Proteolytic Processing of the Open Reading Frame 1b-Encoded Part of Arterivirus Replicase Is Mediated by nsp4 Serine Protease and Is Essential for Virus Replication

LEONIE C. VAN DINTEN,¹ SIETSKE RENSEN,¹ ALEXANDER E. GORBALENYA,^{1,2,3}
AND ERIC J. SNIJDER^{1*}

Department of Virology, Leiden University Medical Center, Leiden, The Netherlands¹; M. P. Chumakov Institute of Poliomyelitis and Viral Encephalitis, Russian Academy of Medical Sciences, 142782 Moscow Region, Russia²; and Advanced Biomedical Computer Center, SAIC/NCI-FCRDC, Frederick, Maryland 21702-1201³

Received 2 September 1998/Accepted 10 December 1998

The open reading frame (ORF) 1b-encoded part of the equine arteritis virus (EAV) replicase is expressed by ribosomal frameshifting during genome translation, which results in the production of an ORF1ab fusion protein (345 kDa). Four ORF1b-encoded processing products, nsp9 (p80), nsp10 (p50), nsp11 (p26), and nsp12 (p12), have previously been identified in EAV-infected cells (L. C. van Dinten, A. L. M. Wassenaar, A. E. Gorbalenya, W. J. M. Spaan, and E. J. Snijder, *J. Virol.* 70:6625–6633, 1996). In the present study, the generation of these four nonstructural proteins was shown to be mediated by the nsp4 serine protease, which is the main viral protease (E. J. Snijder, A. L. M. Wassenaar, L. C. van Dinten, W. J. M. Spaan, and A. E. Gorbalenya, *J. Biol. Chem.* 271:4864–4871, 1996). Mutagenesis of candidate cleavage sites revealed that Glu-2370/Ser, Gln-2837/Ser, and Glu-3056/Gly are the probable nsp9/10, nsp10/11, and nsp11/12 junctions, respectively. Mutations which abolished ORF1b protein processing were introduced into a recently developed infectious cDNA clone (L. C. van Dinten, J. A. den Boon, A. L. M. Wassenaar, W. J. M. Spaan, and E. J. Snijder, *Proc. Natl. Acad. Sci. USA* 94:991–997, 1997). An analysis of these mutants showed that the selective blockage of ORF1b processing affected different stages of EAV reproduction. In particular, the mutant with the nsp10/11 cleavage site mutation Gln-2837→Pro displayed an unusual phenotype, since it was still capable of RNA synthesis but was incapable of producing infectious virus.

Equine arteritis virus (EAV) (12) is a positive-stranded RNA virus (12.7-kb genome) (8) which belongs to the family *Arteriviridae* (for reviews, see references 35 and 42) together with lactate dehydrogenase-elevating virus (LDV), porcine reproductive and respiratory syndrome virus (PRRSV), and simian hemorrhagic fever virus (SHFV). Based on their similar genome organizations and expression strategies and the presumed common ancestry of their replicase proteins (for reviews, see references 11 and 43), the arteriviruses have recently been united with the coronaviruses in the order *Nidovirales* (5). Like many positive-stranded RNA viruses, nidoviruses regulate their gene expression by synthesizing multidomain precursor proteins, which are subsequently processed into smaller subunits by specific virus-encoded proteases (for reviews, see references 6, 11, 13, 20, 37, 42, and 48).

The 5' three-quarters of the arterivirus genome contains two large replicase open reading frames (ORF1a and ORF1b), which are followed by a set of smaller genes encoding mostly structural proteins (10). For EAV, ORF1a encodes a polypeptide of 187 kDa (1,727 amino acids [aa]). ORF1b is expressed upon ribosomal frameshifting (estimated efficiency, 15 to 20% [8]), which results in the production of a 345-kDa ORF1ab polypeptide (3,175 aa). The ORF1a and ORF1ab proteins are proteolytically processed into 8 and 11 nonstructural proteins (nsp's), respectively (Fig. 1) (45, 47, 55, 56). During this process, a number of processing intermediates are generated, and they may also play specific roles during EAV replication.

The ORF1a protein contains three virus-specific proteases. nsp1 and nsp2 both have cysteine autoprotease activities, which are responsible for rapid cleavages at the nsp1/2 and nsp2/3 sites, respectively (Fig. 1) (44, 46). The chymotrypsin-like serine protease (SP) activity of nsp4 was found to be the main protease involved in processing of the ORF1a polyprotein. It is a representative of the subgroup of 3C-like serine proteases (47), which, like the picornavirus 3C-like cysteine proteases, have a substrate specificity for cleavage sites that contain a Gln or Glu at the P1 position and a small amino acid residue (Gly, Ser, or Ala) at the P1' position (2, 13, 17, 47). (The nomenclature for the substrate amino acid residues is Pn, . . . , P2, P1, P1', P2', . . . , Pn', where P1/P1' depicts the cleaved bond). The EAV nsp4 SP was previously shown to cleave dipeptides carrying Glu at the P1 position and Ser or Gly at the P1' position. In the majority of the EAV ORF1a polyproteins, the SP first cleaves at the nsp4/5 site (Glu-1268/Ser), followed by cleavage of the nsp3/4 (Glu-1064/Gly) and nsp7/8 (Glu-1677/Gly) junctions (Fig. 1) (45, 47). Recently, it was shown that processing of the nsp4/5 cleavage site by the nsp4 SP requires the association of cleaved nsp2 with the C-terminal half of the ORF1a protein (nsp3-8) (56). This detailed analysis of EAV ORF1a protein processing revealed an alternative, minor processing pathway in which the nsp4/5 site is not cleaved. Instead, two cleavage sites in the C-terminal region of the ORF1a protein are processed: the nsp5/6 site (Glu-1430/Gly) and the nsp6/7 site (Glu-1452/Ser) (Fig. 1). These cleavages, together with that of the nsp7/8 site, result in the production of an alternative set of processing products (56).

The ORF1b-encoded region of the nidovirus replicase contains a number of highly conserved domains (8). Of these, the putative RNA-dependent RNA polymerase (Pol) domain (25,

* Corresponding author. Mailing address: Department of Virology, Leiden University Medical Center, LUMC P4-26, P.O. Box 9600, 2300 RC Leiden, The Netherlands. Phone: 31 71 5261657. Fax: 31 71 5266761. E-mail: Snijder@Virology.azl.nl.

FIG. 1. Processing scheme for the EAV ORF1a and ORF1ab polyproteins. The three identified protease domains (PCP, CP, and SP; see text) and corresponding cleavage sites (arrowheads) are shown. The arrowheads with question marks indicate the approximate positions of the cleavage sites in the ORF1b protein (55). Hydrophobic domains and conserved replicase ORF1b domains are indicated with black boxes. Abbreviations: PCP, papainlike cysteine protease; CP, cysteine protease; hd, hydrophobic domain; SP, serine protease; POL, putative polymerase domain; M, putative metal-binding region; HEL, putative helicase domain; CTD, conserved C-terminal domain.

36) and the nucleoside triphosphate-binding/helicase (Hel) domain (18, 24) are common to positive-stranded RNA viruses. Only in nidoviruses is the latter domain flanked by an upstream putative metal-binding region (19) and a downstream conserved C-terminal domain (Fig. 1) (8, 41, 43). A previous analysis of the processing of the ORF1b-encoded part of the EAV replicase showed that four end products and a number of processing intermediates are generated in the infected cell. Furthermore, studies of the intracellular localization of the ORF1b-encoded proteins suggested the formation of a membrane-associated replication complex (52, 55). The four ORF1b-encoded processing end products, p80, p50, p26, and p12 (55), have now been named nsp9, nsp10, nsp11, and nsp12, respectively (Fig. 1), as an extension of the numbering of the ORF1a-encoded products (56). The nsp9 N terminus was previously concluded to be generated by processing of the C-terminal-most cleavage site in the ORF1a protein, the nsp7/8 site (Glu-1677/Gly) (47, 55). As a result, the N-terminal 50 aa of nsp9 are identical to nsp8, the C-terminal-most ORF1a-encoded protein (Fig. 1). For simplicity, the part of the replicase consisting of nsp9 to nsp12 will be referred to as the ORF1b protein.

As described above, a set of ORF1b-encoded processing products had been identified in EAV-infected cells (55). However, the protease(s) responsible for their generation, the cleavage sites, and the significance of ORF1b protein processing remained to be determined. In view of the important role of ORF1b-encoded products during replication, this information is a prerequisite for future functional studies of the EAV replicase. This paper describes a detailed study of the proteolytic processing of the EAV ORF1b protein. The ORF1a-encoded nsp4 SP was shown to cleave at three sites in the ORF1b protein, resulting in the generation of the previously identified cleavage products nsp9 to nsp12. Mutagenesis studies of several candidate cleavage sites indicated that Glu-2370/Ser, Gln-2837/Ser, and Glu-3056/Gly represented the nsp9/10, nsp10/11, and nsp11/12 junctions, respectively. The requirement of ORF1b protein processing for viral replication was established by introducing proteolysis-abolishing cleavage site mutations into a recently developed EAV infectious cDNA clone (54).

MATERIALS AND METHODS

Cells, virus, and antisera. Baby hamster kidney (BHK-21) cells and rabbit kidney (RK-13) cells were used for propagation of the EAV Bucyrus strain (12). The ORF1a protein-specific antisera have been described previously (45, 56). The ORF1b protein-specific antisera α B1, α B2, α B3, and α B4 (55) have been renamed α 9, α 10, α 11, and α 12, respectively.

Expression constructs pL1ab, pL1abS1184I, and pL1b. Recombinant plasmids were constructed by standard techniques and procedures (38). Expression vector pL1ab was constructed by extending pL1a, an ORF1a expression vector

previously described by Snijder et al. (47), with the ORF1b sequence containing an engineered *Hind*III site, which has been described before (54). The resulting plasmid is a pBS(-) (Stratagene) derivative containing the complete EAV replicase gene downstream of a T7 promoter and a copy of the encephalomyocarditis virus internal ribosomal entry site, which was used to enhance translation (23). Vector pL1abS1184I was derived from pL1ab by introducing an inactivating mutation into the nsp4 SP (Ser-1184→Ile) (47). To construct a vector expressing nsp9-12 (pL1b), an in-frame ORF1b construct (pL1ab100) was created by mutating nucleotide (nt) A5404 to C and inserting a C between nt G5399 and T5400. Subsequently, vector pL1b was obtained by engineering a translation initiation codon upstream of the Gly-1678 codon (the nsp9 N terminus) in pL1ab100 and by deleting the upstream ORF1a region. Thus, pL1b contained the 3'-terminal 149 nt of the ORF1a sequence in frame with ORF1b and encoded nsp9 to nsp12.

Site-directed mutagenesis of putative cleavage sites. Site-directed PCR mutagenesis was performed according to the procedure of Landt et al. (28). The mutations that were introduced into ORF1b, listed in Table 1, were introduced into pL1b (see above). The mutations specifying the Glu-2370→Pro, Gln-2837→Pro, Gln-2837→Glu, and Glu-3056→Pro substitutions were also transferred to pEAV030, the full-length EAV cDNA clone (54). Clone pEAV030PS was generated by introducing five translationally silent point mutations in the codons for Lys-2836, Gln-2837, and Ser-2838 (Table 1). A clone containing a mutation which was assumed to inactivate the viral RNA-dependent RNA polymerase (pEAV030SGA) was constructed by mutating the conserved Ser-Asp-Asp (SDD) motif (8, 19, 25, 36) to Ser-Gly-Ala (SGA) (Table 1).

Other expression vectors. Construct pL(1065-1268), which encodes the wild-type nsp4 SP, has been described previously (47). This construct was used to create plasmid pL(1065-1268)S1184I, which expresses a proteolytically inactive SP carrying a Ser-1184→Ile mutation (47). Constructs pL(1678-2370), pL(2371-2837), and pL(3057-3175) encoded Gly-1678 to Glu-2370 (nsp9), Ser-2371 to Gln-2837 (nsp10), and Gly-3057 to Val-3175 (nsp12), respectively. To allow translation initiation, an AUG codon was placed upstream of the coding sequence. Therefore, the N-terminal sequence of each of the three expression products was extended with a Met residue. Constructs pL(2839-3055) and pL(2839-3175) encoded Asn-2839 to Gln-3055 (nsp11) and Asn-2839 to Val-3175 (nsp11-12), respectively. The N-terminal sequence of each of these expression products was extended with Met-Ala-Ala. The C terminus of nsp11 expressed from pL(2839-3055) contained a Pro-Leu-Ala-Ser extension. Plasmid pL4(11-12) was constructed by the in-frame fusion of sequences encoding nsp4 and nsp11-12. The Glu-3056→Pro mutation was introduced into pL4(11-12) to obtain construct pL4(11-12)E3056P.

Transient expression, EAV infections, and protein analysis. EAV ORF1ab and ORF1b constructs were transiently expressed in RK-13 cells, using the recombinant vaccinia virus-T7 system (15) as described previously (45). The methods for EAV infection, cell lysis, and immunoprecipitation have been described elsewhere (10). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out essentially by the procedure of Laemmli (26) and was monitored by fluorography (44). Proteins synthesized in either EAV-infected or vaccinia virus-infected and transfected RK-13 cell cultures were labelled from 5 to 8 or 6 to 9 h postinfection, using methionine- and cysteine-free medium containing 200 μ Ci of [35 S]methionine and 80 μ Ci of [35 S]cysteine per ml (Expre 35 S-label; NEN Dupont).

RNA transcription, transfection, and analysis. The methods used for transcription of infectious RNA from EAV full-length cDNA clones and for transfection of BHK-21 cells by electroporation have been described previously (54). For each electroporation, approximately 50 μ g of RNA transcript was used. Procedures for intracellular RNA analysis have been described previously (9). For the direct analysis of RNA transcripts, transcription mixtures were treated with DNase I (Gibco-BRL; 3 U/ μ l) for 30 min at 37°C. Reverse transcription (RT) reactions on DNase I-treated RNA transcripts and RNA from transfected cells were carried out with Moloney murine leukemia virus reverse transcriptase (Gibco-BRL; 6.5 U/ μ l) for 60 min at 42°C. The RT primer was complementary

TABLE 1. Amino acid substitutions introduced into the EAV ORF1b-encoded region

Mutant	aa substitution	Wild-type sequence ^a	Mutated sequence ^{a,b}	Tested in plasmid(s):
D2351P	Asp-2351→Pro	ACC <u>GAC</u> GGA	ACC CCC GGG	pL1b
E2370P	Glu-2370→Pro	TAT <u>GAG</u> AGT	TAT CCG TCG	pL1b + pEAV030
E2800P	Glu-2800→Pro	AAT <u>GAG</u> GGC	AAT CCC GGG	pL1b
D2819P	Asp-2819→Pro	AAG <u>GAT</u> GGT	AAG CCC GGG	pL1b
E2835D	Glu-2835→Asp	TGG <u>GAA</u> AAG	TGG GAT AAG	pL1b
E2835P	Glu-2835→Pro	TGG <u>GAA</u> AAG	TGG CCA AAG	pL1b
E2835Q	Glu-2835→Gln	TGG <u>GAA</u> AAG	TGG CAA AAG	pL1b
Q2837D	Gln-2837→Asp	AAG <u>CAA</u> TCC	AAG GAT TCG	pL1b
Q2837E	Gln-2837→Glu	AAG <u>CAA</u> TCC	AAG GAG TCG	pL1b + pEAV030
Q2837N	Gln-2837→Asn	AAG <u>CAA</u> TCC	AAG AAT TCC	pL1b
Q2837P	Gln-2837→Pro	AAG <u>CAA</u> TCC	AAG CCT TCG	pL1b + pEAV030
E3056P	Glu-3056→Pro	CAA <u>GAG</u> GGT	CAA CCC GGG	pL4(11–12) + pEAV030
PS	Lys-2836/Gln-2837/Ser2838	<u>AAG</u> <u>CAA</u> <u>TCC</u>	<u>AAA</u> <u>CAG</u> <u>AGT</u>	pEAV030
SGA	Asp-2237→Gly-Asp-2238→Ala	TCG <u>GAC</u> <u>GAT</u>	TCG GGC GCC	pEAV030

^a Mutated codons are underlined.^b Nucleotide changes are indicated in boldface.

to nt 9818 to 9833. Subsequently, a PCR was performed with the RT primer and a primer corresponding to either nt 8236 to 8253 or nt 7220 to 7238. The PCR products were used for cloning purposes and direct sequence analysis.

IFAs. BHK-21 cells transfected with transcripts derived from EAV full-length cDNA clones were grown on coverslips at 39.5°C. Indirect IFAs were carried out as described previously (55), using the EAV nsp2-specific antiserum α 2 (45) at a 1:250 dilution and a mouse monoclonal antibody, 93B (tissue culture supernatant), directed against the EAV open reading frame 5 (ORF5)-encoded glycoprotein G_L (16) at a 1:60 dilution. The latter antibody will be referred to as α G_L. A Cy3-conjugated donkey anti-rabbit immunoglobulin G antibody and a fluorescein isothiocyanate-conjugated donkey anti-mouse immunoglobulin G antibody (Jackson ImmunoResearch Laboratories) (1:1,000 and 1:100 dilutions, respectively) were used as secondary antibodies.

RESULTS

The ORF1b-encoded region of the EAV replicase is processed by the nsp4 SP. Previously, proteolytic processing of the ORF1b-encoded region of the EAV replicase had been shown to result in the generation of nsp9 to nsp12 (55). However, the responsible protease remained to be identified. Our preliminary processing scheme (Fig. 1) (55) was based on the esti-

mated sizes of nsp9 to nsp12 and on the locations of the epitopes recognized by the different ORF1b protein-specific peptide antisera. It was assumed that cleavage would not occur within the conserved domains in arteri- and coronavirus replicases (8, 11, 43). Tentative cleavage sites for the SP were identified by computer sequence analysis (55). These dipeptides were conserved among arteriviruses and matched the previously determined substrate specificity of the nsp4 SP (47, 56). To determine whether the SP is indeed responsible for their processing, cleavage of the ORF1b protein was studied by using the recombinant vaccinia virus-T7 expression system. To this end, constructs that expressed the complete ORF1b protein containing either a wild-type SP or an inactivated SP, in which the catalytic nucleophile Ser-1184 had been replaced with Ile, were generated (47).

Figure 2 shows the results of an immunoprecipitation analysis using the α 9 and α 10 sera and cell lysates from transfection experiments (Fig. 2A and B, left panels, lanes 1ab). Although the amount of replicase protein produced in transfected cells

FIG. 2. *trans* cleavage of the ORF1b protein by the nsp4 SP in the recombinant vaccinia virus-T7 system (15). Processing of the ORF1b (in *cis*) and ORF1b (in *trans*) proteins was studied in the presence of a wild-type or mutant SP. Immunoprecipitation analysis was performed with the α 9 (A) and α 10 (B) sera. Cleavage products and precursor proteins are indicated, and the position of nsp2 (61 kDa) is noted as a reference. Abbreviations: M, mock-infected-cell lysate; E, EAV-infected-cell lysate; U, untransfected-cell lysate; 1ab, construct pL1ab; 1abS1184I, construct pL1abS1184I; 1b, construct pL1b; SP, construct pL(1065–2370); S1184I, construct pL(1065–2370)S1184I.

was less than the amount generated in EAV-infected cells, the processing patterns were identical. nsp9 (Fig. 2A, left panel, lane 1ab), nsp10 (Fig. 2B, left panel, lane 1ab), and a number of large precursor proteins (55) were observed. The two smaller ORF1b-derived proteins, nsp11 and nsp12, could not be detected. The low Met and Cys content of nsp11 and nsp12 as well as the relatively poor quality of the α 11 and α 12 antisera might explain this failure. In addition, processing at the nsp11/12 junction might be inefficient in this system. As in our previous experiments, coprecipitation of the ORF1a-derived nsp2, which results from a previously described interaction of this subunit with nsp3-containing precursors, was observed (45, 55, 56). The production of cleaved nsp9 and nsp10 was completely abolished when the SP was inactivated (Fig. 2A and B, left panels, lanes 1abS1184I), which strongly suggested that this protease is responsible for their generation.

In the experiments described above, the protease and substrates were part of the same polypeptide, and because of this, processing could have been mediated in *cis* and/or in *trans*. To test whether the SP was indeed able to cleave the ORF1b protein in *trans*, ORF1b expression construct pL1b and expression vectors encoding either the wild-type or the inactivated SP (47) were cotransfected (Fig. 2A and B, right panels). Upon expression of pL1b, most of the ORF1b protein (calculated size, 163 kDa) migrated as an approximately 175-kDa product and remained uncleaved. Small amounts of lower-molecular-weight products were detected, suggesting that the uncleaved ORF1b protein might be somewhat unstable. The combined expression of the ORF1b protein and the wild-type SP resulted

in the production of cleaved nsp9, nsp10, and the nsp9-10 precursor (Fig. 2A and B, right panels, lanes 1b+SP). Although nsp11 and nsp12 again could not be detected, the nsp11-12 intermediate was observed (data not shown). Cotransfection of pL1b with the mutant SP (Fig. 2A and B, right panels, lanes 1b+S1184I) did not result in cleavage of the ORF1b protein, confirming that the SP is responsible for its processing.

Putative SP cleavage sites in the ORF1b-encoded region of the replicase. The amounts of ORF1b protein and cleaved nsp9 and nsp10 that were produced in the cotransfection assay described above largely exceeded the amounts generated by the pL1ab construct, in which ORF1b expression is downregulated by the (native) ribosomal frameshift site. Therefore, the *trans*-cleavage assay was used for a more precise analysis of the positions of the nsp9/10 and nsp10/11 junctions. Since processing of the nsp11/12 site could not be detected, a different approach was employed to study the cleavage of that junction in more detail (see below).

Previously, candidate cleavage sites for the nsp4 SP in the EAV ORF1b protein had been identified on the basis of arterivirus sequence alignments, the location of conserved domains, the estimated sizes of nsp9 to nsp12, and the SP substrate specificity (see the introduction). Processing was predicted to occur between Glu-2370 and Ser-2371 (nsp9/10 site), Glu-2835 and Lys-2836 (nsp10/11 site), and Glu-3056 and Gly-3057 (nsp11/12 site) (55). These three dipeptides, together with a set of alternative candidates for the nsp9/10 (one) and nsp10/11 (three) junctions (Table 1), were subjected to site-

FIG. 3. Mutational analysis of the nsp9/10 and nsp10/11 cleavage sites in the *trans*-cleavage assay. Shown are the results of immunoprecipitation analysis with the α 9 serum (A) or the α 10 serum (C and D) and of control immunoprecipitations with the α 4 serum, which were performed on all samples to confirm the expression of the nsp4 SP (B). Lanes labeled 1b correspond to cotransfections of the SP expression construct with the wild-type pL1b expression construct. Cotransfections with mutant pL1b expression constructs are indicated by the introduced mutation (see Table 1). nsp2, nsp3-4 (46 kDa), and nsp9 to nsp12 (nsp9-12) are noted as size references. For abbreviations, see the legend to Fig. 2.

directed mutagenesis and tested in the context of the pL1b expression construct, using the cotransfection assay described above.

Mutagenesis of the candidate nsp9/10 junctions. The P1 amino acid of each of the two candidate nsp9/10 sites, Asp-2351/Gly and Glu-2370/Ser, was mutated to Pro (Table 1). Immunoprecipitations with the $\alpha 9$ (Fig. 3A) and $\alpha 10$ (Fig. 3C, lanes D2351P and E2370P) sera were carried out on transfected-cell lysates. Control immunoprecipitations were performed to confirm that the nsp4 SP was indeed expressed (Fig. 3B). Comparison of proteins generated by the Asp-2351→Pro mutant and the wild-type ORF1b protein showed that this mutation did not affect the production of nsp9 and nsp10. In contrast, the Glu-2370→Pro mutation completely abolished cleavage of the nsp9/10 site, resulting in the accumulation of the nsp9-10 precursor. These results strongly suggested that the Glu-2370/Ser dipeptide is the nsp9/10 cleavage site.

Mutagenesis of the candidate nsp10/11 cleavage sites. As for the candidate nsp9/10 cleavage sites, the P1 residue of each of the four candidate nsp10/11 junctions (Table 1) was first replaced by Pro. The results of immunoprecipitations with the $\alpha 10$ serum are shown in Fig. 3C. For two mutants, Glu-2800→Pro and Asp-2819→Pro, the cleavage pattern was identical to that of the wild-type ORF1b protein, although the transfection efficiency of the Asp-2819→Pro mutant was somewhat lower in this experiment. The other two mutations, Glu-2835→Pro and Gln-2837→Pro, severely influenced the processing of the ORF1b protein and almost completely abolished the production of nsp10 and the nsp9-10 precursor. Taken together, these data suggested that of the four mutations tested, the last two affected determinants of the nsp10/11 cleavage site.

In the ORF1b protein, aa 2835 and 2837, which were found to be sensitive to replacement by Pro, are part of two adjoining dipeptides, Glu-2835/Lys-2836 and Gln-2837/Ser-2838. Previously, the importance of the P1 position of the SP cleavage sites was demonstrated by a variety of methods (47, 56), although the influence of other flanking residues on the cleavage efficiency was not evaluated. Given this uncertainty regarding the determinants of SP specificity, the effect of the Pro mutations at positions 2835 and 2837 could be due to replacements at either the P1 and P2' positions (if P1 is at aa 2835) or the P3 and P1 positions (if P1 is at aa 2837). To discriminate between these two possibilities, an extended set of conservative amino acid substitutions of both candidate P1 residues was tested (Table 1). Figure 3D shows the results of immunoprecipitations with the $\alpha 10$ serum after transfection with the new set of mutants.

The two additional mutations tested at the Glu-2835 position (Glu-2835→Asp and Glu-2835→Gln) significantly inhibited the production of nsp10 and the nsp9-10 precursor (Fig. 3D, lanes E2835D and E2835Q), indicating that the cleavage at the nsp10/11 junction was severely compromised in both mutants. The new mutations at the Gln-2837 position had various effects on the processing of the nsp10/11 site. In the Gln-2837→Asp and Gln-2837→Asn mutants, the production of nsp10 was not diminished, although the amount of nsp9-10 precursor was significantly reduced compared to the wild-type situation (Fig. 3D, lanes Q2837D and Q2837N). The Gln-2837→Glu mutant, which now carried a canonical SP cleavage site (Glu/Ser), showed an increased production of nsp10 and, especially, the nsp9-10 intermediate (Fig. 3D, lane Q2837E). Thus, all three mutations at the Gln-2837 position affected nsp10/11 cleavage and indicated that the efficiency of processing at one site (nsp10/11) could modulate cleavage at another (nsp9/10). Remarkably, nsp10 derived from the Gln-

FIG. 4. Analysis of the processing of the nsp11/12 cleavage site. (A) Schematic representation of the nsp4(11-12) expression product. The regions of the ORF1ab polyprotein that were fused are indicated. (B) Immunoprecipitations with the $\alpha 4$, $\alpha 11$, and $\alpha 12$ antisera of cells transfected with the pL(11-12) and pL(11-12)E3056P expression constructs. Precursor protein and cleavage products are indicated. For abbreviations, see the legend to Fig. 2.

2837→Asp and the Gln-2837→Glu mutants migrated slightly more slowly during SDS-PAGE (Fig. 3D, lanes Q2837D and Q2837E), indicating that the presence of an acidic residue at position 2837 apparently affects nsp10's mobility and that this residue must therefore be part of nsp10 rather than nsp11. These data are most consistent with nsp10/11 cleavage occurring at the Gln-2837/Ser dipeptide and imply that in addition to the P1 residue, the P3 aa is an important determinant of cleavage site specificity.

Mutagenesis of the putative nsp11/12 cleavage site. Processing of the nsp11/12 junction was not detected in either of the assays successfully employed for the characterization of the two upstream sites (Fig. 2). To study this processing step, we used a new construct, pL4(11-12), which encoded an nsp4-nsp11-12 fusion protein (Fig. 4A). In this expression product, the nsp11/12 junction is the only naturally occurring SP cleavage site. To test the importance of the Glu-3056/Gly dipeptide for processing of the nsp11/12 cleavage site, Glu-3056 was substituted with Pro. Figure 4B shows the results of immunoprecipitations with the $\alpha 4$, $\alpha 11$, and $\alpha 12$ sera. In the pL4(11-12)-transfected cells, three proteins were observed: the nsp4(11-12) precursor and the nsp4-11 and nsp12 cleavage products. In cells transfected with the mutant construct, pro-

FIG. 5. Comigration analysis of synthetic proteins. Synthetic ORF1b-derived proteins were expressed and their migration was compared with that of native proteins from EAV-infected cells. The numbers over the lanes indicate the synthetic nsp's expressed. nsp2 (EAV-infected cells) and nsp9-12. pL1b-transfected cells) are indicated as size references. N, nucleocapsid protein (14 kDa), coimmunoprecipitated from EAV-infected cells as a result of its affinity for protein A on the surface of *Staphylococcus aureus* cells (50). For other abbreviations, see the legend to Fig. 2.

cessing was not observed, and only the precursor protein was precipitated. These data are compatible with the Glu-3056/Gly dipeptide being the nsp11/12 cleavage site.

Comigration of synthetic proteins with native ORF1b-encoded cleavage products. To gain further support for the identification of the cleavage sites separating the nsp9, nsp10, nsp11, and nsp12 proteins, these four proteins, as well as the nsp11-12 precursor, were expressed by inserting translation initiation and termination codons at the presumed cleavage sites in the ORF1b gene. The constructs were expressed in the recombinant vaccinia virus-T7 expression system, and by using immunoprecipitation and SDS-PAGE, the sizes of synthetic nsp9 to nsp12 were compared with those of the native cleavage products from EAV-infected cells (Fig. 5). Because larger amounts of the nsp11-12 precursor protein could be produced in the recombinant vaccinia virus-T7 expression system (*trans*-cleavage assay), this system was used to produce sufficient amounts of this precursor.

The nsp9, nsp10, nsp12, and nsp11-12 expression products comigrated perfectly with their *in vivo* counterparts. The nsp11 product migrated slightly more slowly than the corresponding protein from EAV-infected cells (Fig. 5), probably because of the presence of additional, pL(2839–3055)-encoded residues at its N and C termini. Together with the mutagenesis data, these results strongly suggest that the Glu-2370/Ser, Gln-2837/Ser, and Glu-3056/Gly dipeptides are the nsp9/10, nsp10/11, and nsp11/12 cleavage sites, respectively.

SP-mediated cleavages in the ORF1b protein are essential for different stages of EAV reproduction. To test the importance of ORF1b protein processing for RNA replication and transcription, the Glu-2370→Pro, Gln-2837→Pro, and Glu-3056→Pro mutations were transferred to an EAV infectious cDNA clone (54). RNA transcripts were transfected into BHK-21 cells, which were fixed for IFA after 12, 24, and 36 h. As positive and negative controls, a wild-type EAV clone (pEAV030H) and a clone encoding a presumably inactivated Pol (pEAV030SGA) were used, respectively. The latter carried mutations at aa 2237 and 2238 of the replicase, in the highly conserved polymerase domain: Ser-Asp-Asp (SDD)→Ser-Gly-Ala (SGA) (Table 1). Cells transfected with pEAV030SGA

RNA did not show detectable RNA replication or subgenomic RNA transcription (data not shown). This result confirmed for the first time the essential nature of this highly conserved nidovirus replicase motif.

Previously, we have shown that EAV RNA synthesis can be monitored indirectly by IFA. The replication of the genomic RNA, the mRNA for the replicase, results in the increasing expression of the viral nsp's and their cleavage products, which can be detected by IFA as early as 3 h postinfection (52). Likewise, IFAs can be used to monitor the expression of the structural proteins from the nested set of subgenomic (sg) mRNAs starting at about 6 h p.i. In earlier experiments, cells were double labeled with antibodies recognizing replicase subunit nsp2 and the ORF5-encoded glycoprotein G_L to detect genome replication and sg mRNA synthesis, respectively (54). This double staining, which is convenient because the two signals do not overlap (52), was used in the present study to monitor the effects of cleavage site mutations that abolished one of the ORF1b protein processing steps. Cells transfected with wild-type pEAV030H transcript showed both nsp2 and G_L signals, as well as spread of the virus to neighboring cells, which confirmed the production of infectious virus (Fig. 6, left panels). The Glu-2370→Pro (nsp9/10 cleavage site) and Glu-3056→Pro (nsp11/12 cleavage site) mutants did not show any nsp2 or G_L signal (data not shown), indicating that cleavage at these junctions is essential for virus reproduction. Even at very late time points, up to 96 h posttransfection, (pseudo)revertants of these mutants could not be detected. Remarkably, a distinct phenotype for the Gln-2837→Pro mutant (nsp10/11 cleavage site) was reproducibly observed in a small fraction of transfected cells (1 to 10%). At 12 h posttransfection, a faint nsp2 signal (but no G_L staining) could be detected in a few cells (Fig. 6, right panels). At later time points, both nsp2 and G_L were detected, indicating that both genomic and sg RNA synthesis was taking place. Surprisingly, spread of the mutant virus to neighboring cells was not observed (even after 96 h), and the medium which was harvested from transfected cells at different time points (12, 24, and 36 h) was not infectious upon passaging in tissue culture (data not shown). These results

FIG. 6. IFA of EAV030H- or EAV030Q2837P-transfected BHK-21 cells. Cells were fixed at 12, 24, and 36 h postelectroporation and subsequently processed for indirect IFA. Cells were double labeled for EAV nsp2 (45) and EAV G_L (16).

indicated that infectious virus was not produced in the EAV030Q2837P-transfected cells.

In conclusion, processing of all three SP cleavage sites in the ORF1b protein proved to be essential for the production of infectious virus. Furthermore, its selective blockage by point mutations appeared to affect different stages of EAV reproduction. In view of the unusual phenotype of the EAV030Q2837P mutant, we decided to characterize this mutant in greater detail.

Mutations at the nsp10/11 junction affect protein function.

Since infectious virions were not produced by the semi-replication-competent Gln-2837→Pro mutant, it was possible that cleavage at the nsp10/11 site was essential for virus maturation. Alternatively, the introduced mutations may have changed an RNA sequence involved in EAV genome encapsidation. An RNA encapsidation signal was described at a comparable position in the ORF1b sequence of the distantly related coronavirus mouse hepatitis virus (14, 53). To test the possibility that the Gln-2837→Pro mutation affected such an RNA structure, two other mutants with mutations at or near replicase codon 2837 were generated and characterized by using the infectious clone (Table 1). First, the Gln-2837→Glu mutation, which

showed efficient cleavage in the recombinant vaccinia virus-T7 expression system (Fig. 3D) and which carried the same number of nucleotide substitutions as the Gln-2837→Pro mutant (Table 1), was transferred to the EAV cDNA clone (pEAV030Q2837E). Second, a mutant which carried five translationally silent point mutations in the Lys-2836, Gln-2837, and Ser-2838 codons was generated (pEAV030PS) (Table 1). After transfection of pEAV030Q2837E and pEAV030PS transcripts into BHK-21 cells, both mutants replicated and spread with wild-type efficiency (data not shown) (compare with wild-type results shown in Fig. 6). Analysis of RNA isolated at 48 h posttransfection (after spread of the virus) confirmed the presence of the original mutations (data not shown). Although minor changes in the fitness of these mutants compared to the wild-type virus cannot be excluded, our data indicate that the pEAV030Q2837P phenotype is not likely caused by an effect at the level of RNA structure.

Analysis of RNA produced in EAV030Q2837P-transfected cells. Since a subset of the EAV030Q2837P-transfected cells showed complete viral RNA synthesis but did not produce infectious virus, we decided to use RNA passaging to further

characterize the properties of the Q2837P mutant and analyze possible changes in the RNA sequence at replicase codon 2837. To this end, intracellular RNA from transfected cells was isolated and transfected into fresh cells. As a control for RNA passaging in the absence of virus production, the EAV030F mutant, which contains a mutation in *nsp10*, was used (54). Although its genomic RNA is replicated efficiently, EAV030F does not produce virus due to a severe deficiency in sg mRNA synthesis (54). Transcripts of wild-type clone pEAV030H were used as a positive control.

RNA transcripts were electroporated into BHK-21 cells, and intracellular RNA was isolated after 24 h (Pr0 RNA, passage 0 progeny). This RNA was used for electroporation of fresh BHK-21 cells, and RNA was again isolated at 24 h posttransfection (Pr1 RNA, passage 1 progeny). Analysis of intracellular EAV030H and EAV030Q2837P RNAs by gel electrophoresis and subsequent hybridization revealed that EAV030Q2837P-transfected cells did indeed produce genomic and sg RNA (data not shown). However, as a result of the small number of Q2837P-transfected cells which showed RNA synthesis (and the apparently reduced level of RNA synthesis in these cells), the amount of RNA was incomparable to that produced in wild-type EAV030H-transfected cells. For pEAV030H and pEAV030F, a twofold increase in transfection efficiency was observed when Pr0 RNA was used instead of the *in vitro*-generated RNA transcript. In contrast, passaging of RNA from cells transfected with the Q2837P mutant resulted in a decrease in the number of cells positive for *nsp2* and *G_L*.

To determine the sequence at the position of mutated codon 2837, an RT-PCR analysis was performed on DNase I-treated RNA transcript, Pr0 RNA, and Pr1 RNA of the Q2837P mutant. Direct sequence analysis of the PCR products revealed that the transcript RNA and the Pr0 RNA exclusively contained a Pro codon at position 2837. However, in the Pr1 RT-PCR product, a mixture was observed for each nucleotide position of the codon (C and G, C and A, and T and G, from the first to the third positions, respectively) (data not shown), which suggested partial conversion of the introduced Pro codon (CCT). Subsequently, the PCR fragments were cloned and subjected to sequence analysis. In the clones from transcript (2 clones) and Pr0 RNA (14 clones), the CCT codon (Pro mutation) was recovered. Of 13 Pr1 RNA clones, only 4 contained the original CCT codon and 9 contained a GAG triplet encoding Glu. Interestingly, the Glu substitution had already been tested in the infectious cDNA clone (pEAV030Q2837E), and this clone had been shown to be infectious and wild-type like (see above). However, in the case of pEAV030Q2837P, the conversion of Pro-2837 to Glu did not result in the production of infectious virus, even at 96 h posttransfection, which suggests that the replacement of the Pro codon must have been accompanied by additional mutations elsewhere in the EAV genome.

DISCUSSION

The *nsp4* SP mediates processing of the ORF1b-encoded region of the arterivirus replicase. The nidovirus ORF1b protein contains (predicted) RNA-dependent RNA polymerase and NTPase/RNA Hel domains which are believed to catalyze viral RNA synthesis (8, 19, 43). Although the nidovirus ORF1b protein's enzymatic activities remain to be characterized, ATPase activity was recently demonstrated for the human coronavirus 229E Hel domain (22). Using an EAV infectious clone, we have shown that replacements of highly conserved residues in the Pol (see Results) and Hel (unpublished data) domains abrogate arterivirus reproduction. As a result of the

pivotal role of ORF1b protein in the viral replication cycle, the regulation of its expression is of vital importance for nidoviruses. Since they are all derived from the same polyprotein, whose expression is downregulated relative to that of the ORF1a protein (8), the Pol and Hel functions and other ORF1b-protein specific activities might act in concert. The expression of ORF1b-encoded domains is further coordinated at the posttranslational level through regulated cleavages at four sites. The processing of the EAV *nsp7/8* site and the three junctions in the ORF1b protein yields four mature proteins and a number of stable intermediates, which may have distinct functions (55). In line with our previous predictions and data recently obtained for coronaviruses (31–34, 49, 57, 58), we have now shown that these cleavages are mediated by the ORF1a-encoded *nsp4* SP. The same protease is responsible for the extensive processing of the C-terminal part of the ORF1a protein (47, 56). Multiple processing intermediates which contain both ORF1a- and ORF1b-encoded domains (55) have previously been identified. Hence, the expression of mature and intermediate products from the two overlapping ORFs is probably coordinated.

We had to solve substantial technical problems when studying the SP-mediated processing of the ORF1b protein. These were caused by low expression levels, inefficient cleavage at some sites, and poor detection of certain cleavage products. To this end, the recombinant vaccinia virus-T7 system was employed to express either full-length ORF1ab, ORF1b alone, or an artificial protein in which *nsp4* was fused in frame to the *nsp11-12* region of ORF1b [*nsp4*(11–12) protein]. We have demonstrated that in this system, *nsp4* is able to process the *nsp9/10* and *nsp10/11* sites, but not the *nsp11/12* junction, *in trans*. It is likely that all three sites are also processed *in cis*, although this was not proven directly. We also believe that in the infected cell, these sites are not equally sensitive to processing by the *nsp4* SP, which would provide a means for the regulated expression of ORF1b-derived products.

Identification of three SP cleavage sites in the EAV ORF1b protein. As a step toward understanding the structural determinants of ORF1b protein processing, the primary structure of the cleavage sites was tentatively determined by a combination of three indirect methods: comparative sequence analysis, site-directed mutagenesis, and comigration assays. In the past, the same approach was proven successful for the identification of SP cleavage sites in the ORF1a protein (47), two of which have recently been confirmed by N-terminal protein sequencing (56). Unfortunately, direct sequencing of ORF1b-derived cleavage products was hampered by the scanty amount of these proteins in EAV-infected cells.

The Glu-2370/Ser, Gln-2837/Ser, and Glu-3056/Gly dipeptides were identified as the probable *nsp9/10*, *nsp10/11*, and *nsp11/12* cleavage sites, respectively (Fig. 7). Of these three, Glu-2370/Ser and Glu-3056/Gly both match the previously determined consensus sequence (Glu/Ser or Glu/Gly) of the EAV SP cleavage sites (47, 56). Surprisingly, Gln-2837/Ser was identified as the probable *nsp10/11* junction, which is thus the first and only SP substrate carrying Gln instead of Glu at the P1 position. Interestingly, in the replicase proteins of the distantly related coronaviruses, cleavages occur after Gln rather than Glu (17, 31–34, 49, 57, 58). This similarity might therefore reflect the evolutionary conservation of a sequence derived from a common nidovirus progenitor. The serine and cysteine chymotrypsin-like proteases of arteri- and coronaviruses, which are employed to process these sites, have been proposed to be related, although their sequences have diverged beyond the level at which common ancestry can be considered proven (for reviews, see references 11, 20, 27, and 42). On the other

SP cleavage sites in the ORF1a protein

nsp3/4	1058	E G G M V F E G L F R S P K
nsp4/5	1262	D G L S N R E S S L S G P Q
nsp5/6	1424	M M K Y F L E G V K E S V
nsp6/7	1446	G K P I T Q E S L T A T L A
nsp7/8	1671	L G K G S Y E G L D Q D K V

SP cleavage sites in the ORF1b protein

nsp9/10		
EAV	2364	F R T K Q Y E S A V C T V C
PRRSV		K L R S H N E G K K F R H C
LDV		N R L S A E E K K K C R T C
nsp10/11		
EAV	2831	A C G W E K Q S N K I S C L
PRRSV		A C S A S L E G S C M P L P
LDV		Q Q L M G L E G T A S P L P
SHFV		L K L D F L E S G S L S P L
nsp11/12		
EAV	3050	A T F Y V Q E G V D A V T S
PRRSV		T A Y F Q L E G L T W S A L
LDV		A T A Y F H E G I R P M E A
SHFV		Q T M Y F Q E G F D Y H T L

FIG. 7. Cleavage sites for the EAV nsp4 SP in the ORF1ab polyprotein. The five previously described ORF1a protein cleavage sites are listed. An alignment of the three ORF1b protein cleavage sites (boldfaced) is shown, including sequences of the arteriviruses EAV, PRRSV, LDV, and SHFV. For SHFV, only a partial ORF1b sequence has been published (40). The amino acid numbers refer to the EAV sequence. Sequences were extracted from the EMBL/GenBank database (accession no. X53459 [EAV], M96262 [PRRSV], U15146 [LDV], and U63121 [SHFV]).

hand, the Gln/Ser dipeptide found at the EAV nsp10/11 junction is not conserved in other arteriviruses, and our own data obtained with the mutant Glu/Ser site in the infectious cDNA clone suggest that the presence of a Gln at this cleavage site is not critical for the basic functions in the EAV life cycle. An interplay between the nsp9/10 and nsp10/11 cleavages was observed, since differences in processing efficiency at one site were found to affect processing of the other site. Previously, an even more-dramatic switch between two SP-mediated processing pathways was described for the EAV ORF1a protein, with nsp2 being involved as a decisive cofactor (56). Additional experiments are required to analyze the functional significance of these modulations in replicase polyprotein processing (see also below).

Given the conserved nature of the EAV SP cleavage sites and the conservation of corresponding sites in other arteriviruses, it was not surprising that processing of the sites in the ORF1b protein was sensitive to replacements at the P1 position. It is interesting that the replacement of the P1 Gln of the nsp10/11 cleavage site by residues with similar properties (Asp, Glu, and Asn) did not adversely affect the production of nsp10 and could even stimulate it (Q2837E [Fig. 3D]). In contrast, the production of the nsp9-10 intermediate was significantly affected, being either diminished (Asn and Asp replacements) or enhanced (Glu replacement). These results again suggest that the particular structure of the nsp10/11 cleavage site may have been selected to ensure a finely balanced production of ORF1b-derived cleavage products and that this may also be the case for other SP sites. In addition to the P1 and P1' residues of the cleavage sites, other amino acids in close proximity to these positions may contain specificity determinants for viral chymotrypsin-like proteases, of which nsp4 SP is a

member (20). One of the best-characterized enzymes, the poliovirus 3C protease, recognizes the P4 position, which is usually occupied by Ala and is sensitive to replacements (4, 7). In the case of the EAV nsp4 SP cleavage sites, only the P1 and P1' positions are occupied by highly conserved residues (56) (Fig. 7). However, the P3 position of the nsp10/11 cleavage site was sensitive to replacements (Fig. 3), indicating that there might be additional determinants of substrate specificity.

This study provides the first experimental evidence for a complete processing map of the arterivirus ORF1b protein and, by including the data obtained for the EAV ORF1a protein (47, 56), also for the entire ORF1ab polyprotein. In addition to being essential for the analysis of the regulation of ORF1b expression, this map is important for the characterization of ORF1b-encoded functions by expressing mature nonstructural proteins separately and in different combinations.

ORF1b protein processing is essential for arterivirus reproduction. We used a recently developed infectious cDNA clone of EAV (54) to initiate the first functional study of the role of replicase processing in nidovirus reproduction. Several of the SP cleavage site mutations which had been characterized in the recombinant vaccinia virus-T7 system were subsequently tested in the context of the infectious clone. All three cleavages in the ORF1b protein were found to be essential for EAV reproduction. The nsp9/10 and nsp11/12 cleavages were crucial for early steps in infection, since RNA synthesis could not be detected when these two cleavages were blocked, and revertants were not recovered.

The results obtained for nsp10/11 cleavage were different. The P1 Pro mutant (Q2837P) displayed a remarkably complex phenotype that may be related to the unique Gln/Ser structure of the nsp10/11 site. In addition, the conversion of the Pro-2837 mutation to Glu, which was observed upon RNA passaging, was a complicating factor in the identification of the genetic determinants of the phenotype. However, the fact that none of the 14 RT-PCR clones derived from Pr0 RNA contained the Pro-2837→Glu conversion strongly suggested that this substitution was not related to the phenotype. Therefore, we assume that most of the Pr0 RNA in the nsp2- and G_L-positive cells at 12 h postelectroporation (Fig. 6) still contained the original Gln-2837→Pro mutation.

Like the Pro mutations introduced at the other two sites, the Gln-2837→Pro substitution was deleterious for virus reproduction. However, in a small fraction of transfected cells, viral RNA and protein synthesis was not completely abolished, suggesting either that a host cell-specific factor might be able to suppress the effect of the mutation or that these cells could overcome the block, e.g., due to transfection with multiple RNA molecules. Quite remarkably, the Q2837P mutant was defective in the production of infectious virions, even though complete RNA and protein synthesis was detected in a subset of the transfected cells. The observed phenotype might be the result of only partial suppression of the effects of the mutation, revealing secondary effects at the level of virus production. Alternatively, ORF1b-derived products may be involved directly in virion morphogenesis, in addition to their role in RNA synthesis. This would be the first indication of such a function for a nidovirus nonstructural protein. For poliovirus, however, convincing genetic data have already implicated the 2C protein (ATPase and putative helicase) in virion assembly (30, 51). It is remarkable that the Gln-2837→Pro mutation also affects the production of a presumed viral ATPase/helicase, EAV nsp10 (8). The possible involvement of this protein in virion assembly may be supported by our recent observation that another nsp10 mutant, which has a substitution in the putative metal-

binding region (His-2414→Cys), has a phenotype which is similar to that of the Q2837P mutant (unpublished data).

The special properties of the Q2837P mutant also became evident upon passing of RNA from transfected cells. After one passage, the Gln→Pro mutation at position 2837 was partially replaced by Glu, the consensus P1 residue for the EAV SP. Interestingly, we have shown in the present study that a mutant carrying a single Gln-2837→Glu substitution has the characteristics of the wild-type virus. Thus, an as-yet-unidentified number of second site mutations must be responsible for the fact that infectious virus is not produced upon Pro-2837→Glu conversion. Furthermore, during a recent, independent repeat of this RNA passing experiment, (pseudo)reversion of the Q2837P mutant to an infectious variant did occur at a late stage (48 h) of the first RNA passage. However, an analysis of the Pr0 RNA from this experiment by means of RT-PCR and sequencing again confirmed the exclusive presence of Pro at position 2837 during the initial transfection experiment. Thus, although the outcome of repeated RNA passing may vary, these results do not alter our conclusion that the Q2837P mutation allows a basic level of RNA synthesis and induces an as-yet-unexplained defect in the production of infectious virus particles.

In summary, we have identified the nsp4 SP as a major factor in the maturation of the EAV ORF1b protein, which contains a number of key enzyme activities for EAV replication. Three probable cleavage sites for the nsp4 SP were determined, and their processing was found to be essential for virus reproduction. Like in alphaviruses (29, 39) and picornaviruses (1, 3, 21), controlled cleavage of nidovirus replicative proteins may result in activation and/or inhibition of specific nonstructural functions serving the irreversible progression of virus infection. Future studies should characterize these activities in molecular terms.

ACKNOWLEDGMENTS

We thank Yvonne van der Meer for assistance with immunofluorescence microscopy and all photographic work; Willy Spaan, Fred Wassenaar, and Johan den Boon for helpful discussions and suggestions; Hans van Tol for technical assistance; and Amy Glaser for the anti-G_L monoclonal antibody.

A.E.G. was supported in part by grants from the Russian Fund for Basic Research and the Netherlands Organization for Scientific Research (N.W.O.) and with federal funds from the National Cancer Institute, National Institutes of Health, under contract no. NO1-CO-5600.

REFERENCES

- Andino, R., G. E. Rieckhof, P. L. Achacoso, and D. Baltimore. 1993. Poliovirus RNA synthesis utilizes an RNP complex formed around the 5'-end of viral RNA. *EMBO J.* **12**:3587-3598.
- Bazan, J. F., and R. J. Fletterick. 1988. Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: structural and functional implications. *Proc. Natl. Acad. Sci. USA* **85**:7872-7876.
- Blair, W. S., X. Li, and B. L. Semler. 1993. A cellular cofactor facilitates efficient 3CD cleavage of the poliovirus P1 precursor. *J. Virol.* **67**:2336-2343.
- Cao, X., and E. Wimmer. 1996. Genetic variation of the poliovirus genome with two VPg coding units. *EMBO J.* **15**:23-33.
- Cavanagh, D. 1997. *Nidovirales*: a new order comprising *Coronaviridae* and *Arteriviridae*. *Arch. Virol.* **142**:629-633.
- 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.
- Charini, W. A., S. Todd, G. A. Gutman, and B. L. Semler. 1994. Transduction of a human RNA sequence by poliovirus. *J. Virol.* **68**:6547-6552.
- 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.
- den Boon, J. A., W. J. M. Spaan, and E. J. Snijder. 1996. Equine arteritis virus subgenomic RNA transcription: UV inactivation and translation inhibition studies. *Virology* **213**:364-372.
- 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.
- de Vries, A. A. F., M. C. Horzinek, P. J. M. Rottier, and R. J. de Groot. 1997. The genome organization of the *Nidovirales*: similarities and differences between arteri-, toro-, and coronaviruses. *Semin. Virol.* **8**:33-47.
- Doll, E. R., J. T. Bryans, W. H. 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.
- 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.
- Fosmire, J. A., K. Hwang, and S. Makino. 1992. Identification and characterization of a coronavirus packaging signal. *J. Virol.* **66**:3522-3530.
- Fuerst, T. R., E. G. Niles, F. W. Studier, and B. Moss. 1986. Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc. Natl. Acad. Sci. USA* **83**:8122-8126.
- Glaser, A. L., A. A. F. de Vries, and E. J. Dubovi. 1995. Comparison of equine arteritis virus isolates using neutralizing monoclonal antibodies and identification of sequence changes in G_L associated with neutralization resistance. *J. Gen. Virol.* **76**:2223-2233.
- 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.
- 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. Sect. D Physicochem. Biol.* **11**:1-84.
- Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov. 1989. Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Res.* **17**:4847-4861.
- Gorbalenya, A. E., and E. J. Snijder. 1996. Viral cysteine proteases. *Perspect. Drug Discov. Des.* **6**:64-86.
- Harris, K. S., S. R. Reddigari, M. J. H. Nicklin, T. Hammerle, and E. Wimmer. 1992. Purification and characterization of poliovirus polypeptide 3CD, a proteinase and a precursor for RNA polymerase. *J. Virol.* **66**:7481-7489.
- Heusipp, G., U. Harms, S. G. Siddell, and J. Ziebuhr. 1997. Identification of an ATPase activity associated with a 71-kilodalton polypeptide encoded in gene 1 of the human coronavirus 229E. *J. Virol.* **71**:5631-5634.
- Jang, S. K., H.-G. Kräusslich, M. J. H. Nicklin, G. M. Duke, A. C. Palmberg, and E. Wimmer. 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during *in vitro* translation. *J. Virol.* **62**:2636-2643.
- Kadaré, G., and A.-L. Haenni. 1997. Virus-encoded RNA helicases. *J. Virol.* **71**:2583-2590.
- Kamer, G., and P. Argos. 1984. Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res.* **12**:7269-7282.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680-685.
- Lai, M. M. C., and D. Cavanagh. 1997. The molecular biology of coronaviruses. *Adv. Virus Res.* **48**:1-100.
- Landt, O., H. P. Grunert, and U. Hahn. 1990. A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene* **96**:125-128.
- Lemm, J. A., T. Rumenapf, E. G. Strauss, J. H. Strauss, and C. M. Rice. 1994. Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus- and plus-strand RNA synthesis. *EMBO J.* **13**:2925-2934.
- Li, J.-P., and D. Baltimore. 1990. An intragenic revertant of a poliovirus 2C mutant has an uncoating defect. *J. Virol.* **64**:1102-1107.
- 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.
- Liu, D. X., S. Shen, H. Y. Xu, and S. F. Wang. 1998. Proteolytic mapping of the coronavirus infectious bronchitis virus 1b polyprotein: evidence for the presence of four cleavage sites of the 3C-like proteinase and identification of two novel cleavage products. *Virology* **246**:288-297.
- Liu, D. X., H. Y. Xu, and T. D. K. Brown. 1997. Proteolytic processing of the coronavirus infectious bronchitis virus 1a polyprotein: identification of a 10-kilodalton polypeptide and determination of its cleavage sites. *J. Virol.* **71**:1814-1820.
- 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.
- Plagemann, P. G. W. 1996. Lactate dehydrogenase-elevating virus and related viruses, p. 1105-1120. *In* B. N. Fields, P. M. Knipe, and P. M. Howley

- (ed.), Fields virology. Lippincott-Raven Publishers, Philadelphia, Pa.
36. 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.
 37. Ryan, M. D., and M. Flint. 1997. Virus-encoded proteinases of the picornavirus supergroup. *J. Gen. Virol.* **78**:699–723.
 38. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Plainview, N.Y.
 39. 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.
 40. Smith, S. L., X. C. Wang, and E. K. Godeny. 1997. Sequence of the 3' end of the simian hemorrhagic fever virus genome. *Gene* **191**:205–210.
 41. 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.
 42. Snijder, E. J., and J. J. M. Meulenbergh. 1998. The molecular biology of arteriviruses. *J. Gen. Virol.* **79**:961–979.
 43. 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, N.Y.
 44. 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.
 45. 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.
 46. 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.
 47. 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.
 48. Strauss, J. H., and E. G. Strauss. 1994. The alphaviruses: gene expression, replication, and evolution. *Microbiol. Rev.* **58**:491–562.
 49. Tibbles, K. W., I. Brierley, D. Cavanagh, and T. D. K. Brown. 1996. Characterization in vitro of an autocatalytic processing activity associated with the predicted 3C-like proteinase domain of the coronavirus avian infectious bronchitis virus. *J. Virol.* **70**:1923–1930.
 50. van Berlo, M. F., J. J. Zeegers, M. C. Horzinek, and B. A. M. van der Zeijst. 1983. Antigenic comparison of equine arteritis virus (EAV) and lactic dehydrogenase virus (LDV); binding of staphylococcal protein A to the nucleocapsid protein of EAV. *Zentbl. Vetmed. Reihe B* **30**:297–304.
 51. Vance, L. M., N. Moscufo, M. Chow, and B. A. Heinz. 1997. Poliovirus 2C region functions during encapsidation of viral RNA. *J. Virol.* **71**:8759–8765.
 52. van der Meer, Y., H. van Tol, J. Krijnse Locker, and E. J. Snijder. 1998. ORF1a-encoded replicase subunits are involved in the membrane association of the arterivirus replication complex. *J. Virol.* **72**:6689–6698.
 53. van der Most, R. G., P. J. Bredenbeek, and W. J. M. Spaan. 1991. A domain at the 3' end of the polymerase gene is essential for encapsidation of coronavirus defective interfering RNAs. *J. Virol.* **65**:3219–3226.
 54. van Dinten, L. C., J. A. den Boon, A. L. M. Wassenaar, W. J. M. Spaan, and E. J. Snijder. 1997. An infectious arterivirus cDNA clone: identification of a replicase point mutation which abolishes discontinuous mRNA transcription. *Proc. Natl. Acad. Sci. USA* **94**:991–996.
 55. van Dinten, L. C., A. L. M. Wassenaar, A. E. Gorbalenya, W. J. M. Spaan, and E. J. Snijder. 1996. Processing of the equine arteritis virus replicase ORF1b protein: identification of cleavage products containing the putative viral polymerase and helicase domains. *J. Virol.* **70**:6625–6633.
 56. Wassenaar, A. L. M., W. J. M. Spaan, A. E. Gorbalenya, and E. J. Snijder. 1997. Alternative proteolytic processing of the arterivirus replicase ORF1a polyprotein: evidence that nsp2 acts as a cofactor for the nsp4 serine protease. *J. Virol.* **71**:9313–9322.
 57. 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.
 58. Ziebuhr, J., and S. G. Siddell. 1999. Processing of the human coronavirus 229E replicase polyproteins by the virus-encoded 3C-like proteinase: identification of proteolytic products and cleavage sites common to pp1a and pp1ab. *J. Virol.* **73**:177–185.