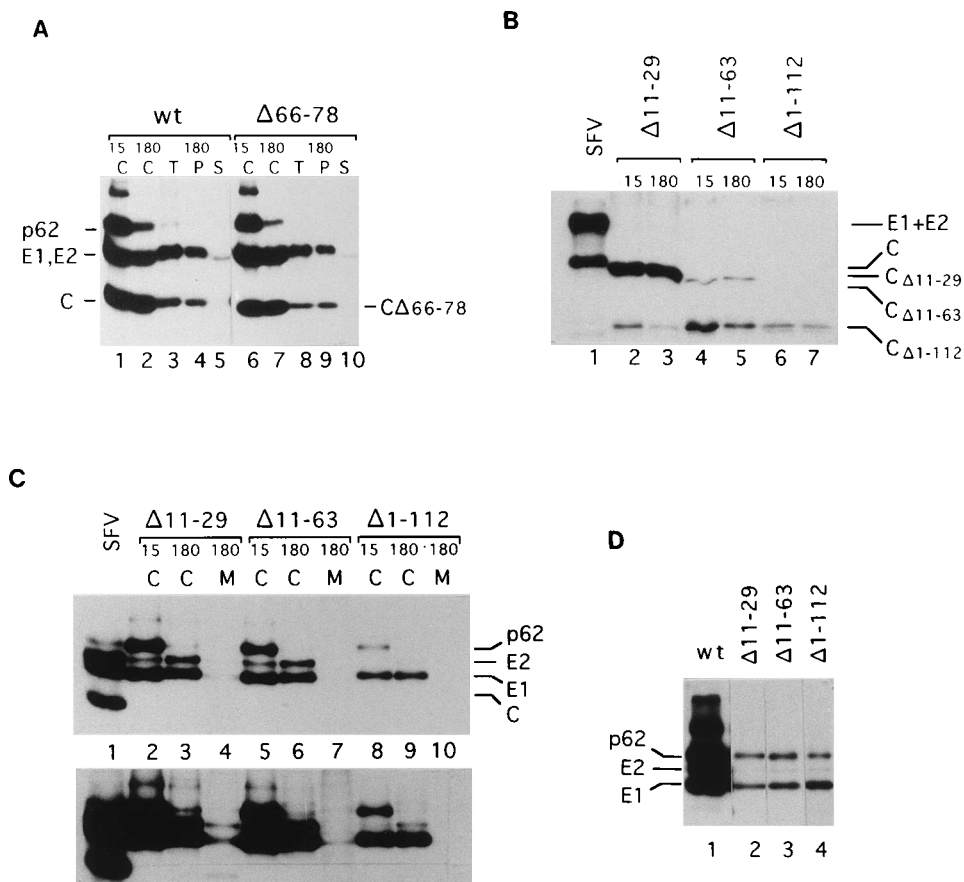


FIG. 2. Viral protein synthesis in cells transfected with capsid deletion variants of SFV. BHK-21 cells were transfected and pulse-labelled for 30 min with [35 S]methionine and chased for 15 min or 3 h before lysis and analysis by SDS-PAGE (10% gel). The medium was also collected and analyzed separately. Viral proteins in cells transfected with SFV wt and SFV Δ 66-78 RNAs are shown in panel A, lanes 1 to 5 and 6 to 10, respectively. The analyses of C protein variants in SFV Δ 11-29, Δ 11-63, and Δ 1-112 RNA-transfected cells are shown in panel B, and corresponding analyses of viral membrane proteins are shown in panel C. Panel D shows a comparison of cell samples (C) and of medium samples with viral particles extracted with anti-E2 antibody (M), total protein of media recovered by TCA precipitation (T), particles pelleted through a 10% sucrose cushion (P), and TCA-precipitable protein left in supernatant after pelleting (S) are shown. Cell samples were extracted with anti-E1 and anti-C antibodies (A), with anti-C antibodies (B) and with anti-E1 antibodies (C and D). SFV wt virus control samples are shown in lanes 1 of panels B to D. All panels represent autoradiographs. Note that samples in panels A and B have been reduced, whereas those in panels C and D have not. When reduced, the E1 and E2 subunits comigrate in SDS-PAGE.

variants which encode internal C protein deletions (Δ 11-29, Δ 11-63, and Δ 66-78) and one (Δ 1-112) in which the gene region encoding the entire NH₂-terminal domain is deleted. In Δ 11-29 RNA, the deletion involves the NH₂-terminal cluster of positively charged residues. In Δ 11-63 RNA, an additional stretch of uncharged amino acid residues is deleted. In Δ 66-78 RNA, most of the middle cluster of positively charged residues is deleted. The SFV C variants were engineered by using plasmid pSP6-SFV4, which contains a full-length cDNA clone of SFV (18). Replication-competent RNA was transcribed from respective plasmid variants and electroporated into BHK-21 cells in order to study the phenotype of virus replication. As all SFV C protein RNA variants except Δ 66-78 RNA expressed protein at levels about 20- to 30-fold lower than that of wt RNA (see below), we will first describe the phenotype of the high-expressing Δ 66-78 RNA variant.

Expression of SFV Δ 66-78 RNA. Figure 2A, lanes 6 and 7, shows the analyses of anti-E1 and anti-C antibody-precipitated proteins of Δ 66-78 RNA-transfected cells which had been metabolically labelled with [35 S]methionine for 30 min and chased for 15 min and 3 h; lanes 1 and 2 show same analyses with wt C protein. It is demonstrated that the mutant C protein

is somewhat smaller than the wt C protein. This size reduction corresponds to the deletion made in the C gene of this variant. Furthermore, it is evident that the mutant RNA directs the synthesis of wt-like p62-E1 complexes. These are precipitated with the monoclonal anti-E1 antibody (compare lanes 1 and 6) and apparently also processed into E2-E1 complexes during the chase (lanes 2, 6, 25, 38, 44). Note that E2 comigrates with E1 in this gel on which reduced samples have been analyzed.

The formation of NC structures from C protein and the viral RNA can be assayed by sedimentation in sucrose gradients following solubilization of infected cells with a mild detergent such as NP-40 (34). Earlier studies have shown that newly synthesized C proteins bind to ribosomes before being complexed with viral RNA (34). The whole process is very fast: within a few minutes after synthesis, most C protein are already assembled into NCs. To analyze the behavior of intracellular Δ 66-78, we therefore subjected lysates of pulse-labelled cells to sedimentation in sucrose gradients at conditions which separated NCs (150S) and ribosome subunits (60S and 40S) into different bands. The result (Fig. 3B) demonstrates that the bulk of the Δ 66-78 protein is found in fast-sediment-

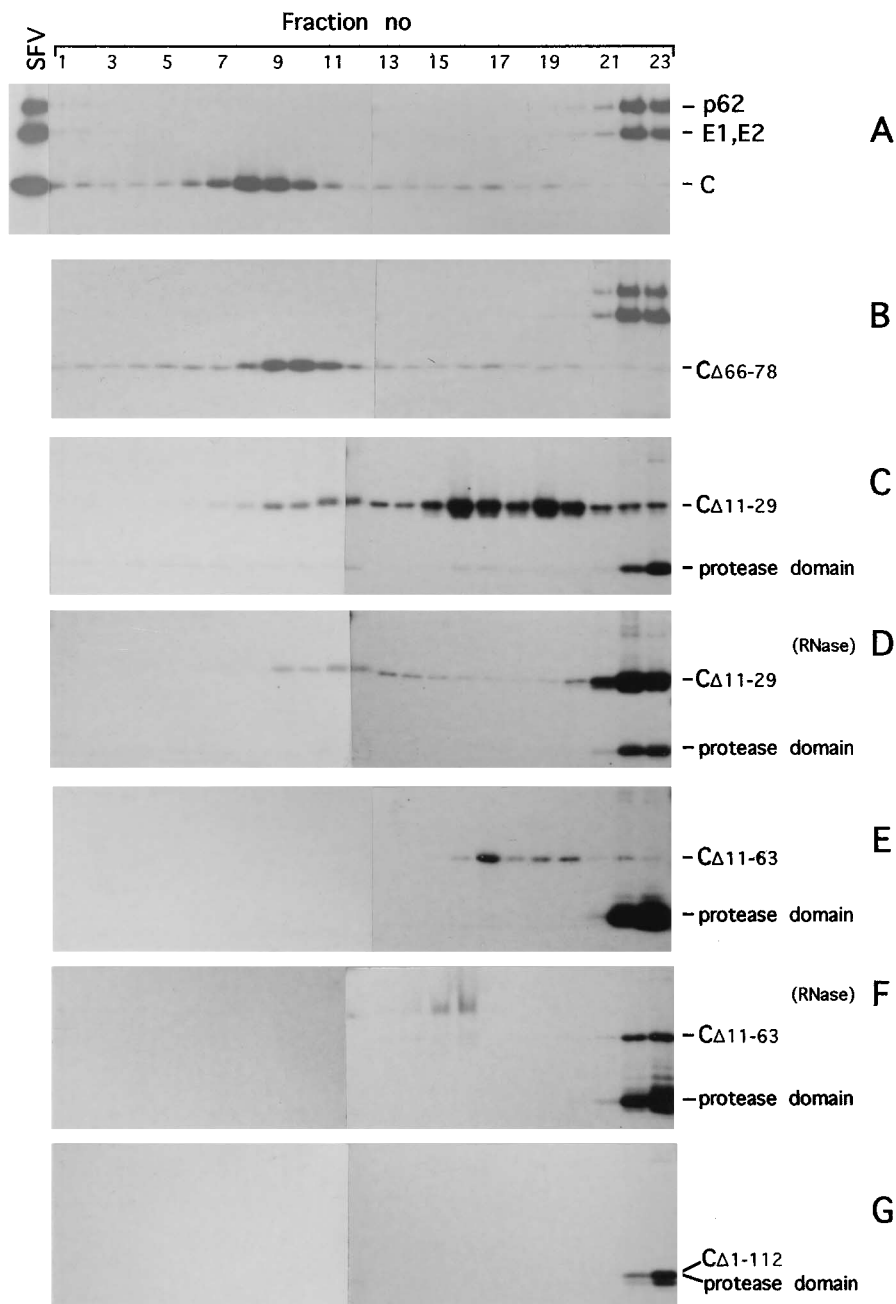


FIG. 3. Sedimentation analyses of lysates of cells transfected with SFV CΔ RNAs. Transfected and pulse-labelled cell lysates were run on 15 to 30% sucrose gradients in an SW41 rotor for 120 min. Gradients were fractionated and analyzed for labelled viral proteins by SDS-PAGE. Analyses of cells transfected with wt SFV RNA (A), SFV CΔ66-78 RNA (B), SFV CΔ11-29 RNA (C and D), SFV CΔ11-63 RNA (E and F), and SFV CΔ1-112 RNA (G) are shown. In the analyses shown in panels D and F, lysates were treated with RNase before sedimentation. In the first lane of panel A, we have analyzed labelled wt SFV by SDS-PAGE as a control. In analysis shown in panels C to G, the viral proteins of each fraction were concentrated by extraction with monoclonal anti-C antibody (top 10 fractions) or TCA precipitation (the other fractions) prior to electrophoresis.

ing structures (fractions 8 to 11). These structures migrate almost as fast as the 150S NCs of wt virus-infected cells (fractions 7 to 10; Fig. 3A). The solubilized membrane protein subunits p62, E2, and E1 are all found at the top of the gradient (fractions 21 to 23). Thus, from these results, it is concluded that the CΔ66-78 protein is able to form NC-like structures. Additional experiments showed that newly synthesized CΔ66-78 is able to assemble into NC slightly more slowly and less efficiently than wt C. This is shown in Fig. 4, which

represents an experiment in which the association of pulse-labelled (5 min) C and CΔ66-78 with ribosome subunits and NCs was monitored after 5- and 15-min chases.

An important question was whether infectious particles were formed in CΔ66-78 RNA-transfected cells. To answer this question, we first analyzed the media of transfected cells for viral proteins that can be sedimented as particles. The results showed clearly that significant amounts of pulse-labelled CΔ66-78 protein and the membrane protein subunits were

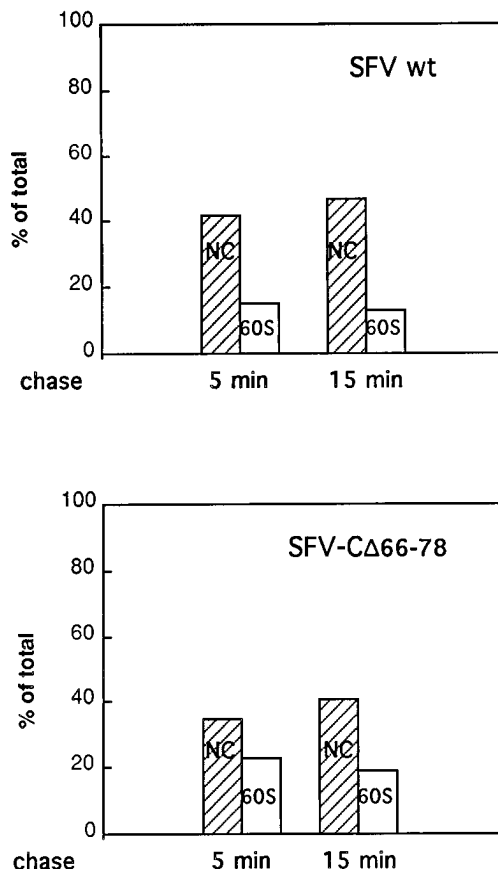


FIG. 4. Assembly of mutant and wt C proteins into NCs. Cell cultures were infected, incubated for 5 h, and then pulse-labelled for 5 min with [³⁵S]methionine. After chase times for 5 and 15 min, cells were lysed and C protein was analyzed by sedimentation on 15 to 30% (wt/wt) sucrose gradients. Fractionation was as shown in Fig. 3. Samples of each fraction were analyzed by SDS-PAGE, and radioactive C proteins in NC bands (fractions 8 to 11 of wt and 9 to 12 of Δ66-78) and ribosome subunit bands (fractions 17 to 20 [60S]) were measured; values are given as percentage of total C protein in gradient fractions.

released into the media as virus-like particles during a 3-h chase period (Fig. 2A, lane 9). A more detailed analyses of release of wt and SFV Δ66-78 particles with time showed that the mutant particles were released with a rate corresponding to 30 to 40% of that of wt particles (Fig. 5). The infectivity of the released mutant virus was measured by plaque and immunofluorescence assays. The results of both assays showed that infectious particles were released from SFV Δ66-78 RNA-transfected cells in an amount which was about 25% of the wt amount (Table 1). The plaques that were produced by SFV Δ66-78 were smaller than wt virus plaques. The specific infectivity (infectivity in relation to released labelled viral protein) calculated for SFV Δ66-78 was found to be very similar to that for wt virus, thus indicating that the released mutant virus was not impaired in entry functions. This mutant virus could also be amplified in tissue culture at least for a few passages without losing its phenotype (see Materials and Methods). Therefore, we conclude that the deletion of residues 66 to 78 from C protein allows a significant assembly of infectious virus particles.

Expression of SFV Δ11-29, Δ11-63, and Δ1-112 RNAs. Figure 2B shows the SDS-PAGE analyses of pulse-labelled and chased C protein variants Δ11-29, Δ11-63, and Δ1-112 that have been extracted from transfected cells with a mono-

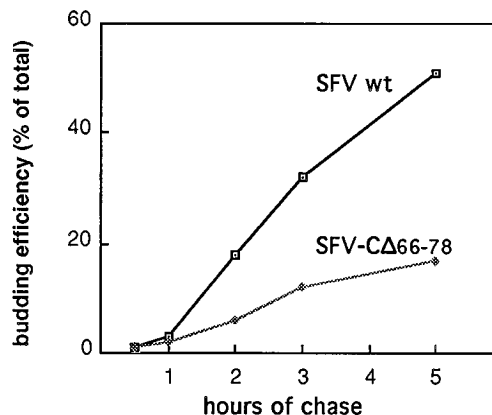


FIG. 5. Kinetics of virus particle release. Cell cultures were infected with SFV Δ66-78 and wt virus and pulse-labelled with [³⁵S]methionine for 30 min. After chases for 0.5, 1, 2, 3, and 5 h, cells were lysed and analyzed for viral proteins by SDS-PAGE. Virus particles in medium samples were pelleted through a 10% (wt/wt) sucrose cushion before analyses by SDS-PAGE. The fraction (percentage) of total virus membrane proteins in cells and medium which was released in particulate form (the budding efficiency) was calculated for each time point.

clonal anti-C antibody. The Δ11-29 protein is seen as a band which migrates slightly faster than wt C protein (lanes 2 and 3). This corresponds to the size of a C protein with the designed deletion. The Δ11-63 material is seen as two bands (lanes 4 and 5), the fainter of which migrates slightly faster than the Δ11-29 material in lane 3. This Δ11-63 protein corresponds to the expected size of the Δ11-63 mutant. The stronger band represents a considerably smaller protein, which comigrates with the Δ1-112 protein (the protease domain) in lanes 6 and 7. The latter proteins migrate as would be expected for a C protein with a 112-residue-long deletion. It therefore appears that Δ11-63 and to a minor extent also Δ11-29 (lanes 2 and 3) are unstable and are degraded into the Δ1-112 form. In other experiments (not shown), we used shorter pulse-labelling (5 min) and chase (2 min) times in order to determine whether the processing of Δ11-63 was a posttranslational event. No accumulation of full-sized Δ11-63 protein could be observed in this sample, suggesting that the protein processing takes place very soon after Δ11-63 chain synthesis (data not shown). We routinely solubilized transfected cells with the detergent NP-40. To exclude the possibility that the Δ11-63 proteins were degraded after the NP-40 solubilization, we also

TABLE 1. Efficiency of protein synthesis, virus budding, and release of infectious virus in SFV Δ RNA-transfected cells

RNA	Translation efficiency ^a (% of wild type)	Budding efficiency ^b (% of wild type)	Infectious units/ml (% of wild type) ^c
Δ11-29	3.5	5.0	1.0 ^d
Δ11-63	4.8	ND ^e	0.01 ^d
Δ66-78	100	40	25 ^f
Δ1-112	4.3	ND	ND

^a Quantitation is based on labeled structural proteins in lysates from transfected cells which had been pulse-labeled for 30 min and chased for 15 min. Values are means from two or more determinations.

^b Quantitation is based on the membrane protein fraction which is released into media after a 3-h chase. Values are means from two or more determinations.

^c Values are means from two or more determinations.

^d Quantitation is based on immunofluorescence assay of 3-h chase media.

^e ND, Not detectable.

^f Quantitation is based on plaque and immunofluorescence assay of 3-h chase media.

used 1% SDS for solubilizing transfected cells. However, the same degradation of C Δ 11-63 protein was observed (data not shown).

Figure 2C shows the membrane protein subunits in SFV C Δ 11-29, Δ 11-63 and Δ 1-112 RNA-transfected cells. It is evident that the anti-E1 antibody precipitates p62-E1 complexes from lysates of cells that have been chased for 15 min (Fig. 2C, lanes 2, 5, and 8) and E2-E1 complexes from the longer chase samples (lanes 3, 6, and 9). Note that in these analyses, E1 and E2 subunits are separated because the samples for gel analysis were not reduced. Note also that in the SFV C Δ 1-112-transfected cells, there seems to be significant degradation of p62-E2 material (lanes 8 and 9). This is probably due to the completely budding negative phenotype of this variant (see below). We have earlier shown that in lack of virus budding, there is rapid degradation of p62-E2 (43). We conclude that all these variant C proteins are able to cleave the C-p62-E1 polyprotein correctly at the C-p62 junction and that the spike protein heterodimers mature as in wt virus. We particularly note that this is also the case with the largest deletion, variant C Δ 1-112. This latter result shows unequivocally that the COOH-terminal serine protease domain can function as a protease completely independently of the NH₂-terminal domain.

In Figure 2D, we have analyzed immunoprecipitated viral membrane protein complexes from corresponding amounts of lysates of cells transfected with wt SFV and SFV C Δ 11-29, Δ 11-63, and Δ 1-112 RNAs in order to compare expression levels. Quantitation shows that about 20 to 30 times more viral proteins are made in wt-transfected cells than in cells transfected with the three deletion variants (Table 1). Other experiments have demonstrated that the reduced expression levels of the three SFV variants are due to the lack of a translational enhancer segment which is located in the 5' region of the C gene (26). This segment is required for efficient translation of this mRNA in infected cells.

Figures 3C through G show the analyses of NC formation in SFV C Δ 11-29, Δ 11-63, and Δ 1-112 RNA-transfected cells. The C Δ 11-29 protein is apparently able to form some NCs that sediment almost as fast as wt NCs (fractions 9 to 11 in Fig. 3C). However, the majority of C Δ 11-29 proteins band in fractions 16 to 20. In control analyses, we have shown that this sedimentation behavior corresponds to that of the ribosomal subunits of a fractionated rabbit reticulocyte lysate (data not shown). The full-length C Δ 11-63 protein also sediments with ribosomal subunits (fractions 17 to 20 in Fig. 3E), whereas the proteolytic fragment of this mutant protein appears at the top of the gradient (fractions 22 and 23). This is also true for the C Δ 1-112 protein (fractions 22 and 23 in Fig. 3G). The nature of the NC-like particles can be tested with RNase treatment. Earlier work has demonstrated that wt NC remains as a complex after such treatment, despite digestion of the encapsidated RNA (16). When the C Δ 11-29 and C Δ 11-63 protein-containing samples were treated with RNase before sedimentation, most of the NC-like material of C Δ 11-29 remained as large sedimenting structures (fractions 9 to 13 in Fig. 3D), whereas all capsid material present in the ribosome subunit fractions (of both C Δ 11-29 and C Δ 11-63) was released into smaller structures that did not migrate significantly into the gradient (fractions 21 to 23 in Fig. 3D and fractions 22 and 23 in Fig. 3F). We conclude that small amounts of wt-like NC structures are found in the case of SFV C Δ 11-29-RNA-transfected cells.

In Fig. 2C, lanes 4, 7, and 10, we have analyzed whether any virus-like particles were released into the media of the SFV C Δ 11-29, C Δ 11-63, C Δ 1-112 RNA-transfected cells. Medium samples of cells that had been pulse-labelled and chased for

180 min were reacted with the anti-E2 monoclonal antibody, and the precipitate was analyzed by SDS-PAGE. The longer exposure of the autoradiograph shows clearly the presence of all viral structural proteins in the medium sample of SFV C Δ 11-29 RNA-transfected cells. This finding suggests that virus particles have been formed in these cells. The budding efficiency was calculated to be about 5% of that of wt virus (Table 1). Very weak viral protein bands are also seen in the analysis of the media from SFV C Δ 11-63 RNA-transfected cells (Fig. 2C, lane 7). No viral proteins were detected in the media of SFV C Δ 1-112 RNA-transfected cells.

To determine whether the media of SFV C Δ 11-29, Δ 11-63, and Δ 1-112 RNA-transfected cells contained infectious particles, we tested clarified culture supernatants for secondary infection of new cell cultures. Infected cells were detected by immunofluorescence staining after 6 h of incubation and used as a measure for infectious particles in the inoculum. This assay was chosen rather than the plaque assay because it would reveal very slowly growing virus variants better. As shown in Table 1, the medium from SFV C Δ 11-29 RNA-transfected cells contained about 10² times fewer infectious particles and that from SFV C Δ 11-63 RNA-transfected cells contained about 10⁴ times fewer infectious particles than medium from wt virus RNA-transfected cells (2.2 × 10⁸ particles per ml).

DISCUSSION

Geigenmüller-Gnirke et al. have shown that the COOH-terminal one-third of the NH₂-terminal domain of the Sindbis virus C protein (amino acid residues 76 to 116) mediates the specific binding of the C protein to the encapsidation signal of the viral RNA genome; this was shown by analyzing the binding of in vitro-synthesized C protein deletion variants to in vitro-transcribed viral RNA that contained the encapsidation signal (12). The corresponding sequence in the SFV C protein is represented by the residues from 77 to 121. In analogy with Sindbis virus, this region most likely specifies the binding of SFV C protein to an encapsidation signal on the SFV RNA genome. In this work, we have used an in vivo approach to analyze the role of the NH₂-terminal domain of the SFV C protein in virus assembly. Our work has concentrated on two-thirds of the NH₂-terminal of this protein domain. Several deletions were introduced into the corresponding C gene region and their effects on virus assembly was analyzed in BHK-21 cells by using our SFV4 cDNA-based expression system (18). We found that the deletion of a considerable amount of the positively charged residues in this part of the C protein is still compatible with the formation of intracellular NCs and with the production of infectious virus particles. The most striking result of this kind was obtained with the SFV C Δ 66-78 variant. In this deletion mutant, eight lysine residues making up the entire middle cluster of positively charged residues have been deleted, and still this mutant C protein is able to take part in the efficient formation of NCs and infectious virus particles. The phenotypes of SFV C Δ 11-29 and C Δ 11-63 are somewhat more difficult to interpret because of the significant reduction in protein expression (20- to 30-fold). Furthermore, these two C deletion mutants were unstable. Naturally, both of these effects will reduce particle assembly considerably, and in this light, we find it very significant that the C Δ 11-29 protein, which has the first cluster of positively charged amino acid residues deleted, and even the C Δ 11-63 protein, in which almost half of the NH₂-terminal domain is deleted, are still able to participate in the formation of some infectious virus. The fact that so much flexibility is allowed in the structure of the NH₂-terminal domain of C suggests that most of this region is involved in

nonspecific interactions with the encapsidated RNA, probably through the positively charged amino acid residues. Thus, in view of these results and those reported by Geigenmüller-Gnirke et al. (12), the alphavirus C protein possesses two types of RNA interactions: one major type of interaction which is nonspecific and another which is specific for the encapsidation signal.

In addition to the internal C deletions described above, we created a deletion which removed the entire NH₂-terminal C protein domain and left only the protease unit linked to the rest of the structural polyprotein. Expression of this mutant showed neither the assembly of NC nor the formation of virus particles. This effect cannot be explained simply by the reduced level of protein expression, which was also displayed by this variant since protein expression was on the same level as that of SFV CΔ11-29, which was able to produce some NC and infectious virus. Therefore, the phenotype of the SFV CΔ1-112 mutant represents clear genetic evidence for the general importance of the NH₂-terminal domain for NC and virus assembly. Furthermore, the fact that this mutant can direct efficient and correct C-p62 cleavage shows that the COOH-terminal protease domain of C can work independently of the rest of C.

Altogether, the assembly reaction of the alphavirus NC might be very similar to those proposed for some simple plant (e.g., brome mosaic virus [BMV]) and bacterial viruses (e.g., MS2, QB, and R17). On the basis of very elegant *in vitro* studies, it has been suggested that the assembly process of R17 is initiated by a specific encapsidation reaction which then stimulates further capsid protein binding to RNA via nonspecific interactions (3). The importance of clusters of basic amino acid residues in capsid protein for RNA packaging has been directly demonstrated in infected cells in the cases of BMV and hepatitis B virus. Sacher and Ahlquist (24) constructed a BMV mutant in which the coat protein was lacking its amino-terminal region with its cluster of basic amino acid residues and showed that the mutant coat protein was unable to encapsidate viral RNA. Similarly, it has been shown that hepatitis B virus capsid needs its basic residues in the COOH-terminal region for nucleic acid encapsidation (15, 21). These residues are grouped into four clusters, one of which is enough for efficient RNA encapsidation. In addition to these *in vivo* studies, there are several *in vitro* studies with the simple plant viruses demonstrating the importance of basic amino acid residues for RNA encapsidation during NC assembly (7, 27, 36, 37).

A very interesting and apparently unique feature of the alphavirus encapsidation process is the involvement of the ribosomes (34). In the infected cell, newly synthesized C polypeptides are found to be associated with the ribosomes shortly before they form NC structures. It is possible that the C polypeptides cannot normally exist in a free form but that the ribosomes act as a carrier or chaperon for the C molecule until it can bind to the RNA. Interestingly, most of the CΔ11-29 and virtually all of the CΔ11-63 variants appear to be associated with ribosomes. This finding suggests that although these C protein variants have lost most (CΔ11-29) or virtually all (CΔ11-63) of their features (sequences) which drive NC assembly, they have retained other features (sequences) that are required for association with ribosomes after synthesis. However, the CΔ1-112 variant, which remained in the top fractions in the gradient analyses, had apparently also lost its capacity to bind to ribosomes.

NC-like structures have been assembled *in vitro* by using isolated monomeric C protein (41). This reaction required the presence of either viral RNA, heterologous single-stranded nucleic acid, polyvinylsulfate, or heparin. These results suggest that simple nonspecific neutralization of the basic charges in

the NH₂-terminal C protein domain is enough to drive NC formation. We have also shown that virus particles can be formed *in vivo* in low amounts if only the SFV structural genome is expressed by a heterologous vaccinia virus-driven expression system (32). These results are difficult to reconcile with the highly selective genome encapsidation in infected cells. However, as has been suggested for the R17 virus encapsidation process (3), this type of selectivity can be obtained if binding to the encapsidation signal is of a slightly higher affinity than to other parts of the genome and if the specific binding stimulates additional nonspecific binding of C.

On the basis of the foregoing discussion, we suggest the following scenario of events leading to formation of alphavirus NC: (i) initiation of synthesis of structural polyprotein from viral 26S mRNA by the ribosome; (ii) binding of the positively charged amino-terminal domain of C protein to ribosome as soon as this part of polypeptide has been completed; (iii) cleavage of the C portion from the rest of the polyprotein when the COOH-terminal protease domain has been made; (iv) transport of C on the ribosome to a viral RNA genome, where specific binding to its encapsidation sequence occurs; (v) stimulation of further C protein to join the complex via C-C interactions and C-RNA interactions of the nonspecific kind; and (vi) completion of RNA encapsidation and NC formation when all 240 C proteins have joined the complex.

We realize that many further studies are required to test this model. In particular, it will be important to identify the C-C interactions. These latter interactions are also likely to be of crucial importance for assembling the RNA and the C protein into an NC structure with a triangulation number which equals 4 (5, 22).

ACKNOWLEDGMENTS

We thank Ingrid Sigurdson typing and Roger Hewson for critical reading of the manuscript.

This work was supported by grant B-AA/BU- 09353-308 from the Swedish Natural Science Research Council.

REFERENCES

- Aliperti, G., and M. J. Schlesinger. 1978. Evidence for an autoprotease activity of Sindbis virus capsid protein. *Virology* **90**:366-369.
- Barth, B.-U., J. M. Wahlberg, and H. Garoff. The oligomerization reaction of the Semliki Forest virus membrane protein subunits. *J. Cell Biol.*, in press.
- Beckett, D., H.-N. Wu, and O. C. Uhlenbeck. 1988. Roles of operator and non-operator RNA sequences in bacteriophage R17 capsid assembly. *J. Mol. Biol.* **204**:939-947.
- Boere, W. A. M., T. Harmsen, J. Vinje, B. J. Benaissa-Trouw, C. A. Kraaijeveld, and H. Snippe. 1984. Identification of distinct antigenic determinants on Semliki Forest virus by using monoclonal antibodies with different antiviral activities. *J. Virol.* **52**:575-582.
- Choi, H.-K., L. Tong, W. Minor, P. Dumas, U. Boege, M. G. Rossmann, and G. Wengler. 1991. Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion. *Nature (London)* **354**:37-43.
- de Curtis, L., and K. Simons. 1988. Dissection of Semliki Forest virus glycoprotein delivery from the trans-Golgi network to the cell surface in permeabilized BHK cells. *Proc. Natl. Acad. Sci. USA* **85**:8052-8056.
- 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.
- Garoff, H. Unpublished data.
- 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.
- Garoff, H., C. Kondor-Koch, and H. Riedel. 1982. Structure and assembly of alphaviruses. *Curr. Top. Microbiol. Immunol.* **99**:1-50.
- Garoff, H., K. Simons, and B. Dobberstein. 1978. Assembly of Semliki Forest virus membrane glycoproteins in the membrane of the endoplasmic reticulum *in vitro*. *J. Mol. Biol.* **124**:587-600.
- Geigenmüller-Gnirke, U., H. Nitschko, and S. Schlesinger. 1993. Deletion analysis of the capsid protein of Sindbis virus: identification of the RNA

- binding region. *J. Virol.* **67**:1620–1626.
13. Greiser-Wilke, I., V. Moennig, O.-R. Kaaden, and L. T. M. Figueiredo. 1989. Most alphaviruses share a conserved epitopic region on their nucleocapsid protein. *J. Gen. Virol.* **70**:743–748.
 14. 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.
 15. Hatton, T., S. Zhou, and D. N. Standring. 1992. RNA- and DNA-binding activities in hepatitis B virus capsid protein: a model for their roles in viral replication. *J. Virol.* **66**:5232–5241.
 16. Kääriäinen, L., and H. Söderlund. 1971. Properties of Semliki Forest virus nucleocapsid. I. Sensitivity to pancreatic ribonuclease. *Virology* **43**:291–299.
 17. Liljeström, P., and H. Garoff. 1991. A new generation of animal cell expression vectors based on the Semliki Forest virus replicon. *Bio/Technology* **9**:1356–1361.
 18. Liljeström, P., S. Lusa, D. Huylebroeck, and H. Garoff. 1991. In vitro mutagenesis of a full-length cDNA clone of Semliki Forest virus: the 6,000-molecular-weight membrane protein modulates virus release. *J. Virol.* **65**:4107–4113.
 19. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 20. Melancon, P., and H. Garoff. 1987. Processing of the Semliki Forest virus structural polyprotein: role of the capsid protease. *J. Virol.* **61**:1301–1309.
 21. Nassal, M. 1992. The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. *J. Virol.* **66**:4107–4116.
 22. Paredes, A. M., D. T. Brown, R. Rothnagel, W. Chiu, R. J. Schoepp, R. E. Johnston, and B. V. V. Prasad. 1993. Three-dimensional structure of a membrane-containing virus. *Proc. Natl. Acad. Sci. USA* **90**:9095–9099.
 23. 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.
 24. 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.
 25. Schlesinger, S. S., and M. J. Schlesinger. 1986. Formation and assembly of alphavirus glycoproteins, p. 121–148. *In* S. S. Schlesinger and M. J. Schlesinger (ed.), *The Togaviridae and Flaviviridae*. Plenum Press, New York.
 26. Sjöberg, M., M. Suomalainen, and H. Garoff. 1994. A significantly improved Semliki Forest virus expression system based on translation enhancer segments from the viral capsid gene. *Bio/Technology* **12**:1127–1131.
 27. 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.
 28. 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.
 29. Strong, R. K., and S. C. Harrison. 1990. Proteolytic dissection of Sindbis virus core protein. *J. Virol.* **64**:3992–3994.
 30. Suomalainen, M., M. Baron, and H. Garoff. 1990. The E2 signal sequence of rubella virus remains part of the capsid protein and confers membrane association in vitro. *J. Virol.* **64**:5500–5509.
 31. Suomalainen, M., and H. Garoff. 1994. Incorporation of homologous and heterologous proteins into the envelope of Moloney murine leukemia virus. *J. Virol.* **68**:4879–4889.
 32. Suomalainen, M., P. Liljeström, and H. Garoff. 1992. Spike protein-nucleocapsid interactions drive the budding of alphaviruses. *J. Virol.* **66**:4737–4747.
 33. Takkinen, K. 1986. Complete nucleotide sequence of the nonstructural protein genes of Semliki Forest virus. *Nucleic Acids Res.* **14**:5667–5682.
 34. Ulmanen, I., H. Söderlund, and L. Kaariainen. 1976. Semliki Forest virus capsid protein associates with the 60S ribosomal subunit in infected cells. *J. Virol.* **20**:203–210.
 35. Vaux, D. J. T., A. Helenius, and I. Mellman. 1988. Spike-nucleocapsid interaction in Semliki Forest virus reconstructed using network antibodies. *Nature (London)* **336**:36–42.
 36. 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.
 37. Vriend, G., B. J. M. Verduin, and M. A. Hemminga. 1981. 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.
 38. Wahlberg, J. M., W. A. Boere, and H. Garoff. 1989. The heterodimeric association between the membrane proteins of Semliki Forest virus changes its sensitivity to mildly acidic pH during virus maturation. *J. Virol.* **63**:4991–4997.
 39. Weiss, B., U. Geigenmüller-Gnirke, and S. Schlesinger. 1994. Interactions between Sindbis virus RNAs and a 68 amino acid derivative of the viral capsid protein further defines the capsid binding site. *Nucleic Acids Res.* **22**:780–786.
 40. 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.
 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. Wengler, G., D. Würkner, and G. Wengler. 1992. Identification of a sequence element in the alphavirus core protein which mediates interaction of cores with ribosomes and the disassembly of cores. *Virology* **191**:880–888.
 43. Zhao, H., and H. Garoff. 1992. The role of cell surface spikes for alphavirus budding. *J. Virol.* **66**:7089–7095.
 44. Ziemiecki, A., H. Garoff, and K. Simons. 1980. Formation of the Semliki Forest virus membrane glycoprotein complexes in the infected cell. *J. Gen. Virol.* **50**:111–123.