JOURNAL OF VIROLOGY, Oct. 1996, p. 6625–6633 Vol. 70, No. 10
0022-538X/96/$04.00+0

Processing of the Equine Arteritis Virus Replicase ORF1b Protein: Identification of Cleavage Products Containing the Putative Viral Polymerase and Helicase Domains

LEONIE C. VAN DINTEN,[1] ALFRED L. M. WASSENAAR,[1] ALEXANDER E. GORBALENYA,[2,3] WILLY J. M. SPAAN,[1] AND ERIC J. SNIJDER[1]*

Department of Virology, Institute of Medical Microbiology, Leiden University, 2333 AA Leiden, The Netherlands,[1] and M. P. Chumakov Institute of Poliomyelitis and Viral Encephalitides, Russian Academy of Medical Sciences, 142782 Moscow Region,[2] and A. N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, 119899 Moscow,[3] Russia

Received 13 February 1996/Accepted 25 June 1996

The replicase open reading frame 1b (ORF1b) protein of equine arteritis virus (EAV) is expressed from the viral genome as an ORF1ab fusion protein (345 kDa) by ribosomal frameshifting. Processing of the ORF1b polyprotein was predicted to be mediated by the nsp4 serine protease, the main EAV protease. Several putative cleavage sites for this protease were detected in the ORF1b polyprotein. On the basis of this tentative processing scheme, peptides were selected to raise rabbit antisera that were used to study the processing of the EAV replicase ORF1b polyprotein (158 kDa). In immunoprecipitation and immunoblotting experiments, processing products of 80, 50, 26, and 12 kDa were detected. Of these, the 80-kDa and the 50-kDa proteins contain the putative viral polymerase and helicase domains, respectively. Together, the four cleavage products probably cover the entire ORF1b-encoded region of the EAV replicase, thereby representing the first complete processing scheme of a coronaviruslike ORF1b polyprotein. Pulse-chase analysis revealed that processing of the ORF1b polyprotein is slow and that several large precursor proteins containing both ORF1a- and ORF1b-encoded regions are generated. The localization of ORF1b-specific proteins in the infected cell was studied by immunofluorescence. A perinuclear staining was observed, which suggests association with a membranous compartment.

Equine arteritis virus (EAV) (8) is the prototype of the arterivirus family (for a review, see reference 27), which also includes lactate dehydrogenase-elevating virus (15), porcine reproductive and respiratory syndrome virus (26), and simian hemorrhagic fever virus (16). The EAV RNA genome is 12.7 kb in length and is of positive polarity. The replicase gene covers the 5′-terminal three-quarters of the genome and is composed of two open reading frames, ORF1a and ORF1b (Fig. 1A) (8), that are expressed from the genomic RNA. ORF1b is expressed by means of a ribosomal frameshifting mechanism which produces a 345-kDa ORF1ab fusion protein of 3,175 amino acids. In the context of a reporter gene construct, the in vivo frameshifting efficiency was estimated to be between 15 and 20% (8).

The EAV replicase is processed extensively by at least three viral proteases (Fig. 1B). A preliminary processing scheme in which the 187-kDa ORF1a protein (1,727 amino acids) is cleaved into six major end products, nonstructural proteins 1 to 6 (nsp1 to nsp6), has been published (35). The N-terminal nsp1 contains a papainlike cysteine protease which cleaves the nsp1/2 site (34). The adjacent nsp2 contains a second cysteine protease activity that cleaves the nsp2/3 site (36). Both the nsp1 and nsp2 autoproteases act rapidly and probably exclusively in *cis*. The remaining nsp3456 precursor (96 kDa) is processed by a chymotrypsinlike serine protease (SP). This protease is located in nsp4 and was shown to be a representative of the relatively unique subgroup of 3C-like serine proteases (37). The SP cleaves the nsp4/5 site, a step which precedes the processing of the nsp3/4 and nsp5/6 sites (35, 37). Recently, a more detailed analysis of the processing of the EAV ORF1a protein has revealed that the nsp5 cleavage product from the scheme in Fig. 1B can be cleaved internally at three additional sites (42). Two of these additional cleavages have been attributed to the nsp4 SP. The protease carrying out the third novel cleavage within nsp5 remains to be identified (Fig. 1B).

The arterivirus replicase gene and replication strategy are evolutionarily related to those of coronaviruses and toroviruses (33). This is illustrated by the presence of a set of homologous replicase domains, the use of ribosomal frameshifting to express ORF1b, and the generation of similar nested sets of mRNAs to express the structural proteins encoded in the 3′-terminal region of the genome (Fig. 1A). The ORF1b-encoded part of the replicase contains the domains which are most conserved between arteriviruses and the members of the two genera of the coronavirus family, the coronaviruses and the toroviruses (8, 15, 33). Of these domains (Fig. 1B), the putative viral polymerase (28) and nucleoside triphosphate-binding/helicase motifs (18) are common to all positive-stranded RNA viruses. A possible metal-binding region (8, 19) is located between the polymerase and helicase motifs, and a fourth conserved domain that is unique for coronaviruslike viruses is found in the C-terminal region of the ORF1b protein (8, 32). The ORF1b polyprotein can thus be predicted to play a crucial role in the replication of coronaviruslike viruses, and knowledge about its posttranslational processing will be essential for its functional analysis.

By using a tentative processing scheme of the ORF1b protein, which was based on comparative sequence analysis, four

* Corresponding author. Mailing address: Department of Virology, Institute of Medical Microbiology, Leiden University, AZL gebouw 1, L4-Q, Postbus 9600, 2300 RC Leiden, The Netherlands. Phone: 31 71 5261657. Fax: 31 71 5266761. Electronic mail address: E.J.Snijder @Microbiology.MedFac.LeidenUniv.NL.

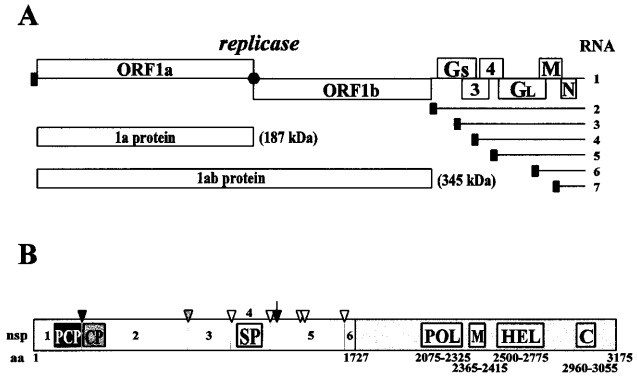

FIG. 1. Schematic representation of the EAV genome and replicase. (A) EAV genome organization and expression. The replicase 1a and 1ab polyproteins are expressed from the genomic RNA. The filled circle indicates the ribosomal frameshift site. The structural proteins are expressed from a nested set of subgenomic mRNAs, RNAs 2 to 7, that contain a common 5′ leader sequence (filled box). (B) Schematic representation of the replicase ORF1ab polyprotein. The ORF1a-encoded part of the protein is depicted in white. The three previously identified protease domains (see text) and corresponding cleavage sites (arrowheads) are shown. The arrow represents a cleavage by an as yet unidentified protease. The ORF1a-encoded cleavage products nsp1 to nsp6 (35) are indicated. The ORF1b-encoded part of the protein is shaded, and the locations of conserved domains (8) are shown. Amino acid (aa) numbers are indicated at the bottom. Abbreviations: PCP, papainlike cysteine protease; CP, cysteine protease; POL, putative polymerase domain; M, putative metal-binding region; HEL, putative helicase domain; C, conserved C-terminal domain.

ORF1b-specific peptide antisera have been raised. Using these tools, we have now been able to identify a set of cleavage products derived from the ORF1b-encoded part of the arterivirus replicase. Pulse-chase experiments revealed that cleavage of the EAV ORF1b polyprotein is extremely slow and yields many large processing intermediates. However, four end products of 80, 50, 26, and 12 kDa, which probably cover the entire ORF1b polypeptide, were detected. Immunofluorescence studies revealed that the ORF1b-encoded replicase proteins are located in the perinuclear region of the EAV-infected cell, where they are probably associated with a membranous compartment.

MATERIALS AND METHODS

Cells and virus. Baby hamster kidney (BHK-21) and rabbit kidney (RK-13) cells were used for propagation of the EAV Bucyrus strain (11) as described by de Vries et al. (9). EAV-infected Vero (African green monkey kidney), BHK-21, and RK-13 cells were used for immunofluorescence assays.

Preparation of antigens. The ORF1b-encoded amino acids of the EAV replicase will be indicated by their positions in the ORF1ab fusion protein (residues 1728 to 3175). Four EAV ORF1b protein-specific peptides (B1, B2, B3, and B4 [Table 1]) were synthesized and coupled to bovine serum albumin (BSA) as described by Snijder et al. (35). Each synthetic peptide was used to immunize two rabbits as described before (35). Rabbits were boosted repeatedly by using BSA-coupled or free peptide, after which they were bled. These final bleeds will be referred to as antisera αB1, αB2, αB3, and αB4, which were used in all experiments described in this report. The preimmune sera and antisera were tested in immunoprecipitation (Fig. 3) and immunoblotting (Fig. 4) analyses using lysates of EAV-infected and mock-infected RK-13 cells.

Radioactive labeling of EAV nonstructural proteins. RK-13 cells were infected with EAV (multiplicity of infection of 50) or were mock infected as described by de Vries et al. (9). Subsequently, they were incubated at 39.5°C and starved in medium without methionine and cysteine for 15 min prior to labeling. Proteins were labeled from 6 to 9 h postinfection (p.i.) with 200 μCi of [^{35}S]methionine and 80 μCi of [^{35}S]cysteine (Tran^{35}S-label; ICN Biomedicals Inc.) per ml of medium. For the immunoprecipitation experiment using the αB3 and the αB4 sera, 200 μCi of [^{35}S]cysteine (ICN Biomedicals) per ml of cysteine-free medium was used. For pulse-chase experiments, proteins were labeled at 8 h p.i., using a 30-min pulse with 1 mCi of [^{35}S]methionine and 400 μCi of [^{35}S]cysteine per ml of medium. Following the pulse, cells were washed twice with phosphate-buffered saline (PBS) and incubated in medium containing 2 mM each unlabeled methionine and cysteine.

Immunoprecipitation, SDS-PAGE, and Western blotting (immunoblotting). Cells were lysed as described by de Vries et al. (9) in lysis buffer containing the protease inhibitors phenylmethylsulfonyl fluoride (400 μM) and leupeptin (10 μM). The methods for immunoprecipitation, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), fluorography, and Western blotting were described previously (35).

Immunofluorescence assay. Vero, RK-13, and BHK-21 cells were grown on coverslips, infected with EAV (multiplicity of infection of approximately 1), and incubated at 39.5°C. Cells were fixed with 3% paraformaldehyde in PBS and washed with 10 mM glycine in PBS. Following permeabilization with 0.1% Triton X-100, indirect immunofluorescence assays were carried out with the αB1, αB2, αB3, or αB4 antiserum at a 1:100 dilution in PBS containing 5% fetal calf serum to reduce background fluorescence. A Cy3-conjugated donkey anti-rabbit immunoglobulin G (1:800 dilution; Jackson ImmunoResearch Laboratories) was used as the secondary antibody.

Fluorescence microscopy. Confocal fluorescence microscopy was performed with the confocal scanning laser beam microscope developed at the European Molecular Biology Laboratory, Heidelberg, Germany (38). The 529-nm laser line of an argon ion laser (Spectra-Physics), an Axiophot microscope (Zeiss), and a 100× Plan-neofluar objective (Zeiss) were used for imaging. The laser power, photomultiplier sensitivity, and number of averages (usually 32) were adjusted to generate images with sufficient contrast. Computer-stored images were recorded on Kodak TMX-100 film by using a freeze-frame camera.

RESULTS

Tentative ORF1b protein processing scheme and selection of peptides. Antisera raised against synthetic peptides were used as a sensitive tool to study the proteolytic processing of the EAV replicase ORF1b protein in infected cells. A tentative processing scheme for the ORF1b protein was generated. To this end, the primary structures of the ORF1b polyproteins of the related arteri-, corona-, and toroviruses (see the introduction) were compared. Details of this analysis will be presented elsewhere. These polyproteins are assumed to be related by common ancestry (8, 15, 33) and share the same organization of conserved domains. They have been predicted to be cleaved by the 3C-like protease encoded by ORF1a (15, 19, 21, 33). Three conserved potential cleavage sites for the EAV nsp4 SP were identified in the ORF1b protein sequence (Fig. 2), which were not identical to those previously predicted by Godeny et al. (15). By using the PHD program (29, 30), the secondary structure and water-accessible areas were predicted for the four cleavage products that were expected to be generated upon processing of the EAV ORF1b polyprotein. Peptides from potentially water-accessible and disordered areas were selected to raise the antibodies used in this study (Fig. 3B and Table 1). Peptide B1 represented the first 23 ORF1b-encoded residues of the ORF1ab polyprotein, encoded immediately downstream of the ORF1a/1b frameshift site. Peptides B2 and B3 were derived from the regions on either side of the predicted cleavage site downstream of the helicase domain (Glu-2837 ↓ Lys-2838). Peptide B4 represented replicase sequences from the C-terminal region of the ORF1b protein.

Production of ORF1b protein-specific rabbit antisera. Synthetic peptides (Table 1) were coupled to BSA and used to

TABLE 1. Overview of the antigens used to raise EAV ORF1b-specific antisera

Serum	Antigen[a]	Position in the ORF1ab protein	Amino acid sequence of the antigen
α5[b]	FP	1508–1727	RDAAR-212 amino acids-VNQLN
αB1	P	1728–1750	LRAPHIFPGDVGRRTFADSKDKG
αB2	P	2812–2827	YTHVPIKDGVIHSYPN
αB3	P	2875–2889	VVSNDRYPNCLQITL
αB4	P	3108–3125	EWALSTEPPPAGYAIVRR

[a] FP, bacterial glutathione *S*-transferase fusion protein; P, synthetic peptide.
[b] Described previously by Snijder et al. (35).

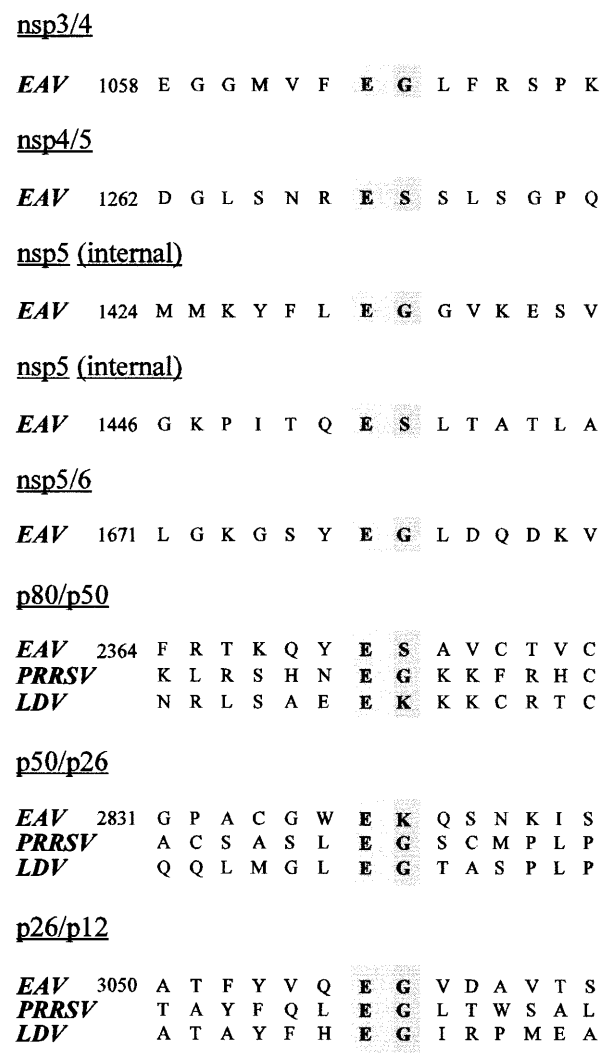

FIG. 2. Cleavage sites for the EAV 3C-like nsp4 protease in the ORF1ab polyprotein. The five previously described (37, 42) ORF1a protein cleavage sites are listed. Also shown is an alignment of the most probable cleavage sites in the ORF1b protein sequences of the arteriviruses EAV, porcine reproductive and respiratory syndrome virus (PRRSV), and lactate dehydrogenase-elevating virus (LDV). Sequences were extracted from the EMBL/GenBank database (accession numbers X53549, M96262, and U15146, respectively).

raise rabbit antisera directed against the ORF1b polyprotein. Each antigen was injected into two rabbits. For peptides B1, B2, and B4, the antisera obtained from the two different rabbits that had been immunized with the same antigen produced identical results. However, the signal/noise ratio varied depending on the serum and assay used. For peptide B3, only one of the two injected rabbits produced an ORF1b protein-specific antiserum. The EAV-specific proteins recognized by the antisera are discussed below.

The antisera were tested in immunoprecipitation (Fig. 3A) and Western blot (Fig. 4) analyses using EAV-infected RK-13 cell lysates and the conditions previously used for the analysis of the ORF1a protein processing (35). In both assays, preimmunization sera and uninfected cell lysates were used as negative controls. The previously described ORF1a protein-specific α5 serum (35) was used as a positive control. The most prominent bands recognized by this serum are nsp5 and the nsp3456 and nsp56 processing intermediates (Fig. 3A). Furthermore, under the mildly denaturing conditions used in this assay, the antiserum coimmunoprecipitates nsp2 (61 kDa), a property which is attributed to a previously described interaction between nsp2 and precursor proteins containing the nsp3 region (35).

Detection of ORF1b-encoded cleavage products. The proteins recognized by the ORF1b-specific antisera were analyzed in immunoprecipitation using cell lysates from a 6- to 9-h p.i. interval labeling (Fig. 3A). Furthermore, they were tested in Western blot analysis (Fig. 4) using unlabeled 9-h p.i. cell lysates. In both assays, each of the antisera recognized one protein that was not detected by any of the other sera. These cleavage products will be discussed in order of the position of the peptides in the ORF1b polyprotein sequence (Table 1). A number of precursor proteins that were recognized by multiple sera used for the experiments shown in Fig. 3A and 4 are discussed below.

The αB1 serum recognized an approximately 80-kDa protein, p80 (Fig. 3A and 4). Only in Western blot analysis (Fig. 4), this serum detected additional proteins of approximately 49, 47, and 30 kDa. Sera αB2, αB3, and αB4 recognized products of 50 kDa (p50), 26 kDa (p26), and 12 kDa (p12), respectively (Fig. 3A and 4). With the ORF1a protein-specific α5 serum, the proteins described above were not detected and were therefore concluded to be ORF1b encoded. The detection of a 61-kDa product after immunoprecipitation (Fig. 3A) was explained as coprecipitation of nsp2 by all ORF1b protein-specific antisera (see Discussion), a property previously described for a number of ORF1a protein-specific sera as well (35).

The tentative position in the ORF1b polyprotein of the products described above is indicated in Fig. 3B. Protein p80, which was recognized by serum αB1 but not by serum α5, was assumed to extend approximately 80 kDa into the ORF1b polyprotein and should therefore contain the putative polymerase domain (see also Discussion). In the C-terminal half of the ORF1b protein, the B2 and B3 peptides are separated by a segment of only 47 amino acid residues which contains one of the predicted cleavage sites (Fig. 2). Thus, it was not unexpected that the αB2 and αB3 sera each recognized a product (p50 and p26, respectively) that was not recognized by the other serum. Logically, most of the 50-kDa protein had to be derived from the region upstream of the B2 peptide. This product was therefore concluded to contain the putative helicase domain and is likely to be adjacent to p80 (Fig. 3B). Likewise, most of p26 had to be derived from the region downstream of the B3 peptide, and this product was assumed to adjoin p50. Finally, the 12-kDa protein, detected by the αB4 serum, was concluded to be the most C-terminal cleavage product of the ORF1b polyprotein (Fig. 3B). The origin of the 49-, 47-, and 30-kDa proteins recognized in immunoblotting by the αB1 serum remains unclear. These proteins were not recognized by the αB1 serum in immunoprecipitation.

Detection of ORF1ab-encoded precursor proteins. In immunoprecipitation (Fig. 3A) and immunoblot (Fig. 4) analyses, several precursor proteins were detected. The largest, of approximately 210 kDa (p210), was recognized by the four ORF1b protein-specific sera and by serum α5. The most prominent precursor protein in Fig. 3A, of about 190 kDa (p190), was recognized by the same set of antisera. In view of the coding capacity of ORF1b (158 kDa) and the recognition of these proteins by the α5 serum, it was concluded that these products contained the full-length ORF1b protein and a part of the ORF1a protein (see Discussion). Several minor precursor proteins, migrating just below p190, were assumed to be ORF1ab-encoded processing intermediates that were lacking either the C-terminal region of the ORF1b polyprotein or a

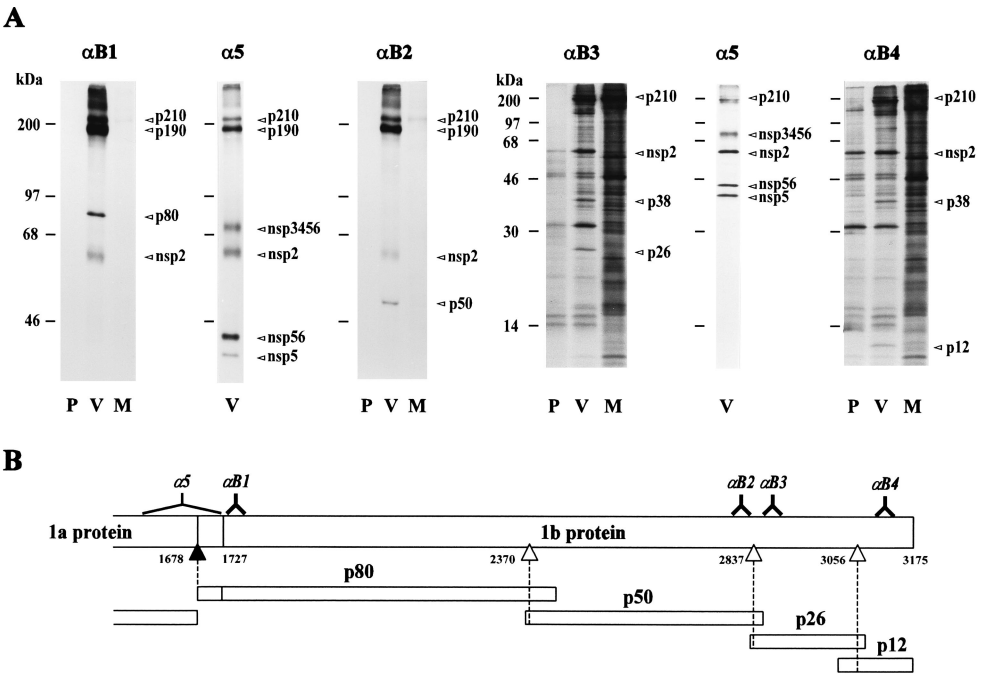

FIG. 3. (A) Immunoprecipitation analysis of EAV-infected and mock-infected RK-13 cell lysates, using the αB1, αB2, αB3, and αB4 sera. The positions of the molecular mass markers used during SDS-PAGE and of several EAV-specific proteins (see text) are indicated. The ORF1a protein-specific α5 serum was included as a positive control. For the α5 lanes, the autoradiographs were exposed for 10-times-shorter periods of time. Abbreviations: P, preimmune serum used on EAV-infected cell lysate; V, immune serum used on EAV-infected cell lysate; M, immune serum used on mock-infected cell lysate. (B) Provisional map of the ORF1b protein-specific cleavage products. The locations of the protein sequences used to raise the αB1, αB2, αB3, αB4, and α5 antisera are depicted. The white arrowheads indicate the putative cleavage sites in the ORF1b protein (Fig. 2), and the numbers of their P1 residues are given. The black arrowhead indicates the most C-terminal cleavage site in the ORF1a-encoded region of the ORF1ab polyprotein. For each of the newly detected cleavage products, the N- and C-terminal borders of the region from which they could originate are indicated. These boundaries were based on the estimated size (SDS-PAGE) of the proteins and the locations of the peptides used to raise the antisera with which the proteins were detected.

part of the ORF1a-encoded domain at their N termini (see Discussion). A protein of approximately 38 kDa (p38) was recognized by only two of the antisera, αB3 and αB4 (Fig. 3A and 4). This protein was therefore concluded to be a precursor of p26 and p12.

Pulse-chase analysis of ORF1ab polyprotein processing. To study the kinetics of ORF1ab polyprotein processing and to analyze precursor-product relationships, pulse-chase experiments were carried out with two of the ORF1b-specific sera. EAV-infected cells were pulse-labeled for 30 min at 8 h p.i. Subsequently, the labeled proteins were chased for up to 4 h and immunoprecipitated with sera αB1 and αB2 (Fig. 5B and C). The ORF1a protein-specific α5 serum was used to analyze ORF1ab-encoded precursor proteins (Fig. 5A).

Directly after the pulse-labeling, the p80 cleavage product was present (Fig. 5B). With time, the amount of p80 increased, and this increase coincided with the processing of precursor proteins migrating in the top of the gel (see below). Similar results were obtained with serum αB2 (Fig. 5C), although p50 was first observed at a later time point (1 h postpulse) than p80. Furthermore, small amounts of several products migrating between 40 and 47 kDa were precipitated by both the αB1 and αB2 sera (Fig. 5). This was attributed to coimmunoprecipitation of the ORF1a-encoded proteins nsp34, nsp56, and nsp5 (35). For p26 and p12, a pulse-chase analysis was not feasible because of the low incorporation of radioactivity.

After the pulse-labeling, most of the label was present in the largest precursor protein, p210. The amount of p210 first increased slightly and then slowly decreased, although it was still detectable after a 4-h chase period. The most abundant precursor band in Fig. 3A and 4, p190, turned out to be a doublet. Both p190 and the slightly smaller one, p180, were detected directly after the pulse. Their quantities first increased somewhat before decreasing slowly again (Fig. 5). Furthermore, we observed several precursor proteins that had not been detected in previous experiments (Fig. 3A and 4). The largest of these proteins, approximately 130 kDa (p130), was recognized by both the αB1 and αB2 sera (Fig. 5B and C). It was first observed after a 1-h chase period, and in view of its size, it was assumed to be a precursor of p80 and p50. A protein of approximately 120 kDa (p120) was observed with the α5 and αB1 sera (Fig. 5A and B) and was assumed to be a precursor protein containing nsp5 (41 kDa) and p80. Finally, a protein of approximately 90 kDa (p90 [Fig. 5C]) was detected directly after the pulse with the αB2 serum. We consider this product a potential precursor of the p50, p26, and p12 products.

Intracellular localization of ORF1b-encoded proteins. The intracellular localization of proteins containing the αB1 and αB2 epitopes was studied by confocal immunofluorescence microscopy. Three different cell lines, Vero, RK-13, and BHK-21, were used for this analysis, and the results obtained with two of them are shown in Fig. 6. For all three cell lines, the results obtained with the αB1 and αB2 sera were identical. A time course experiment showed that the first traces of ORF1b-encoded proteins could be detected at 4 h p.i. (data not shown). At 10 h p.i., a strong, somewhat punctate labeling of the perinuclear region in Vero cells (Fig. 6A and B) was observed, usually slightly polarized on one side of the nucleus. In EAV-infected RK-13 cells, this polarization of the signal was

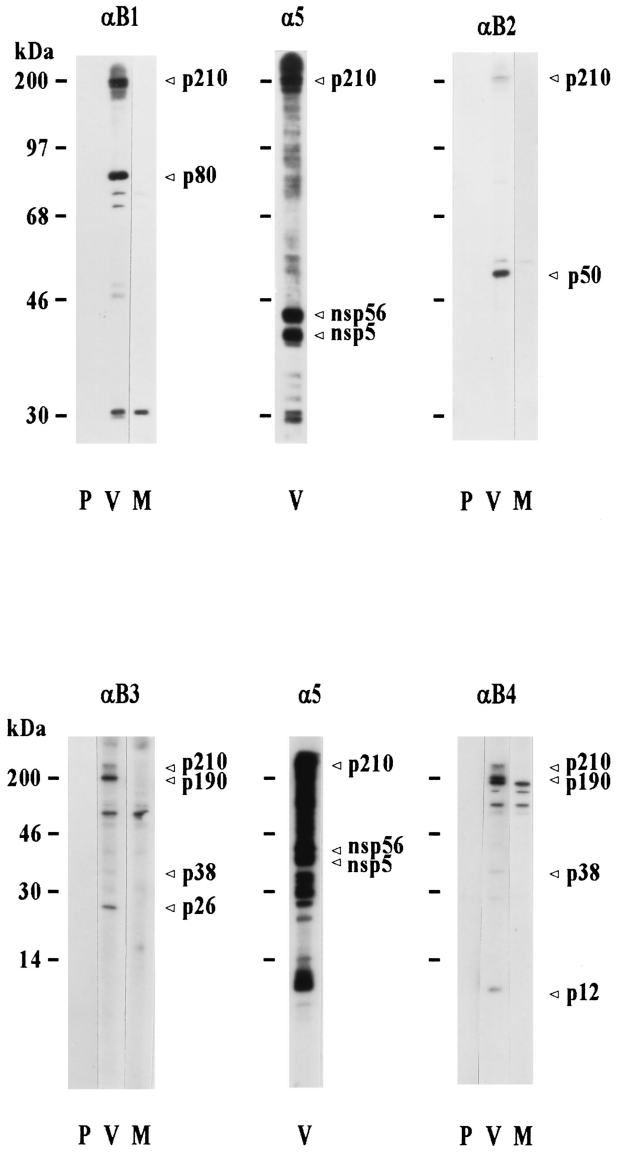

FIG. 4. Immunoblot analysis of EAV-infected and mock-infected RK-13 cell lysates, using the αB1, αB2, αB3, and αB4 sera. The ORF1a protein-specific α5 serum was included as a positive control. The positions of the molecular mass markers used during SDS-PAGE and of several EAV-specific proteins (see text) are indicated. Abbreviations are as for Fig. 3.

much more pronounced (Fig. 6C and D). In BHK-21 cells, relatively small but strongly labeled areas were seen on either side of the nucleus (data not shown). The observed pattern suggests that the ORF1b-encoded proteins are associated with a membranous compartment of the host cell, most likely the endoplasmic reticulum and/or intermediate compartment. The use of the αB3 serum gave rise to a nonspecific pattern throughout the cell (data not shown). A faint but specific signal, identical to that observed with the αB1 and αB2 sera, was obtained when we used the αB4 serum on EAV-infected RK-13 cells (data not shown).

DISCUSSION

Provisional map of mature processing products encoded by ORF1b. Studies on the proteolytic cleavage of the EAV replicase have revealed many features of the ORF1a polyprotein (35–37). In contrast, experimental data on the processing of the ORF1ab replicase protein were completely lacking. Using a tentative processing scheme and four novel antisera, we have now detected a set of processing end products that, in principle, covers the entire ORF1b-encoded region of the EAV replicase. The tentative position of p80, p50, p26, and p12 in the ORF1b polyprotein (Fig. 3B) was deduced from their estimated sizes and from the fact that each of these products reacted with only one of the four ORF1b protein-specific antisera. The provisional map, presented in Fig. 3B, was later supported by the sizes of a number of processing intermediates and by their reactivity with the various antisera (Fig. 5). Furthermore, the location of the predicted cleavage sites in the ORF1b protein (Fig. 2) is in agreement with the processing scheme presented in Fig. 3B and 7. Although only three candidate cleavage sites were identified and the sizes of the four processing end products add up to the approximate coding capacity of ORF1b, we cannot formally exclude the existence of additional, small cleavage products. The latter could be derived from parts of the polyprotein that are located between the four cleavage products described in this report.

As expected, processing of the ORF1b-encoded part of the EAV replicase does not occur within the well-defined conserved domains of the protein (Fig. 7) (8). Furthermore, the processing scheme of the EAV ORF1a protein (35) appears to apply to the ORF1a-encoded part of the ORF1ab frameshift protein as well. The nsp5/6 junction (between Glu-1677 and Gly-1678) (37) is assumed to be the most C-terminal cleavage site in the ORF1a protein. When the ORF1ab protein is cleaved at this site, products that are largely ORF1b encoded but contain the 50 most C-terminal residues of the ORF1a protein at their N termini will be generated. We assume that p80 and the p130 intermediate, which were detected with the αB1 serum, fall into this category (Fig. 7). In principle, these products could have been recognized by the α5 serum that was raised by using the C-terminal 220 amino acids of the ORF1a protein (35). The fact that this was not the case indicates that the major epitopes recognized by this serum are not located in the C-terminal 50 residues of the ORF1a protein. Assuming that Gly-1678 indeed is the N-terminal residue of p80, its size nicely fits into a processing scheme in which the p80 C terminus is generated by a cleavage between Glu-2370 and Ser-2371 (Fig. 2 and 3B).

Theoretically, p50 could occupy any position in the ORF1b protein that includes the B2 peptide (residues 2812 to 2827). However, p50 was not recognized by serum αB3, which was raised by using a sequence just downstream of the B2 peptide. The latter is therefore located close to the C terminus of p50. The p50 N terminus is located upstream of the helicase domain and is likely to be generated by the same cleavage that generates the C terminus of p80 (Fig. 3B). The region of the replicase downstream of p50 yields two small cleavage products, p26 and p12. Furthermore, a relatively abundant precursor of 38 kDa (p38) was observed (Fig. 3A and 7). Both this precursor and p26 contain the conserved C-terminal domain of unknown function that is specific for arteri-, corona-, and toroviruses.

Proteolytic cleavage of the ORF1ab polyprotein is slow and yields many processing intermediates. Our previous pulse-chase analysis of ORF1a protein processing (35) revealed that proteolysis of the nsp1/2 site was complete within a 15-min pulse-labeling period and that the liberation of nsp2 was only marginally slower. Subsequently, the nsp4/5 site in the nsp3456 precursor was cleaved at an intermediate rate. The nsp3/4 and nsp5/6 junctions were processed very slowly, and these cleavages were far from complete after a 3-h chase period. Thus, it

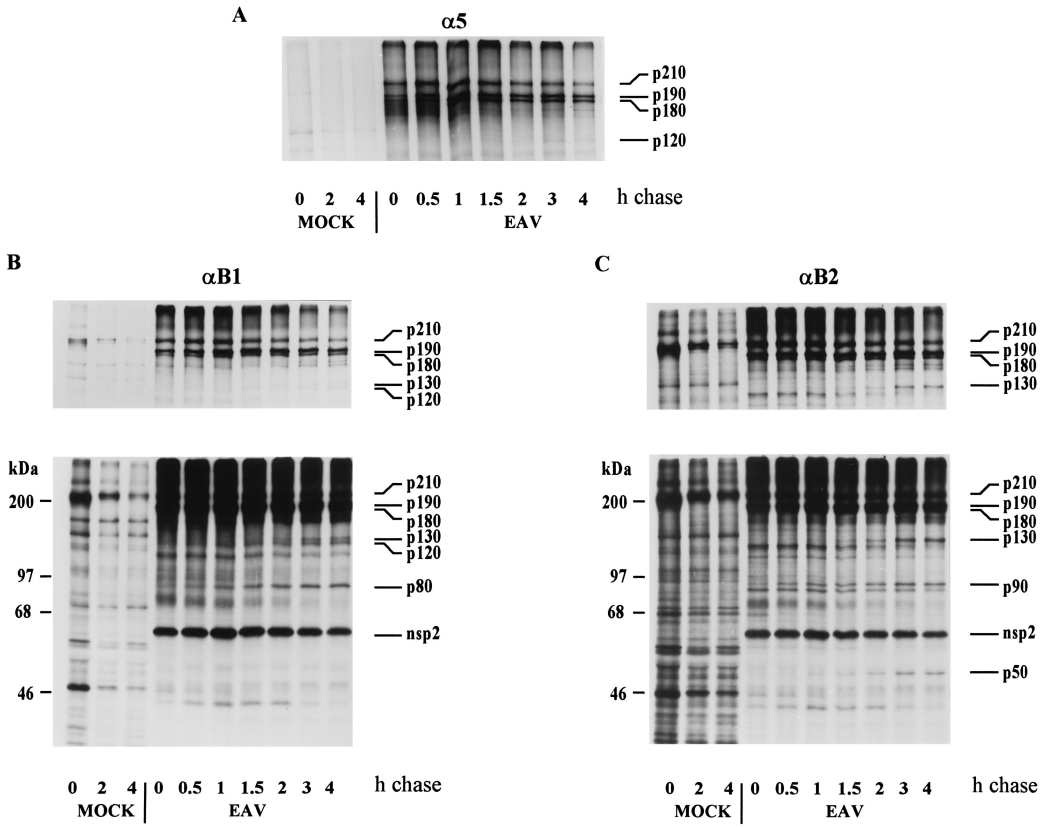

FIG. 5. Pulse-chase analysis of EAV ORF1ab protein processing. EAV- or mock-infected RK-13 cells were pulse-labeled for 30 min, and the label was chased for the various times indicated. (A) Immunoprecipitation using the ORF1a protein-specific α5 serum. Only the high-molecular-weight region of the gel is shown. (B and C) Immunoprecipitation analyses of the same samples used for panel A, using the αB1 and αB2 sera, respectively. The upper panel shows a short exposure of the high-molecular-weight region of the gel shown in the lower part.

could be predicted that ribosomal frameshifting and ORF1b translation (158 kDa) would yield extended versions of three polypeptides that contain the C terminus of the ORF1a protein: nsp3456, nsp56, and nsp6 (Fig. 7).

On the basis of its size and the results of pulse-chase experiments (Fig. 5), the p210 precursor most likely represents nsp3456 extended with the full-length ORF1b polyprotein (Fig. 7). There is a substantial difference between the calculated size of this product (255 kDa, starting at Gly-832) (37) and its estimated size derived from SDS-PAGE. Of course, a certain inaccuracy is inherent to estimating protein sizes from gel. However, it should be noticed that parts of the C-terminal half of the ORF1a protein are known to induce aberrant migration during SDS-PAGE (35, 37), which is exemplified by the migration of nsp3456 (estimated size, 70 kDa; calculated size, 96 kDa) in the gel shown in Fig. 3 (serum α5).

On the basis of our knowledge of ORF1a protein processing, the largest of the remaining set of precursor proteins, p190, most likely represents nsp56 extended with the full-length ORF1b protein sequence (Fig. 7) (35). A number of the less abundant precursors migrating just below p190, including p180 (Fig. 3A and 5), are likely to be generated by the additional processing of the C-terminal region of the ORF1a and ORF1ab proteins (37, 42). In view of the uncertainty about the exact number and locations of cleavage sites, we have, for the moment, refrained from numbering the ORF1b protein-derived cleavage products. The generation of a number of smaller processing intermediates (p130, p120, p90, and p38) could be explained by cleavages at the same three sites that yield the four ORF1b-encoded end products (Fig. 7).

Our pulse-chase analysis (Fig. 5) revealed that processing of the ORF1b polypeptide is slow and thus resembles the cleavage of some of the sites in the ORF1a protein (35). Moreover, the cleavages in the ORF1b protein do not appear to take place in a specific order. As a result, many processing intermediates arise, and several proteins containing one or more conserved domains are present in infected cells (Fig. 7). In view of the results obtained with other viral systems (20, 31, 39, 43), it is possible that these differentially processed replicase subunits fulfill different roles in the viral life cycle. Thus, proteolytic cleavages may either activate or inactivate specific functions during viral replication and transcription.

Proteolytic processing of the coronaviruslike replicase. The replicases of arteriviruses and coronaviruses contain a number of conserved domains, the presence of which formed the basis for the proposal of a coronaviruslike superfamily (8, 33). Despite the approximately twofold size difference between the replicase polyproteins of coronaviruses and arteriviruses, the similarities in replicase organization and expression are obvious: the order of the conserved domains is identical, and both virus groups utilize ribosomal frameshifting and extensive proteolytic processing during replicase expression. Most of the conserved domains are located in the ORF1b-encoded part of the replicase. The best-conserved domain in the ORF1a protein is a 3C-like protease, which is thought to be the main protease involved in replicase processing (33). Despite the fact

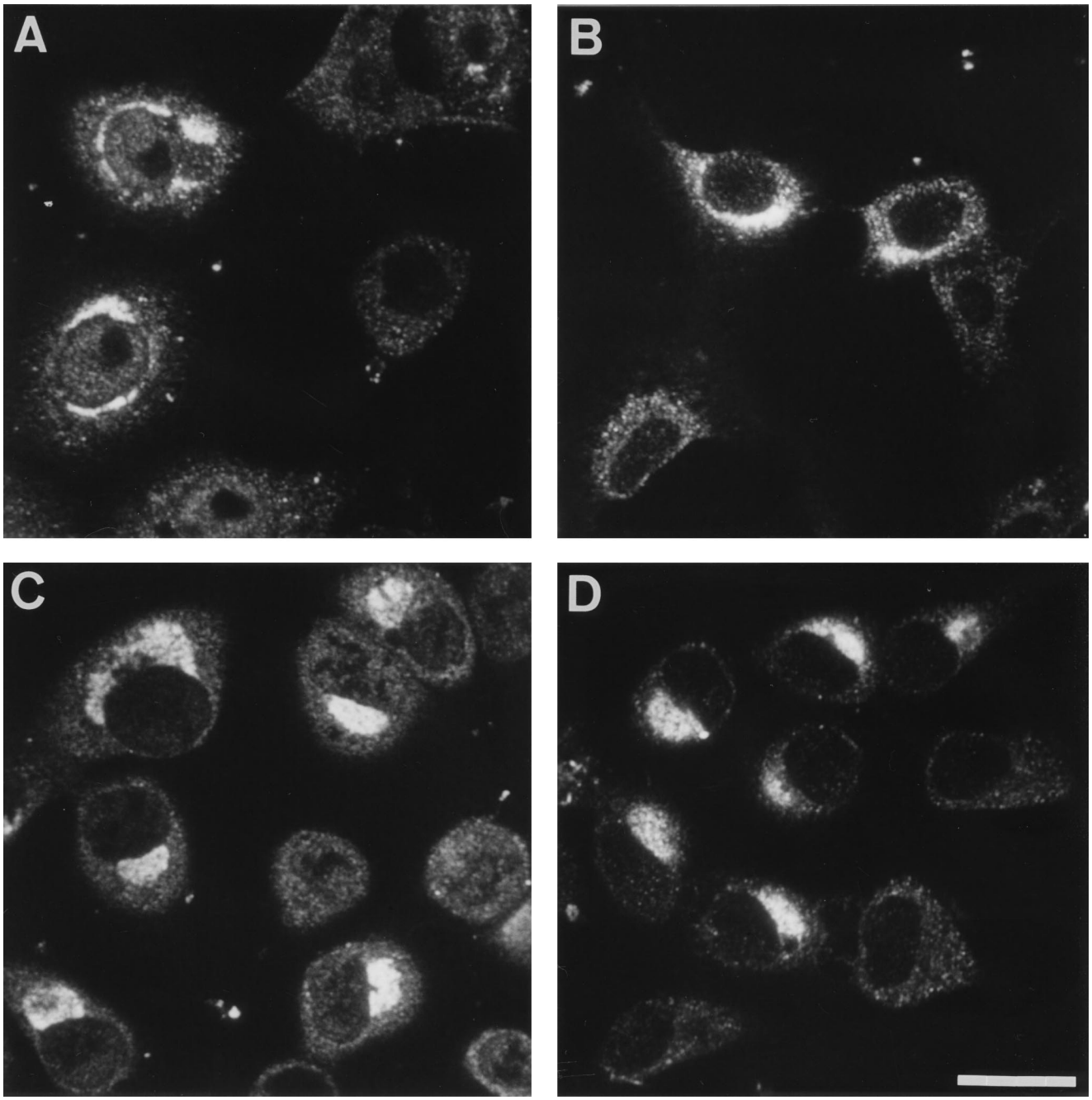

FIG. 6. Confocal immunofluorescence analysis of EAV-infected Vero (A and B) and RK-13 (C and D) cells. Labeling was carried out with the αB1 (A and C) or αB2 (B and D) serum. The bar represents approximately 22 μm.

that the principal catalytic residue of this 3C-like protease is Cys in coronaviruses and Ser in arteriviruses, it is likely that these proteolytic functions are related by common ancestry (37). The cleavage site specificity of both enzymes has recently been shown to be similar to that of the prototype 3C-like enzymes, the picornavirus 3C Cys proteases (for a review, see reference 12). The coronavirus 3C-like Cys protease can cleave at Gln ↓ Ala/Ser/Gly dipeptides (23, 24, 44), a finding which is consistent with previous cleavage site predictions (13, 17, 21). Evidence that the EAV nsp4 3C-like Ser protease cleaves five Glu ↓ Gly/Ser/Ala junctions in the C-terminal half of the ORF1a protein has recently been obtained (Fig. 1 and 2) (37, 42). Conserved amino acid residues in the substrate-binding region of 3C-like proteases are thought to determine their specificity for cleavage sites with Gln or Glu at the P1 position (25, 37).

No putative protease domains have been detected in the ORF1b polyproteins of arteriviruses and coronaviruses. For infectious bronchitis virus, the prototype coronavirus, a largely ORF1b-encoded 100-kDa polypeptide that contains the putative viral polymerase domain has been identified (22). It was shown that this product probably results from cleavages carried

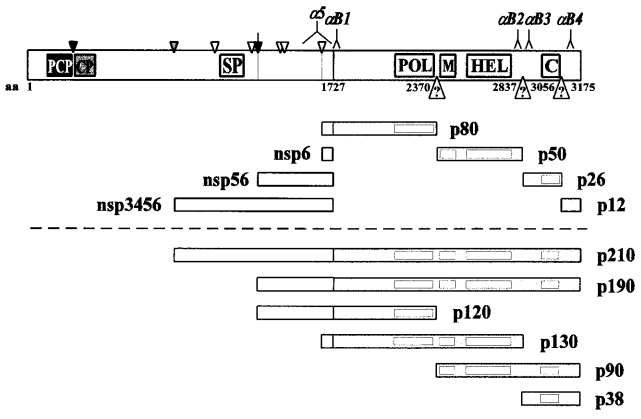

FIG. 7. Tentative processing scheme for the EAV ORF1ab polyprotein. See the legend to Fig. 1 for symbols and abbreviations. The arrowheads with question marks indicate putative cleavage sites (see text), and the numbers of their P1 residues are given. The locations of three cleavage products that contain the ORF1a C terminus (nsp3456, nsp56, and nsp6) are indicated. The putative positions in the lab polyprotein of the processing products identified by using the αB1, αB2, αB3, and αB4 sera are shown. Below the dotted line, the putative positions of several ORF1ab-encoded processing intermediates are indicated.

out by the 3C-like Cys protease. Two potential Gln ↓ Ser cleavage sites were identified just upstream of the ORF1a/1b frameshift region and downstream of the putative polymerase domain (22, 23). Its presumed location in the infectious bronchitis virus replicase implies that this 100-kDa coronavirus cleavage product is the equivalent of the EAV p80 protein described in this report. Furthermore, preliminary results from coexpression of the wild-type and inactivated EAV nsp4 protease and the ORF1b polyprotein have confirmed that also in arteriviruses, the 3C-like protease is responsible for processing of the ORF1b polyprotein (41).

The information summarized above enabled us to use comparative sequence analysis (Fig. 2) to generate a tentative processing scheme for the EAV ORF1b protein, which contains three putative cleavage sites for the nsp4 SP. This model was subsequently used to select peptides to raise antibodies against ORF1b-encoded cleavage products. The four processing end products that were described in this report fit perfectly into the tentative processing scheme (Fig. 7). The candidate cleavage sites are currently being tested by site-directed mutagenesis and expression of cDNA constructs. Our predictions imply that the total number of cleavages in the EAV replicase is (at least) 11 and that 8 of these processing steps are carried out by the 3C-like SP in nsp4.

Localization of the EAV replicase complex. The subcellular localization of ORF1b-encoded proteins was studied by confocal immunofluorescence microscopy using the αB1 and αB2 sera (Fig. 6). The observed perinuclear signal suggests that proteins containing the putative viral polymerase and helicase domains (precursors, end products, or both) are associated with a membranous compartment, most likely (parts of) the endoplasmic reticulum or intermediate compartment. Since the proteins recognized by the αB1 and αB2 sera are likely to play a key role in viral replication and transcription, we predict that their localization is identical to that of the viral replication complex in the infected cell.

The replication complexes of many animal and plant positive-stranded RNA viruses have been shown to be membrane associated (2, 5, 7, 14), but information on coronaviruslike viruses was lacking thus far. Different intracellular membranes are used and modified by different viruses, probably to generate a structural scaffold for the formation of the replication complex. Furthermore, membrane association may create a protective microenvironment that facilitates the various replicative processes that have to be carried out by viral nonstructural proteins. A particularly well studied replication complex is that of the picornavirus poliovirus (1, 3, 4, 6, 10). Poliovirus infection stimulates lipid synthesis and induces the accumulation of virus-specific vesicles that carry the replication complex. Various subunits of the poliovirus polyprotein have been shown to interact with membranes. When these proteins are expressed independently, they can induce vesicle formation and can severely affect membrane trafficking in the host cell.

Computer analysis of the ORF1b-encoded polyproteins of arteriviruses and coronaviruses did not predict any transmembrane hydrophobic domains. However, such domains are present in several of their ORF1a-derived cleavage products. In the case of EAV, they are found in nsp2 and in nsp3 and nsp5, on either side of the 3C-like nsp4 protease domain (35). Interestingly, similar hydrophobic regions flank the 3C-like protease domain of coronaviruses. With EAV ORF1a-specific antisera, an immunofluorescence pattern identical to the signal obtained with the new ORF1b-specific sera is obtained (Fig. 6) (40). This finding implies that most replicase cleavage products assemble into a large complex that is probably anchored in the membrane through the hydrophobic regions in the ORF1a protein. Furthermore, coimmunoprecipitation of several ORF1a-derived (nsp2, nsp34, nsp56, and nsp5 [Fig. 5]) proteins by the ORF1b protein-specific sera indicates formation of complexes containing ORF1a- and ORF1b-derived proteins. Experiments to characterize the EAV replication complex in more detail and to analyze the role of individual cleavage products are in progress.

ACKNOWLEDGMENTS

We thank Johan den Boon for technical assistance and useful comments. We thank the staff of the test animal facility (CPV) of the Leiden Academic Hospital. We thank Jan Wouter Drijfhout (Department of Immunohaematology, Leiden University) for peptide synthesis. E.J.S. thanks the Griffiths laboratory at the European Molecular Biology Laboratory (Heidelberg, Germany) for their hospitality. We are indebted to Yvonne van der Meer for assistance with immunofluorescence microscopy and related photographic work.

A.E.G. was supported in part by a grant from the Russian Foundation for Basic Research.

REFERENCES

1. **Barco, A., and L. Carrasco.** 1995. A human virus protein, poliovirus protein 2BC, induces membrane proliferation and blocks the exocytic pathway in the yeast *Saccharomyces cerevisiae*. EMBO J. **14:**3349–3364.
2. **Bienz, K., D. Egger, and T. Pfister.** 1994. Characteristics of the poliovirus replication complex. Arch. Virol. Suppl. **9:**147–157.
3. **Bienz, K., D. Egger, T. Pfister, and M. Troxler.** 1992. Structural and functional characterization of the poliovirus replication complex. J. Virol. **66:** 2740–2747.
4. **Bienz, K., D. Egger, M. Troxler, and L. Pasamontes.** 1990. Structural organization of poliovirus RNA replication is mediated by viral proteins of the P2 genomic region. J. Virol. **64:**1156–1163.
5. **Chambers, T. J., C. S. Hahn, R. Galler, and C. M. Rice.** 1990. Flavivirus genome organization, expression, and replication. Annu. Rev. Microbiol. **44:**649–688.
6. **Cho, M. W., N. Teterina, D. Egger, K. Bienz, and E. Ehrenfeld.** 1994. Membrane rearrangement and vesicle induction by recombinant poliovirus 2C and 2BC in human cells. Virology **202:**129–145.
7. **de Graaff, M., and E. M. J. Jaspers.** 1994. Plant viral RNA synthesis in cell-free systems. Annu. Rev. Phytopathol. **32:**311–335.
8. **den Boon, J. A., E. J. Snijder, E. D. Chirnside, A. A. F. de Vries, M. C. Horzinek, and W. J. M. Spaan.** 1991. Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. J. Virol. **65:**2910–2920.
9. **de Vries, A. A. F., E. D. Chirnside, M. C. Horzinek, and P. J. M. Rottier.** 1992. Structural proteins of equine arteritis virus. J. Virol. **66:**6294–6303.
10. **Doedens, J. R., and K. Kirkegaard.** 1995. Inhibition of cellular protein

secretion by poliovirus proteins 2B and 3A. EMBO J. **14:**894–907.
11. **Doll, E. R., J. T. Bryans, W. H. M. McCollum, and M. E. Wallace.** 1957. Isolation of a filterable agent causing arteritis of horses and abortion of mares. Its differentiation from the equine (abortion) influenza virus. Cornell Vet. **47:**3–41.
12. **Dougherty, W. G., and B. L. Semler.** 1993. Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. Microbiol. Rev. **57:**781–822.
13. **Eleouet, J., D. Rasschaert, P. Lambert, L. Levy, P. Vende, and H. Laude.** 1995. Complete sequence (20 kilobases) of the polyprotein-encoding gene 1 of transmissible gastroenteritis virus. Virology **206:**817–822.
14. **Froshauer, S., J. Kartenbeck, and A. Helenius.** 1988. Alphavirus RNA replicase is located on the cytoplasmic surface of endosomes and lysosomes. J. Cell Biol. **107:**2075–2086.
15. **Godeny, E. K., L. Chen, S. N. Kumar, S. L. Methven, E. V. Koonin, and M. A. Brinton.** 1993. Complete genomic sequence and phylogenetic analysis of the lactate dehydrogenase-elevating virus (LDV). Virology **194:**585–596.
16. **Godeny, E. K., L. Zeng, S. L. Smith, and M. A. Brinton.** 1995. Molecular characterization of the 3′ terminus of the simian hemorrhagic fever virus genome. J. Virol. **69:**2679–2683.
17. **Gorbalenya, A. E., A. P. Donchenko, V. M. Blinov, and E. V. Koonin.** 1989. Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. A distinct protein superfamily with a common structural fold. FEBS Lett. **243:**103–114.
18. **Gorbalenya, A. E., and E. V. Koonin.** 1993. Comparative analysis of the amino acid sequences of the key enzymes of the replication and expression of positive-strand RNA viruses. Validity of the approach and functional and evolutionary implications. Sov. Sci. Rev. D Physicochem. Biol. **11:**1–84.
19. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov.** 1989. Coronavirus genome: prediction of putative functional domains in the nonstructural polyprotein by comparative amino acid sequence analysis. Nucleic Acids Res. **17:**4847–4861.
20. **Jore, J., B. De Geus, R. J. Jackson, P. H. Pouwels, and B. E. Enger-Valk.** 1988. Poliovirus protein 3CD is the active protease for processing of the precursor protein P1 in vitro. J. Gen. Virol. **69:**1627–1636.
21. **Lee, H. J., C. K. Shieh, A. E. Gorbalenya, E. V. Koonin, N. La Monica, J. Tuler, A. Bagdzhadzhyan, and M. M. Lai.** 1991. The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. Virology **180:**567–582.
22. **Liu, D. X., I. Brierley, K. W. Tibbles, and T. D. K. Brown.** 1994. A 100-kilodalton polypeptide encoded by open reading frame (ORF) 1b of the coronavirus infectious bronchitis virus is processed by ORF 1a products. J. Virol. **68:**5772–5780.
23. **Liu, D. X., and T. D. K. Brown.** 1995. Characterisation and mutational analysis of an ORF 1a-encoding proteinase domain responsible for proteolytic processing of the infectious bronchitis virus 1a/1b polyprotein. Virology **209:**420–427.
24. **Lu, Y., X. Lu, and M. R. Denison.** 1995. Identification and characterization of a serine-like proteinase of the murine coronavirus MHV-A59. J. Virol. **69:**3554–3559.
25. **Matthews, D. A., W. W. Smith, R. A. Ferre, B. Condon, G. Budahazi, W. Sisson, J. E. Villafranca, C. A. Janson, H. E. McElroy, and C. L. Gribskov.** 1994. Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polyprotein. Cell **77:**761–771.
26. **Meulenberg, J. J. M., M. M. Hulst, E. J. de Meijer, P. L. Moonen, A. den Besten, E. P. De Kluyver, G. Wensvoort, and R. J. M. Moormann.** 1993. Lelystad virus, the causative agent of porcine epidemic abortion and respiratory syndrome (PEARS), is related to LDV and EAV. Virology **192:**62–72.
27. **Plagemann, P. G. W., and V. Moennig.** 1992. Lactate dehydrogenase-elevating virus, equine arteritis virus, and simian hemorrhagic fever virus: a new group of positive-strand RNA viruses. Adv. Virus Res. **41:**99–192.
28. **Poch, O., I. Sauvaget, M. Delarue, and N. Tordo.** 1989. Identification of four conserved motifs among the RNA dependent polymerase encoding elements. EMBO J. **8:**3867–3874.
29. **Rost, B.** 1996. PHD: predicting one-dimensional protein structure by profile based neural networks. Methods Enzymol. **266:**525–539.
30. **Rost, B., and C. Sander.** 1994. Conservation and prediction of solvent accessibility in protein families. Proteins **20:**216–226.
31. **Shirako, Y., and J. H. Strauss.** 1994. 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.
32. **Snijder, E. J., J. A. den Boon, P. J. Bredenbeek, M. C. Horzinek, R. Rijnbrand, and W. J. M. Spaan.** 1990. The carboxyl-terminal part of the putative Berne virus polymerase is expressed by ribosomal frameshifting and contains sequence motifs which indicate that toro- and coronaviruses are evolutionarily related. Nucleic Acids Res. **18:**4535–4542.
33. **Snijder, E. J., and W. J. M. Spaan.** 1995. The coronaviruslike superfamily, p. 239–255. *In* S. G. Siddell (ed.), The Coronaviridae. Plenum Press, New York.
34. **Snijder, E. J., A. L. M. Wassenaar, and W. J. M. Spaan.** 1992. The 5′ end of the equine arteritis virus replicase gene encodes a papainlike cysteine protease. J. Virol. **66:**7040–7048.
35. **Snijder, E. J., A. L. M. Wassenaar, and W. J. M. Spaan.** 1994. Proteolytic processing of the replicase ORF1a protein of equine arteritis virus. J. Virol. **68:**5755–5764.
36. **Snijder, E. J., A. L. M. Wassenaar, W. J. M. Spaan, and A. E. Gorbalenya.** 1995. The arterivirus Nsp2 protease. An unusual cysteine protease with primary structure similarities to both papain-like and chymotrypsin-like proteases. J. Biol. Chem. **270:**16671–16676.
37. **Snijder, E. J., A. L. M. Wassenaar, L. C. van Dinten, W. J. M. Spaan, and A. E. Gorbalenya.** 1996. The arterivirus nsp4 protease is the prototype of a novel group of chymotrypsin-like enzymes, the 3C-like serine proteases. J. Biol. Chem. **271:**4864–4871.
38. **Stelzer,E. H., I. Wacker, and J. R. DeMey.** 1991. Confocal fluorescence microscopy in modern cell biology. Semin. Cell Biol. **2:**145–152.
39. **Strauss, J. H., and E. G. Strauss.** 1994. The alphaviruses: gene expression, replication, and evolution. Microbiol. Rev. **58:**491–562.
40. **van der Meer, Y., G. J. van Tol, G. Griffiths, J. Krijnse Locker, and E. J. Snijder.** Unpublished data.
41. **van Dinten, L. C., A. E. Gorbalenya, W. J. M. Spaan, and E. J. Snijder.** Unpublished data.
42. **Wassenaar, A. L. M., W. J. M. Spaan, and E. J. Snijder.** Unpublished data.
43. **Ypma-Wong, M. F., P. G. Dewalt, V. H. Johnson, J. G. Lamb, and B. L. Semler.** 1988. Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. Virology **166:**265–270.
44. **Ziebuhr, J., J. Herold, and S. G. Siddell.** 1995. Characterization of a human coronavirus (strain 229E) 3C-like proteinase activity. J. Virol. **69:**4331–4338.