Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Functions of Alphavirus Nonstructural Proteins in RNA Replication

LEEVI KÄÄRIÄINEN AND TERO AHOLA

Institute of Biotechnology, Biocenter Viikki FIN-00014 University of Helsinki, Finland

I. Introduction

Alphaviruses are enveloped positive-strand RNA viruses transmitted to vertebrate hosts by mosquitoes. Several alphaviruses are pathogenic to humans or domestic animals, causing serious central nervous system infections or milder infections with, e.g., arthritis, rash, and fever (*1*). The structure and replication of *Semliki Forest virus* (SFV) and *Sindbis virus* (SIN) have been studied extensively during the last 30 years (*2*).

0079-6603/02 $35.00

Alphaviruses have been important probes in cell biology to study translation, glycosylation, folding, and transport of membrane glycoproteins, as well as endocytosis and membrane fusion mechanisms (3). A new organelle, the intermediate compartment, operating between the endoplasmic reticulum and the Golgi complex was found by the aid of SFV (4). During the last 10 years, alphavirus replicons have been increasingly used as expression vectors for basic research, for generation of vaccines, and for production of recombinant proteins in industrial scale. The first attempts to use them as gene therapy vectors, and even in cancer therapy, have already been reported (5, 6).

Taken together with some well-studied alphavirus-like plant viruses, notably *Brome mosaic virus* (BMV) and *Tobacco mosaic virus* (TMV), alphaviruses form the most advanced model system to study eukaryotic positive-strand RNA virus replication. This review focuses on the functions and properties of alphavirus nonstructural or replicase proteins, which form the viral RNA-dependent RNA polymerase complex. A deeper understanding of RNA replication also provides insight to a fundamental stage in the history of life, as RNA replication, currently the exclusive property of viruses, is thought to antedate development of DNA-based genetic information. From a practical point of view, it is needed to combat viral diseases and to improve the properties of RNA virus expression vectors.

The main approaches of our laboratory in the recent years have been twofold. On one hand, we have discovered and characterized the enzymatic activities of the individual replicase proteins, and on the other hand, we have studied the localization, membrane association, and other cell biological aspects of the replication complex.

II. Replication Cycle of Alphaviruses

The alphavirus particles are icosahedral, with a diameter of 700 Å. The virus envelope consists of dimers of transmembrane glycoproteins E1 and E2, which form 80 projections, each consisting of three heterodimers following the symmetry $T = 4$. In SFV a third extrinsic glycoprotein, E3, is associated with the heterodimers. The envelope surrounds the icosahedral nucleocapsid consisting of 240 capsid proteins, arranged in pentamers and hexamers, and a single-stranded RNA molecule of about 11.5 kb (7).

After attachment to cellular receptors, the virus is internalized by adsorptive endocytosis (8). The acid milieu induces a conformational change in the virus envelope, resulting in fusogenic E1 homotrimers, which mediate the fusion of the virus envelope with the endosomal membranes (3). As a result, the nucleocapsid enters the cytoplasm, where uncoating of the genome RNA is carried out by ribosomes (9). The 5′ two-thirds of the SFV 42S RNA genome is translated into a large polyprotein of 2432 aa, designated P1234, which is then autocatalytically

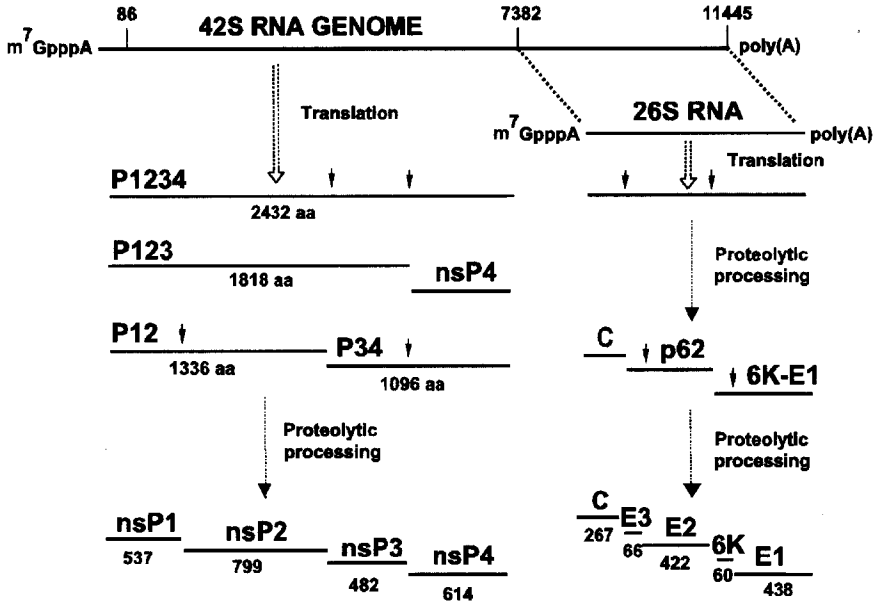

FIG. 1. Protein synthesis directed by SFV mRNA molecules. The primary translation products of the genomic 42S RNA and subgenomic 26S RNA and their proteolytical processing intermediates and individual nonstructural and structural proteins are shown.

cleaved to yield nonstructural protein 4 (nsP4) and P123 (Fig. 1). This is the "early RNA polymerase" responsible for the synthesis of full-size complementary RNA (42S RNA minus strand). After further cleavage of P123 into nsP1, nsP2, and nsP3, the minus strand in turn is used as template for the synthesis of new 42S RNA plus strands, as well as subgenomic 26S mRNA of the structural proteins.

The 26S RNA of SFV, which corresponds to the 3′ third of 42S RNA, is translated to structural polyprotein of 1250 aa consisting of capsid protein, E3, E2, 6K, and E1 proteins (Fig. 1). Translation of 26S RNA starts on free ribosomes, but nascent autocatalytic cleavage of the capsid protein releases the signal peptide at the N terminus of E3, which guides the ribosomes to the endoplasmic reticulum membrane. The capsid protein associates transiently with the 60S subunit of the translating ribosome, and is then transferred to the 42S RNA genome during the assembly of the nucleocapsid (*10*). The assembly of the nucleocapsid is poorly understood, but RNA–protein interactions probably play an important role (*11–13*). E3 plus E2 are translocated as a precursor protein p62, which is co-translationally glycosylated. E1 is preceded by a hydrophobic 6K protein, the N terminus of which serves as a signal peptide. The cellular

signal peptidase cleaves between p62 and 6K, and between 6K and E1. The E1 and p62, which acquire N-linked complex glycans (*14, 15*), are transported via the Golgi complex to the plasma membrane. p62 is cleaved into E3 and E2 by a furin-like protease during transport from the *trans*-Golgi network to the plasma membrane. In SIN-infected cells, E3 is secreted to the medium, whereas SFV E3 remains associated with the E1–E2 heterodimer and is incorporated into virions during their budding at the plasma membrane.

III. Alphavirus-like Superfamily

Positive-strand RNA viruses can be divided into large groups termed superfamilies, the members of which share common features in their encoded replicase proteins, genome organization, and replication strategies. Thus, the superfamily concept is biologically useful, although superfamilies have no officially recognized taxonomic status. A comprehensive analysis of RNA virus sequences suggests that the superfamilies might be as few as three in number, i.e., the picornavirus-like, alphavirus-like, and flavivirus-like viruses (*16*), although it could be argued that the coronavirus-like viruses and positive-strand RNA bacteriophages could also be placed in their own groups. A striking feature of the superfamilies is that although the replicase proteins share sequence similarity, and therefore a common descendent within the superfamily, the viruses may exhibit widely variable structures (nonenveloped and enveloped) and infect different kinds of hosts (plants and animals). Even within the superfamilies, the replicase proteins often show rearrangements and acquisition or deletion of domains.

Members of the alphavirus-like superfamily of positive-strand RNA viruses include the animal viruses of genus *Alphavirus,* and *Rubella virus* (comprising family *Togaviridae*) and *Hepatitis E virus,* as well as the insect viruses of *Tetraviridae* family, and numerous groups of plant viruses, including *Bromoviridae, Closteroviridae,* and the genera *Tobamo-, Tobra-, Hordei-, Furo-, Beny-, Capillo-, Tymo-, Carla-,* and *Potexvirus,* and other plant virus groups. The genome organization of three superfamily members is illustrated in Fig. 2 to highlight some of the similarities and differences. The replicase proteins always feature three conserved domains (*16*), an RNA-dependent RNA polymerase module, a conserved helicase-like domain of helicase superfamily SF1, and a module termed methyltransferase, or more recently capping enzyme. The presence of the methyltransferase/capping enzyme is the distinguishing hallmark of the alphavirus-like superfamily, as it is always present in superfamily members, and on the other hand, its relationship to other viral or cellular polypeptides is exceedingly distant (*17, 18*). The three conserved domains are always organized in the order methyltransferase–helicase–polymerase (in cases where all three

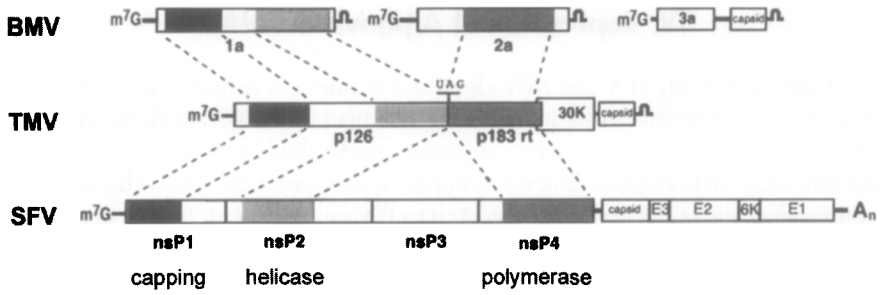

FIG. 2. Comparison of the genome structures of three alphavirus superfamily members: SFV, BMV (brome mosaic virus), and TMV (tobacco mosaic virus). Conserved replicase domains (capping enzyme, helicase, and polymerase) are indicated. Most of the alphavirus-like plant viruses do not encode a protease function, and none contains an nsP3-like domain. The tRNA-like structures at the 3′ end of TMV and BMV genome RNAs are indicated by a schematic structure, and the translational readthrough site of the TMV nonstructural protein (UAG) is marked.

are encoded by the same RNA), although they may be separated by other domains or be finally present in separate polypeptides as a result of proteolytic processing.

Other similarities shared by members of the alphavirus superfamily include (i) membrane-associated replication, (ii) presence of cap0 structure at the 5′ end of positive-strand mRNAs, (iii) an extra untemplated G residue at the 3′ end of minus strand, (iv) asymmetry of replication producing an excess of plus-strand RNA, (v) shut-off of minus-strand RNA synthesis late in infection, and (vi) regulated production of subgenomic mRNA molecules by internal initiation. In many other respects the viruses can be highly dissimilar. For instance, the genome can either consist of a single RNA or be segmented. It can end in a poly(A) or a tRNA-like structure, and the nonstructural proteins may or may not be subject to proteolytic processing (*19*).

The pairwise sequence identity of, for instance, SFV and BMV methyltransferase, helicase, and polymerase domains at the amino acid level is only 15–18%. As the remaining sequence similarity of even the conserved portions of the replicase is so low, it would be surprising if they had not adapted in many ways to different functions during the evolution of different lineages within the superfamily. Nevertheless, the basic replication complex with the conserved capping enzyme, helicase and polymerase domains is likely to be similar throughout the alphavirus-like superfamily, as all viruses have to perform a closely similar set of coordinated reactions. These similarities, when experimentally discovered and proven, will validate the virus superfamily concept. We hope that the information presented here on alphaviruses will prove useful and stimulating for investigators working on other virus groups.

IV. Replication of Alphavirus RNAs

Early work with SFV and SIN identified in infected cells two major single-stranded RNAs sedimenting in sucrose gradients at 42S (49S in SIN) and 26S (*20–22*). In addition, there is a heterogenous, partly double-stranded replicative intermediate (RI) RNA sedimenting between 20S and 29S. When this heterogenous material is isolated and subjected to RNase treatment, double-stranded RNAs arise. The replicative form RFI consists of a complete duplex of full-size plus and minus strands, whereas RFIII represents 26S RNA plus strand together with a respective portion of minus strand. RFII is a double-stranded form of the 5′ two-thirds of the genome. After short pulses with tritiated uridine, label is found in RFI and RFIII, whereas labeling of RFII takes a much longer time. These results lead to the deduction that the genome RNA is synthesized via a replicative intermediate RI_a, which after ribonuclease treatment yields RFI, whereas another structure RI_b is involved in the synthesis of the subgenomic 26S RNA. RNase treatment of RI_b yields both RFII and RFIII. The slow labeling of RFII was interpreted to mean a slow interconversion between RI_a and RI_b, i.e., between synthesis of genomic and subgenomic RNAs (*23, 24*) (Fig. 3).

Numerous studies have shown that virus-specific RNA polymerase activity is associated with cytoplasmic membranes together with heterogenous RNA (*25–27*). The role of 26S RNA as messenger for the structural proteins was finally proven by its *in vitro* translation in the presence of ER membranes, which yields all structural proteins (*28*). The role of the 42S RNA genome as the messenger for nonstructural proteins of SFV has been established (*29–31*) and confirmed by determination of the complete nucleotide sequence of SFV (*32*) and SIN (*33*).

A. Synthesis of RNA Complementary to the Genome

The genomes of alphaviruses have four conserved sequence motifs, two at the 5′ end within the first about 200 nucleotides, one of about an 20-nt-long motif at the 3′ end preceding the poly(A) sequence, and one of about 20 nt, which in minus-strand RNA serves as the promoter for the synthesis of 26S RNA (*34*). In addition, the genome has a recognition signal for its incorporation into nucleocapsid particles, the so-called encapsidation signal. In SIN RNA it is located in the coding region of nsP1 (*35*); and in SFV, in the region coding for nsP2 (*36*). The conserved 5′ and 3′ end sequences plus the encapsidation sequences have been regularly found in defective-interfering RNAs (DI-RNAs) of alphaviruses, although often in rearranged order (*37–40*). The extreme 5′ end of some DI-RNAs differs from that of the genome RNA (*41, 42*). This suggests that the conserved sequences downstream of the extreme 5′ terminus act as the RNA replication signals (*43, 44*) (Fig. 3).

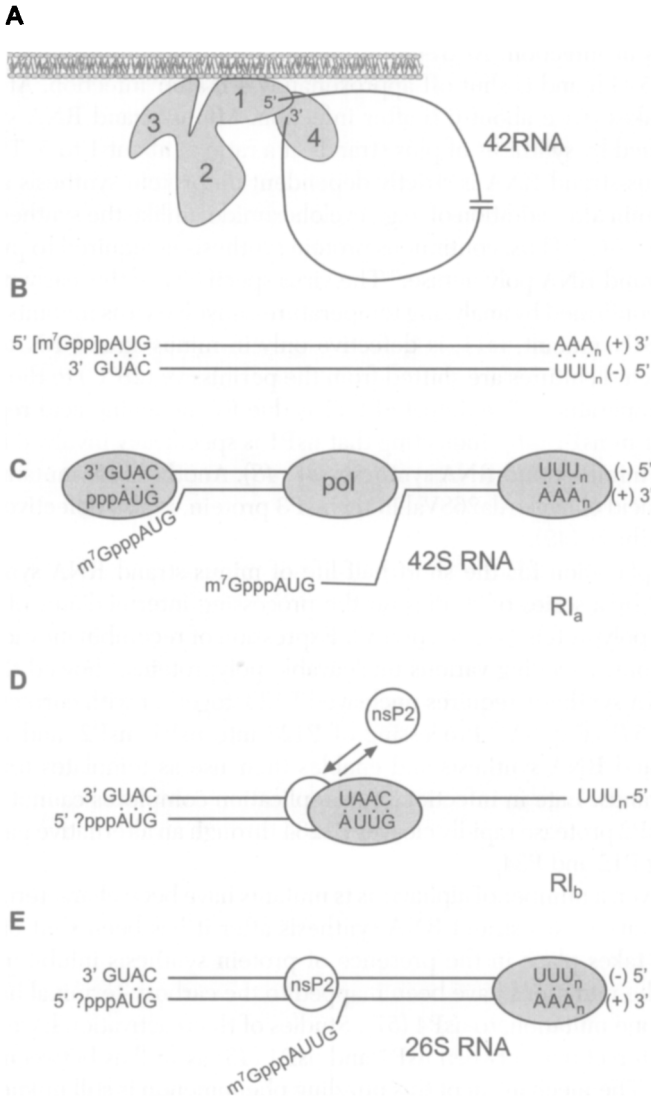

FIG. 3. Models for different stages in the replication of alphavirus RNA. (A) Initiation of minus-strand synthesis. Membrane-bound P123 + nsP4 recognize both ends of the template, which start to cycle through the polymerase. (B) A complete minus strand with a 3′ extra G, annealed with the plus strand, is shown. A puzzling observation indicates that double-stranded RNA in alphavirus-infected cells lacks a cap structure, shown in brackets (*156*), but it is not clear whether this form is an intermediate in replication or an aberrant dead-end product. (C) Classical model of the replicative intermediate RI_a, during the synthesis of genomic plus-strand RNAs with polymerase complexes (pol) consisting of nsP1–nsP4, engaged in initiation, elongation, and termination. (D and E) Initiation and termination of subgenomic RNA synthesis. A regulatory transcription factor (nsP2), interacting with the subgenomic promoter, is shown.

Synthesis of SFV and SIN 42S RNA minus strands takes place within the first hours of infection. At 37–40°C, minus-strand RNA synthesis is maximal around 2.5–3 h and is shut off approximately 4 h after infection. At 28°C, the shut-off takes place about 6 h after infection. Minus-strand RNA synthesis is accompanied by synthesis of plus strands at a ratio of about 1 to 5. The synthesis of minus-strand RNA is strictly dependent on protein synthesis and ceases within 15 min after addition of, e.g., cycloheximide, unlike the synthesis of RNA plus strands (*45*). Thus, continuous protein synthesis is required to produce the "minus-strand RNA polymerase." The virus specificity of this early polymerase has been confirmed by analyzing temperature-sensitive virus mutants. An RNA-negative SIN mutant, ts11, is defective only in minus-strand RNA synthesis, when infected cultures are shifted from the permissive (28°C) to the restrictive (39°C) temperature. The defect of ts11 is due to one amino acid replacement Ala348Thr in nsP1 (*46*), indicating that nsP1 is specifically involved in the regulation of minus-strand RNA synthesis (*47, 48*). Another SIN mutant, ts4, with an amino acid change Ala268Val in the nsP3 protein, is also defective in minus-strand synthesis (*49*).

An explanation for the short half-life of minus-strand RNA synthesis was generated by a series of studies on the processing intermediates of SIN nonstructural polyprotein (see Section V). Expression of recombinant vaccinia virus constructions, encoding various uncleavable polyproteins, showed that minus-strand RNA synthesis requires uncleaved P123, together with correctly cleaved nsP4 (*50–54*) (Fig. 3A). Processing of P123 into nsP1, nsP2, and nsP3 stops minus-strand RNA synthesis and enables their use as templates for synthesis of plus strands. Late in infection, new replication complexes cannot be assembled, as nsP2 protease rapidly cleaves P1234 through an alternative pathway first producing P12 and P34.

However, a number of alphavirus ts mutants have been characterized, which can reactivate minus-strand RNA synthesis after it has been shut off. The reactivation takes place in the presence of protein synthesis inhibitors (*55–58*). Many of the mutations have been mapped to the carboxy-terminal half of nsP2 (*59*); and one mutation, to nsP4 (*57*). Studies of the reactivation have suggested specific interactions between nsP1 and nsP4 (*48*) as well as between nsP2 and nsP4 (*58*). The mechanism of this puzzling phenomenon is still unknown, but it implies that in these mutants the late RNA polymerase, in which all proteolytic cleavages have taken place, can rearrange into the conformation of the early RNA polymerase after a temperature shift from 28 to 39°C. Thus, the structural differences between the two functions of the RNA polymerase must be small.

B. Plus-Strand RNA Synthesis

Quantitation by [^{32}P]orthophosphate equilibrium labeling has shown that about 200,000 molecules of both 42S and 26S RNAs have been synthesized at

8 h after infection in an SFV-infected BHK21 cell (*60*). As the number of minus strands has been estimated by dot-blotting to be approximately 5000 (*48*), RNA replication as a whole is highly asymmetric. The rate of 42S and 26S RNA synthesis is almost linear between 4 and 8 h after infection and is not affected by inhibition of protein synthesis (*61*), indicating that stable RNA replication complexes are continuously producing the plus-strand RNAs in replicative intermediates RI_a and RI_b. Interestingly, plus-strand RNA synthesis is not increased in mutant-infected cells, in which minus-strand RNA synthesis is not shut off (*56*). Neither is RNA synthesis affected by overproduction of viral nonstructural proteins (*62*).

The site of late RNA synthesis has been localized to alphavirus-specific structures, designated cytoplasmic vacuoles (CPVs) (*63–66*). Their diameter varies from 0.2 to 1 μm, and their surface consists of small invaginations, or spherules, with a diameter of about 50 nm (Fig. 4). The four nonstructural proteins as well as nascent RNA molecules are associated with CPVs and more closely associated with the spherules (*66*). The time of appearance of CPVs depends on the amount of infecting virus. With 10–20 plaque-forming units per cell, they first appear 3 h after infection, and their number starts to decline at 6 h. Addition of cycloheximide at 2 h after infection reduces the number of CPVIs, whereas addition after 4 h has no effect (*67*). This time scale coincides with the cessation of minus-strand RNA synthesis (*45*) and the cessation of the synthesis of nsPs

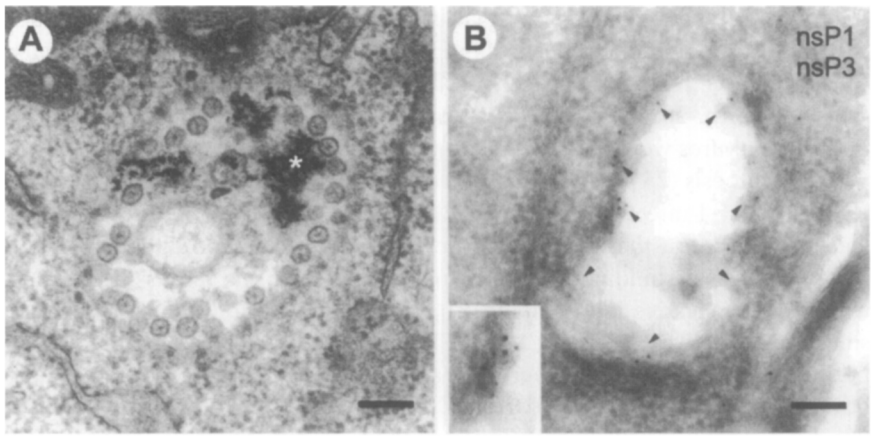

FIG. 4. Electron microscopic images of alphavirus-specific cytoplasmic vacuoles. (A) In an Epon section, invaginations or spherules line the membrane of the vacuole, whose endosomal nature is shown by endocytosed gold-labeled bovine serum albumin (asterisk). (B) Cryo-immunoelectron-microscopic image of spherules (arrowheads) double-labeled with antibodies against nsP1 (small gold particles) and nsP3 (large gold particles). Scale bars = 200 nm. Reproduced from Kujala *et al.*, *J. Virol.*, **75,** 3873–3884 (2001), with permission by the American Society for Microbiology.

(*68*), suggesting that CPVs arise through the action of nsPs like the minus-strand RNA. Later in infection, the synthesis of plus strands continues in association with stable CPVs using the previously synthesized minus strands as templates. CPVs stain for endosomal and lysosomal markers, indicating that they are modified endosomes and lysosomes (*64–66*). Similar structures have been described in rubella virus-infected cells (*69, 70*).

Cytoplasmic membrane fractions, from SFV- and SIN-infected cells sedimenting at 15,000g contain essentially all virus-specific RNA polymerase activity of the cell (*26, 71*). The polymerase activity is associated with a smooth membrane fraction (*27, 71, 72*), which synthesises 42S and 26S RNAs and replicative intermediate RNAs. Attempts to isolate template-free RNA polymerase have so far failed. Expression of P123 plus nsP4 yields membrane fractions, which can support double-stranded RNA synthesis after addition of suitable positive-sense template RNAs containing the conserved 5′ and 3′ sequences of the genome RNA (*54*).

C. Synthesis of 26S mRNA of the Structural Proteins

At a time when the function of 26S RNA was not known, evidence accumulated that its synthesis was not coupled with the synthesis of 42S RNA. Puromycin, given early in SIN infection, inhibits only 26S RNA production (*45, 73*). RNA-negative ts mutants of SFV and SIN, which are unable to replicate at 39°C, have been isolated and characterized (*74–77*). With some of the mutants, the synthesis of only 26S RNA stops after shift to the restrictive temperature. One particular SFV mutant, ts4, turned out to be interesting in this respect. When ts4-infected cells were incubated at 28°C for 6 h to start the RNA synthesis, followed by a shift to 39°C in the presence of cycloheximide, the synthesis of 26S RNA was specifically shut off. However, it was resumed, once the cultures were shifted back to 28°C. The reversal was independent of protein synthesis (*78*). Careful analysis of the labeled RIs and double-stranded RNAs from ts4-infected cells revealed a reversible shift of RI_b to RI_a. This leads us to suggest that a virus-specific protein regulates the initiation of 26S RNA synthesis by binding reversibly to a promoter on the minus-strand RNA within the replicative intermediate RI_b (*79*) (Fig. 3D and 3E). SFV ts4 has a single-amino acid replacement Met781Thr in nsP2. Reversible cross-linking experiments suggested that an increased amount of nsP2 was associated with the other nsPs at 28°C as compared to 39°C (*59*). All SIN ts mutants with a defect in 26S RNA synthesis have been mapped to the C-terminal half on nsP2 (*46, 59*). However, direct proof that nsP2 interacts with the 26S RNA promoter is still lacking.

The subgenomic RNA promoter in the 42S minus strand has been carefully mapped for SIN RNA (*80*). The minimal region is 19 nucleotides upstream and 5 nucleotides downstream of the initiation site of 26S RNA transcription.

However, a larger region containing nucleotides −98 to +14 is required for optimal transcription (*81*). More detailed mapping revealed that optimal transcription could be obtained with sequences −40 to −20 and +6 to +14 together with the core promoter (*82*).

V. Processing of Alphavirus Nonstructural Polyprotein P1234

Short-lived processing intermediates P123, P12, and P34 have been identified as precursors for the four nonstructural proteins, later designated nsP1–nsP4 (*2, 83–85*) (Fig. 1). Proteins of similar size were also identified in SIN-infected cells (*76, 86*). Sequencing of the SIN (*33*) and SFV (*32*) genome RNAs revealed differences in the translation strategies of the nonstructural proteins. In SFV-infected cells, the entire nonstructural region coding for 2432 aa is translated as polyprotein P1234. In the SIN genome, there is an opal termination codon close to the carboxy terminus of nsP3. This is occasionally suppressed, giving rise to P1234 (2513 aa). Thus, in SIN-infected cells, an excess of P123 is produced, and nsP4 can only be produced by processing of P1234 (*34*). *In vitro* translation of SIN genome RNA has been used to study the processing of P123 and P1234, produced from an RNA where the opal codon was mutated to a codon for cysteine. Both P123 and P1234 are processed *in vitro* autocatalytically due to protease activity of nsP2 (*87–90*). This complex process was approached by using constructions in which the cleavage sites were mutated either alone or in combination with substrates in which the protease activity had been eliminated (*34, 89*). The experiments resulted in cleavage "rules": (i) P1234 can cleave autocatalytically to produce P123 and nsP4; (ii) P123 cannot cleave in *cis* but can be cleaved only in a bimolecular reaction; (iii) P12 can slowly cleave autoproteolytically to nsP1 and nsP2; (iv) all three sites can be cleaved *in trans* (*34, 91*). However, the most important result from this work was the possibility to express stable precursor proteins, which revealed the nature of the "early RNA polymerase" (P123 plus nsP4) (*52, 53*).

Many of these experiments have been repeated with SFV nonstructural polyproteins translated *in vitro*. The polyproteins were also expressed alone and in combinations in insect cells using baculovirus vectors. Coexpression of protease P23 with proteolytically defective P1234 or P123 yields P12 in addition to nsP3 from P123 and nsP3 and nsP4 from P1234. Expression of P12 or P23 in insect cells and *in vitro* resulted in their cleavage to nsP1 and nsP2 and to nsP2 and nsP3, respectively. All attempts to cleave proteolytically defective P12 or P123 *in trans*, using, e.g., P123 as the protease, have failed. P34 and P1234 yielded regularly nsP4 by P23 protease, indicating that cleavage at the 3/4 site takes place *in trans* (*92*; Merits *et al.*, unpublished results).

VI. nsP1: A Unique RNA-Capping Enzyme and Membrane Anchor

In recent years, nsP1 has been one of the main objects of our studies. We have discovered a novel kind of coupled methyltransferase and guanylyltransferase activity needed in the capping of virus-specific mRNA molecules and conserved in the alphavirus-like superfamily. On the other hand, we have studied the membrane interaction of nsP1, which is mediated by covalent palmitoylation and by direct interaction of the polypeptide chain with anionic membrane lipids (Fig. 5). Genetic evidence obtained by us and others indicates that nsP1 is involved in the synthesis of minus-strand RNAs.

A. Methyltransferase and Guanylyltransferase Activities

NsP1 is a guanine-7-methyltransferase transferring a methyl group from S-adenosyl-methionine (AdoMet) to GTP (*93, 94*), and it forms a covalent complex with 7-methyl-GMP (m^7GMP) in a guanylyltransferase-like reaction (*95*). These results indicate that nsP1 is an enzyme acting in the formation of m^7GpppA cap structures on virus-specific genomic and subgenomic mRNAs. The substrate specificity of nsP1 differs from cellular RNA-capping enzymes. In the methyltransferase reaction, nsP1 prefers GTP (and dGTP) as methyl-accepting substrates, whereas cellular enzymes methylate cap analogs and unmethylated guanosine-capped RNAs (*94, 96*) (Fig. 6). In the guanylyltransferase reaction, nsP1 forms a covalent complex exclusively with m^7GMP, whereas cellular enzymes form a GMP–enzyme complex. These specificities demonstrate that alphaviruses possess a novel RNA-capping pathway, where GTP is first methylated and only thereafter forms a covalent m^7GMP intermediate with the enzyme (*95*) (Fig. 6). In capping of eukaryotic nuclear RNA, guanylyltransferase first transfers GMP to RNA via a covalent intermediate, and 7-methylation is the final

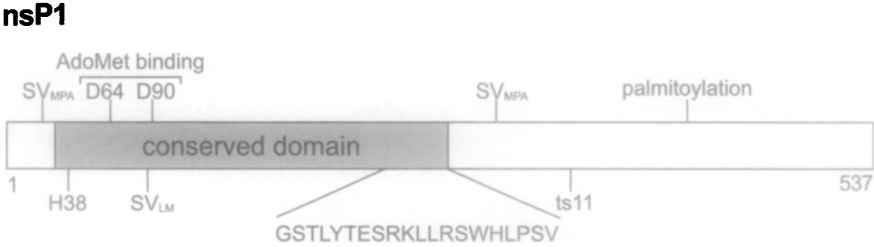

FIG. 5. Scheme of nsP1, showing the location of mutations and other features discussed in the text. SV_{LM} is a SIN mutant resistant to low methionine, and SV_{MPA} is a mutant resistant to mycophenolic acid. The sequence of the lipid-binding peptide, which is located at the C terminus of the conserved domain, is given in single-letter code. The numbering follows SFV sequence.

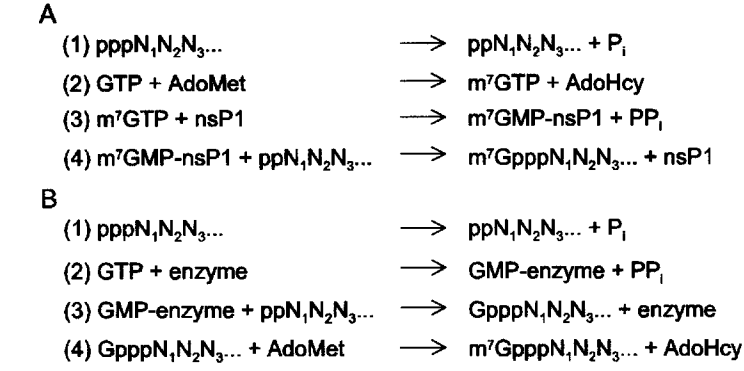

FIG. 6. (A) RNA-capping reactions catalyzed by alphavirus proteins nsP2 (reaction 1, RNA triphosphatase) and nsP1 (reactions 2–4, guanine-7-methyltransferase and mRNA 7-methyl guanylyltransferase). Although nsP1 can release m^7GTP to solution (2), the synthesis of the covalent complex m^7GMP–nsP1 requires methylation and complex formation (reactions 2 and 3) to take place concomitantly, without the release of m^7GTP (18). The transfer of the bound nucleotide to RNA acceptor (step 4) has not been demonstrated. (B) RNA-capping reactions catalyzed by cellular RNA triphospahatase (1), mRNA guanylyltransferase (2–3), and mRNA guanine-7-methyltransferase (4) enzymes.

step taking place on the RNA molecule (for reviews, see Refs. 97 and 98). In both patways, these reactions are preceded by RNA 5′ triphosphatase, which removes the 5′ gamma phosphate of the nascent RNA. In alphaviruses, nsP2 has RNA triphosphatase activity (see Section VII,B).

NsP1-related proteins encoded by members of the alphavirus-like superfamily catalyze similar virus-specific RNA-capping reactions. Methyltransferase activity and covalent complex formation exclusively with m^7GMP have been demonstrated for TMV p126 (*99*), BMV 1a (*100, 101*), and the N-terminal fragments of bamboo mosaic potexvirus (*102*) and HEV replicase proteins (*103*). Thus, all members of the alphavirus-like superfamily are likely to share a similar conserved RNA-capping activity. Since it differs from cellular RNA capping, it offers a target for virus-specific inhibitors of RNA replication. So far, it has been shown that some cap analogs can inhibit SFV nsP1 and HEV-capping enzyme (*96, 103*). Importantly, in this respect, it has been demonstrated by mutational analysis that the capping enzyme activities are essential for alphavirus and BMV replication (*104, 105*). In a BMV replication system in yeast cells, it was possible to directly demonstrate the role of 1a protein in viral RNA capping *in vivo*, as mutation of the host exonuclease XRN1 permitted replication and thus the study of uncapped viral RNAs (*105*).

Some mechanistic details of the reactions catalyzed by nsP1 and related proteins have been studied. The guanylyltransferase reaction absolutely requires

divalent cations (Mg^{2+} or Mn^{2+}), whereas the methyltransferase reaction of at least some family members can proceed in the presence of EDTA (*94, 95, 100, 103*). For BMV 1a, it has been directly demonstrated by NMR spectroscopy analysis of reaction products that methylation takes place only on the 7-position of guanylate (*100*), which is in keeping with the fact that alphavirus-like viral mRNAs are capped with cap0 structure m^7GpppN lacking further methyl groups. Although nsP1 exclusively forms a covalent complex with m^7GMP, it cannot accept m^7GTP directly as a substrate for complex formation. Instead, GTP and AdoMet are required, indicating that methylation and covalent complex formation are normally coupled (*18*). dGTP is a good substrate for methylation by both nsP1 and BMV 1a, but yet these enzymes preferentially form a covalent complex with m^7GMP, as compared to m^7dGMP, indicating that further selection takes place at the guanylyltransfer step (*95, 100, 101*). BMV 1a with a covalently bound guanylate appears to be inhibited in further methyltransferase reactions (*100*). SFV nsP1 has to be associated with anionic phospholipids in order to be active in the capping reactions (*106;* see below). This feature may not be shared among all members of the alphavirus-like superfamily (*103*).

As indicated above, the RNA-capping pathway is conserved within the alphavirus-like superfamily, and the methyltransferase and guanylyltransferase activities are associated with the conserved methyltransferase/capping enzyme domain, which is the hallmark of the alphavirus-like superfamily. It was initially thought that this domain would bear no resemblance to any cellular proteins (*17*). Only when the structures of several cellular methyltransferases became available, it was demonstrated that the predicted secondary structure of a portion of nsP1 and related viral proteins was similar to known methyltransferase structures (*18*). It was also shown by cross-linking that two conserved residues Asp64 and Asp90 in two adjacent loops of SFV nsP1, which correspond to the AdoMet-binding region of structurally studied methyltransferases, are needed for AdoMet binding and methyltransferase activity (*18, 100*). The same region is implicated in AdoMet binding by a double-mutation Arg87Leu and Ser88Cys in SIN nsP1, which enables virus growth in methionine-deprived cells and lowers the K_M of the methyltransferase for AdoMet (*107, 108*). Together, these results strongly suggest that nsP1-related proteins structurally resemble methyltransferases and pinpoint the AdoMet-binding site (Fig. 5).

In contrast, it is less clear where the methyl-accepting GTP substrate binds, as mutations in conserved residues do not abrogate GTP cross-linking to the protein (*18*). However, a SIN mutant resistant to mycophenolic acid and ribavirin, compounds which lower intracellular GTP concentration by inhibiting inosine monophosphate dehydrogenase, appears highly interesting in this respect (*109*). An nsP1 double-mutation Ser23Asn and Val302Met is required for resistance (*110, 111*), which may indicate that these two regions of the protein

flanking the conserved methyltransferase domain may come together in the three-dimensional structure to form the GTP-binding site.

NsP1 does not seem to have any similarity to the cellular guanylyltransferase family of proteins (*18*). In cellular enzymes, a conserved lysine residue forms the covalent linkage with GMP. There are no conserved lysines in nsP1-like proteins, although the properties of the covalent linkage resemble that of a phosphoamide-type bond (*95*). Instead, an absolutely conserved histidine (His38 in SFV nsP1) is a good candidate for covalent binding of m^7GMP, in that mutations of this residue permit retention of methyltransferase activity but completely abolish covalent nucleotide binding (*18, 100*). Thus, joining of guanylyltransferase with the methyltransferase region and coupling of these two activities may represent an evolutionary innovation, which took place in a progenitor of the alphavirus superfamily. It should be noted that although the region showing high conservation within the alphavirus-like superfamily encompasses only 200 aa (*17*), at least alphavirus nsP1s require approximately 500 aa for enzymatic activities (*18, 104*). Considerable structural variation may therefore exist in the capping enzymes of alphavirus-like viruses.

Finally, the active site of nsP1-related proteins may have an additional role in viral RNA replication. Mutations of BMV 1a active site residues can either abolish or enhance the recruitment of viral RNA to a membrane-bound stable form prior to replication. This is a complex pathway requiring host proteins in addition to 1a, and even 1a appears to have several roles in the process. Based on the mutation data, one of these roles could either be influenced by binding of substrates involved in RNA capping or more directly involve a direct recognition of the RNA cap structure by 1a, as one step of recruiting RNA from translation to replication (*105*).

B. Membrane Association

NsP1 is tightly associated with membranes both in alphavirus-infected cells and when expressed alone in mammalian cells (*112, 113*). It can only be released to solution by detergents or highly alkaline solutions (pH 12) (*112, 114*). The tight membrane interaction of nsP1 is mediated by covalent palmitoylation of one or more of the cysteine residues 418–420 in SFV nsP1, and a single-cysteine 420 in SIN nsP1 (*113, 114*) (Fig. 5). Conserved cysteine(s) are found in this position in all alphavirus nsP1s, suggesting that palmitoylation is a conserved feature. However, if the palmitoylation sites of nsP1 are removed by mutagenesis, the protein still retains a weak peripheral affinity for membranes (*114*). Nonpalmitoylated nsP1 can also function in SFV or SIN replication in complexes that appear normal in their localization and morphology. The resultant viruses show only very slightly reduced growth rates in cell culture, but neuropathogenesis of at least SFV encoding nonpalmitoylated nsP1 is abolished in mice (*113*).

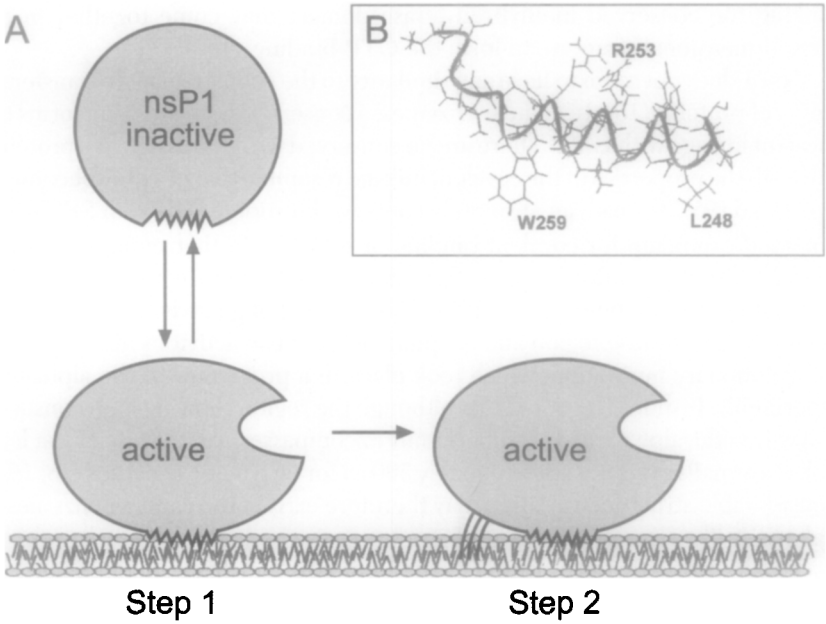

FIG. 7. (A) A schematic model for the membrane binding of nsP1. Association of the lipid-binding peptide with membranes containing anionic phopholipids changes the conformation of nsP1 and activates it as capping enzyme. Association via the binding peptide is weak (step 1) and reversible by high salt, whereas palmitoylated form is bound tightly (step 2) resembling integral membrane proteins. (B) NMR solution structure of the synthetic binding peptide (aa 245–264) of nsP1. W259, which points downwards in the image, intercalates deeply into the lipid bilayer. Figure 7B reproduced from Lampio *et al., J. Biol. Chem.* **275,** 37,853–37,859 (2000), with permission by the American Society for Biochemistry & Molecular Biology.

Nonpalmitoylated nsP1 binds to membranes through direct interaction of the polypeptide chain with anionic phospholipids (Fig. 7). This interaction can be observed for nsP1 expressed in *Escherichia coli* or by *in vitro* translation and is mainly mediated by amino acids 245–264, where both positively charged and hydrophobic residues are required for binding (*106*). As studied by NMR spectroscopy, the respective synthetic peptide forms an amphipathic alpha-helix, which can interact with liposomes containing acidic phospholipids. Trp259 is crucial for association and intercalates deep in the bilayer (*115*). These results are consistent with monotopic binding of nsP1 to membranes, mediated by an amphipathic alpha-helical peptide using both ionic and hydrophobic interactions. Such a mechanism appears to be common for several diverse groups of proteins (*116*). The nsP1 polypeptide segment 245–264 is among the most highly conserved in the alphavirus-like superfamily (*17*) and could play a role in the membrane association of other superfamily members. All nsP1-related proteins

studied to date are associated with membranes (*100, 102, 103*), although the binding mechanism has not been studied in other cases. The palmitoylation site is not conserved in other superfamily members and, indeed, they do not appear to be palmitoylated.

Association with anionic membrane phospholipids is also required for the enzymatic activities of nsP1 (*106*) (Fig. 7). The methyltransferase and guanylyltransferase activities of nonpalmitoylated nsP1 are strongly inhibited by detergents and reactivated by added phospholipid vesicles or mixed micelles containing anionic phospholipids. Phospholipids appear to act by inducing a conformational change is nsP1, as cross-linking of the substrate AdoMet is also inhibited by detergents and reactivated by anionic lipids (*106*). The behavior of palmitoylated nsP1 is slightly more complex. It is also inactivated by several detergents, but activated by deoxycholate and octylglucoside (*94*), which may interact favorably with the palmitoylated form of the protein and permit maintenance of the active conformation. These results indicate that nsP1 is designed to function in a membraneous environment containing anionic phospholipids.

When expressed alone in mammalian cells, both palmitoylated and nonpalmitoylated forms of nsP1 localize predominantly to the cytoplasmic surface of the plasma membrane and also to that of endosomes and lysosomes (*112, 114*). The plasma membrane and the endosomal apparatus, which are connected by vesicle transport, are rich in anionic phospholipids, and therefore the hypothesis arises that the affinity of nsP1 to these lipids directs its intracellular localization (*66*). Palmitoylation could serve a secondary role by fixing the protein tightly to membranes after the initial interaction. However, only palmitoylated forms of nsP1 expressed in animal cells either by themselves by transfection or in the context of alphavirus infection are capable of inducing thin filopodia-like protrusions of the cell surface and selectively disrupting the actin cytoskeleton (*117*). The significance of these phenotypic changes is not clear, but it is intriguing to speculate that they would be connected with the increased pathogenicity of the wild-type virus, as compared to virus-encoding palmitoylation-deficient nsP1 (*113*).

C. Role in Minus-Strand RNA Synthesis

So far, only one of the alphavirus ts mutants, SIN ts11, where Ala348 is changed to Thr, has been mapped to nsP1 (*46*). The phenotype of ts11 has been highly informative, as it exhibits a rapid and selective cessation of minus-strand RNA synthesis but does not interfere with plus-strand synthesis catalyzed by stable replication complexes (*47*). Ts11 suppresses nsP4 mutant ts24 (Gln191Lys), which in otherwise wild-type background allows reactivation of minus-strand synthesis by the stable replication complexes. This is interpreted to mean that nsP1 always directly participates in minus-strand synthesis during either initiation or elongation (*48*). A mutation of the adjacent amino acid in SIN nsP1, Thr349Lys, is able to suppress the effect of a normally nonviable change in nsP4

of the N-terminal Tyr to Ala, Arg, or Leu. As an alteration of the nsP4 amino terminus appears to also cause a defect in minus-strand synthesis, it can be hypothesized that an interaction of these regions of nsP1 and nsP4 is required for minus-strand synthesis or promoter recognition (*118*) (Fig. 3A).

VII. nsP2: A Multifunctional Enzyme and Regulatory Protein

NsP2 is the largest replicase protein consisting of 794–807 aa residues among different alphaviruses. Several SIN RNA-negative ts mutants of group A have been mapped to various parts of nsP2 (*46, 59, 119*). They indicate that nsP2 functions in both the regulation of subgenomic RNA synthesis and the regulation of shut-off of minus-strand synthesis (see Section IV). According to sequence comparisons, the N-terminal half of nsP2 has sequence motifs typical for RNA and DNA helicases, whereas the C-terminal part resembles papain-like cysteine proteinases (Fig. 8) (*88, 120, 121*).

A. Nucleosidetriphosphatase and RNA Helicase Activities

Direct evidence of nucleosidetriphosphatase (NTPase) activity has been demonstrated for purified nsP2 preparations expressed in *E. coli* (*122*). The activity is associated also with the 470 N-terminal amino acids of the protein. NTPase activity is stimulated by poly(A), poly(U), oligo(A), and tRNA. Mutation of the putative NTP-binding site GVPGSG**K**$_{192}$S to GVPGSG**N**$_{192}$S inhibits the NTPase activity. Both nsP2 and nsP2-N bind to single-stranded RNA and can be cross-linked to oxidized ATP, GTP, CTP, and UTP. When the GNS mutation was introduced to the SFV genome, no virus replication could be detected. After longer incubation, a revertant virus appeared, in which GNS had been back-mutated to GKS (*122, 123*).

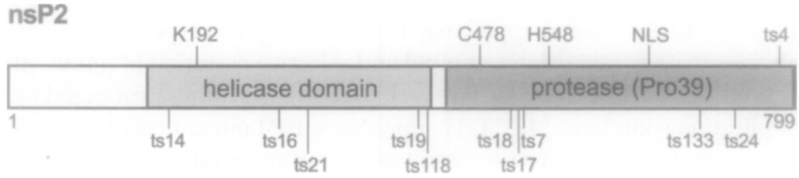

FIG. 8. Scheme of nsP2 showing the conserved helicase domain as well as the protease domain (Pro39). Some of the active site residues are indicated together with the numerous ts mutations mapped to nsP2. All mutants, except SFV ts4 are SIN mutants. NLS is the nuclear localization signal of nsP2.

NsP2 is also an RNA helicase, which can unwind partially double-stranded RNA preparations in the presence of NTPs and dNTPs (*124*). The reaction is inhibited by the GNS mutation suggesting that NTPase activity is necessary to drive the unwinding (*125*). The helicase activity is inhibited in NaCl concentrations higher than 100 m*M*. So far, helicase activity has not been demonstrated for the nsP2-N fragment.

After demonstration of RNA helicase activity for nsP2, it could be presumed that all homologous proteins encoded by members of the alphavirus-like superfamily have helicase activity. The role of helicase in the replication cycle of positive-strand RNA viruses remains mysterious, although it is often thought to either separate double-stranded RNAs or remove secondary structures from template RNAs. A ts mutant within the helicase domain of BMV 1a encodes a protein defective in all forms of RNA synthesis, consistent with this line of thinking (*126*). On the other hand, the phenotype of SIN RNA-negative ts mutants located in the helicase is not very clear-cut but has been interpreted to support a role in the conversion from negative-strand to positive-strand synthesis (*119*). Finally, mutations in the helicase motifs of BMV 1a completely prevent minus-strand synthesis but also disrupt severely the initial recruitment of viral RNAs to the replication complex (*105*). Thus, the helicase proteins are likely to have multiple functions at different stages of the replication cycle, which need to be untangled in further experiments.

B. RNA Triphosphatase Activity

As described above, the methyltransferase and guanylyltransferase activities needed in the capping of alphavirus RNAs are carried out by nsP1 (see Section VI,A). However, nsP1 lacks RNA triphosphatase activity essential for mRNA capping (*97, 98*). Using short γ-^{32}P-labeled RNA molecules for screening, nsP2 has been identified as an RNA triphosphatase for both SFV and SIN. This activity, which removes specifically only the outmost γ-phosphate, is also associated with the nsP2-N fragment, previously shown to have NTPase activity (*122, 127*). Unlike the helicase activity, the RNA triphosphatase activity tolerates relatively high-salt concentrations. There is an absolute requirement for divalent cations, indicating that the alphavirus enzyme belongs to the metal-dependent RNA triphosphatases (*98*). The GNS mutation in the NTP-binding site inactivated the triphosphatase activity, suggesting that NTPase and RNA triphosphatase reactions take place in the same reaction center. If so, NTPs should competitively inhibit the RNA triphosphatase reaction, as shown for, e.g., vaccinia virus-capping enzyme (*128, 129*). However, this is not the case, as addition of GTP rather enhanced the triphosphatase activity, suggesting that hydrolysis of γ-phosphate of NTPs and that of RNA take place in different reaction centers. This mode has been suggested for flavivirus NS3 protein, which is a 68-kDa multifunctional protein with NTPase/RNA triphosphatase/protease/RNA

helicase activities (*130*). It must be remembered that triphosphatase activity is needed only once during the synthesis of an RNA molecule, while RNA helicase/NTPase activities may be required throughout the synthesis.

C. Protease Activity

Both mapping of active site residues (*90*) and assaying the sensitivity to inhibitors (*131*) strongly support the hypothesis that the C-terminal part of nsP2 is a papain-like cysteine protease. To better understand the complex enzymatic processing of the nonstructural polyprotein (see Section V), nsP2 and a soluble C-terminal fragment, Pro39, consisting of amino acid residues 459–799 of SFV nsP2, was recently expressed and purified (*131*). *In vitro* synthesized, labeled, proteolytically inactive polyproteins $P12^{CA}$, $P2^{CA}3$, $P12^{CA}$, P34, $P12^{CA}3$, $P12^{CA}34$, with the Cys478Ala mutation in the predicted active site of the enzyme, were used as substrates for Pro39 and nsP2. Cleavage was observed at all sites but with different efficiences. Site 3/4 (between nsP3 and nsP4) was cleaved most readily, site 1/2 next, and site 2/3 poorly (Fig. 1).

In order to eliminate possible shielding effects within the polyprotein, short cleavage site sequences were joined to a thioredoxin carrier. By this means, it was possible to study the proteolysis *in vitro* with purified reagents. The cleavage products were isolated, and their masses and N-terminal sequences were determined. These results show that (i) the cleavage takes place at the same sites as those determined in virus-infected cells (*132, 133*); (ii) the cleavage at site 3/4 is most efficient; (iii) about 5000-fold more enzyme is required for complete cleavage at site 1/2, whereas cleavage at site 2/3 remains poor even with huge excess of Pro39-enzyme (*131*). These results are compatible with the previous observations that processing at site 3/4 takes readily place *in trans*. The poor cleavability of site 1/2 cannot be due to inaccessibility for the enzyme in this *in vitro* system. Rather, it suggests that some cofactor(s) is needed. This hypothetical "factor" is evidently present in polyproteins P12, P123, and P1234, in which the cleavage at site 1/2 takes place *in cis*. The very inefficient cleavage at site 2/3 is surprising, since P23 is cleaved rapidly after expression in insect cells and after *in vitro* translation (92; Merits *et al.*, unpublished). In the polyprotein precursors P1234 and P123 the cleavage of 2/3 site must be preceded by cleavage at the site 1/2 yielding P23 or P234 (*89, 92*). Thus, cleavage at site 2/3 must also require a "cofactor," which is probably different from that needed in the cleavage at site 1/2. The specificity of the cleavage at all three sites is determined solely by the protease moiety of nsP2, but the efficiency of the cleavage requires unidentified cofactors.

D. Nuclear Transport and Neuropathogenicity

Immunofluorescence microscopy has shown that of the nsPs, only nsP2 is transported to the nuclei and nucleoli of SFV-infected cells (*134*). Cell fractionation and immunoprecipitation experiments showed that about 25% of labeled nsP2 was associated with cytoplasmic membranes sedimenting at 15,000*g*, 25%

remained in the supernatant fraction, and about 50% in the nuclear fraction. The transport took place already early in infection, and nsP2 was detectable in the nuclear matrix fraction within 5 min in pulse–chase experiments, reaching 50% in 20 min. NsP2 had to be cleaved from its precursors (P12 and P123) before transport took place. When nsP2 was expressed alone in BHK21 cells using an SV40-based vector, almost all of the protein was transported to the nucleus.

A deletion analysis was carried out to find the putative nuclear localization signal (NLS) of nsP2. We identified a pentapeptide region $PR_{648}RRV$ responsible for the nuclear transport of nsP2 (Fig. 8). Mutations changing arginines 648–649 to aspartic acids (RDR, DRR, and DDR) resulted in cytoplasmic nsP2. A fusion protein, consisting of β-galactosidase and 232 C-terminal amino acid residues of nsP2, was transported to the nucleus only if it had an intact PRRRVsequence. Further sequences, especially within region 470–490, were required for nucleolar targeting (*135*). The NLS of nsP2 was also functional in the yeast *Saccharomyces cerevisiae,* since expression of SFV P1234 resulted in the accumulation of nsP2 in the nucleus (*136*).

When the RDR mutation was introduced to infectious cDNA, transfection of cells with the 42S transcript resulted in virus production (SFV-RDR), showing that functional NLS of nsP2 is not vital for SFV replication. A closer analysis of the synthesis of nsP2 and its precursors showed that the synthesis of nsPs was delayed early in infection. Later in infection, there was a considerable delay in the processing of the polyproteins (P1234, P123, and P12). There was also a clear delay in the virus release from the SFV-RDR mutant-infected cells. However, the final yields were the same as those with the wild type. Normally, SFV inhibits cellular protein and RNA and DNA syntheses efficiently (*137*). There is a clear difference in the inhibition of host cell DNA synthesis at 6 h after infection between wild type (13% of uninfected control cells) and SFV-RDR (about 50%). No significant difference in the inhibition of host protein synthesis was found between the wild type and the mutant. Most interestingly, SFV-RDR has lost its pathogencity for adult mouse injected intraperitoneally (*123*).

Both wild-type and SFV-RDR are neuroinvasive after intraperitoneal inoculation. However, wild-type SFV spreads rapidly, infecting cells throughout the brain, while SFV-DRD infection is confined to small foci of cells. When mice deficient in type I interferon responses were infected with SFV-RDR, the virus was distributed widely in the brain causing death (*137a*). These results stress the importance of the interferon system as a host defence mechanism. They also show that a component of RNA polymerase complex can be vital for the neuropathogenicity of SFV. The slower processing of P1234 strongly suggests that the RDR mutation affects the protease moiety of nsP2, causing a delay in early replication events. This delay may be vital for the host defense system to be able to limit the infection to the primary foci of infected cells. Thus, the nuclear localization of nsP2 may reflect the structural requirements of the multifunctional nsP2 protein rather than a separate nuclear function.

Mutation of the nsP1 palmitoylation site also results in a viable but apathogenic SFV mutant, which causes viremia in mice, but no infectious virus is detectable in the brain (see Section VI,B). In HeLa and BHK cells, there is a delay in virus replication during infection (*113*). Furthermore, deletion of 50 amino acid residues from the variable region of SFV nsP3 results in viable but apathogenic virus mutant (see Section VIII,B). There is a clear reduction in the RNA synthesis rate between 2.5 and 6 h after infection, which results in a delay of virus release from the cells, although final virus yields are same as those with the wild type (*138*). One simple possibility is that, with all these mutants, the delay in virus replication is enough for the host to successfully limit the infection of brain cells, which usually causes the death of the animals in wild-type SFV infection. However, further studies may reveal that each of these proteins has specific functions in the pathogenesis of alphaviruses.

VIII. nsP3: An Ancient Conserved Protein and Phosphoprotein

Alphavirus nsP3 is an intriguing hybrid protein. Its N-terminal domain of unknown function exhibits a high degree of similarity to some cellular proteins, whereas its carboxy terminus is hypervariable in length and sequence even among alphaviruses. Approximately 320 N-terminal aa of nsP3 show conservation between different alphaviruses. These are followed by a "tail" region of 150–250 aa, rich in acidic residues serine and threonine and devoid of predicted secondary structure (Fig. 9A). A main line of investigation by us and others has been the characterization of nsP3 "tail" phosphorylation.

A. Sequence Conservation

The most N-terminal 160 aa of alphavirus nsP3 form a small domain, which shows an unusual pattern of conservation. Related sequences are found not only in the genomes of rubella virus and hepatitis E virus but also in otherwise unrelated coronaviruses (but not in alphavirus-like plant viruses or coronavirus-like arteriviruses) (*120*). Recently, it has become apparent that this domain is widely, although not universally, distributed in bacteria, archae, and eukaryotes (*139*). It usually exists on its own as a small protein, but the human genome also contains open reading frames in which this domain is repeated. A gene termed *BAL*, which contains a duplicated nsP3-like domain, is a highly expressed risk factor in certain aggressive lymphomas (*140*). The conservation between viral and cellular homologs is unusually high for an RNA virus protein, reaching up to 35–40% sequence identity (*139*). Furthermore, an unusual class of histone variants found in animal cells predominantly associated with the inactivated X chromosome, the macrohistones H2A, contains a more distantly related nsP3-like domain following the core histone domain (*139*). The biochemical function

of the nsP3 domain remains unknown, but the high conservation points to a basic and essential role in catalysis or binding. A possible clue is provided by the yeast gene YBR022w, the product of which was shown to hydrolyze ADP-ribose-1″phosphate (*141*). However, the yeast homolog is much less related to human and bacterial proteins than the latter are to each other or to nsP3, pointing to a possible divergence of function. Furthermore, a more general role for such an unusual nucleotide compound may be difficult to envision. Nevertheless, nsP3-like proteins could potentially be phosphoesterases hydrolyzing an as yet unidentified substrate.

B. Phosphorylation

NsP3 is phosphorylated on threonine and more heavily on serine, but not on tyrosine (*142, 143*). Deletion mapping suggested that the phosphorylation sites of SIN and SFV nsP3 are mainly located in the variable C-terminal region (*138, 144*). This has been proven in the case of SFV nsP3, where the phosphorylation sites have been mapped by mass spectrometric methods (*145*), supported by mutation analysis (*138*). The phosphorylation pattern is complex, as a total of 16 possible sites (Ser320, 327, 332, 335, 347, 348, 349, 352, 356, 359, 362, 367, and Thr344, 345, 350, 354) can become phosphorylated. Differentially phosphorylated forms of the protein exist, since forms containing from 7 up to the 12 possible phosphates in the region Thr344–Ser367 have been detected. Based on point mutational analysis, Thr344/Thr345 appear to be the major threonine phosphorylation sites. Thus, phosphorylation is confined to a small subregion of the tail (Fig. 9A), the deletion of which abolishes phosphorylation both in the cells expressing nsP3 alone and in the context of virus infection (*138, 145*).

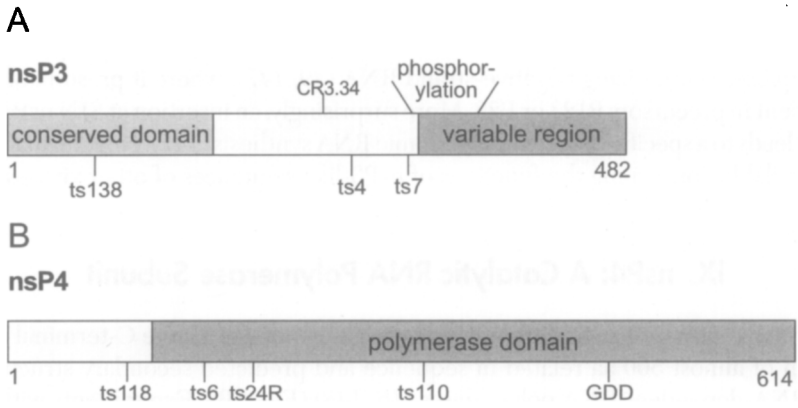

FIG. 9. Schemes of nsP3 (A) and nsP4 (B). Domains and mutations discussed in the text are indicated. CR3.34 is the linker insertion mutation after aa 252 of SIN nsP3, and GDD denotes the most conserved polymerase motif in nsP4.

The kinases responsible for phosphorylation of nsP3 remain unknown. Although a kinase resembling casein kinase II in its biochemical properties can phosphorylate SIN nsP3 (*143*), the presence of multiple sites with dissimilar sequence contexts suggests phosphorylation by multiple cellular kinases (*138*). Similarly, the role of phosphorylation in RNA replication is not clear. Although phosphorylation is not essential, as deletions reducing or abolishing phosphorylation are tolerated, the deletions have biological effects. Deletions in SIN nsP3 tail reduce plaque formation in mosquito cells (*144*), and deletions in SFV nsP3 tail reduce the level of RNA synthesis in BHK cells, although virus growth remains unchanged (*138*). Most interestingly, a SFV variant devoid of nsP3 phosphorylation also exhibits greatly reduced mouse pathogenicity (*138*). Another study has stressed the importance of the C-terminal variable region of nsP3 for neurovirulence of SFV in mice (*146*). Phosphorylation and the entire variable tail of nsP3, which has been subject to rapid alteration during alphavirus evolution, may be involved in the optimization and fine-tuning of replication in diverse host cell types.

C. Other Features

NsP3 has a weak peripheral affinity for membranes (*138*) and it may contribute to the membrane association of the replication complex. When expressed alone in mammalian cells, nsP3 associates with unidentified cytoplasmic vesicles, which do not contain endosomal or lysosomal markers (*65*). Membrane association is mediated by the region of nsP3 conserved within alphaviruses and it does not require phosphorylation (*138*). However, in SFV-infected cells, nsP3 found in the membrane-associated replication complex fraction is phosphorylated to a greater extent than soluble nsP3 (*142*).

Genetic experiments have also failed to define a precise role for nsP3, although it is essential for RNA synthesis (*46*). Both classical and linker insertion ts mutants indicate that nsP3 functions in the formation of early replication complexes synthesizing negative-strand RNA (*49, 147*), where it presumably is present in precursors P123 or P23. More surprisingly, an insertion at SIN nsP3 aa 252 leads to a specific defect in subgenomic RNA synthesis (*147*). No information is available concerning the functions of nsP3-like sequences of other viruses.

IX. nsP4: A Catalytic RNA Polymerase Subunit

NsP4 (607–614 aa in different alphaviruses) contains a large C-terminal domain of almost 500 aa related in sequence and predicted secondary structure to RNA-dependent RNA polymerases (*16, 148*) (Fig. 9B). Experiments with ts mutants strongly support the contention that nsP4 is the catalytic polymerase subunit. SIN mutations Gly153Glu (ts6) and Gly324Glu (ts110) map to nsP4 (*149*). Ts6 causes a rapid cessation of all RNA synthesis when cells are shifted to

nonpermissive temperature (*47, 76*), and ts110 has a similar but milder phenotype (*149*). The ts6 defect is reproduced in polymerase extracts prepared from infected cells, which enabled an elegant demonstration that the defect affects elongation of RNA chains first initiated at the permissive temperature (*71*). SIN mutations causing resistance to pyrazofurin, a compound lowering cellular UTP and CTP levels, also map to nsP4 and may increase the affinity of the polymerase for pyrimidine nucleotides (*150*). Finally, a mutation Gln191 Lys in SIN nsP4 causes a reactivation of minus-strand synthesis at the restrictive temperature, a phenotype similar to many nsP2 mutants, indicating that nsP4 is also involved in the regulation of minus-strand synthesis shut-off (57) (see Section IV,B).

The small N-terminal extension of nsP4 has no conserved counterparts in other viruses and it has no known enzymatic activity (*92*). It may be involved in interactions with other components of the replication complex. A double-mutation Val425Ala in nsP2 and Gln93Arg in nsP4 is required for the phenotype of SIN ts118, which might among other possibilities suggest that the helicase domain of nsP2 interacts with the polymerase to form a functional complex (*149*). The N-terminal amino acid of mature nsP4 is normally a tyrosine, and only an aromatic amino acid or histidine is functional at the N terminus of SIN nsP4 (*51, 151*). Alteration of the nsP4 N terminus to normally nonviable residues appears to cause a defect in minus-strand RNA synthesis, which can be suppressed by a mutation in nsP1 (see Section VI,C). This is interpreted to support a role for nsP1–nsP4 interaction in minus-strand synthesis or promoter interaction (*118*).

The bulk of nsP4 is unstable in alphavirus-infected cells, although a stable fraction presumably associated with the polymerase complex also exists (*85, 86, 152, 153*). NsP4 is degraded by the N-end rule pathway, as it bears a destabilizing N-terminal Tyr residue (*154*), and it is accordingly stabilized by inhibitors of proteosome activity (*92*). The amount of nsP4 can thus be regulated at multiple levels. Translation of nsP4 requires a readthrough of a termination codon in many alphaviruses, production of active polymerase complexes requires cleavage of nsP4 from the nonstructural polyprotein, and nsP4 is susceptible for degradation by the proteosome system.

X. The Replication Complex

After alphavirus entry and uncoating, which may take place in the close vicinity of endosomes, the genome RNA is subjected to translation. As a single virus particle can cause infection (*155*), we have to think that one RNA molecule is first translated, followed by its recruitment as a template for minus-strand synthesis. These primary events cannot be detected by available techniques. Experiments described in Section IV,A indicate that P1234, cleaved at site 3/4, yields the minus-strand RNA polymerase, P123 plus nsP4, although it is not

known whether one heterodimer constitutes an active complex. Our recent results with expression of uncleavable polyproteins P1234 and P123 in insect and mammalian cells have shown that they both bind to intracellular membranes and become palmitoylated (Salonen *et al.*, unpublished). Thus, we can expect that the minus-strand RNA polymerase is membrane bound, probably by a mechanism similar to that demonstrated for nsP1 (see Section VI,B). We assume that several rounds of translation take place, resulting in a small pool of membrane-bound polyprotein.

A polymerase complex recognizes the conserved sequences at the 5′ and 3′ ends of the parental 42S RNA genome, possibly through joint activity of the nsP1 portion of P123, and nsP4 (Fig. 3A). A complete uncapped complementary RNA is synthesized, which has an unpaired extra G residue at the 3′ end of the minus strand, the origin of which is not known (*156, 157*). A similar extra G residue at the 3′end of cucumber mosaic satellite virus minus strand RNA is essential for RNA replication (*158*). As there is no evidence of the existence of free 42S RNA minus-strands, a double-stranded RNA may be synthesized (*23, 24, 45, 47, 55, 159*). The polymerase may utilize the plus-strand template for only one round of minus-strand synthesis. This would explain why only about 20% of synthesized RNA is of negative polarity early in infection (*45*). The occurrence of homologous recombination and the ease of creation of defective-interfering RNAs indicate that template switching can take place, presumably during minus-strand synthesis (*37, 160*).

The shift from minus-strand to plus-strand RNA synthesis requires the proteolytic processing of P123 (see Section IV,B). The 42S RNA minus strand acts as a template, but whether it is part of either double- or single-stranded RNA is not definitively known. It can be deduced that the synthesis of the plus strand starts from the penultimate base. The presence of replicative intermediate RNA with nascent plus strands has suggested a model, according to which a single minus strand is copied by several polymerase molecules (*23, 24*) (Fig. 3C). According to this model, the double-stranded replicative forms RFI, II, and III arise during RNA isolation by hybridization of the excess plus strands to the minus-strand template. The same replication complex can alternate between synthesis of 42S and 26S RNA, regulated by an extra "soluble" nsP2 transcription factor (*59, 79*) (Fig. 3D and 3E). An important feature of the late replication complex is its stability, as it continues to make both 42S RNA and 26S RNA for several hours using the same minus strand as template.

We have recently shown that all four nsPs as well as nascent RNA are associated with small invaginations, or spherules, of alphavirus-specific large cytoplasmic vacuoles, CPVs (Fig. 4) (*66*). CPVs are modified endosomes and lysosomes (*64–66, 112*). We have suggested that each spherule represents one unit of replication, carrying a 42S RNA minus-strand template as well as a set of late RNA polymerases, each consisting of nsP1–nsP4 (*66*). This hypothesis is supported

by electron microscopic studies. Autoradiography showed a clear association of nascent virus-specific RNA with spherules (*63*), while on the other hand protrusion of RNA-like structures from the spherules was described (*64*).

The membrane association of the replication complex has interesting consequences. Since the polymerase is fixed to the membrane, the template RNA must move. If a single spherule constitutes a unit of replication, the template RNA of about 11.5 kb must be packed within the spherule of a diameter of about 50 nm. If the template were double-stranded RNA, the packing would be even tighter. In order to achieve continuous synthesis of plus strands, the template should preferably be circular. The classical model of RI, with one template occupied by several polymerases (Fig. 3C), is difficult to reconcile with polymerases that are fixed to the membrane. The problem could be solved, assuming that the same template is successively passing from one membrane-associated polymerase to another in a circular manner (Fig. 10). As both 42S RNA and 26S RNA have a cap structure at their 5′ end, the coordinated capping reactions catalyzed by

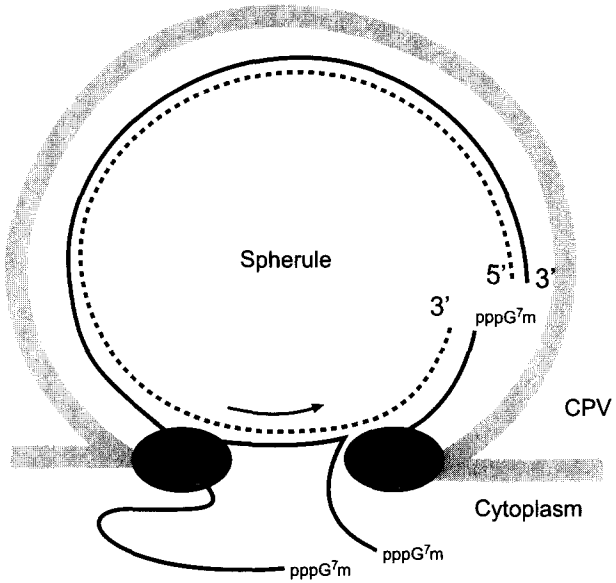

FIG. 10. A speculative model of the alphavirus spherule-associated unit of replication synthesizing positive 42S RNA strands. The minus-strand template (dotted line) is shown in association with a complementary plus strand (solid line). The late RNA polymerase molecules (pol), consisting of nsP1–nsP4, localized to the orfice of the spherule, are fixed to the membrane. The dsRNA template rotates through two different polymerases (in this image) in the direction shown by the long arrow. Nascent, capped, positive RNA strands, arise by semiconservative replication. Upon reaching the cytoplasm they are utilized for translation early in infection and later for the assembly of nucleocapsids (*64*).

nsP1 and nsP2 (Fig. 6) must presumably take place soon after initiation. Anionic membrane lipids are a vital cofactor for nsP1 at this stage. So far, there is no direct experimental data concerning the role of the RNA helicase activity of nsP2.

The presence of spherules also on the outer surface of the plasma membrane, in addition to CPVs (*63, 64, 66*), has led us to propose the following hypothesis. The primary assembly of the replication complex takes place in association with endosomes, or the plasma membrane, through the affinity of the lipid-binding peptide of nsP1 to phosphatidyserine-rich lipids (Fig. 7). Only a fraction of the replication complexes assemble correctly and manage to recruit a template. After proteolytic cleavages, the components of unsuccessful complexes dissociate from each other, and nsP1 associates permanently with the plasma membrane. NsP2 is transported to the nucleus, and nsP3 and nsP4 remain distributed in the cytoplasm. The correctly assembled complexes participate in the endosomal circulation, redistributing the complexes to new endosomes and finally to lysosomes (*66*).

We have aimed at emphasizing throughout this review how genetic, virological, biochemical, and cell biological approaches have complemented each other in providing a better understanding of alphavirus RNA replication and the replication complex itself. Even though progress has been made, our knowledge of the molecular mechanisms of RNA replication is far from complete. The three-dimensional structures of the nsPs are required to further define their roles in RNA replication. An even more demanding task is to isolate, or reconstitute, the unit of replication in functional form.

ACKNOWLEDGMENTS

We acknowledge the financial support we have received for the studies described from the Academy of Finland (grants no 8397, 62496, 78763) the Finnish Technology Development Center (Tekes), the Sigrid Juselius Foundation, the Helsinki University Foundation, and the Centre for International Mobility. We thank Dr. Marja Makarow for critical reading of the manuscript. LK has been a Biocentrum Helsinki fellow from 1995–2000.

REFERENCES

1. R. E. Johnston and C. J. Peters, Alphaviruses, *in* "Fields Virology," 3rd ed. (B. N. Fields, D. M. Knipe, and P. M. Howley, eds.), pp. 843–898. Lippincott-Raven, Philadelphia, 1996.
2. L. Kääriäinen and H. Söderlund, Structure and replication of alphaviruses. *Curr. Top. Microbiol. Immunol.* **82,** 15–69 (1978).
3. A. Helenius, Alphavirus and flavivirus glycoproteins: Structures and functions. *Cell* **81,** 651–653 (1995).

4. J. Saraste and E. Kuismanen, Pre- and post-Golgi vacuoles operate in the transport of Semliki Forest virus membrane glycoproteins to the cell surface. *Cell* **38,** 535–543 (1984).
5. P. Berglund, I. Tubulekas, and P. Liljeström, Alphaviruses as vectors for gene delivery. *Trends Biotechnol.* **14,** 130–134 (1996).
6. S. Schlesinger and T. W. Dubensky, Alphavirus vectors for gene expression and vaccines. *Curr. Opin. Biotechnol.* **10,** 434–439 (1999).
7. E. J. Mancini, M. Clarke, B. E. Gowen, T. Rutten, and S. D. Fuller, Cryo-electron microscopy reveals the functional organization of an enveloped virus, Semliki Forest virus. *Mol. Cell* **5,** 255–266 (2000).
8. M. Marsh and A. Helenius, Virus entry into animal cells. *Adv. Virus Res.* **36,** 107–151 (1989).
9. I. Singh and A. Helenius, Role of ribosomes in Semliki Forest virus nucleocapsid uncoating. *J. Virol.* **66,** 7049–7058 (1992).
10. I. Ulmanen, H. Söderlund, and L. Kääriäinen, Role of protein synthesis in the assembly of Semliki forest virus nucleocapsid. *Virology* **99,** 265–276 (1979).
11. H. Söderlund, L. Kääriäinen, and C. H. von Bonsdorff, Properties of Semliki Forest virus nucleocapsid. *Med. Biol.* **53,** 412–417 (1975).
12. I. Ulmanen, Assembly of Semliki Forest virus nucleocapsid: Detection of a precursor in infected cells. *J. Gen. Virol.* **41,** 353–365 (1978).
13. K. M. Coombs and D. T. Brown, Form-determining functions in Sindbis virus nucleocapsids: Nucleosomelike organization of the nucleocapsid. *J. Virol.* **63,** 883–891 (1989).
14. M. Pesonen and O. Renkonen, Serum glycoprotein-type sequence of monosaccharides in membrane glycoproteins of Semliki Forest virus. *Biochim. Biophys. Acta* **455,** 510–525 (1976).
15. M. Pesonen, Sequence analysis of lactosamine type glycans of individual membrane proteins of Semliki Forest virus. *J. Gen. Virol.* **45,** 479–487 (1979).
16. E. V. Koonin and V. V. Dolja, Evolution and taxonomy of positive-strand RNA viruses: Implications of comparative analysis of amino acid sequences. *Crit. Rev. Biochem. Mol. Biol.* **28,** 375–430 (1993).
17. M. N. Rozanov, E. V. Koonin, and A. E. Gorbalenya, Conservation of the putative methyltransferase domain: a hallmark of the "Sindbis-like" supergroup of positive-strand RNA viruses. *J. Gen. Virol.* **73,** 2129–2134 (1992).
18. T. Ahola, P. Laakkonen, H. Vihinen, and L. Kääriäinen, Critical residues of Semliki Forest virus RNA capping enzyme involved in methyltransferase and guanylyltransferase-like activities. *J. Virol.* **71,** 392–397 (1997).
19. K. W. Buck, Comparison of the replication of positive-stranded RNA viruses of plants and animals. *Adv. Virus Res.* **47,** 159–251 (1996).
20. R. M. Friedman, H. B. Levy, and W. B. Carter, Replication of Semliki Forest virus: Three forms of viral RNA produced during infection. *Proc. Natl. Acad. Sci. U.S.A.* **56,** 440–446 (1966).
21. J. A. Sonnabend, E. M. Martin, and E. Mecs, Viral specific RNAs in infected cells. *Nature* **213,** 365–367 (1967).
22. L. Kääriäinen and P. J. Gomatos, A kinetic analysis of the synthesis in BHK 21 cells of RNAs specific for Semliki Forest virus. *J. Gen. Virol.* **5,** 251–265 (1969).
23. D. T. Simmons and J. H. Strauss, Replication of Sindbis virus: I. Relative size and genetic content of 26S and 49S RNA. *J. Mol. Biol.* **71,** 599–613 (1972).
24. D. T. Simmons and J. H. Strauss, Replication of Sindbis virus: II. Multiple forms of double-stranded RNA isolated from infected cells. *J. Mol. Biol.* **71,** 615–631 (1972).
25. R. M. Friedman and I. K. Berezesky, Cytoplasmic fractions associated with Semliki Forest virus ribonucleic acid replication. *J. Virol.* **1,** 374–383 (1967).
26. M. Ranki and L. Kääriäinen, Solubilized RNA replication complex from Semliki Forest virus-infected cells. *Virology* **98,** 298–307 (1979).

27. P. J. Gomatos, L. Kääriäinen, S. Keränen, M. Ranki, and D. L. Sawicki, Semliki Forest virus replication complex capable of synthesizing 42S and 26S nascent RNA chains. *J. Gen. Virol.* **49,** 61–69 (1980).
28. H. Garoff, K. Simons, and B. Dobberstein, Assembly of the Semliki Forest virus membrane glycoproteins in the membrane of the endoplasmic reticulum in vitro. *J. Mol. Biol.* **124,** 587–600 (1978).
29. N. Glanville, M. Ranki, J. Morser, L. Kääriäinen, and A. E. Smith, Initiation of translation directed by 42S and 26S RNAs from Semliki Forest virus *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* **73,** 3059–3063 (1976).
30. N. Glanville, B.-E. Lachmi, A. E. Smith, and L. Kääriäinen, Tryptic peptide mapping of the nonstructural proteins of Semliki Forest virus and their precursors. *Biochim. Biophys. Acta* **518,** 497–506 (1978).
31. P. Lehtovaara, I. Ulmanen, L. Kääriäinen, S. Keränen, and L. Philipson, Synthesis and processing of Semliki Forest virus-specific nonstructural proteins *in vivo* and *in vitro*. *Eur. J. Biochem.* **112,** 461–468 (1980).
32. K. Takkinen, Complete nucleotide sequence of the nonstructural protein genes of Semliki Forest virus. *Nucleic Acids Res.* **14,** 5667–5682 (1986).
33. E. G. Strauss, C. M. Rice, and J. H. Strauss, Complete nucleotide sequence of the genomic RNA of Sindbis virus. *Virology* **133,** 92–110 (1984).
34. J. H. Strauss and E. G. Strauss, The alphaviruses: Gene expression, replication, and evolution. *Microbiol. Rev.* **58,** 491–562 (1994).
35. E. Frolova, I. Frolov, and S. Schlesinger, Packaging signals in alphaviruses. *J. Virol.* **71,** 248–258 (1997).
36. C. L. White, M. Thomson, and N. J. Dimmock, Deletion analysis of a defective interfering Semliki Forest virus RNA genome defines a region in the nsP2 sequence that is required for efficient packaging of the genome into virus particles. *J. Virol.* **72,** 4320–4326 (1998).
37. P. Lehtovaara, H. Söderlund, S. Keränen, R. F. Pettersson, and L. Kääriäinen, 18S defective interfering RNA of Semliki Forest virus contains a triplicated linear repeat. *Proc. Natl. Acad. Sci. U.S.A.* **78,** 5353–5357 (1981).
38. P. Lehtovaara, H. Söderlund, S. Keränen, R. F. Pettersson, and L. Kääriäinen, Extreme ends of the genome are conserved and rearranged in the defective interfering RNAs of Semliki Forest virus. *J. Mol. Biol.* **156,** 731–748 (1982).
39. S. S. Monroe and S. Schlesinger, Common and distinct regions of defective-interfering RNAs of Sindbis virus. *J. Virol.* **49,** 865–872 (1984).
40. M. Thomson and N. J. Dimmock, Common sequence elements in structurally unrelated genomes of defective interfering Semliki Forest virus. *Virology* **199,** 354–365 (1994).
41. R. F. Pettersson, 5′-Terminal nucleotide sequence of Semliki forest virus 18S defective interfering RNA is heterogeneous and different from the genomic 42S RNA. *Proc. Natl. Acad. Sci. U.S.A.* **78,** 115–119 (1981).
42. S. S. Monroe and S. Schlesinger, RNAs from two independently isolated defective interfering particles of Sindbis virus contain a cellular tRNA sequence at their 5′ ends. *Proc. Natl. Acad. Sci. U.S.A.* **80,** 3279–3283 (1983).
43. H. G. M. Niesters and J. H. Strauss, Defined mutations in the 5′ nontranslated sequence of Sindbis virus RNA. *J. Virol.* **64,** 4162–4168 (1990).
44. H. G. M. Niesters and J. H. Strauss, Mutagenesis of the conserved 51-nucleotide region of Sindbis virus. *J. Virol.* **64,** 1639–1647 (1990).
45. D. L. Sawicki and S. G. Sawicki, Short-lived minus-strand polymerase for Semliki Forest virus. *J. Virol.* **34,** 108–118 (1980).
46. Y. S. Hahn, E. G. Strauss, and J. H. Strauss, Mapping of RNA$^-$ temperature-sensitive mutants of Sindbis virus: Assignment of complementation groups A, B, and G to nonstructural proteins. *J. Virol.* **63,** 3142–3150 (1989).

47. D. L. Sawicki, S. G. Sawicki, S. Keränen, and L. Kääriäinen, Specific Sindbis virus-coded function for minus-strand RNA synthesis. *J. Virol.* **39,** 348–358 (1981).
48. Y.-F. Wang, S. G. Sawicki, and D. L. Sawicki, Sindbis virus nsP1 functions in negative-strand RNA synthesis. *J. Virol.* **65,** 985–988 (1991).
49. Y.-F. Wang, S. G. Sawicki, and D. L. Sawicki, Alphavirus nsP3 functions to form replication complexes transcribing negative-strand RNA. *J. Virol.* **68,** 6466–6475 (1994).
50. J. A. Lemm and C. M. Rice, Assembly of functional Sindbis virus RNA replication complexes: Requirement for coexpression of P123 and P34. *J. Virol.* **67,** 1905–1915 (1993).
51. J. A. Lemm and C. M. Rice, Roles of nonstructural polyproteins and cleavage products in regulating Sindbis virus RNA replication and transcription. *J. Virol.* **67,** 1916–1926 (1993).
52. J. A. Lemm, T. Rümenapf, E. G. Strauss, J. H. Strauss, and C. M. Rice, Polypeptide requirements for assembly of functional Sindbis virus replication complexes: A model for the temporal regulation of minus- and plus-strand RNA synthesis. *EMBO J.* **13,** 2925–2934 (1994).
53. Y. Shirako and J. H. Strauss, Regulation of Sindbis virus RNA replication: Uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. *J. Virol.* **68,** 1874–1885 (1994).
54. J. A. Lemm, A. Bergqvist, C. M. Read, and C. M. Rice, Template-dependent initiation of Sindbis virus RNA replication *in vitro*. *J. Virol.* **72,** 6546–6553 (1998).
55. S. G. Sawicki, D. L. Sawicki, L. Kääriäinen, and S. Keränen, A Sindbis virus mutant temperature-sensitive in the regulation of minus-strand RNA synthesis. *Virology* **115,** 161–172 (1981).
56. S. G. Sawicki and D. L. Sawicki, The effect of loss regulation of minus-strand RNA synthesis on Sindbis virus replication. *Virology* **151,** 339–349 (1986).
57. D. L. Sawicki, D. B. Barkhimer, S. G. Sawicki, C. M. Rice, and S. Schlesinger, Temperature sensitive shut-off of alphavirus minus strand RNA synthesis maps to a nonstructural protein, nsP4. *Virology* **174,** 43–52 (1990).
58. D. L. Sawicki and S. G. Sawicki, A second nonstructural protein functions in the regulation of alphavirus negative-strand RNA synthesis. *J. Virol.* **67,** 3605–3610 (1993).
59. J. Suopanki, D. L. Sawicki, S. G. Sawicki, and L. Kääriäinen, Regulation of alphavirus 26S mRNA transcription by replicase component nsP2. *J. Gen. Virol.* **79,** 309–319 (1998).
60. K. Tuomi, L. Kääriäinen, and H. Söderlund, Quantitation of Semliki Forest virus RNAs in infected cells using ^{32}P equilibrium labelling. *Nucleic Acids Res.* **2,** 555–565 (1975).
61. G. Wengler and Gi. Wengler, Studies on the synthesis of viral RNA-polymerase-template complexes in BHK 21 cells infected with Semliki Forest virus. *Virology* **66,** 322–326 (1975).
62. S. G. Sawicki and D. L. Sawicki, The effect of overproduction of nonstructural proteins on alphavirus plus-strand and minus-strand RNA synthesis. *Virology* **152,** 507–512 (1986).
63. P. M. Grimley, I. K. Berezesky, and R. M. Friedman, Cytoplasmic structures associated with an arbovirus infection: Loci of viral ribonucleic acid synthesis. *J. Virol.* **2,** 1326–1338 (1968).
64. S. Froshauer, J. Kartenbeck, and A. Helenius, Alphavirus RNA replicase is located on the cytoplasmic surface of endosomes and lysosomes. *J. Cell Biol.* **107,** 2075–2086 (1988).
65. J. Peränen and L. Kääriäinen, Biogenesis of type I cytopathic vacuoles in Semliki Forest virus-infected BHK cells. *J. Virol.* **65,** 1623–1627 (1991).
66. P. Kujala, A. Ikäheimonen, N. Ehsani, H. Vihinen, P. Auvinen, and L. Kääriäinen, Biogenesis of the Semliki Forest virus RNA replication complex. *J. Virol.* **75,** 3873–3884 (2001).
67. P. M. Grimley, J. G. Levin, I. K. Berezesky, and R. M. Friedman, Specific membranous structures associated with the replication of group A arboviruses. *J. Virol.* **10,** 492–503 (1972).
68. B.-E. Lachmi and L. Kääriäinen, Control of protein synthesis in Semliki Forest virus infected cells. *J. Virol.* **22,** 142–149 (1977).
69. D. Magliano, J. A. Marshall, D. S. Bowden, N. Vardaxis, J. Meanger, and J.-Y. Lee, Rubella virus replication complexes are virus-modified lysosomes. *Virology* **240,** 57–63 (1998).

70. P. Kujala, T. Ahola, N. Ehsani, P. Auvinen, H. Vihinen, and L. Kääriäinen, Intracellular distribution of rubella virus nonstructural protein P150. *J. Virol.* **73,** 7805–7811 (1999).
71. D. J. Barton, S. G. Sawicki, and D. L. Sawicki, Demonstration *in vitro* of temperature-sensitive elongation of RNA in Sindbis virus mutant *ts6*. *J. Virol.* **62,** 3597–3602 (1988).
72. D. J. Barton, S. G. Sawicki, and D. L. Sawicki, Solubilization and immunoprecipitation of alphavirus replication complexes. *J. Virol.* **65,** 1496–1506 (1991).
73. C. M. Scheele and E. R. Pfefferkorn, Inhibition of interjacent ribonucleic acid (26S) synthesis in cells infected by Sindbis virus. *J. Virol.* **4,** 117–122 (1969).
74. S. Keränen and L. Kääriäinen, Isolation and basic characterization of temperature-sensitive mutants from Semliki Forest virus. *Acta Pathol. Microbiol. Scand. B* **82,** 810–820 (1974).
75. E. G. Strauss, E. M. Lenches, and J. H. Strauss, Mutants of Sindbis virus. I. Isolation and partial characterization of 89 new temperature-sensitive mutants. *Virology* **74,** 154–168 (1976).
76. S. Keränen and L. Kääriäinen, Functional defects of RNA-negative temperature-sensitive mutants of Sindbis and Semliki Forest viruses. *J. Virol.* **32,** 19–29 (1979).
77. D. L. Sawicki and S. G. Sawicki, Functional analysis of the A complementation group mutants of Sindbis HR virus. *Virology* **144,** 20–34 (1985).
78. J. Saraste, L. Kääriäinen, H. Söderlund, and S. Keränen, RNA synthesis directed by a temperature-sensitive mutant of Semliki Forest virus. *J. Gen. Virol.* **37,** 399–406 (1977).
79. D. L. Sawicki, L. Kääriäinen, C. Lambek, and P. J. Gomatos, Mechanism for control of synthesis of Semliki Forest virus 26S and 42S RNA. *J. Virol.* **25,** 19–27 (1978).
80. R. Levis, S. Schlesinger, and H. V. Huang, Promoter for Sindbis virus RNA-depedent subgenomic RNA transcription. *J. Virol.* **64,** 1726–1733 (1990).
81. R. Raju and H. V. Huang, Analysis of Sindbis virus promoter recognition *in vivo*, using novel vectors with two subgenomic mRNA promoters. *J. Virol.* **65,** 2501–2510 (1991).
82. M. M. Wielgosz, R. Raju, and H. V. Huang, Sequence requirements for Sindbis virus subgenomic mRNA promoter function in cultured cells. *J. Virol.* **75,** 3509–3519 (2001).
83. B.-E. Lachmi and L. Kääriäinen, Sequential translation of nonstructural proteins in cells infected with a Semliki Forest virus mutant. *Proc. Natl. Acad. Sci. U.S.A.* **73,** 1936–1940 (1976).
84. L. Kääriäinen, D. L. Sawicki, and P. J. Gomatos, Cleavage defect in the non-structural polyprotein of Semliki Forest virus has two separate effects on virus RNA synthesis. *J. Gen. Virol.* **39,** 463–473 (1978).
85. S. Keränen and L. Ruohonen, Nonstructural proteins of Semliki Forest virus: Synthesis, processing, and stability in infected cells. *J. Virol.* **47,** 505–551 (1983).
86. W. R. Hardy and J. H. Strauss, Processing the nonstructural polyproteins of Sindbis virus: Study of the kinetics *in vivo* by using monospecific antibodies. *J. Virol.* **62,** 998–1007 (1988).
87. M. Ding and M. J. Schlesinger, Evidence that Sindbis virus nsP2 is an autoprotease which processes the virus nonstructural polyprotein. *Virology* **171,** 280–284 (1989).
88. W. R. Hardy and J. H. Strauss, Processing the nonstructural polyproteins of Sindbis virus: Nonstructural proteinase is in the C-terminal half of nsP2 and functions both in cis and in *trans*. *J. Virol.* **63,** 4653–4664 (1989).
89. Y. Shirako and J. H. Strauss, Cleavage between nsP1 and nsP2 initiates the processing pathway of Sindbis virus nonstructural polyprotein P123. *Virology* **177,** 54–64 (1990).
90. E. G. Strauss, R. J. deGroot, R. Levinson, and J. H. Strauss, Identification of the active site residues in the nsP2 proteinase of Sindbis virus. *Virology* **191,** 932–940 (1992).
91. E. ten Dam, M. Flint, and M. D. Ryan, Virus-coded proteinases of the Togaviridae. *J. Gen. Virol.* **80,** 1879–1888 (1999).
92. A. Merits, L. Vasiljeva, T. Ahola, L. Kääriäinen, and P. Auvinen, Proteolytic processing of Semliki Forest virus-specific non-structural polyprotein by nsP2 protease. *J. Gen. Virol.* **82,** 765–773 (2001).

93. S. Mi and V. Stollar, Expression of Sindbis virus nsP1 and methyltransferase activity in *Escherichia coli*. *Virology* **184,** 423–427 (1991).
94. P. Laakkonen, M. Hyvönen, J. Peränen, and L. Kääriäinen, Expression of Semliki Forest virus nsP1-specific methyltransferase in insect cells and in *Escherichia coli*. *J. Virol.* **68,** 7418–7425 (1994).
95. T. Ahola and L. Kääriäinen, Reaction in alphavirus mRNA capping: Formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. *Proc. Natl. Acad. Sci. U.S.A.* **92,** 507–511 (1995).
96. A. Lampio, T. Ahola, E. Darzynkiewicz, J. Stepinski, M. Jankowska-Anyszka, and L. Kääriäinen, Guanosine nucleotide analogs as inhibitors of alphavirus mRNA capping enzyme. *Antiviral Res.* **42,** 35–46 (1999).
97. Y. Furuichi and A. J. Shatkin, Viral and cellular mRNA capping: Past and prospects. *Adv. Virus Res.* **55,** 135–184 (2000).
98. S. Shuman, Structure, mechanism, and evolution of the mRNA capping apparatus. *Prog. Nucleic Acid Res. Mol. Biol.* **66,** 1–40 (2000).
99. A. Merits, R. Kettunen, K. Mäkinen, A. Lampio, P. Auvinen, L. Kääriäinen, and T. Ahola, Virus-specific capping of tobacco mosaic virus RNA: Methylation of GTP prior to formation of covalent complex p126-m7GMP. *FEBS Lett.* **455,** 45–48 (1999).
100. T. Ahola and P. Ahlquist, Putative RNA capping activities encoded by brome mosaic virus: Methylation and covalent binding of guanylate by replicase protein 1a. *J. Virol.* **73,** 10,061–10,069 (1999).
101. F. Kong, K. Sivakumaran, and C. Kao, The N-terminal half of the brome mosaic virus 1a protein has RNA capping-associated activities: Specificity for GTP and S-adenosylmethionine. *Virology* **259,** 200–210 (1999).
102. Y. I. Li, Y. J. Chen, Y. H. Hsu, and M. Meng, Characterization of the AdoMet-dependent guanylyltransferase activity that is associated with the N terminus of bamboo mosaic virus replicase. *J. Virol.* **75,** 782–788 (2001).
103. J. Magden, N. Takeda, T. Li, P. Auvinen, T. Ahola, T. Miyamura, A. Merits, and L. Kääriäinen, Virus-specific mRNA capping enzyme encoded by hepatitis E virus. *J. Virol.* **75,** 6249–6255 (2001).
104. H.-L. Wang, J. O'Rear, and V. Stollar, Mutagenesis of the Sindbis virus nsP1 protein: Effects on methyltransferase activity and viral infectivity. *Virology* **217,** 527–531 (1996).
105. T. Ahola, J. A. den Boon, and P. Ahlquist, Helicase and capping enzyme active site mutations in brome mosaic virus protein 1a cause defects in template recruitment, negative-strand RNA synthesis, and viral RNA capping. *J. Virol.* **74,** 8803–8811 (2000).
106. T. Ahola, A. Lampio, P. Auvinen, and L. Kääriäinen, Semliki Forest virus mRNA capping enzyme requires association with anionic membrane phospholipids for activity. *EMBO J.* **18,** 3164–3172 (1999).
107. S. Mi, R. K. Durbin, H. V. Huang, C. M. Rice, and V. Stollar, Association of the Sindbis virus RNA methyltransferase activity with the nonstructural protein nsP1. *Virology* **170,** 385–391 (1989).
108. L. M. Scheidel, R. K. Durbin, and V. Stollar, SV_{LM21}, a Sindbis virus mutant resistant to methionine deprivation, encodes an altered methyltransferase. *Virology* **173,** 408–414 (1989).
109. L. M. Scheidel, R. K. Durbin, and V. Stollar, Sindbis virus mutants resistant to mycophenolic acid and ribavirin. *Virology* **158,** 1–7 (1987).
110. L. M. Scheidel and V. Stollar, Mutations that confer resistance to mycophenolic acid and ribavirin on Sindbis virus map to the nonstructural protein nsP1. *Virology* **181,** 490–499 (1991).
111. C. I. Rosenblum, L. M. Scheidel, and V. Stollar, Mutations in the nsP1 coding sequence of Sindbis virus which restrict viral replication in secondary cultures of chick embryo fibroblasts prepared from aged primary cultures. *Virology* **198,** 100–108 (1994).

112. J. Peränen, P. Laakkonen, M. Hyvönen, and L. Kääriäinen, The alphavirus replicase protein nsP1 is membrane-associated and has affinity to endocytic organelles. *Virology* **208,** 610–620 (1995).

113. T. Ahola, P. Kujala, M. Tuittila, T. Blom, P. Laakkonen, A. Hinkkanen, and P. Auvinen, Effects of palmitoylation of replicase protein nsP1 on alphavirus infection. *J. Virol.* **74,** 6725–6733 (2000).

114. P. Laakkonen, T. Ahola, and L. Kääriäinen, The effects of palmitoylation on membrane association of Semliki Forest virus RNA capping enzyme. *J. Biol. Chem.* **271,** 28,567–28,571 (1996).

115. A. Lampio, I. Kilpeläinen, S. Pesonen, K. Karhi, P. Auvinen, P. Somerharju, and L. Kääriäinen, Membrane binding mechanism of an RNA virus-capping enzyme. *J. Biol. Chem.* **275,** 37,853–37,859 (2000).

116. J. E. Johnson and R. B. Cornell, Amphitropic proteins: Regulation by reversible membrane interactions. *Mol. Membr. Biol.* **16,** 217–235 (1999).

117. P. Laakkonen, P. Auvinen, P. Kujala, and L. Kääriäinen, Alphavirus replicase protein Nsp1 induces filopodia and rearrangement of actin filaments. *J. Virol.* **72,** 10,265–10,269 (1998).

118. Y. Shirako, E. G. Strauss, and J. H. Strauss, Suppressor mutations that allow Sindbis virus RNA polymerase to function with nonaromatic amino acids at the N-terminus: Evidence for interaction between nsP1 and nsP4 in minus-strand RNA synthesis. *Virology* **276,** 148–160 (2000).

119. I. Dé, S. G. Sawicki, and D. L. Sawicki, Sindbis virus RNA-negative mutants that fail to convert from minus-strand to plus-strand synthesis: Role of the nsP2 protein. *J. Virol.* **70,** 2706–2719 (1996).

120. A. E. Gorbalenya, E. V. Koonin, and M.-C. Lai, Putative papain-related thiol proteases of positive-strand RNA viruses. Identification of rubi- and aphtovirus proteases and delineation of a novel conserved domain associated with proteases of rubi-, alpha- and coronaviruses. *FEBS Lett.* **288,** 201–205 (1991).

121. A. E. Gorbalenya and E. V. Koonin, Helicases: Amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Cell Biol.* **3,** 419–429 (1993).

122. M. Rikkonen, J. Peränen, and L. Kääriäinen, ATPase and GTPase activities associated with Semliki Forest virus nonstructural protein nsP2. *J. Virol.* **68,** 5804–5810 (1994).

123. M. Rikkonen, Functional significance of the nuclear-targeting and NTP-binding motifs of Semliki Forest virus nonstructural protein nsP2. *Virology* **218,** 352–361 (1996).

124. M. Gomez de Cedrón, N. Ehsani, M. L. Mikkola, J. A. García, and L. Kääriäinen, RNA helicase activity of Semliki Forest virus replicase protein NSP2. *FEBS Lett.* **448,** 19–22 (1999).

125. L. E. Bird, H. S. Subramanya, and D. B. Wigley, Helicases: A unifying structural theme?. *Curr. Opin. Struct. Biol.* **8,** 14–18 (1998).

126. P. A. Kroner, B. M. Young, and P. Ahlquist, Analysis of the role of brome mosaic virus 1a protein domains in RNA replication, using linker insertion mutagenesis. *J. Virol.* **64,** 6110–6120 (1990).

127. L. Vasiljeva, A. Merits, P. Auvinen, and L. Kääriäinen, Identification of a novel function of the *Alphavirus* capping apparatus. RNA 5′ triphosphatase activity of Nsp2. *J. Biol. Chem.* **275,** 17,281–17,287 (2000).

128. J. R. Myette and E. G. Niles, Characterization ot the vaccinia virus RNA 5′-triphosphatase and nucleoside triphosphate phosphohydrolase activities. Demonstration that both activities are carried out at the same active site. *J. Biol. Chem.* **271,** 11,945–11,952 (1996).

129. C. K. Ho, Y. Pei, and S. Shuman, Yeast and viral RNA 5′ triphosphatases comprise a new nucleoside triphosphatase family. *J. Biol. Chem.* **273,** 34,151–34,156 (1998).

130. G. Wengler and Gi. Wengler, The NS3 nonstructural protein of flaviviruses contains an RNA triphosphatase activity. *Virology* **197,** 265–273 (1993).

131. L. Vasiljeva, L. Valmu, L. Kääriäinen, and A. Merits, Site-specific protease activity of the carboxyl-terminal domain of Semliki Forest virus replicase protein nsP2. *J. Biol. Chem.* **276,** 30,786–30,793 (2001).
132. N. Kalkkinen, M. Laaksonen, H. Söderlund, and H. Jörnvall, Radio-sequence analysis of in vivo multilabeled nonstructural protein ns86 of Semliki Forest virus. *Virology* **113,** 188–195 (1981).
133. N. Kalkkinen, Radio-sequence analysis: An ultra-sensitive method to align protein and nucleotide sequences. *in* "Advanced Methods in Protein Microsequence Analysis" (B. Wittmann-Liebold, J. Salnikow, and V. A. Erdmann, Eds.), pp. 194–206. Springer-Verlag, Berlin, 1986.
134. J. Peränen, M. Rikkonen, P. Liljeström, and L. Kääriäinen, Nuclear localization of Semliki Forest virus-specific nonstructural protein nsP2. *J. Virol.* **64,** 1888–1896 (1990).
135. M. Rikkonen, J. Peränen, and L. Kääriäinen, Nuclear and nucleolar targeting signals of Semliki Forest virus nonstructural protein nsP2. *Virology* **189,** 462–473 (1992).
136. P. Russo, P. Laakkonen, T. Ahola, and L. Kääriäinen, Synthesis of Semliki Forest virus RNA polymerase components nsP1 through nsP4 in *Saccharomyces cerevisiae* by expression of cDNA encoding the nonstructural polyprotein. *J. Virol.* **70,** 4086–4089 (1996).
137. L. Kääriäinen and M. Ranki, Inhibition of cell functions by RNA-virus infections. *Annu. Rev. Microbiol.* **38,** 91–109 (1984).
137a. J. K. Fazakerley, A. Boyd, M. L. Mikkola, and L. Kääriäinen, A single amino acid change in the nuclear localization sequence of the nsP2 protein affects the neurovirulence of Semliki Forest virus. *J. Virol.* **76,** 392–396, (2002).
138. H. Vihinen, T. Ahola, M. Tuittila, A. Merits, and L. Kääriäinen, Elimination of phosphorylation sites of Semliki Forest virus replicase protein nsP3. *J. Biol. Chem.* **276,** 5745–5752 (2001).
139. J. R. Pehrson and R. N. Fuji, Evolutionary conservation of histone macroH2A subtypes and domains. *Nucleic Acids Res.* **26,** 2837–2842 (1998).
140. R. C. Aguiar, Y. Yakushijin, S. Kharbanda, R. Salgia, J. A. Fletcher, and M. A. Shipp, BAL is a novel risk-related gene in diffuse large B-cell lymphomas that enhances cellular migration. *Blood* **96,** 4328–4334 (2000).
141. M. R. Martzen, S. M. McCraith, S. L. Spinelli, F. M. Torres, S. Fields, E. J. Grayhack, and E. M. Phizicky, A biochemical genomics approach for identifying genes by the activity of their products. *Science* **286,** 1153–1155 (1999).
142. J. Peränen, K. Takkinen, N. Kalkkinen, and L. Kääriäinen, Semliki Forest virus-specific nonstructural protein nsP3 is a phosphoprotein. *J. Gen. Virol.* **69,** 2165–2178 (1988).
143. G. Li, M. W. LaStarza, W. R. Hardy, J. H. Strauss, and C. M. Rice, Phosphorylation of Sindbis virus nsP3 *in vivo* and *in vitro*. *Virology* **179,** 416–427 (1990).
144. M. W. LaStarza, A. Grakoui, and C. M. Rice, Deletion and duplication mutations in the C-terminal nonconserved region of Sindbis virus nsP3: Effects on phoshorylation and on virus replication in vertebrate and invertebrate cells. *Virology* **202,** 224–232 (1994).
145. H. Vihinen and J. Saarinen, Phosphorylation site analysis of Semliki Forest virus nonstructural protein 3. *J. Biol. Chem.* **275,** 27,775–27,783 (2000).
146. M. T. Tuittila, M. G. Santagati, M. Röytta, J. A. Määttä, and A. E. Hinkkanen, Replicase complex genes of Semliki Forest virus confer lethal neurovirulence. *J. Virol.* **74,** 4579–4589 (2000).
147. M. W. LaStarza, J. A. Lemm, and C. M. Rice, Genetic analysis of the nsP3 region of Sindbis virus: Evidence for roles in minus-strand and subgenomic RNA synthesis. *J. Virol.* **68,** 5781–5791 (1994).
148. E. K. O'Reilly and C. C. Kao, Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. *Virology* **252,** 287–303 (1998).

149. Y. S. Hahn, A. Grakoui, C. M. Rice, E. G. Strauss, and J. H. Strauss, Mapping of RNA$^-$ temperature-sensitive mutants of Sindbis virus: Complementation group F mutants have lesions in nsP4. *J. Virol.* **63,** 1194–1202 (1989).

150. Y. H. Lin, P. Yadav, R. Ravatn, and V. Stollar, A mutant of Sindbis virus that is resistant to pyrazofurin encodes an altered RNA polymerase. *Virology* **272,** 61–71 (2000).

151. Y. Shirako and J. H. Strauss, Requirement for an aromatic amino acid or histidine at the N terminus of Sindbis virus RNA polymerase. *J. Virol.* **72,** 2310–2315 (1998).

152. G. Li and C. M. Rice, Mutagenesis of the in-frame opal termination codon preceding nsP4 of Sindbis virus: Studies of translational readthrough and its effect on virus replication. *J. Virol.* **63,** 1326–1337 (1989).

153. K. Takkinen, J. Peränen, and L. Kääriäinen, Proteolytic processing of Semliki Forest virus-specific non-structural polyprotein. *J. Gen. Virol.* **72,** 1627–1633 (1991).

154. R. J. deGroot, T. Rümenapf, R. J. Kuhn, E. G. Strauss, and J. H. Strauss, Sindbis virus RNA polymerase is degraded by the N-end rule pathway. *Proc. Natl. Acad. Sci. U.S.A.* **88,** 8967–8971 (1991).

155. P.-Y. Cheng, Purification, size, and morphology of a mosquito-borne animal virus, Semliki Forest virus. *Virology* **14,** 124–131 (1961).

156. G. Wengler, Gi. Wengler, and H. J. Gross, Replicative form of Semliki Forest virus RNA contains an unpaired guanosine. *Nature* **282,** 754–756 (1979).

157. G. Wengler, Gi. Wengler, and H. J. Gross, Terminal sequences of Sindbis virus-specific nucleic acids: Identity in molecules synthesized in vertebrate and insect cells and characteristic properties of the replicative form RNA. *Virology* **123,** 273–283 (1982).

158. G. Wu and J. M. Kaper, Requirement of 3′-terminal guanosine in (-)-stranded RNA for *in vitro* replication of cucumber mosaic virus satellite RNA by viral RNA-dependent RNA polymerase. *J. Mol. Biol.* **238,** 655–657 (1994).

159. D. L. Sawicki and P. J. Gomatos, Replication of Semliki Forest virus: Polyadenylate in plus-strand RNA and polyuridylate in minus-strand RNA. *J. Virol.* **20,** 446–464 (1976).

160. K. R. Hill, M. Hajjou, J. Y. Hu, and R. Raju, RNA-RNA recombination in Sindbis virus: Roles of the 3′ conserved motif, poly(A) tail, and nonviral sequences of template RNAs in polymerase recognition and template switching. *J. Virol.* **71,** 2693–2704 (1997).