JOURNAL OF VIROLOGY, Jan. 1991, p. 147–154
0022-538X/91/010147-08$02.00/0

Vol. 65, No. 1

Internally Located Cleavable Signal Sequences Direct the Formation of Semliki Forest Virus Membrane Proteins from a Polyprotein Precursor

PETER LILJESTRÖM* AND HENRIK GAROFF

Department of Molecular Biology, Center for Biotechnology, Karolinska Institute, Novum, S-141 52 Huddinge, Sweden

Received 9 July 1990/Accepted 3 October 1990

The proteolytic processes involved in the cotranslational production of the Semliki Forest virus proteins p62, 6K, and E1 from a common precursor polypeptide were analyzed by an in vitro translation-translocation assay. By studying the behavior of wild-type and mutant variants of the polyprotein, we show that the signal sequences responsible for membrane translocation of the 6K and E1 proteins reside in the C-terminal regions of p62 and 6K, respectively. We present evidence suggesting that the polyprotein is processed on the luminal side by signal peptidase at consensus cleavage sites immediately following the signal sequences. Our results also lead us to conclude that the 6K protein is a transmembrane polypeptide with its N terminus on the luminal side of the membrane (type I). Thus, the production of all three membrane proteins is directed by alternating signal and stop-transfer (anchor) sequences that function in translocation and cleavage of the virus precursor polyprotein. This also shows conclusively that internally located signal sequences can be cleaved by signal peptidase.

Enveloped mammalian viruses have learned to utilize the biosynthetic transport pathway of their host cells as a means of synthesizing their own membrane components. Because of this feature, these viruses have become very useful model systems with which to study membrane formation in mammalian cells. For instance, the viral models have been involved in much of the pioneering work on membrane protein structure and synthesis at the endoplasmic reticulum (ER). Viral membrane proteins were among the very first examples of monotopic (influenza virus HA, vesicular stomatitis virus G) and polytopic (coronavirus E1) membrane proteins that were characterized in structural and topological detail (2, 24, 32, 45). Some of the monotopic membrane proteins were found to be inserted into the membrane of the ER in an N-terminus-in/C-terminus-out orientation through the sequential expression of an N-terminal signal sequence followed by a stop-transfer (or anchor) sequence (type I membrane proteins such as influenza virus HA and vesicular stomatitis virus G) (36, 38). Other monotopic membrane proteins were found to be inserted in the opposite orientation via combined signal-anchor sequences usually located at or close to the N-terminal end of the polypeptide chain (type II membrane proteins such as influenza virus NA) (7, 33).

We have been interested in the way in which the membrane proteins of Semliki Forest virus (SFV), an alphavirus, are synthesized. In contrast to most other membrane proteins, those of SFV are generated from a common coding unit on a 26S (4.1-kb) mRNA which includes all structural proteins of the virus. The cytoplasmic nucleocapsid (NC) protein C (33 kDa) is made first, followed by the membrane proteins p62 (62 kDa), 6K (6 kDa), and E1 (50 kDa) (for reviews, see references 18 and 40). This polyprotein system introduces several interesting questions about membrane protein topogenesis, such as (i) the nature of the signal sequences which are able to initiate or reinitiate polypeptide chain translocation at internal positions of the (nascent) polyprotein precursor and (ii) the kinds of cleavage events that are required to separate the individual proteins on the polyprotein sequence.

To this end, we and others have shown that the C protein is released from the nascent precursor chain by autoproteolysis (1, 22, 23, 35), thereby revealing an N-terminal signal sequence which is used for p62 chain translocation (6, 16, 19). Furthermore, we recently characterized a hydrophobic peptide in the C-terminal region of the 6K protein which can direct the translocation of the E1 protein as well as that of a heterologous protein (34). In the present work we have shown that this hydrophobic peptide is also required for E1 translocation in the context of the complete structural polyprotein and that the 6K-E1 cleavage site is sensitive to changes violating the consensus features of the signal peptidase cleavage site. Results are also presented which suggest that the 6K protein carries a functional stop-transfer signal and that the C-terminal region of the p62 protein (which precedes the 6K protein on the polyprotein sequence) functions as a signal sequence for the 6K protein. Therefore, we suggest that the 6K protein represents a type I transmembrane molecule like the two larger SFV membrane proteins and that in addition to the N-terminal signal of p62, two internal cleavable signal sequences are involved in the generation of the SFV membrane proteins.

MATERIALS AND METHODS

Bacteria, phage, and plasmids. *Escherichia coli* strains used were DH5α (Bethesda Research Laboratories), which is *recA endA1 gyrA96 thi-1 hsdR17 supE44 relA1* Δ(*lacZYA-argF*)*U169* φ80d*lacZ*Δ(M15); GM161 (*E. coli* CGSC 6476), which is *thr-1 leuB6 dam-4 thi-1 hsdS1 lacY1 tonA21 supE44*; DH5αF'IQ, which is *endA1 hsdR1 supE44 thi-1 recA1 gyrA96 relA1* φ80d*lac*Δ(M15) Δ(*lacZYA-argF*)*U169*/F' *proAB lacI*^{q} *lacZ*Δ(M15) TN5 (Bethesda Research Laboratories); and RZ1032 (30), which is Hfr KL16 *dut-1 ung-1 thi-1 relA1 supE44 zbd-279*::Tn*10*. Bacteriophage M13mp19 (46) and plasmids pBR322 (5) and pGEM1 (Promega Corp.) have been described previously.

Materials. Most restriction enzymes, DNA polymerase I,

* Corresponding author.

Klenow fragment, calf intestinal phosphatase, T4 DNA ligase, and T4 polynucleotide kinase were from Boehringer, Mannheim, Federal Republic of Germany. *Sph*I, *Stu*I, and *Kpn*I, together with RNase inhibitor (RNasin) and SP6 polymerase, were from Promega Biotec, Madison, Wis. Sequenase (Modified T7 polymerase) was from United States Biochemical, Cleveland, Ohio. Proteinase K was from Merck, Darmstadt, Federal Republic of Germany. Ribonucleotides, deoxyribonucleotides, dideoxyribonucleotides, and the cap analog $m^7G(5')ppp(5')G$ were from Pharmacia. Oligonucleotides were produced by using an Applied Biosystems synthesizer 380B followed by high-pressure liquid chromatography (HPLC) and NAP-5 (Pharmacia) purification. Spermidine, phenylmethylsulfonyl fluoride, diethylpyrocarbonate, bovine serum albumin (BSA), creatine phosphate, and creatine phosphokinase were from Sigma, St. Louis, Mo. Pansorbin was from Calbiochem, La Jolla, Calif. Agarose was purchased from FMC Bioproducts, Rockland, Maine, and acrylamide was from Bio-Rad, Richmond, Calif. The glycosylation acceptor peptide *N*-benzoyl-Asn-Leu-Thr-*N*-methylamide and the nonacceptor peptide *N*-benzoyl-Asn-Leu-(*allo*)Thr-*N*-methylamide were synthesized as described previously (14). Reticulocyte lysate was prepared from rabbits as described previously (25). Canine pancreas rough microsomes (50 A_{280} units/ml) were prepared as described previously (26). ^{14}C-labeled proteins used as molecular weight markers were from DuPont, Dreieich, Federal Republic of Germany. L-[^{35}S]methionine and [α-^{35}S]dATP-α-S were from Amersham.

DNA methods. Plasmids were isolated essentially by using the alkali-sodium dodecyl sulfate (SDS) method (4). All basic DNA procedures were done essentially as described previously (39). DNA fragments isolated from agarose gels (3) were purified by benzoyl-naphthoyl-DEAE (BND)-cellulose chromatography (41). Plasmids were purified by sedimentation through 1 M NaCl followed by banding in CsCl (39). Competent cells were prepared as described previously (10). DNA sequencing was performed by following the United States Biochemicals protocol for using Sequenase.

Construction of mutants. pGEM1-SFV plasmid (also called pG-SFV-15/5) (34) carrying the complete SFV 26S cDNA region coding for all structural proteins was used in production of wild-type RNA for translation (Fig. 1). For the construction of most mutant variants, we transferred the *Cla*I fragment spanning the relevant DNA region into M13 for mutagenesis. Since both *Cla*I sites in the SFV cDNA are followed by a C residue, they become methylated and cannot be cleaved when the plasmid is isolated from a normal bacterial host; therefore, the plasmid was grown in the *dam*-negative host strain GM161 to overcome this methylation. The *Cla*I fragment spanning the p62-6K-E1 joint regions was isolated from an agarose gel and recloned into the *Cla*I site of pBR322 to obtain unique flanking restriction endonuclease sites to be used for cloning into M13 (data not shown). From this plasmid the *Eco*RI-*Hin*dIII fragment was isolated and recloned into *Eco*RI-*Hin*dIII-cut M13mp19 replicative form (RF). The RF was transformed into RZ1032, and phage was grown for 6 h in the presence of uridine to incorporate uracil residues into the DNA (30). The phage was precipitated with polyethylene glycol, and the single-stranded DNA was isolated by phenol extraction. Appropriate phosphorylated oligonucleotides were hybridized and mutagenesis was performed by primer extension with Sequenase as described previously (30, 43). Mutant, in vitro-synthesized heteroduplex RFs were transformed into DH5αF'IQ, and the resulting phages were analyzed for the presence of mutations by sequencing the complete insert. The RFs of the correct mutants was isolated and cut with *Eco*RI and *Hin*dIII, and the DNA fragments were gel purified and recloned into pGEM1-SFV as a *Cla*I fragment via intermediate *Eco*RI-*Hin*dIII cloning in pBR322 and growth in the *dam* host GM161. The presence of the correct mutation was once more confirmed by sequencing the concerned regions of the pGEM1-SFV variants. The oligonucleotide 5'-TGTCGAATGTTCGTATGCGTGCGCCCGCGG was used to delete the 6K region, the oligonucleotide 5'-CACACTAG CAAAGTGCGCCCG was used to change the -1 alanine residue of the 6K signal sequence to a phenylalanine, the oligonucleotide 5'-GTAAGCTCTGTTGGTTGCCCC changed the -3 alanine residue of the E1 signal sequence to a phenylalanine, and, finally, oligonucleotide 5'-GGCGGTTGCCCC GCGGCTCAGTAGCACTCTAAAAGAAAGGCT changed the leucine 5 and the leucine 10 residues of the E1 signal sequence to arginines.

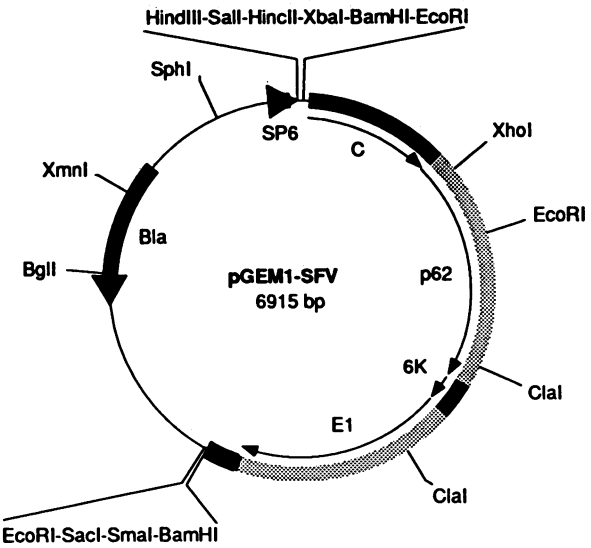

FIG. 1. Plasmid used for transcription of SFV membrane protein genes. pGEM1-SFV contains the complete subgenomic cDNA of the SFV, coding for all structural proteins (in the transcriptional order C-p62-6K-E1) of the virus. In vitro transcription starts at the promoter for the SP6 polymerase. The region flanked by the two *Cla*I sites was recloned for in vitro mutagenesis.

In vitro transcription and translation. Uncut CsCl-banded plasmid DNA was used as a template for in vitro transcription. RNA was synthesized at 37°C for 1 h in 10- to 50-μl reaction mixtures containing 40 mM Tris hydrochloride (pH 7.6), 6 mM $MgCl_2$, 2 mM spermidine hydrochloride, 5 mM dithiothreitol, 100 μg of nuclease-free BSA per ml, 1 mM each ATP, CTP, and UTP, 500 μM GTP, 1 U of RNasin per μl, and 100 to 500 U of SP6 RNA polymerase per ml. For production of capped transcripts, the analog $m^7G(5')ppp(5')G$ was included in the reaction at 1 mM (29). In vitro translation reactions with a rabbit reticulocyte lysate was performed at 30°C essentially as described previously (34): 1.5 μl of the in vitro-synthesized RNA was translated in a total volume of 15 μl with or without 2 μl of dog pancreas ER microsomal membranes. Potassium, magnesium, and spermidine concentrations were 100, 1.2, and 0.375 mM, respectively. Competing (acceptor) or control (nonacceptor) peptides for N-linked protein glycosylation reactions were added to a final concentration of 200 μM (16). For protease

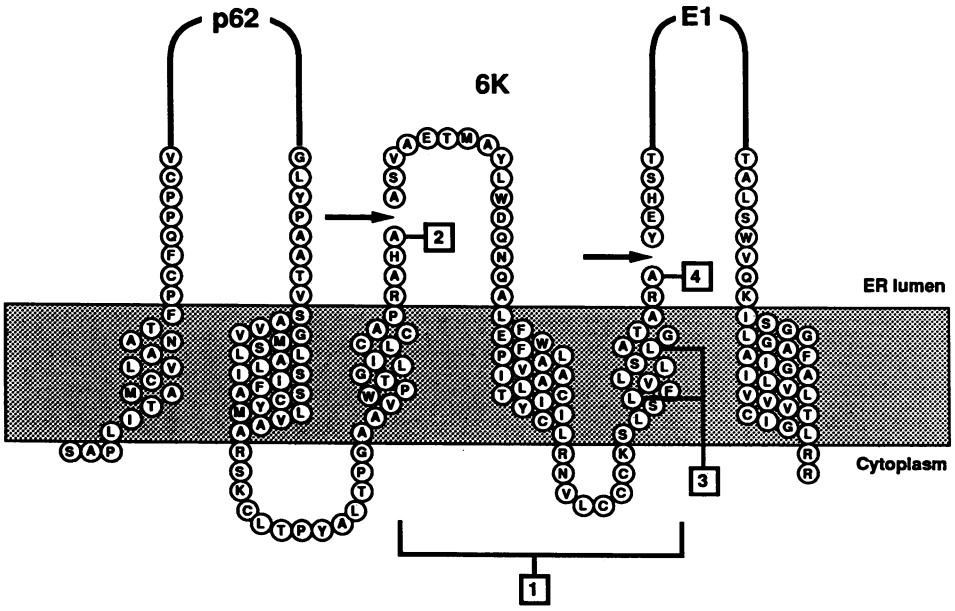

FIG. 2. Schematic presentation of the proposed cleavage events leading to the generation of the SFV structural proteins and their final topology in the ER membrane. The amino acid residues that were changed by site-directed mutagenesis in this study are marked by mutation numbers 1 to 4. Signal peptidase cleavages are indicated by arrows. Most of the luminal portions of the p62 and E1 proteins are drawn as solid lines and are not to scale.

protection experiments, proteinase K was added to a final concentration of 0.5 mg/ml and the samples were incubated at 0°C for 30 min in the presence or absence of 1% Triton X-100. Proteolysis was stopped by adding phenylmethylsulfonyl fluoride to 400 μg/ml, and samples were kept at 0°C for an additional 5 min before being processed for electrophoresis.

Gel electrophoresis. Samples for sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) were prepared and run on 10% separating gels with a 5% stacking gel as previously described (11). For resolving the 6K peptide, a 10 to 20% linear acrylamide gradient gel was used. The gels were fixed in 10% acetic acid–30% methanol for 1 h before being exposed to Kodak XAR-5 film. When the gel was prepared for fluorography (8), it was washed for 30 min in 30% methanol after fixation and then soaked in 1 M sodium salicylate–30% methanol for 30 min before being dried. Nucleic acids were run on agarose gels with 50 mM Tris-borate–2.5 mM disodium EDTA as the buffer. For staining, 0.2 μg of ethidium bromide per ml was included in the buffer and gel during the run. Sequencing gels were run as described previously (39) with Tris-borate as the buffer.

RESULTS

The C-terminal region of p62 contains a signal sequence. Since our model (35) predicted that the 6K protein is integrated into the ER membrane, we wanted to define the region of the polyprotein responsible for its translocation. The site of cleavage between p62 and 6K has been defined by amino acid sequence analysis (27), and this, together with the hydrophobic character of the tail of p62 (15), suggested that a signal sequence resides in the C-terminal part of p62. We reasoned that if this was the case and if we were to delete 6K by fusing the p62 and E1 coding regions in frame, translocation of E1 should still occur since the tail of p62 now would function as a signal sequence for E1. Therefore, using site-directed mutagenesis, we deleted the whole 6K gene from plasmid pGEM1-SFV (Fig. 2, mutation 1).

We first analyzed the behavior of the wild-type SFV polyprotein during in vitro translation. In vitro-synthesized RNA was translated in the absence of membranes, and two proteins with apparent sizes of about 100 and 30 kDa were produced (Fig. 3, lane 1). This corresponds well to the unglycosylated precursor protein consisting of p62, 6K, and E1 (107 kDa), while the smaller protein corresponds to the size of the capsid protein (C; 33 kDa). We conclude that faithful translation and cleavage of the nascent chain had occurred. Translation in the presence of microsomes produced protein species that corresponded closely to the expected sizes of glycosylated p62 (62 kDa), E1 (50 kDa), and 6K (6 kDa) (Fig. 3, lane 6; Fig. 4, lane 1). Addition of proteinase K showed that the p62 and E1 molecules had indeed been translocated, since their luminal parts were protected by the microsomal membrane (Fig. 3, lane 7). The protease treatment resulted in a small decrease in the size of p62, an indication of the trimming of its 31-amino-acid cytoplasmic tail. E1 did not change in size, since it has only two arginine residues exposed outside the microsomal vesicle, and therefore possible trimming could not be visualized by this assay. The untranslocated capsid protein was completely sensitive to added protease. The same was true for the remaining precursor form, since it was totally sensitive to protease, indicating that it was not translocated. The resistance of the membrane proteins to protease was due to inaccessibility, since protease treatment in the presence of Triton X-100 completely degraded all protein species (lane 8). During in vitro translation, p62 is glycosylated at four Asn positions while E1 is glycosylated only once (18). We used this fact to verify translocation by adding an excess of a competing acceptor peptide for N-linked glycosylation to the translation-translocation reaction mixture. Since distinct smaller forms corresponding to the sizes of unglycosylated

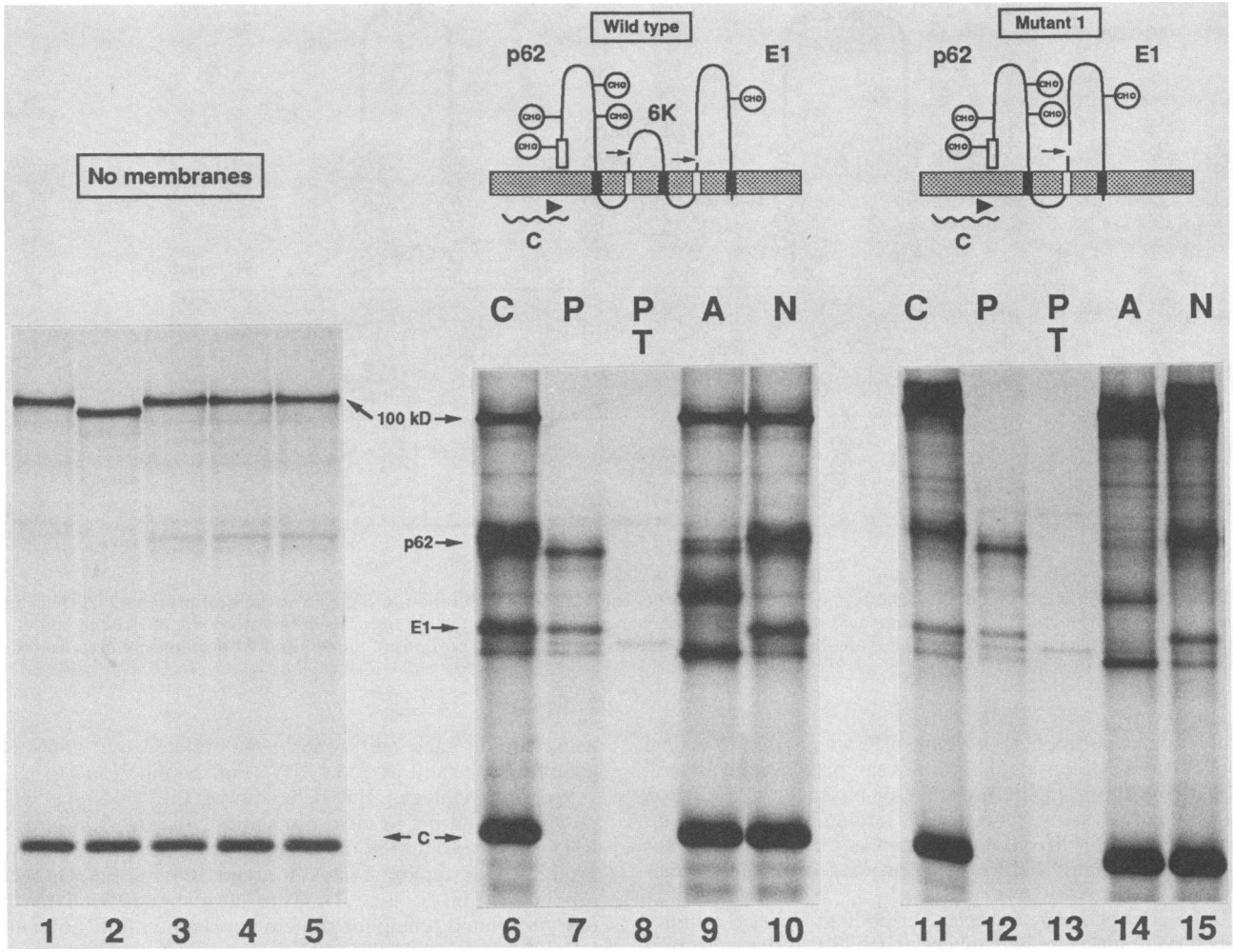

FIG. 3. In vitro translation-translocation of pGEM1-SFV derivatives. SFV 26S cDNA was transcribed in vitro (SP6), and the RNA was translated (rabbit reticulocyte lysate) in the absence or presence of microsomal membranes. The samples were analyzed by SDS-PAGE (10% acrylamide) and autoradiography. Treatment with protease (P) or protease in the presence of Triton X-100 (P/T) is indicated. Translations were also performed in the presence of an acceptor (A) or nonacceptor (N) peptide. Bands corresponding to the 100-kDa precursor, p62, E1, and C proteins are marked. Note that of the two faint bands below E1, the smaller represents unglycosylated E1 (compare, e.g, lanes 6 and 9) and the larger represents a tRNA-dependent (but ribosome-independent) addition of methionine to a rabbit reticulocyte lysate protein (25). This is resistant to protease digestion (see, e.g., lane 8). This makes the band an excellent internal marker, especially when judging the shifts in molecular weight of E1. Above each set of translation-translocation assays, the interpreted end result is depicted. Luminal arrows indicate signal peptidase cleavage. Glycosylation of p62 and E1 is marked by CHO. Symbols: ▶, cytoplasmically occurring capsid protease cleavage; ■, membrane anchors; □, signal sequences; ▨, E1 signal variant containing two arginine residues.

chains were found, we conclude that p62 and E1 were glycosylated and hence also translocated (lane 9). This inhibition was specific, since addition of a control peptide had no effect (lane 10).

In vitro-synthesized RNA from the 6K deletion construct was also translated in the absence of membranes. As expected, this construct gave a slightly smaller precursor protein, whereas production of the capsid protein was unaffected (Fig. 3, lane 2). In the presence of membranes, p62, E1, and C proteins of wild-type sizes were produced (Fig. 3, lane 11), whereas the 6K peptide was absent (Fig. 4, lane 2). In this case the p62 and E1 proteins were translocated and cleaved somewhat less efficiently than the wild-type construct. As in the case of the wild-type construct, translocation of the p62 and E1 species could be judged from their resistance to added protease (lanes 12 and 13) as well as their glycosylation (lanes 14 and 15). Some of the precursor form appeared to be translocated, since a minor portion of it had a slightly lower mobility than the rest of the precursor (lane 11), indicating that it had become glycosylated. By contrast, translation in the presence of an acceptor peptide resulted in a decrease in the apparent size of this precursor form (lane 14). Since neither of the precursors was totally protected from added protease, we conclude that the translocated (but uncleaved) portions became trimmed within the cytoplasmic loop domain of the p62 tail, resulting in protein species of p62 and E1 sizes comigrating with the appropriately cleaved (by signal peptidase) ones (lane 12).

Signal peptidase cleaves between p62 and 6K. Since it was evident from these results that the p62 tail does contain a functional signal sequence, it was of interest to define the mode of cleavage between p62 and 6K in the processing of

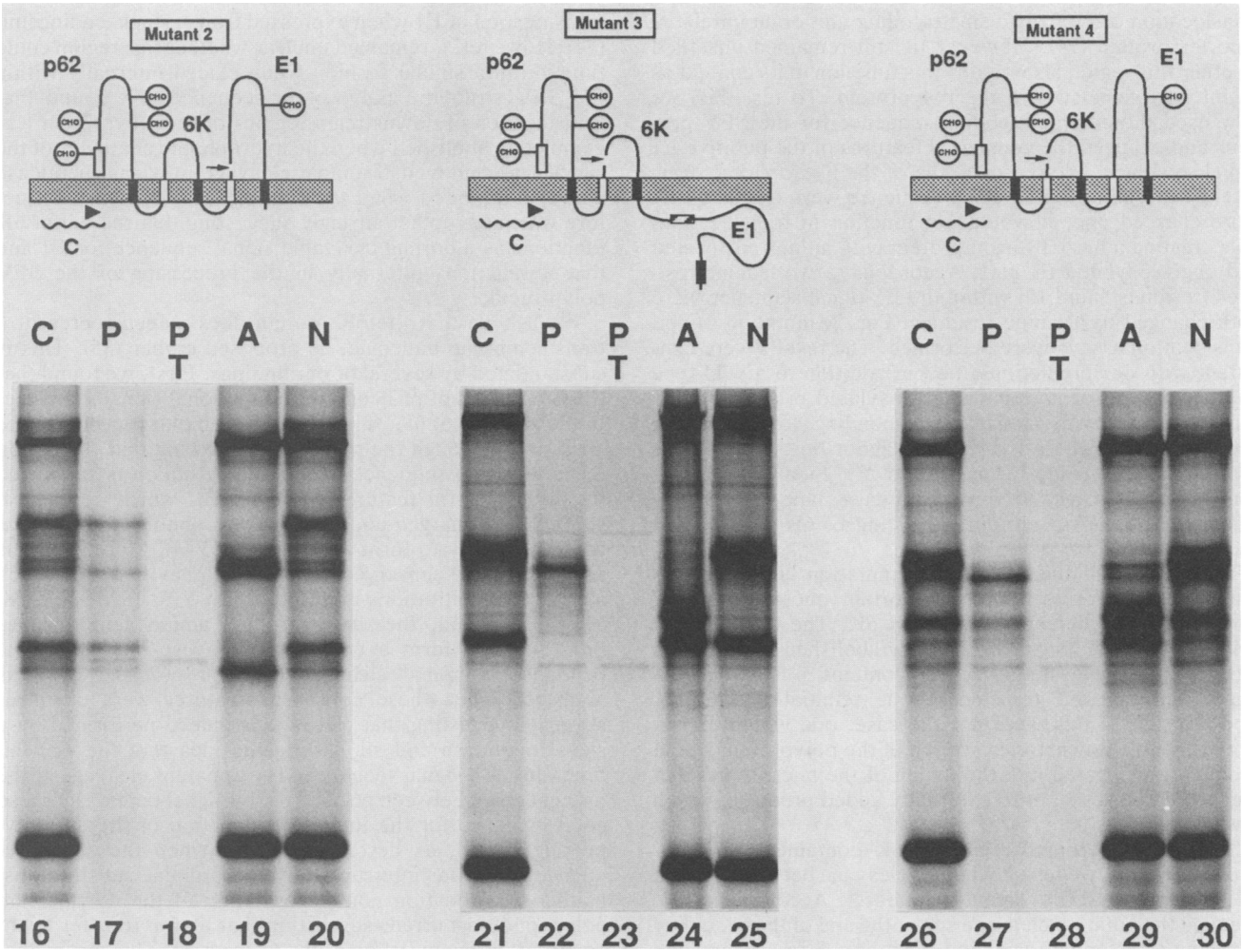

FIG. 3—*Continued.*

the SFV precursor protein. To this end, the ultimate C-terminal alanine residue of p62 was changed to a phenylalanine (Fig. 2, mutation 2). If cleavage is indeed performed by signal peptidase, this mutation should prohibit cleavage since the Phe residue is at position −1 of the putative signal peptidase cleavage site, a position that contains only small amino acid species (44). When translated in the absence of membranes, this mutant construct encoded 100- and 30-kDa protein species corresponding to p62-6K-E1 and C, respectively, suggesting that the substitution had no effect on translation per se (Fig. 3, lane 3). However, in the presence of membranes a p62 protein which was somewhat larger than the wild-type protein appeared, indicating that cleavage between p62 and 6K had not occurred (lane 16). Moreover, production of the 6K protein species could not be detected (Fig. 4, lane 3). Production of E1 was not affected by the mutation. Small portions of both the fusion protein and E1 remained unglycosylated (Fig. 3, lane 16), although they evidently had been translocated, since they were resistant to protease (lane 17). Protease treatment did not cleave the fusion protein, suggesting that since the p62 tail now was linked to the luminal part of 6K, it was perhaps pulled close to the microsomal membrane and hence not accessible to protease digestion. Translocation also resulted in glycosylation of the proteins, as verified by the addition of an acceptor peptide which inhibited glycosylation, resulting in smaller products (lanes 19 and 20).

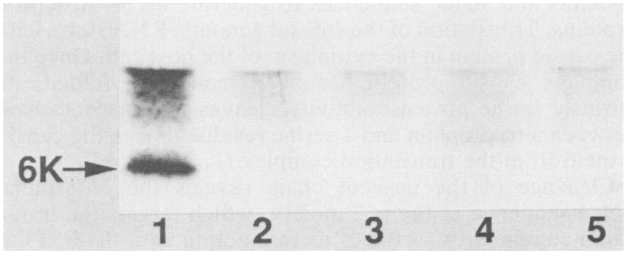

FIG. 4. In vitro translation-translocation analyzed on an SDS–10 to 20% gradient gel. Only the region below the globin area is shown. The position of the 6K protein is indicated. The wild-type construct (lane 1) and mutant variants 1 to 4 (lanes 2 to 5, respectively) are shown.

The 6K tail functions internally in the polyprotein as a signal sequence for E1. Having defined the mode of cleavage at the N-terminal border of 6K, we next wanted to characterize the signal sequence of E1. It was shown previously that the C-terminal region of the 6K protein can function as a signal recognition particle (SRP)-dependent signal sequence for the

translocation of E1 from a single coding unit or an unrelated reporter protein (34). However, it still remained unsettled whether this region also retains this function in its context as an internal signal within the polyprotein. To test this, we now used the complete coding sequence for the SFV proteins and changed the consensus features of the putative E1 signal sequence. We reasoned that if the hydrophobic character of the putative E1 signal sequence were changed to a more charged one, it would not function in translocation, thus creating a 6K-E1 fusion protein with an untranslocated and unglycosylated E1 part. Accordingly, two leucine residues (residues 5 and 10) within the E1 signal sequence were both changed to arginine residues (Fig. 2, mutation 3) and translocation assays were performed. The results were consistent with our predictions; i.e., in addition to a wild-type p62, an untranslocated and unglycosylated protein species was produced with a size corresponding to that of an unglycosylated 6K-E1 hybrid with about the same size as glycosylated E1 (Fig. 3, lanes 21 to 25). Finally, since we could not detect any 6K protein (Fig. 4, lane 4), we concluded that a fusion protein consisting of 6K and E1 had been produced.

The double arginine mutant (from mutation 3 in Fig. 2) was also used to address another important question, that of stop-transfer proficiency exerted by 6K. The fact that the mutation in the E1 signal completely abolished downstream translocation suggested that 6K contains a stop-transfer signal and that E1 translocation is reinitiated after 6K stop-transfer. If this were not the case, one would expect that the whole downstream stretch of the polyprotein would become translocated into the lumen of the microsomes and thus would be resistant to externally added protease, which it was not.

Since the C-terminal region of 6K contained signal sequence activity, we tested whether cleavage between 6K and E1 is performed by signal peptidase. Accordingly, we changed the putative cleavage site at the end of the 6K tail by changing the alanine residue at position −3 to a phenylalanine (Fig. 2, mutation 4). Translation in the absence of membranes showed that the mutant construct was faithfully translated (Fig. 3, lane 5). Membrane translocation assays showed that signal peptidase cleavage between 6K and E1 was indeed abolished. Translation in the presence of membranes produced a wild-type-size p62 while the E1 protein band increased in size, suggesting that 6K now was fused to E1 (Fig. 3, lane 26). At the same time, we could not detect any 6K protein (Fig. 4, lane 5). Again, translocation was verified by testing for protease resistance and occurrence of glycosylation (Fig. 3, lanes 27 to 30). The p62 tail was trimmed in a normal fashion, whereas the E1-6K fusion protein remained totally resistant. This result was somewhat surprising, since the external (putative) loop region between 6K and E1 should consist of 8 amino acids (Fig. 2) and we expected this loop to be cleaved. Apparently, the close vicinity to the membrane blocked digestion.

DISCUSSION

By using an in vitro transcription-translation-translocation system, we analyzed the cotranslational processing of the SFV structural polyprotein. More specifically, by studying the phenotypic expression of a panel of site-specific mutations, we defined the mode of events that lead to the cleavage and formation of the p62, 6K, and E1 proteins.

The C-terminal region of the 6K protein was previously shown to contain a signal sequence that is able to initiate translocation of E1 when expressed from a single coding unit (34). However, it remained unclear whether this region could function in a similar fashion when placed internally within the SFV structural polyprotein sequence. We found that translocation of downstream regions of the polyprotein was completely abolished when the hydrophobic character of the signal was removed. Luminal cleavage by signal peptidase was also inhibited when the cleavage site consensus structure was interrupted, strongly suggesting that the tail of 6K functions as a normal cleavable signal sequence for E1 and that it functions internally in the processing of the SFV polyprotein.

We believe that the 6K protein does indeed represent a transmembrane molecule, as proposed earlier (35). This is substantiated by several of our findings. First, we found that the C terminus of p62 contains a functional signal sequence. In the absence of 6K, this region was able to translocate the E1 protein through the membrane, showing that the region contains signal sequence activity. This notion is supported by the structural features of this area, which contains a stretch of uncharged and mostly hydrophobic residues that have the ability to form an α-helix (20, 44). Moreover, this region is highly conserved within the alphaviruses, in which almost all substitutions are conservative (9, 13, 15, 21, 28, 31, 42). Second, the three terminal amino acid residues preceding 6K form a consensus cleavage site for signal peptidase. Luminal cleavage between p62 and 6K was abolished when these consensus features were changed, strongly suggesting that cleavage is indeed performed by a signal peptidase and supporting the idea that the p62 tail functions as a signal sequence. The apparent cleavage of the nascent chain between p62 and 6K by signal peptidase is also good evidence for the initial translocation of the 6K polypeptide. Third, as described above, when the signal sequence of E1 was interrupted by two arginine substitutions, neither translocation nor glycosylation of the downstream polypeptide occurred, suggesting that a stop-transfer event occurred during insertion of the 6K protein. This means that the E1 signal would have to reinitiate E1 translocation. Further evidence stems from the fact that the 6K-E1 fusion protein generated by mutation 4 (cleavage site) was totally resistant to protease digestion; therefore, the major part of 6K must have been translocated through the membrane. All in all, we think that the apolar region between the glutamate (6K residue 21) and the next arginine (6K residue 37), together with the four hydrophobic residues (LFWL) preceding the glutamate, constitutes the anchor region of 6K (44).

The results from this work and earlier studies now allow us to propose a complete scenario by which SFV utilizes both viral and host proteases in combination with anchor domains and signal sequences to generate all its structural proteins. Translation of the 26S subgenomic RNA starts with the capsid protein in the cytoplasm of the host cell. Once the complete capsid protein has been made and folded, its intrinsic serine protease activity cleaves the nascent chain between a tryptophan and a serine residue to free the capsid protein from the translation complex (1, 22, 23, 35).

Cleavage of the nascent chain reveals the N-terminal signal sequence of the p62 moiety, which targets the translation complex (by virtue of its interaction with the SRP) to the ER membrane, initiating the cotranslational translocation of the polypeptide through the membrane (6, 19). The p62 signal sequence is unusual in the sense that it is not cleaved off by signal peptidase. Instead, it is released into the lumen of the ER, where it becomes glycosylated long

before synthesis of p62 has been completed (16). This shows that the signal sequence is required only for chain translocation initiation, but not for completion.

Translocation of the p62 protein is arrested when its anchor sequence inserts into the membrane (11, 12, 15, 17). The C-terminal region of p62, which is left on the external (cytoplasmic) side, contains a second signal sequence which reinitiates translocation of the 6K moiety of the growing polypeptide (see above). This process could be SRP independent, since the ribosome complex is already at the membrane and should not need membrane targeting. This notion is supported by the fact that a deletion mutant of the SFV polyprotein, lacking all downstream residues from residue 25 of 6K, can efficiently translocate and can be processed correctly between p62 and 6K (11). Since the ribosome covers some 40 residues of a growing polypeptide chain (a stretch which in this case spans the 6K signal as well as the first 24 residues of 6K), the translation complex would dissociate (and thus prohibit the SRP from binding) before the signal sequence of 6K emerges from the ribosome interior to react with the SRP, since at least 50 residues are needed before SRP-mediated translational arrest occurs (37).

The signal sequence of 6K is cleaved by signal peptidase (see above), and translocation continues until arrested by the insertion of the 6K anchor sequence into the membrane. Soon after this, translocation of E1 is reinitiated by the E1 signal sequence resident in the 6K tail (34) (see above). During translocation of E1, the signal sequence is cleaved off by signal peptidase (see above), and finally translocation is arrested when the E1 anchor reaches the membrane. The end result will be the generation of three transmembrane proteins, p62, 6K, and E1, with the topology shown in Fig. 2.

ACKNOWLEDGMENTS

We acknowledge Päivi Nieminen and Raija Jönsson for excellent technical assistance. We are indebted to Anne Peters and Maarit Närvä for production of the oligonucleotides used in this study.

This work was supported by the Swedish Medical Research Council (grant B88-12X-0872-01A), the Swedish National Board for Technical Development (grant 87-02750P), and the Swedish Natural Science Research Council (grant B-BU 9353-301).

REFERENCES

1. **Aliperti, G., and M. J. Schlesinger.** 1978. Evidence for an autoprotease activity of Sindbis virus capsid protein. Virology **90**:366–369.
2. **Armstrong, J., H. Niemann, S. Smeekens, P. Rottier, and G. Warren.** 1984. Sequence and topology of a model intracellular membrane protein, E1 glycoprotein, from a coronavirus. Nature (London) **308**:751–752.
3. **Benson, S. A.** 1984. A rapid procedure for isolation of DNA fragments from agarose gels. BioTechniques **2**:66–68.
4. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
5. **Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, and H. W. Boyer.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95–113.
6. **Bonatti, S., G. Migliaccio, G. Blobel, and P. Walter.** 1984. Role of the signal recognition particle in the membrane assembly of Sindbis viral glycoprotein. Eur. J. Biochem. **140**:499–502.
7. **Bos, T. J., A. R. Davis, and D. P. Nayak.** 1984. NH_2-terminal hydrophobic region of influenza virus neuraminidase provides the signal function in translocation. Proc. Natl. Acad. Sci. USA **81**:2327–2331.
8. **Chamberlain, J. P.** 1979. Fluorographic detection of radioactivity in polyacrylamide gels with water-soluble fluor, sodium salicylate. Anal. Biochem. **98**:132–135.
9. **Chang, G.-J. J., and D. W. Trent.** 1987. Nucleotide sequence of the genome region encoding the 26S mRNA of eastern equine encephalomyelitis virus and the deduced amino acid sequence of the viral structural proteins. J. Gen. Virol. **68**:2129–2142.
10. **Chung, C. T., and R. T. Miller.** 1988. A rapid and convenient method for the preparation and storage of competent bacterial cells. Nucleic Acids Res. **16**:3580.
11. **Cutler, D. F., and H. Garoff.** 1986. Mutants of the membrane-binding region of Semliki Forest virus E2 protein. I. Cell surface transport and fusogenic activity. J. Cell Biol. **102**:889–901.
12. **Cutler, D. F., P. Melançon, and H. Garoff.** 1986. Mutants of the membrane-binding region of Semliki Forest virus E2 protein. II. Topology and membrane binding. J. Cell Biol. **102**:902–910.
13. **Dalgarno, L., C. M. Rice, and J. H. Strauss.** 1983. Ross River virus 26S RNA: complete nucleotide sequence and deduced sequence of the encoded structural proteins. Virology **129**:170–187.
14. **Erickson, B. W., and R. B. Merrifield.** 1976. Solid phase peptide synthesis, p. 255–527. *In* R. L. Hill and H. Neurath (ed.), The proteins. Academic Press, Inc. (London), Ltd., London.
15. **Garoff, H., A.-M. Frischauf, K. Simons, H. Lehrach, and H. Delius.** 1980. Nucleotide sequence of cDNA coding for Semliki Forest virus membrane glycoproteins. Nature (London) **288**:236–241.
16. **Garoff, H., D. Huylebroeck, A. Robinson, U. Tillman, and P. Liljeström.** 1990. The signal sequence of the p62 protein of Semliki Forest virus is involved in initiation but not in completing chain translocation. J. Cell Biol. **111**:867–876.
17. **Garoff, H., C. Kondor-Koch, R. Pettersson, and B. Burke.** 1983. Expression of Semliki Forest virus proteins from cloned complementary DNA. II. The membrane-spanning glycoprotein E2 is transported to the cell surface without its normal cytoplasmic domain. J. Cell Biol. **97**:652–658.
18. **Garoff, H., C. Kondor-Koch, and H. Riedel.** 1982. Structure and assembly of alphaviruses. Curr. Top. Microbiol. Immunol. **99**:1–50.
19. **Garoff, H., K. Simons, and B. Dobberstein.** 1978. Assembly of Semliki Forest virus membrane glycoproteins in the membrane of the endoplasmic reticulum in vitro. J. Mol. Biol. **124**:587–600.
20. **Gierasch, L. M.** 1989. Signal sequences. Biochemistry **28**:923–930.
21. **Hahn, C. S., S. Lustig, E. G. Strauss, and J. H. Strauss.** 1988. Western equine encephalitis virus is a recombinant virus. Proc. Natl. Acad. Sci. USA **85**:5997–6001.
22. **Hahn, C. S., E. G. Strauss, and J. H. Strauss.** 1985. Sequence analysis of three Sindbis virus mutants temperature-sensitive in the capsid protein autoprotease. Proc. Natl. Acad. Sci. USA **82**:4648–4652.
23. **Hahn, C. S., and J. H. Strauss.** 1990. Site-directed mutagenesis of the proposed catalytic amino acids of the Sindbis virus capsid protein autoprotease. J. Virol. **64**:3069–3073.
24. **Hurtley, S. M., and A. Helenius.** 1989. Protein oligomerization in the endoplasmic reticulum. Annu. Rev. Cell Biol. **5**:277–307.
25. **Jackson, R. J., and T. Hunt.** 1983. Preparation and use of nuclease-treated rabbit reticulocyte lysate for the translation of eukaryotic messenger RNA. Methods Enzymol. **96**:50–74.
26. **Kaderbhai, M. A., and B. M. Austen.** 1984. Dog pancreatic microsomal-membrane polypeptides analysed by two-dimensional gel electrophoresis. Biochem. J. **217**:145–147.
27. **Kalkkinen, N.** 1980. Carboxy-terminal sequence analysis of the four structural proteins of Semliki Forest virus. FEBS Lett. **115**:163–166.
28. **Kinney, R. M., B. J. B. Johnson, J. B. Welch, K. R. Tsuchiya, and D. W. Trent.** 1989. The full-length nucleotide sequences of the virulent Trinidad donkey strain of Venezuelan equine encephalitis virus and its attenuated vaccine derivative, strain TC-83. Virology **170**:19–30.
29. **Krieg, P. A., and D. A. Melton.** 1987. In vitro RNA synthesis with SP6 RNA polymerase. Methods Enzymol. **155**:397–415.
30. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection.

Methods Enzymol. **154**:367–382.
31. **Levinson, R. S., J. H. Strauss, and E. G. Strauss.** 1990. Complete sequence of the genomic RNA of O'Nyong-nyong virus and its use in the construction of alphavirus phylogenetic trees. Virology **175**:110–123.
32. **Machamer, C., and J. K. Rose.** 1987. A specific transmembrane domain of a coronavirus E1 glycoprotein is required for its retention in the Golgi region. J. Cell Biol. **105**:1205–1214.
33. **Markoff, L., B.-C. Lin, M. M. Sveda, and C.-J. Lai.** 1984. Glycosylation and surface expression of the influenza virus neuraminidase requires the N-terminal hydrophobic region. Mol. Cell. Biol. **4**:8–16.
34. **Melançon, P., and H. Garoff.** 1986. Reinitiation of translocation in the Semliki Forest virus structural polyprotein: identification of the signal for the E1 glycoprotein. EMBO J. **5**:1551–1560.
35. **Melançon, P., and H. Garoff.** 1987. Processing of the Semliki Forest virus structural polyprotein: role of the capsid protease. J. Virol. **61**:1301–1309.
36. **Min Jou, W., M. Verhoeyen, R. Devos, E. Saman, R. Fang, D. Huylebroeck, W. Fiers, G. Threlfall, C. Barber, M. Carey, and S. Emtage.** 1980. Complete structure of the hemagglutinin gene from the human influenza A/Victoria/3/75 (H3N2) strain as determined from cloned DNA. Cell **19**:683–696.
37. **Okun, M. M., E. M. Eskridge, and D. Shields.** 1990. Truncations of a secretory protein define minimum lengths required for binding to signal recognition particle and translocation across the endoplasmic reticulum membrane. J. Biol. Chem. **265**:7478–7484.
38. **Rose, J. K., W. J. Welch, B. M. Sefton, F. S. Esch, and N. C. Ling.** 1980. Vesicular stomatitis virus glycoprotein is anchored in the viral membrane by a hydrophobic domain near the COOH-terminus. Proc. Natl. Acad. Sci. USA **77**:3884–3888.
39. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
40. **Schlesinger, S. S., and M. J. Schlesinger.** 1986. Formation and assembly of alphavirus glycoproteins, p. 121–148. *In* S. S. Schlesinger and M. J. Schlesinger (ed.), The Togaviridae and Flaviviridae, Plenum Press, New York.
41. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions, p. 303. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
42. **Strauss, E. G., C. M. Rice, and J. H. Strauss.** 1984. Complete nucleotide sequence of the genomic RNA of Sindbis virus. Virology **133**:92–110.
43. **Su, T.-Z., and M. R. El-Gewely.** 1988. A multisite-directed mutagenesis using T7 DNA polymerase: application for reconstructing a mammalian gene. Gene **69**:81–89.
44. **von Heijne, G.** 1988. Transcending the impenetrable: how proteins come to terms with membranes. Biochim. Biophys. Acta **947**:307–333.
45. **Wiley, D. C., and J. J. Skehel.** 1987. The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. Annu. Rev. Biochem. **56**:365–394.
46. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.