JOURNAL OF VIROLOGY, June 1993, p. 3288–3294
0022-538X/93/063288-07$02.00/0

Vol. 67, No. 6

Processing of the Envelope Glycoproteins of Pestiviruses

TILLMANN RÜMENAPF,[1] GABI UNGER,[2] JAMES H. STRAUSS,[1] AND HEINZ-JÜRGEN THIEL[2]*

Division of Biology, California Institute of Technology, Pasadena, California 91125,[1] and Federal Research Center for Virus Diseases of Animals, Paul Ehrlich Strasse 28, D-7400 Tübingen, Germany[2]

Received 23 December 1992/Accepted 2 March 1993

The genomic RNA of pestiviruses is translated into a large polyprotein that is cleaved into a number of proteins. The structural proteins are N terminal in this polyprotein and include three glycoproteins called E0, E1, and E2 on the basis of the order in which they appear in the polyprotein. Using pulse-chase experiments, we show that a pestiviral glycoprotein precursor, E012, is formed that is processed into E0, E1, and E2 in an ordered fashion. Processing is initiated by a nascent cleavage between the capsid and the translocated E012 followed by cleavage at the C terminus of E2. E012 is then rapidly cleaved to form E01 and E2. After E2 is released from the precursor, E01 is processed into E0 and E1. To identify the sites of cleavage, the N termini of the glycoproteins of the pestivirus classical swine fever virus (formerly termed hog cholera virus) were sequenced after expression in the vaccinia virus system. The N termini are Glu-268 for E0 (gp44/48), Leu-495 for E1 (gp33) and Arg-690 for E2 (gp55). The sequences around the cleavage sites capsid/E0 and E1/E2 conform to the rules known for cellular signal proteases, as does the sequence at the presumed C terminus of E2. The sequence upstream of the E0/E1 cleavage site also shows sequence characteristics of signalase processing sites but lacks the typical hydrophobic signal peptide; this cleavage site has characteristics in common with a site in flaviviruses that is also cleaved in a delayed fashion. The absence of any membrane-spanning region results in the shedding of E0 by infected cells, and E0 can be detected in the virus-free supernatant. Comparison of the sequences around the cleavage sites of pestiviruses suggests a general processing scheme for the structural glycoproteins. Comparison of the pesti- and flaviviral structural glycoproteins suggests analogies between E012 and prM-E.

Pestivirus, a genus of the family *Flaviviridae* (23), consists of three distinct pathogens of farm animals, classical swine fever virus (CSFV), bovine viral diarrhea virus (BVDV), and border disease virus of sheep. CSFV was formerly termed hog cholera virus and abbreviated HCV. We have reverted to the older name CSFV for this virus in order to avoid confusion with hepatitis C virus, which constitutes another genus within the family *Flaviviridae* and is also abbreviated HCV. Pestiviruses are enveloped, single-stranded RNA viruses that encode a single polyprotein of about 4,000 amino acids that is processed co- and posttranslationally by both host cell and virus-encoded proteases (17, 24, 25). The structural proteins are located within the N-terminal third of the polyprotein (17). Recently, we reported the structural protein composition of CSFV (18). Purified virions consist of p14, believed to be the nucleocapsid protein (C), and three envelope-associated glycoproteins (E proteins), E0 (gp44/48), E1 (gp33), and E2 (gp55), in the order of their arrangement in the polyprotein. The pestiviral capsid protein is preceded in the polyprotein by a nonstructural protein, p23, which is a putative protease with autoproteolytic activity (18, 24) and which has no counterpart in flaviviruses.

Antibodies that neutralize virus infectivity are directed against epitopes located on E2 and E0, as shown recently (21, 22). E1 is believed to be buried in the viral envelope (22), and hardly any anti-E1 antibodies have been described. The glycoproteins form disulfide-linked complexes, an E0 homodimer with a size of 100 kDa, an E1-E2 heterodimer with a size of 75 kDa, and an E2 homodimer with a size of 100 kDa, all of which are found in infected cells as well as in virions (18).

* Corresponding author.

On the basis of sequence analysis, processing of the structural proteins was suggested to be mediated by host cell signalase (17); typical hydrophobic signal sequences were identified upstream of the putative N termini of E0 and E2 by employing von Heijne's algorithm (20). No predictions for the cleavage that releases the N terminus of E1 could be made. In the experiments reported in this paper, we determined the N termini of the three glycoproteins after expression in the vaccinia virus system and analyzed the cleavage events involved in their production.

MATERIALS AND METHODS

Cells and viruses. CVI and BHK-21 cells were obtained from the American Type Culture Collection (Rockville, Md.). The pig lymphoma cell line $38A_1D$ was kindly provided by W. Schäfer (Max-Planck-Institut für Virusforschung, Tübingen, Germany) (13). Cell lines were grown in Dulbecco's modified Eagle medium with 10% fetal calf serum. CSFV Alfort was from B. Liess (Veterinary School, Hannover, Germany). The vaccinia virus-CSFV recombinant VAC3.8 contains the entire region encoding the structural proteins of CSFV in the vaccinia virus thymidine kinase gene under control of the p7.5 promoter and has been described previously (14).

Metabolic labeling of cells and viruses. $38A_1D$ cells (10^7) were labeled for 16 h with 0.5 mCi of [^{3}H]glucosamine (Amersham Buchler, Braunschweig, Germany) beginning at 48 h after infection by CSFV. The labeling medium contained 20 mM fructose instead of glucose. After the labeling period, cells and tissue culture supernatant were harvested and stored at −70°C.

Preparation of labeled CSFV and virus-free supernatant. Supernatant from infected $38A_1D$ cells was centrifuged at

5,400 × *g* for 20 min to remove cells, nuclei, and large cell debris. After the virus had been pelleted in a Beckman SW50.1 rotor at 40,000 rpm for 1 h, the resuspended pellet and the virus-free supernatant were stored at −70°C.

Pulse-chase labeling of cells. CVI cells were infected with vaccinia virus recombinant VAC3.8 (14) at a multiplicity of infection of 10, and beginning at 16 h after infection, the cells were labeled for 30 min with a mixture of [^{35}S]methionine and [^{35}S]cysteine (Amersham Buchler) (0.5 mCi/ml) in medium lacking the respective amino acids. The labeling medium contained no cysteine and no methionine. The chase was initiated by removal of the labeling medium and the addition of Dulbecco's modified Eagle medium containing methionine and cysteine (2.5 mM). Cells were removed 0, 15, 30, and 90 min after the beginning of the pulse and stored at −70°C.

Radioimmunoprecipitation and SDS-PAGE. Cell and virus extracts as well as tissue culture supernatants were prepared for immunoprecipitation as described previously (13) and incubated with 100 μl of hybridoma supernatant. Precipitates were formed with cross-linked *Staphylococcus aureus*, analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (15) under reducing conditions, and processed for fluorography by using EnHance (New England Nuclear).

N-terminal radiosequencing. BHK-21 cells (5 × 10^6) were infected with VAC3.8 at a multiplicity of infection of 10 and labeled with 0.3 mCi of [^{35}S]cysteine (ICN) or 0.5 mCi of [^{3}H]isoleucine (Amersham) from 8 to 22 h postinfection in Dulbecco's modified Eagle medium lacking the respective amino acid. The CSFV specific glycoproteins were purified by immunoprecipitation with monoclonal antibodies (MAbs) 24/16 and A18. Under nonreducing conditions, MAb A18 reacts with the E1-E2 heterodimer, which permits efficient copurification of E1. The immunoprecipitated proteins were separated on an SDS-containing 10% polyacrylamide gel and electrotransferred to a polyvinylidene difluoride membrane (Immobilon-P, Millipore) essentially as described previously (6). The proteins were localized by autoradiography of the dried membrane, the parts of the membrane containing the proteins of interest were excised, and the proteins were subjected to automatic Edman degradation. The amino acid derivatives released at each cycle were collected, and the radioactivity was assayed by liquid scintillation counting.

RESULTS

Standardized nomenclature for pestivirus glycoproteins. Pestiviral glycoproteins have mostly been named according to their apparent molecular masses, and thus the analogous proteins from different strains have different names. The apparent differences in molecular mass found among pestivirus strains probably result from variations in the amino acid sequences, the degree of glycosylation, and the utilization of different gel systems. A standardized nomenclature for the E proteins facilitates comparisons within the *Pestivirus* genus as well as comparisons within the family *Flaviviridae*. We suggest that the E proteins be called E0, E1, and E2 to reflect their order along the genome. By using gene order, the assignment of the glycoprotein name will ensure that analogous proteins in different pestiviruses will always be given the same name. The names 0, 1, and 2 were chosen because we believe that E1 and E2 of pestiviruses correspond to E1 and E2 of hepatitis C virus, respectively (see below). In this terminology, E0 represents gp44/48 of CSFV Alfort, gp42 of CSFV Brescia, and gp48 of BVDV; E1 represents gp33 of CSFV Alfort, gp31 of CSFV Brescia, and gp25 of BVDV; and E2 represents gp55 of CSFV Alfort, gp51-54 of CSFV Brescia, and gp53 of BVDV (NADL) (Fig. 1).

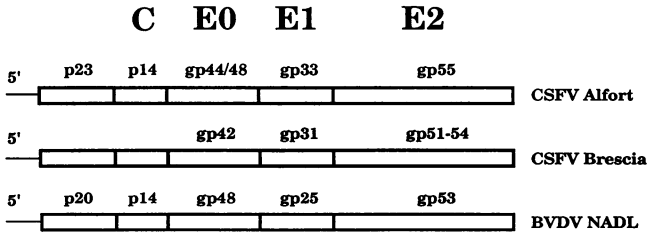

FIG. 1. Suggested nomenclature and comparative representation of the structural proteins of three pestiviruses. In most cases, different designations have been used for the analogous proteins. The genome localization of the putative capsid protein p14 was shown for CSFV Alfort and BVDV (18). We propose the names C, E0, E1, and E2 according to the order of the proteins in the polyprotein.

Kinetics of formation of the glycoproteins. To study the kinetics of biosynthesis of pestivirus glycoproteins, CV1 cells were infected with recombinant vaccinia virus VAC3.8, which expresses CSFV structural proteins (14), pulse-labeled for 30 min with a mixture of [^{35}S]methionine and cysteine, and chased for various times after addition of excess unlabeled methionine and cysteine. Cell extracts were subjected to immunoprecipitation by using MAbs directed against E0 (αE0; MAb 24/16) (18) or against E2 (αE2; MAb A18) (22), and the results are shown in Fig. 2. In addition to the three final glycoprotein products, E0, E1, and E2, two precursors were observed. One is a polyprotein containing the primary sequences of all three glycoproteins, referred to as E012 (130 kDa), and reacts with either antibody. The second polyprotein, referred to as E01 (73 kDa), is precipitated with MAb 24/16 (αE0) only (Fig. 2, lane 0) and, according to its molecular mass, most likely consists of the sequences of E0 and E1. The polyprotein E012 was surprisingly stable, having a half-life of several minutes. Because no polyprotein E12 was detectable, it appears that the cleavage between E01 and E2 occurs prior to the processing of E01. Thus, E012 is cleaved to form E01 and E2, and E01 is subsequently cleaved to form E0 and E1. During the course of the experiment, both E0 and E2 showed a significant reduction in their apparent molecular weights due to trimming of the carbohydrate chains on their route through the endoplasmic reticulum and the Golgi apparatus. The final destination of the glycoproteins is not known precisely, but viral assembly is believed to take place at intracellular membranes.

The results are largely consistent with data on BVDV proteins from Collett's group (2, 5). A low-abundance precursor with a size of about 120 kDa (gp116) corresponding to E012 gave rise to gp53 (E2) and gp62 (E01); the latter was further processed to produce gp48 (E0) and gp25 (E1). In addition, a precursor with an apparent molecular mass of 78 kDa (gp78) was described for BVDV, which was assumed to consist of the nucleocapsid protein p14, gp48, and gp25. However, in the CSFV system, we did not detect a precursor C-E01. The 73-kDa protein precursor identified in VAC3.8-infected cells as E01 apparently comprises only E0 and E1, because it cannot be precipitated with antibodies against the nucleocapsid protein p14 (data not shown). E012 is also not recognized by anti-p14 antibodies (data not

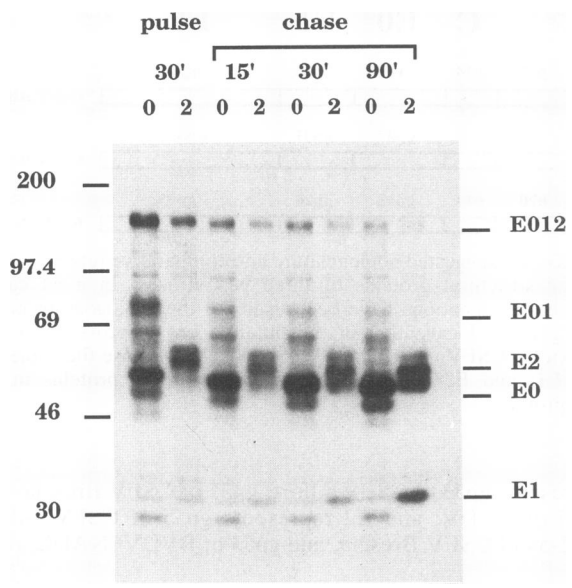

FIG. 2. Pulse-chase experiment with VAC3.8-infected cells. CVI cells were labeled with [^{35}S]methionine-cysteine for 30 min at 16 h after infection by VAC3.8 at a multiplicity of infection of 10. The chase was initiated by addition of unlabeled methionine-cysteine. Cells were collected 0, 15, 30, and 90 min after the beginning of the chase, and extracts were analyzed by immunoprecipitation with MAb 24/16 (αE0) (lanes 0) or MAb A18 (αE2) (lanes 2). Precipitates were subjected to SDS-PAGE (8% acrylamide) under reducing conditions. Note the coprecipitation of E1 and E2, which is detectable only after a long chase. The numbers on the left refer to molecular mass markers (Amersham) in kilodaltons.

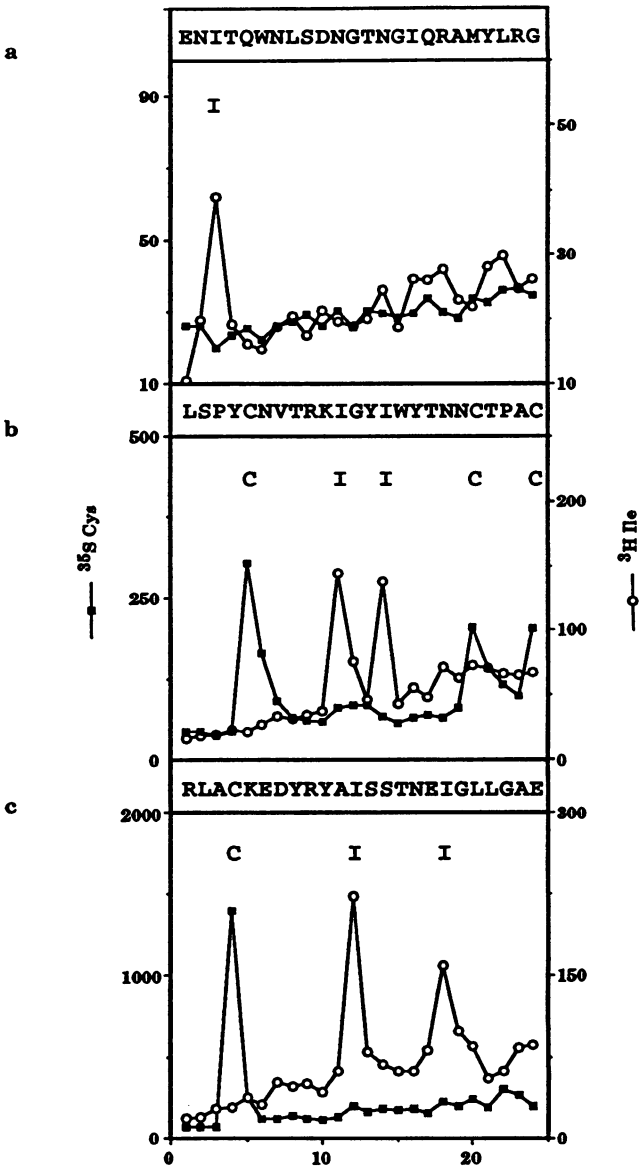

FIG. 3. Determination of the N termini of CSFV glycoproteins E0 (a), E1 (b), and E2 (c) from VAC3.8-infected cells. The amino acid sequence at the top of each panel is a sequence from the CSFV polyprotein beginning with the N-terminal amino acid determined in this study. Radiosequencing results are shown in the graphs in each panel. Open circles represent the radioactivity (in counts per minute) released during each cycle for [^{3}H]Ile-labeled protein, and solid squares represent [^{35}S]Cys-labeled protein (in counts per minute).

shown). The lower-molecular-mass protein with a size of about 64 kDa (Fig. 2) was not observed in pulse-chase experiments with CSFV-infected cells (unpublished data) and has not been further characterized; it may represent a cross-reacting vaccinia virus protein. The reason for the observed differences between the BVDV and CSFV systems remains to be determined.

The experiment shown in Fig. 2 also confirms the fact that E1 and E2 form a stable complex that can be precipitated with antibodies directed against E2 (22). Note that this complex forms slowly; only after a 90-min chase are significant amounts of the complex demonstrable.

Identification of the N termini of the glycoproteins. We wished to identify the cleavage sites used to process the glycoproteins in order to more fully characterize these proteins and because the nature of these cleavage sites might define the proteases involved in the processing of pestivirus E proteins. N-terminal sequencing of CSFV glycoproteins was conducted with radiolabeled CSFV E proteins which were produced in BHK-21 cells by recombinant vaccinia virus VAC3.8 (14). The N termini determined were Glu-268 for E0, Leu-495 for E1, and Arg-690 for E2 (Fig. 3a through c). N-terminal sequencing of purified unlabeled CSFV proteins yielded identical results, which confirms the authentic processing of VAC3.8-expressed proteins (18a).

Properties of the E proteins. The N terminus of E0 was found to be Glu-268 (Fig. 3a). With an assumed C terminus at Ala-494, E0 consists of 227 amino acids with a calculated molecular mass of 25.7 kDa. E0 is heavily glycosylated, suggesting that all nine possible N-glycosylation sites are occupied. E0 does not appear to possess a membrane anchor, and after its release from E01 presumably remains associated with the budding virus only through noncovalent associations with the other glycoproteins. In order to test for the stability of the association of E0 with the other glycoproteins, the culture medium covering CSFV-infected cells was collected, virus was purified from it, and the purified virus preparation and the virus-free medium were examined for the presence of glycoproteins by immunoprecipitation. Figure 4 shows an immunoprecipitation with MAb 24/16 (αE0) and MAb A18 (αE2), in which a lysate from CSFV-

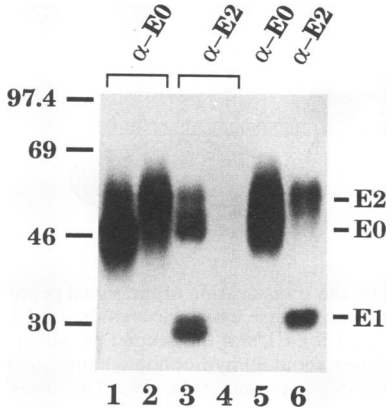

FIG. 4. Secretion of E0 into the medium. $38A_1D$ cells were labeled with [^{3}H]glucosamine for 16 h, beginning at 48 h after infection by CSFV. Cell lysates (lanes 1 and 3), virus-free supernatant (lanes 2 and 4), and pelleted virions (lanes 5 and 6) were subjected to immunoprecipitation with MAb 24/16 (αE0) or MAb A18 (αE2) as indicated in the figure. E0 is clearly detectable as a soluble protein in the virus-free medium. The precipitates were subjected to SDS-PAGE (8% acrylamide) under reducing conditions. The numbers on the left refer to molecular mass markers (Amersham) in kilodaltons.

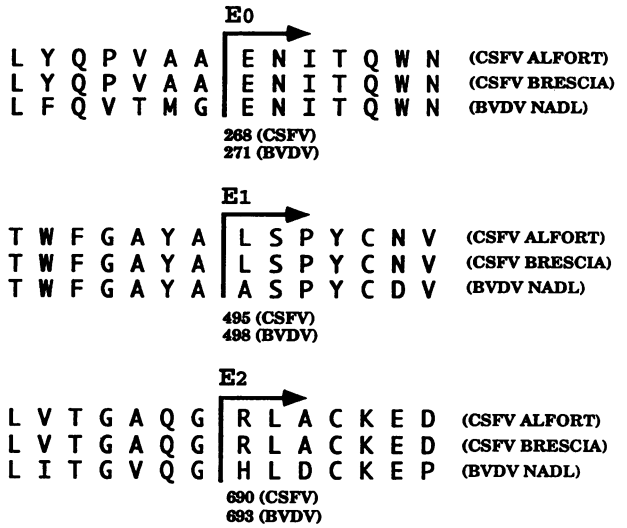

FIG. 5. Alignment of the amino acid sequences of CSFV Alfort and Brescia and BVDV NADL in the region of the cleavage sites determined for CSFV Alfort. The arrows indicate the start of each protein, and the numbers give the positions of the N-terminal amino acid in the polyproteins of CSFV or BVDV.

infected cells (Fig. 4, lanes 1 and 3) is compared with purified virions (lanes 5 and 6) and virion-depleted medium (lanes 2 and 4) from the same experiment. Significant amounts of soluble E0 were present in the virus-free medium as well as in the cell lysates and in purified virions. In contrast, E2 was only present in the cell lysate and in virions. The shedding of E0 into the medium is consistent with the absence of a membrane anchor in E0. Note the increase in the molecular masses of E0, E1, and E2 present in virions (Fig. 4, lanes 5 and 6) and virus-depleted supernatant (Fig. 4, lane 2) compared with those from cells. Most likely, changes in the carbohydrate composition during transit through the Golgi apparatus account for this shift. We have also found that soluble E0 is present in large amounts in the supernatant of VAC3.8-infected cells (data not shown).

Leu-495 (Fig. 3b) forms the N terminus of E1. If the protein extends to a C-terminal Gly-689, the size of E1 is 195 amino acids and its unmodified mass is 21.8 kDa. There are three sites for N glycosylation, all of which are probably occupied in CSFV Alfort. Two quite hydrophobic sequences in the protein probably serve as membrane anchors. The first anchor covers the sequence Leu-548 to Pro-579, and the second one covers Thr-659 to Gly-689. Presumably, the first anchor functions as a stop-transfer signal for E1 and the second anchor functions as a signal peptide for the translocation of E2, analogous to the situation in other enveloped viruses, such as those of the families *Togaviridae* and *Flaviviridae*, that also contain more than one membrane-spanning envelope protein.

The N terminus of E2 is formed by Arg-690 (Fig. 3c). The C terminus of E2 is not defined precisely, but functional studies with C-terminal truncations suggest that the C terminus is located around amino acid 1060 (12a, 19). From this estimate, E2 consists of about 370 amino acids (calculated molecular mass, 41 kDa), which conforms to previous studies on the size of the nonglycosylated E2 of BVDV (2). This would also indicate that the C-terminal domain of E2 contains at least 40 hydrophobic amino acids which probably form one or more membrane-spanning anchors.

Characteristics of the cleavage sites. The sequence surrounding the E1/E2 cleavage site containing Arg-690 conforms to a typical site cleaved by signal peptidase. There is a 16-amino-acid membrane-spanning hydrophobic region followed by residues defined by the −1, −3 rule (20). In computer-aided analysis (7), this processing site was predicted with a score of 8.8, which indicates there is a high probability that the site is in fact cleaved by signal peptidase. In view of the finding that E012 appears in VAC3.8-infected cells, cleavage at this site appears to be slightly delayed, rather than instantaneous as is the case with most signal peptidase cleavages.

In contrast to the situation with E2, the N terminus of E1 was not predicted correctly. The sequence upstream of Leu-495 contains the residues required by the −1, −3 rule for signalase cleavage but lacks the hydrophobic domain which commonly defines a signal peptide. It could represent an incomplete signal sequence, which accounts for the atypical, delayed signalase cleavage, or the sequence could be cleaved by a different enzyme, possibly a Golgi apparatus-associated protease.

The cleavage site Ala-267/Glu-268 that generates the N terminus of E0 is preceded by a hydrophobic sequence of 18 amino acids and resembles a typical signal peptidase cleavage site (20). Cleavage at this site appears to be analogous to the cleavage separating the flavivirus capsid and prM proteins, which is also believed to be catalyzed by signalase (11).

Pestiviruses show a high degree of sequence conservation among the sequenced strains. Figure 5 illustrates the peptide sequences of CSFV Alfort (9) and Brescia (10) and BVDV NADL (3) in the region of the processing sites determined by us for CSFV Alfort. The overall similarity in these regions suggests that processing of the E proteins of the other pestiviruses occurs at the homologous sites. Interestingly, the highest degree of conservation is found around the E0/E1 cleavage site, while the hydrophobic signal peptides upstream of E0 and E2 exhibit more divergence.

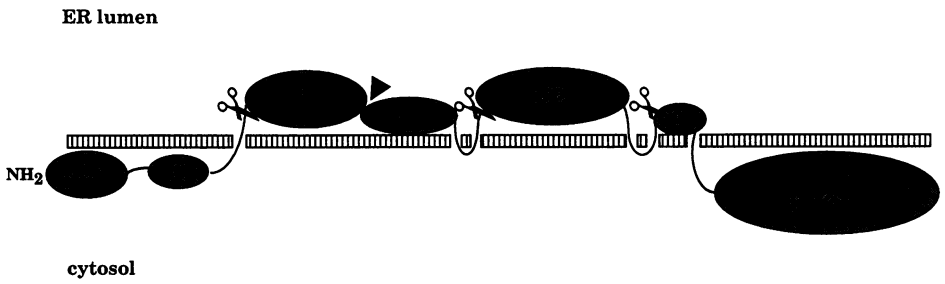

FIG. 6. Proposed processing scheme of pestiviral E proteins. Processing is initiated by the translocation of the signal peptide upstream of E0. Scissors indicate signalase cleavage; the solid wedge indicates cleavage by an unknown protease which cleaves the E0/E1 site. E1 bears two putative membrane-spanning regions, the first of which acts as a stop-transfer signal for E1 and the second of which functions as a translocation signal for E2. Preliminary studies indicate that the C terminus of E2 includes about 30 hydrophobic amino acids which could function as a stop-transfer signal for E2 and as a translocation signal for the downstream protein (12a, 19). The translation product downstream of E2 is hypothetical and has not yet been identified; Collett suggested that the N terminus of the nonstructural p125 was near amino acid 1130 (4), which would leave space for a 70-amino-acid peptide. The character of this sequence is hydrophobic, which probably renders the protein membrane associated. ER, endoplasmic reticulum.

DISCUSSION

Synthesis and processing of the glycoproteins. Three cleavage events in the formation of the glycoproteins, namely capsid/E0, E1/E2, and the cleavage at the C terminus of E2, are most likely mediated by host cell signalase. The composition of the sequences upstream of these cleavage sites conforms to the rules proposed by von Heijne (20). The −1, −3 rule postulates a small amino acid at position −1, usually Ala or Gly (50% and 24%, respectively). At the −3 position, the rule excludes aromatic, charged, or large polar residues. Here, Ala is most abundant (47%), followed by Val, Ser, Cys, and Thr. The −1 position of the CSFV (Alfort) E protein cleavage sites is Ala or Gly, and from amino acid sequence homology, Ala or Gly is also expected to be at the −1 position for the two other pestiviruses whose sequences are known. The −3 position is occupied in these sequences by Ala, Val, or Thr (Fig. 5).

We can now suggest a processing scheme for the production of pestiviral glycoproteins (Fig. 6). The first step is the translocation of the internal signal peptide downstream of the capsid protein. The cleavage between the capsid and the nascent E012 occurs rapidly upon translocation and is effected by signalase. Whether the signal sequence represents an integral part of the capsid protein, rendering it an "anchored core" as described for flaviviruses (11), is currently under investigation. Transfer of the nascent E012 is interrupted by two hydrophobic stretches at the end of E1—the first acts as a stop-transfer signal, and the second acts as a signal sequence to begin transfer of E2. The hydrophobic sequences also form an anchor for E1. Translocation of E2 then follows until interrupted by a set of stop/start transfer signals at the putative C terminus which anchors E2 and may also direct translocation of the next protein in the polyprotein. The cleavage between E012 and the downstream translation product apparently occurs rapidly, because there is no evidence for a larger precursor (Fig. 2). The internal processing of the E012 precursor is initiated by the release of E2. E012 is readily detectable in VAC3.8-infected cells (Fig. 2), and the time required for cleavage at the E1/E2 site appears to be slightly delayed, which is surprising because signalase cleavage is normally considered to occur instantly upon translocation. The release of E2 then allows cleavage at the E0/E1 site. It is unclear whether this hierarchical cleavage pattern occurs because E01 serves as a substrate for signalase only after the release of E2 permits a particular structure to be formed or whether E01 is transported to a different compartment, such as the Golgi apparatus, where it is processed by an as yet undefined protease. The E0/E1 cleavage has characteristics in common with the cleavage at the flavivirus NS1/NS2A site. The cleavage occurs downstream of a sequence similar to those commonly employed by signalase but lacks an upstream hydrophobic domain and is delayed. The flavivirus cleavage has been hypothesized to be effected by an uncharacterized protease in the Golgi apparatus (12). The pestivirus cleavage site does not have the consensus sequence of the flavivirus site, however, and further study is required to determine the enzyme responsible for cleavage.

It is interesting that the formation of a stable E1-E2 heterodimer occurs late, apparently after trimming of carbohydrate chains and thus after the E proteins have left the endoplasmic reticulum (Fig. 2). This could indicate that E1 and E2 associate only late, possibly at the location of virus assembly. Alternatively, E1 and E2 could form an unstable complex early which is later stabilized by formation of disulfide bonds or by formation of stronger noncovalent bonds. The glycoproteins present in the virion are exclusively hetero- or homodimeric; thus, dimerization could be important in the assembly of the virus. While E1 and E2 are almost certainly anchored in the lipid bilayer of the envelope, the mechanism by which E0 is associated with the virion remains unclear. Noncovalent interactions with the E1-E2 dimer or hydrophobic interactions with the lipid bilayer, which could conceivably include covalently linked lipids or glycolipids, could link E0 to the envelope. The fact that considerable amounts of E0 are shed into the medium indicates that the association of E0 with the envelope is more tenuous than is the association of E1 and E2.

Similarities of pestivirus, flavivirus, and hepatitis C virus glycoproteins. The properties of the E proteins of pestiviruses show analogies to those of other members of the family *Flaviviridae*. In the genus *Flavivirus*, there are two envelope proteins, prM and E, positioned immediately downstream of the capsid protein in the polyprotein precursor. The capsid protein is followed by an internal signal peptide that directs translocation of prM and results in an anchored core after signalase cleavage generates the N terminus of prM. Following prM is a stop/start transfer sequence analogous to that between pestivirus E01 and E2 which directs translocation of E and anchors prM after signalase cleavage generates the

N terminus of E. prM is later processed to M, which is C terminal in prM, and an N-terminal fragment, often referred to as pr, which lacks an anchor and is released into the medium; this cleavage occurs late, possibly in conjunction with virus maturation. While only M and E are considered structural proteins, prM is also detectable in "immature" virions as unprocessed prM (11). The flaviviral E protein shares functional similarity with E2 as the major antigen, and in its hydrophobicity M is reminiscent of E1. We suggest therefore that pestivirus E01 is analogous to prM of flaviviruses and that pestivirus E2 is analogous to flavivirus E. Although the cleavages of E01 and prM occur late and may be associated with virus maturation, different enzymes are responsible for the cleavages. Cleavage of prM occurs downstream of a dibasic sequence and is believed to be effected by a cathepsin-like protease present in the Golgi apparatus (1). We speculate that a Golgi-associated protease may also be responsible for the cleavage of E01, but the enzyme cleaves after amino acids with short side chains similar to those preferred by signalase. However, there are precedents for atypical signalase cleavages, and it remains possible that signalase is responsible for the E01 cleavage.

Following E2 in pestiviruses is a very hydrophobic protein reminiscent of the NS2A-NS2B complex of flaviviruses. In our model, then, there is no equivalent to flavivirus NS1 in pestiviruses. NS1 is a glycoprotein but is not incorporated into the virion. It is present on the cell surface and elicits a cytotoxic immune response. It is possible that flaviviral NS1 initially was an envelope glycoprotein that lost its function as a structural protein but remained indispensable for virus assembly or for some other required function.

The putative envelope proteins of hepatitis C virus, presently called E1 and E2/NS1, were recently identified by using different expression systems (8, 16). Initially termed E and NS1 on the basis of comparisons of the hydrophobicity plots of hepatitis C virus and of flavivirus proteins, both glycoproteins show features of envelope proteins. E2 of hepatitis C virus is very likely analogous to the pestiviral E2, and E1 of hepatitis C virus may be analogous to pestiviral E1. There appears to be no equivalent of pestiviral E0 in hepatitis C virus. In this model, hepatitis C virus also has no equivalent of flaviviral NS1. In the absence of amino acid sequence homology, it is not possible to determine which of the envelope proteins may have been derived from the same ancestral protein or to be certain whether the proteins E1 and E2 are functionally equivalent to flaviviral M and E. Collett et al. have suggested, for example, that E2 of the pestiviruses is equivalent to or arose from NS1 (4). Structural studies to resolve these issues will be very interesting.

While pestiviruses, flaviviruses, and hepatitis C virus undoubtedly share a common heredity, they show considerable divergence in their structural proteins, probably reflecting adaptation for replication in different host tissues and in response to immunological pressure. Divergence may have arisen in part because of a general permissiveness for the introduction of point mutations in this region, but it is also probable that recombinational events led to some of the apparent differences in the envelope proteins from the different genera of the family *Flaviviridae*.

ACKNOWLEDGMENTS

We thank E. Strauss for critical review of the manuscript and Petra Ullrich and Silke Mauritz for excellent technical assistance.

This work was supported by the Bundesministerium für Forschung und Technologie and Intervet International BV (project 0319028A), a fellowship from Deutsche Forschungsgemeinschaft (T.R.), and grant AI 20612 from the NIH to J.H.S.

REFERENCES

1. **Chambers, T. J., C. S. Hahn, R. Galler, and C. M. Rice.** 1990. Flavivirus genome organization, expression, and replication. Annu. Rev. Microbiol. **44**:649–699.
2. **Collett, M. S., R. Larson, S. Belzer, and E. Retzel.** 1988. Proteins encoded by bovine viral diarrhea virus: the genome organization of a pestivirus. Virology **165**:200–208.
3. **Collett, M. S., R. Larson, C. Gold, D. Strick, D. K. Anderson, and A. F. Purchio.** 1988. Molecular cloning and nucleotide sequence of the pestivirus bovine viral diarrhea virus. Virology **165**:191–199.
4. **Collett, M. S., V. Moennig, and M. C. Horzinek.** 1989. Recent advances in pestivirus research. J. Gen. Virol. **70**:253–266.
5. **Collett, M. S., M. A. Wiskerchen, E. Welniak, and S. K. Belzer.** 1991. Bovine viral diarrhea virus genomic organization. Arch. Virol. **3**:19–27.
6. **de Groot, R. J., T. Rümenapf, R. J. Kuhn, E. G. Strauss, and J. H. Strauss.** 1991. Sindbis virus RNA polymerase is degraded by the N-end rule pathway. Proc. Natl. Acad. Sci. USA **88**:8967–8971.
7. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
8. **Hijikata, M., N. Kato, Y. Ootsujama, M. Nakagawa, and K. Shimotono.** 1991. Gene mapping of the putative structural region of the hepatitis C virus genome by in vitro processing analysis. Proc. Natl. Acad. Sci. USA **88**:5547–5551.
9. **Meyers, G., T. Rümenapf, and H.-J. Thiel.** 1989. Molecular cloning and nucleotide sequence of the genome of hog cholera virus. Virology **171**:555–567.
10. **Moormann, R. J. M., P. A. M. Warmerdam, B. Van der Meer, W. M. M. Schaaper, G. Wensvoort, and M. M. Hulst.** 1990. Molecular cloning and nucleotide sequence of hog cholera virus strain Brescia and mapping of the genomic region encoding envelope glycoprotein E1. Virology **177**:184–198.
11. **Nowak, T., P. Färber, G. Wengler, and G. Wengler.** 1989. Analyses of the terminal sequences of West Nile virus structural proteins and of the in vitro translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavages involved in their synthesis. Virology **169**:365–376.
12. **Pethel, M., B. Falgout, and C. J. Lai.** 1992. Mutational analysis of the octapeptide sequence motif at the NS1-NS2A cleavage junction of dengue type 4 virus. J. Virol. **66**:7225–7231.

12a. **Rümenapf, T.** Unpublished observations.

13. **Rümenapf, T., G. Meyers, R. Stark, and H.-J. Thiel.** 1989. Hog cholera virus—characterization of specific antiserum and identification of cDNA clones. Virology **171**:18–27.
14. **Rümenapf, T., R. Stark, G. Meyers, and H.-J. Thiel.** 1991. Structural proteins of hog cholera virus expressed by vaccinia virus: further characterization and induction of protective immunity. J. Virol. **65**:589–597.
15. **Schägger, H., and G. von Jagow.** 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. **166**:368–379.
16. **Spaete, R. R., D. Alexander, M. E. Rugroden, Q. L. Choo, K. Berger, K. Crawford, C. Kou, S. Leng, S. Lee, R. Ralston, K. Thudium, J. W. Tung, and M. Houghton.** 1992. Characterization of the hepatitis C virus E2/NS1 gene product expressed in mammalian cells. Virology **188**:819–830.
17. **Stark, R., T. Rümenapf, G. Meyers, and H.-J. Thiel.** 1990. Genomic localization of hog cholera virus glycoproteins. Virology **174**:286–289.
18. **Thiel, H.-J., R. Stark, E. Weiland, T. Rümenapf, and G. Meyers.** 1991. Hog cholera virus: molecular composition of virions from a pestivirus. J. Virol. **65**:4705–4712.

18a. **Unger, G., and H.-J. Thiel.** Unpublished observations.

19. **van Zijl, M., G. Wensvoort, E. de Kluyver, M. Hulst, H. van der Gulden, A. Gielkens, A. Berns, and R. Moormann.** 1991. Live attenuated pseudorabies virus expressing envelope glycoprotein

E1 of hog cholera virus protects swine against both pseudorabies and hog cholera. J. Virol. **65**:2761–2765.
20. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14**:4683–4690.
21. **Weiland, E., R. Ahl, R. Stark, F. Weiland, and H.-J. Thiel.** 1992. A second envelope glycoprotein mediates neutralization of a pestivirus, hog cholera virus. J. Virol. **66**:3677–3682.
22. **Weiland, E., R. Stark, B. Haas, T. Rümenapf, G. Meyers, and H.-J. Thiel.** 1990. Pestivirus glycoprotein which induces neutralizing antibodies forms part of a disulfide-linked heterodimer. J. Virol. **64**:3563–3569.
23. **Wengler, G.** 1991. Family *Flaviviridae*, p. 223–233. *In* R. I. B. Francki, C. M. Fauquet, D. L. Knudson, and F. Brown (ed.), Classification and nomenclature of viruses. Fifth report of the International Committee on Taxonomy of Viruses. Springer Verlag, Berlin.
24. **Wiskerchen, M. A., S. K. Belzer, and M. S. Collett.** 1991. Pestivirus gene expression: the first protein product of the bovine viral diarrhea virus large open reading frame, P20, possesses proteolytic activity. J. Virol. **65**:4508–4514.
25. **Wiskerchen, M. A., and M. S. Collett.** 1991. Pestivirus gene expression: protein p80 of bovine viral diarrhea virus is a proteinase involved in polyprotein processing. Virology **184**: 341–350.