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

VIROLOGY 182, 655–663 (1991)

Another Triple-Spanning Envelope Protein among Intracellularly Budding RNA Viruses: The Torovirus E Protein

JOHAN A. DEN BOON,[1] ERIC J. SNIJDER,[1] JACOMINE KRIJNSE LOCKER, MARIAN C. HORZINEK, AND PETER J. M. ROTTIER[2]

Institute of Virology, Veterinary Faculty, State University of Utrecht, Yalelaan 1, 3584 CL, Utrecht, The Netherlands

Received October 15, 1990; accepted January 11, 1991

The nucleotide sequence of the Berne virus envelope (E) protein gene was determined and its 26.5K translation product was identified by *in vitro* transcription and translation. Computer analysis of the protein sequence revealed the characteristics of a class III membrane protein lacking a cleaved signal sequence but containing three successive transmembrane α-helices in the N-terminal half, much the same as the coronavirus membrane (M) protein. The disposition of the E protein in the membrane was studied by *in vitro* translation in the presence of microsomes and by subsequent proteinase K digestion. Only small portions of either end of the polypeptide were found to be exposed on opposite sides of the vesicle membranes. Experiments with a hybrid E protein (EM) containing the C-terminal tail of a coronavirus M protein, to which an anti-peptide serum was available, showed that this C-terminus was present at the cytoplasmic side of the membrane, which is another similarity to the coronavirus M protein. Immunofluorescence experiments indicated that the EM protein, expressed by a recombinant vaccinia virus, accumulated in intracellular membranes, predominantly those of the endoplasmic reticulum. The common features of the torovirus E and the coronavirus M protein support our hypothesis that an evolutionary relationship exists between these groups of intracellularly budding viruses.

INTRODUCTION

Berne virus (BEV) is the prototype of a recently described group of enveloped positive-stranded RNA viruses, the toroviruses (Horzinek and Weiss, 1984; Horzinek *et al.*, 1987), which have recently been assigned a generic status (Francki, 1991). Over the past few years we have studied the BEV genome organization and replication strategy. A substantial part of the genome has been cloned and the structural genes as well as part of the nonstructural genes of the virus have been sequenced. These studies have revealed the basic features of the toroviral transcription and translation strategy.

The BEV genome is a polygenic RNA with an estimated size of 25–30 kb. The replication strategy of the virus involves the generation of a 3′-coterminal nested set of four subgenomic mRNAs to express the viral structural genes (Snijder *et al.*, 1990a). Apart from the smallest mRNA, which is monocistronic, the subgenomic mRNAs are structurally polycistronic but only their respective 5′-terminal open reading frames are translated (Snijder *et al.*, 1988). By these characteristics the expression of the BEV genome is very similar to the mechanism used by coronaviruses (Spaan *et al.*, 1988). In addition, a clear but distant ancestral relationship between the polymerase genes of both virus groups has been established (Snijder *et al.*, 1990c).

Toroviruses are characterized by the unique tubular morphology of their helical nucleocapsid and by the presence of club-shaped spikes (peplomers) on their envelope (Weiss *et al.*, 1983). The most abundant structural protein of BEV is the 18.3K nucleocapsid (N) protein (Snijder *et al.*, 1989) which constitutes about 84% of the total virion protein mass (Horzinek *et al.*, 1985). Three putative membrane-associated proteins with M_r of 75–100K, 37K, and 22K have been described (Horzinek *et al.*, 1985). Of these, the 75–100K moiety represents the N-glycosylated peplomer (P) protein which constitutes the BEV surface projections (Snijder *et al.*, 1990b). The virus specificity of the 37K protein has become questionable since a corresponding gene has not been identified in the 3′ part of the viral genome where the information for the structural proteins is clustered (Snijder *et al.*, 1990a).

This paper focuses on the smallest and most abundant membrane protein, the 22K envelope (E) protein. This unglycosylated polypeptide (Horzinek *et al.*, 1986) accounts for about 13% of the virion protein mass (Horzinek *et al.*, 1985). *In vitro* studies have shown that the E protein is translated from BEV RNA 3 (Snijder *et*

[1] Present address: Institute of Medical Microbiology, Department of Virology, State University of Leiden, AZL, Postbus 320, 2300 AH Leiden, The Netherlands.

[2] To whom correspondence and reprint requests should be addressed. The nucleotide sequence data reported in this paper have been submitted to the EMBL nucleotide sequence database and have been assigned Accession No. X52505.

0042-6822/91 $3.00

al., 1988). In this report, we present the primary structure and identification of the BEV E protein gene; in addition, its membrane topology and intracellular localization have been studied.

MATERIALS AND METHODS

cDNA synthesis and cloning

The preparation of cDNA libraries from poly(A)-selected RNA from BEV-infected cells and the screening of cDNA clones have been described before (Snijder *et al.*, 1990a).

DNA and protein sequence analysis

The sequence analysis of the most 3'-terminal 10 kb of the BEV genome has been described previously (Snijder *et al.*, 1990a). Direct sequencing of plasmid DNA was carried out using the technique described by Chen and Seeburg (1985). Protein sequences were analyzed using the methods of von Heijne (1981) and Eisenberg *et al.* (1984) and the PEPTIDESTRUCTURE option of the software provided by the Computer Genetics Group/University of Wisconsin (Version 5, 1989; Devereux *et al.*, 1984). The FASTA program (Pearson and Lipman, 1988) was used to search for sequence similarities among the entries of the NBRF Protein Identification Resource (Release 22.0).

Construction of pBSE1 and pBSE2

An 1182-bp *Bam*HI–*Bgl*II fragment from BEV clone 108 (Snijder *et al.*, 1990a) containing the 3' end of the BEV P protein gene and the entire E protein gene was cloned into the *Bam*HI site of transcription vector pBS (Stratagene). To remove the upstream P protein gene sequences, 1 μg of this construct was cut with *Bam*HI and *Kpn*I and then treated for $5\frac{1}{2}$ min at 23° with 40 U exonuclease III in a 25-μl reaction buffer containing 50 m*M* Tris–HCl (pH 8.0), 5 m*M* $MgCl_2$, 10 μg/ml tRNA, and 10 m*M* 2-mercaptoethanol. Following eightfold dilution in ice-cold mung bean nuclease buffer [30 m*M* NaAc (pH 5.0), 50 m*M* NaCl, 1 m*M* $ZnCl_2$, and 5% glycerol], the sample was heated for 15 min at 68° and subsequently incubated with 5 U mung bean nuclease for 30 min at 30°. The nuclease was inactivated by adding 8 μl 10% SDS, 10 μl 1 *M* Tris–HCl (pH 9.5), and 20 μl 8 *M* LiCl. The DNA was purified, made blunt with the Klenow fragment of *Escherichia coli* DNA polymerase I (Maniatis *et al.*, 1982), religated, and used to transform *E. coli* cells. The constructs designated pBSE1 and pBSE2 were selected by direct plasmid DNA sequencing.

Construction of pBSEM

An *Xho*I linker (5' CCTCGAGG 3') was inserted into the *Sac*I site of pBSE1, which is located upstream of the E protein gene (pBSE*Xho*). Construct pBSM (Snijder *et al.*, 1990c) was used to prepare the 3'-terminal fragment of the mouse hepatitis virus (MHV) membrane (M) protein gene. First, the *Bam*HI site from the pBS multiple cloning region was removed by digestion with *Sma*I and *Sal*I, filling in of the *Sal*I sticky ends, and religation. Then a *Bam*HI linker (5' CGCGGATCCGCG 3') was inserted into the filled-in *Sty*I site of the M gene (position 693; Fig. 1, Armstrong *et al.*, 1984) resulting in pBSM*Bam.* Finally, the fragment was obtained by digestion with *Bam*HI and filling in of the sticky ends, followed by *Pst*I digestion and purification. The fragment was cloned between the filled-in *Acc*I site of the BEV E protein gene (Fig. 1, nucleotide position 705) and the pBS *Pst*I site of pBSE*Xho.* The BEV E/MHV M junction in this hybrid EM gene (pBSEM) was checked by sequence analysis.

In vitro transcription and translation

In vitro transcription using linearized pBSE1, pBSE2, or pBSEM was performed as described previously (Snijder *et al.*, 1989). One microgram of each transcript was translated in an 11.5-μl system containing 7.5 μl rabbit reticulocyte lysate (N.90; Amersham), 25 μCi [^{35}S]methionine, and 20 U RNasin, either in the presence or in the absence of 0.01 A_{280} units of salt-extracted, nuclease-treated dog pancreatic rough microsomes (Blobel and Dobberstein, 1975; kindly provided by D. I. Meyer). Translation reaction mixtures were incubated for 30 min at 30°.

Proteinase K digestion

Post-translational proteinase K digestions were performed by adding proteinase K (final concentration: 0.25 μg/μl) to 2.5 μl of *in vitro* translation reaction mixture, either in the absence of detergent or in the presence of 0.05% saponin or 1% Triton X-100. Samples were incubated at room temperature for 60 min and the enzyme was inactivated by adding phenylmethylsulfonyl fluoride to a final concentration of 2.5 μg/μl.

Generation of recombinant vaccinia viruses

The EM hybrid gene was cut out from pBSEM with *Xho*I and the filled-in fragment was cloned into the *Sma*I site of transfer vector pSC11 (Chakrabarti *et al.*, 1985). Vaccinia virus-infected HeLa cells were transfected with the plasmid DNA to recombine the EM gene into the thymidine kinase locus of the vaccinia virus genome (Mackett *et al.*, 1984). The recombinant

designated vEM was selected by blue/white screening (Chakrabarti *et al.*, 1985) and plaque-purified. The generation of the vaccinia virus recombinant expressing the MHV M protein (vM) will be described elsewhere.

Radioactive labeling of BEV proteins

BEV-infected embryonic mule skin cells (10^6 cells; m.o.i. of 10; Snijder *et al.*, 1988) were labeled with 100 μCi/ml [^{35}S]methionine from 9 to 15 hr postinfection (p.i.). Cells were lysed in 300 μl lysis buffer [10 m*M* Tris–HCl (pH 7.4), 1 m*M* EDTA, 150 m*M* NaCl, 1% Triton X-100]. Nuclei were removed by centrifugation at 10,000 *g* for 30 min and the lysate was stored at −70°.

Immunoprecipitation and SDS–PAGE

In vitro translation products of pBSEM and proteins from lysates of vEM-infected cells were immunoprecipitated (Snijder *et al.*, 1990c) using an anti-peptide serum directed against the 18 most C-terminal amino acid residues of the MHV-A59 M protein. Translation reaction mixtures, infected cell lysates, and immunoprecipitated proteins were analyzed in 20% SDS–polyacrylamide gels followed by fluorography.

Immunofluorescence assay

Indirect immunofluorescence microscopy of recombinant vaccinia virus-infected COS cells was carried out essentially as described by Rose and Bergmann (1982). Semiconfluent cell monolayers, grown on cover slips, were infected with an m.o.i. of 10 and fixed with paraformaldehyde after 16 hr of incubation. Following permeabilization with 1% Triton X-100, cells were incubated with the MHV M-specific anti-peptide serum (see above) at a 1:150 dilution in the presence of 5% fetal calf serum to reduce background fluorescence. Subsequently, cells were stained with fluorescein-conjugated goat anti-rabbit IgG (1:150; Kallestad, Chaska, MN). Cover slips were mounted in phosphate-buffered saline (pH 8.6), containing 90% glycerol and 25 mg/ml Dabco (Sigma). Fluorescence was viewed using an Olympus BHS-F microscope.

RESULTS

Nucleotide sequence of BEV ORF 3

The nucleotide sequence analysis of the region of the BEV genome which encodes the structural proteins has been described recently (Snijder *et al.*, 1990a). Four open reading frames (ORFs; numbered 2 to 5) were identified in the 3′-terminal 7 kb of the genome. On the basis of their position in the sequence and the lengths of the subgenomic BEV RNAs the expression of ORFs 2 to 5 was assigned to RNAs 2 to 5 (Snijder *et al.*, 1990a). Recently, ORFs 2 and 5 have been identified as the BEV P and N protein genes, respectively (Snijder *et al.*, 1989, 1990b).

The nucleotide sequence of the 699-nt ORF 3 is presented in Fig. 1. The sequence corresponds to nucleotide positions 1909 through 1211 of the BEV genome relative to the 3′ poly(A) tail. A so-called "core promoter" sequence occurs upstream and downstream of the coding region. This conserved sequence has been found to precede each BEV ORF and has been postulated to be a regulating element in the transcription process that generates the subgenomic mRNAs (Snijder *et al.*, 1990a).

BEV ORF 3 encodes the E protein

The nucleotide sequence of BEV mRNA 3 comprises the 3′-terminal 2 kb of the genome and contains ORFs 3, 4, and 5 (Snijder, 1990a). *In vitro* translation of RNA 3 resulted in the synthesis of a polypeptide with the electrophoretic mobility of the viral E protein (Snijder *et al.*, 1988). Consequently, it was inferred that ORF 3, the 5′-terminal ORF in this mRNA, encodes this structural protein. Although the M_r of the BEV E protein had been estimated at about 22K (Horzinek *et al.*, 1985), the calculated size of the predicted 233-amino-acid ORF 3 product is 26.5K. This discrepancy might be the result of aberrant electrophoretic migration. Alternatively, translation of the mRNA could initiate at the second instead of the first in-frame AUG codon (nucleotides 176–178 in Fig. 1), yielding a 187-amino-acid 21.2K predicted polypeptide.

To identify the BEV E protein gene, both the complete ORF 3 sequence and a truncated form, lacking the first 36 nucleotides but still containing the second in-frame ATG, were cloned into transcription vector pBS. *In vitro* transcripts of the two resulting plasmids (pBSE1 and pBSE2, respectively) were translated in a reticulocyte cell-free system. The products were analyzed by polyacrylamide gel electrophoresis. As is shown in Fig. 2 (lane 4) the truncated ORF 3 construct pBSE2 directed the synthesis of one major polypeptide running ahead of the BEV E protein synthesized in infected cells (lane 1). A polypeptide with a similar mobility (M_r of about 17K) was also obtained as a minor product from the full-size construct pBSE1, probably resulting from internal initiation at the second AUG. The main product of pBSE1, however, was a polypeptide with the electrophoretic mobility of the BEV E protein.

Theoretical analysis of the BEV E protein

The amino acid sequence of the BEV E protein is shown in Fig. 1. The secondary structure of the poly-

FIG. 1. Nucleotide sequence and deduced amino acid sequence of BEV ORF 3. Initiation (>>>) and termination (<<<) codons for translation are underlined. Methionine residues in the amino acid sequence are boxed. Putative core promoter sequences for the synthesis of BEV RNAs 3 and 4 are indicated. The only potential N-glycosylation site is marked with a triangled asterisk. The nucleotide sequence data presented in this figure have been submitted to the EMBL nucleotide sequence database and have been assigned Accession No. X52505.

peptide was analyzed using the method of von Heijne (1981). This analysis is based on the calculation of the free energy cost of burying successive stretches of amino acid residues into a membrane. Using a 21-residue moving window the pattern of Fig. 3A was generated. The plot shows a striking resemblance with the ones obtained earlier for the small envelope glycoprotein of several coronaviruses (e.g., Boursnell *et al.*, 1984; Rottier *et al.*, 1986). For comparison the profile of the M (formerly E1; see Cavanagh *et al.*, 1990) protein of the coronavirus mouse hepatitis virus is shown in Fig. 3B. Like the M protein, the predicted BEV polypeptide is generally rather hydrophobic except for its hydrophilic N-terminus. These features are indicative of a membrane protein with an internal signal sequence which may be present in the region between residues 30 and 115, where the plot in Fig. 3A suggests 3 membrane-spanning domains. Interestingly, as in most coronavirus M proteins a so-called surface helix is predicted immediately adjacent to the presumed third transmembrane domain. According to the method described by Eisenberg *et al.* (1984) residues 115 through 128 adopt this structure (values of mean hydrophobicity and mean hydrophobic moment approx. 0.1 and 0.65, respectively). Notably, this is the only region where the sequences of the BEV E protein and the coronavirus M proteins show similarity. As illustrated in Fig. 4 this is most obvious when comparing with MHV and bovine coronavirus (9 identities and 2 conserving substitutions within the 13-residue boxed domain). For the rest of the molecule the similarity was low and even strikingly absent in the region immediately C-terminal to the predicted surface helix. This is surprising since this domain has been extremely well conserved among coronaviral M proteins: except for the human coronavirus (HCV 229E) a stretch of 9 residues is essentially identical in all sequences known to

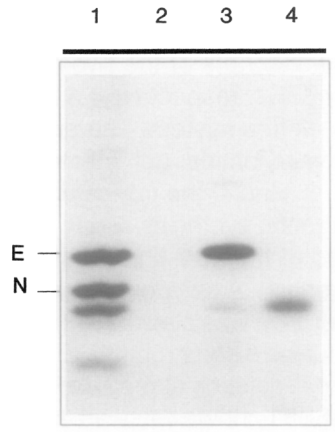

FIG. 2. Identification of the BEV E protein gene. Lane 1 shows a direct analysis of a BEV-infected cell lysate. The positions of the BEV E and N protein are indicated. The *in vitro* translation products of transcripts from the constructs pBSE1 and pBSE2 were analyzed in lanes 3 and 4, respectively. Lane 2 shows a control translation to which no RNA was added.

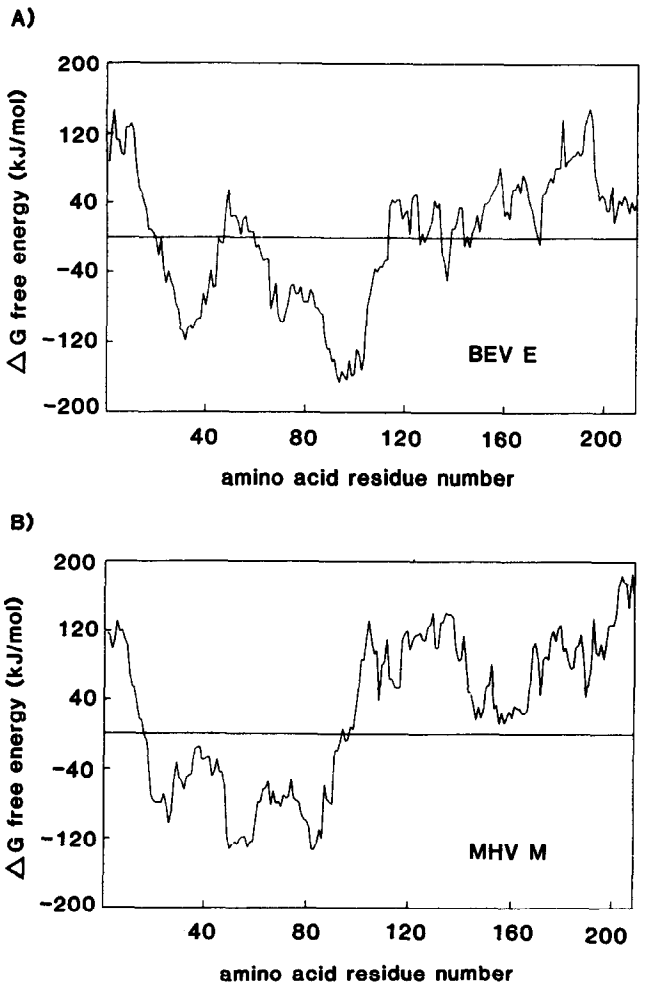

FIG. 3. Free energy values corresponding to the insertion of successive 21-amino-acid protein segments from an aqueous environment into a lipid bilayer. Calculations were done according to von Heijne (1981) and plotted as a function of the first amino acid residue in the segment. (A) The BEV E protein. (B) The MHV M protein.

date (Fig. 4). With the exception of the small motif described above, no significant sequence similarities between the BEV E protein and other proteins (including the entries in the NBRF protein database) were detected.

The C-terminal half of the BEV E protein is of an amphipathic nature. A possible N-glycosylation site is predicted at residue number 208 (Fig. 1). This site is probably not used, however, since earlier studies of the BEV envelope proteins have indicated that only the P protein is N-glycosylated (Horzinek *et al.*, 1986).

Membrane topology of the BEV E protein

To analyze the disposition of the E protein in membranes we used the protease protection assay that proved successful earlier in establishing the topology of the coronavirus MHV M protein (Rottier *et al.*, 1984). Transcripts of pBSE1 were translated in a reticulocyte lysate in the presence of dog pancreas rough microsomes and samples of the translation reaction were treated with proteinase K. As is clear from Fig. 5B (cf. lanes 3 and 4) the E protein was inserted into the microsomal membranes such that only about 2.2K was accessible to the enzyme. When the digestion was done after membrane permeabilization with saponin an additional fragment of approximately 1.3K was removed (lane 5). These results indicate that the two termini of the protein end up at opposite sides of the membrane, are exposed, and can be removed whereby a 18.4K protease-resistant polypeptide is left. When the proteinase K digestion was carried out after solubilization of the lipids with Triton X-100 a fragment of about 12.7K unexpectedly survived the treatment (lane 6). A resistant fragment of similar size was obtained when the translation had been done in the absence of microsomal membranes (lane 7), suggesting that part of the E protein is folded into a protease-resistant conformation also in the absence of membrane integration. Another important observation in this experiment was that translation of the pBSE1 RNA yielded the same product in the presence and in the absence of membranes (cf. lanes 2 and 3). Therefore, no cleavage of a signal sequence seems to occur upon integration of the E protein in the lipid bilayer.

As these results do not allow conclusions about the orientation of the membrane-assembled E protein, a construct (pBSEM) was engineered in which the 3′ end of the BEV E protein gene was replaced by sequences encoding the carboxy-terminal tail of the MHV-A59 M protein, to which an anti-peptide serum was available (Fig. 5A). The net 16-amino-acid C-terminal extension

IBV	M	FACLSFVGY	WIQSIRLFKRCRS	WWSFNPESN	118
FIPV	M	VTFALWMMY	FVRSVQLYRRTKS	WWSFNPETN	154
TGEV	M	VTFVLWIMY	FVRSIQLYRRTNS	WWSFNPETK	154
HCV	M	STLVMWVMY	FANSFRLFRRART	FWAWNPEVN	116
BCV	M	VAIIMWIVY	FVNSIRLFIRTGS	WWSFNPETN	122
MHV	M	VSIVIWIMY	FVNSIRLFIRTGS	WWSFNPETN	123
BEV	E	ILLIMMFVK	FVNFVKLFYRTGS	FAIAIRGPI	137

FIG. 4. Sequence alignment of the region adjacent to the third transmembrane helix of the BEV E protein and six coronaviral M proteins. A domain of 13 amino acids where the BEV sequence shows similarity to the M proteins is boxed. Note that this region is flanked by a domain that is extremely well conserved among M proteins but lacks any similarity to the BEV E protein. Abbreviations: IBV, infectious bronchitis virus (Boursnell *et al.*, 1984); FIPV, feline infectious peritonitis virus (H. Vennema, unpublished results); TGEV, transmissible gastroenteritis virus (Laude *et al.*, 1987); HCV, human coronavirus 229E (Raabe and Siddell, 1989); BCV, bovine coronavirus (Lapps *et al.*, 1987); MHV, mouse hepatitis virus (Armstrong *et al.*, 1984).

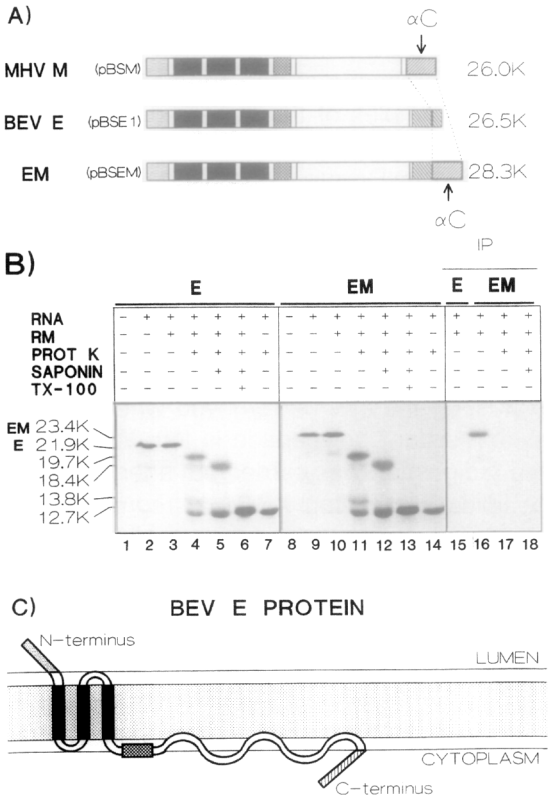

FIG. 5. Membrane topology of the BEV E protein. (A) Composition of the EM hybrid protein. The 3′-terminal sequence of the BEV E gene in pBSE1 was replaced by the 3′-terminal sequence of the MHV M protein gene from pBSM. Dotted and hatched areas represent N- and C-terminal sequences, respectively. The transmembrane domains and the domain with sequence similarity adjacent to the third transmembrane domain (Fig. 4) are indicated by filled and cross-hatched boxes, respectively. (B) Proteinase K protection analysis of wild-type E (lanes 1–7) and hybrid EM (lanes 8–14) proteins synthesized *in vitro* in the presence and absence of rough microsomes (RM). The presence (+) or absence (−) of the various components in the translation and in the protease assay is indicated. Lanes 15–18 show the results of immunoprecipitations (IP) with the anti-peptide serum recognizing the C-terminus of the MHV M protein. M_r values indicate the sizes of the E and EM proteins and their digestion products as they were calculated from their electrophoretic mobility. The positions of the primary translation products E and EM are indicated. (C) Topological model of the BEV E protein.

of the polypeptide resulting from this replacement was observed as a slight mobility shift upon electrophoretic analysis of the pBSEM translation product (Fig. 5B). The protease protection assay of the pBSEM product synthesized in the presence of microsomal membranes demonstrated that the integration of the protein in the lipid bilayer was unaltered. Moreover, it demonstrated that the C-terminal extension occurred at the outside of the membranes as it was removed by the enzyme from intact microsomal vesicles. The latter treatment also abolished recognition of the protein through the appended coronaviral peptide by the specific antiserum (cf. lanes 16 and 17), confirming the exposed status of the C-terminus.

Intracellular localization (and transport) of the EM protein

Although the exact site of budding has not yet been established, toroviruses are known to assemble at intracellular membranes (Weiss and Horzinek, 1986; Fagerland *et al.*, 1986). Viral membrane proteins are considered to be important determinants in localizing the budding of enveloped viruses. Therefore, it was of interest to examine the E protein's intracellular transport. As we did not succeed in obtaining an antiserum to the protein suitable for use in immunofluorescence, we decided to exploit the hybrid EM protein described above, which contains a coronaviral tail. Recombinant vaccinia virus vEM was generated to express the EM protein in COS cells. Its correct expression was confirmed by immunoprecipitation using the MHV M specific anti-peptide serum (data not shown). The location of the EM protein was studied by indirect immunofluorescence on permeabilized COS cells expressing the protein and was compared to that of the MHV-A59 M protein which has been shown to accumulate in the Golgi apparatus in these cells (Rottier and Rose, 1987). As shown in Fig. 6A the EM protein was localized intracellularly. Prominent reticular staining extending from the perinuclear region throughout the cell was observed. Also the nuclear membrane stained positive. The intracellular perinuclear staining remained stable even after incubating the cells for several hours in the presence of cycloheximide before analysis (data not shown). The more polarized fluorescence seen in vM-infected cells (Fig. 6B) is clearly different and typical for the localization of the MHV M protein in the Golgi complex. The faint staining over the nucleus is nonspecific as is clear from the control (Fig. 6C).

DISCUSSION

Initially considered a unique group of enveloped RNA viruses due mainly to their distinct morphological and serological characteristics, toroviruses have been shown through recent work to share several important features with coronaviruses. One interesting common feature is the intracellular maturation of these viruses, which is probably governed by specific properties of the envelope proteins. The M protein of coronaviruses has several features that distinguish it from the majority of membrane proteins (Armstrong *et al.*, 1984; Boursnell *et al.*, 1984; Rottier *et al.*, 1984, 1986); it was implicated to play a crucial role in the budding process of these viruses (Dubois-Dalcq *et al.*, 1982; Holmes *et al.*, 1981; Rottier *et al.*, 1981). Here we present evi-

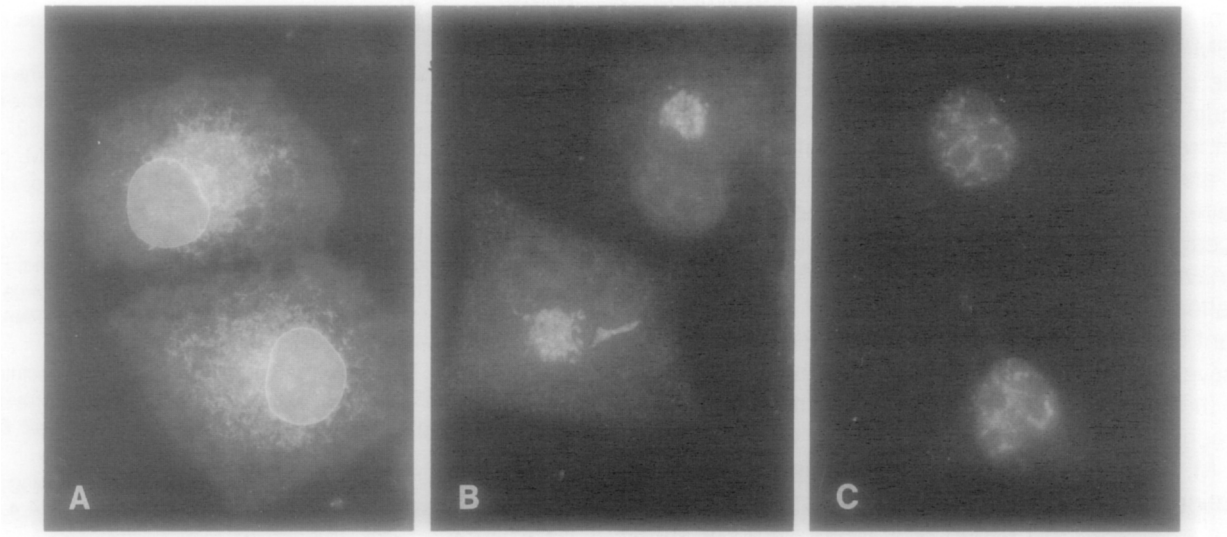

FIG. 6. Indirect immunofluorescence of recombinant vaccinia virus-infected COS cells. Cells expressing the EM protein (A), the coronavirus M protein (B), and control cells infected with wild-type vaccinia virus (C) were fixed at 16 hr p.i. and stained using the anti-peptide serum recognizing the C-terminus of the MHV M protein and fluorescein-conjugated goat anti-rabbit IgG.

dence that toroviruses contain a membrane protein with a structure similar to that of the coronaviral M protein.

Berne virus ORF 3 was identified as the viral E protein gene. Its product exhibited the features of the E protein which is expressed from subgenomic mRNA 3 in infected cells. However, its apparent molecular weight upon electrophoresis in polyacrylamide gel was smaller than the M_r predicted from the sequence. This was shown not to result from initiation of translation at a downstream AUG codon (Fig. 2). Since processing of the protein can be excluded, we conclude that the E protein migrates aberrantly in gels possibly due to its extremely hydrophobic nature. Similar observations have been made for the coronavirus M protein (Armstrong *et al.*, 1984).

Analysis of the E protein sequence revealed the characteristics of a class III membrane protein containing three membrane spanning α-helices in its N-terminal part. Our experiments showed that the protein is not proteolytically processed upon membrane insertion, which implies that one of the hydrophobic transmembrane domains functions as an internal signal sequence.

The proteinase K assays indicated that about 85% of the E protein becomes resistant to proteolysis when it is inserted into the lipid bilayer. Only small portions of either end of the polypeptide appear to be exposed. However, roughly half of the E protein (12.7K) folds into a conformation that resists proteolytic attack irrespective of the presence of membranes. This fragment could represent either the N-terminal hydrophobic core or a tightly assembled C-terminal part of the protein. However, examination of the primary structure of the protein reveals that no methionine residues occur in the C-terminal half of the molecule (Fig. 1), which excludes this part as being the 12.7K fragment. Altogether these considerations lead us to the model shown in Fig. 5C, which is similar to the model that has been put forward for coronavirus M proteins (Rottier *et al.*, 1986; Kapke *et al.*, 1988).

The proposed disposition of the E protein in microsomal membranes implies that it is the C-terminal domain that will interact with the viral nucleocapsid at the cytoplasmic face of the membrane during the act of budding. Consequently, in virions the N-terminus will be exposed at the outside while the C-terminus will be buried inside the particle. The BEV E protein is not N-glycosylated upon *in vivo* (data not shown) or *in vitro* expression, which can now be explained by its topology in the membrane: the only potential N-glycosylation site (amino acid position 208) is oriented towards the cytoplasm.

By electron microscopy toroviral budding has been shown to occur at various intracellular membranes (Weiss and Horzinek, 1986; Fagerland *et al.*, 1986). The Golgi system was described as the most prominent site of budding but virus particles were also found to bud into non-Golgi smooth membranes and, at late stages of infection, even into the ER and the perinuclear space (Weiss and Horzinek, 1986). We assume that the toroviral envelope proteins (E and P) play an important role in virus assembly. This implies that their intracellular localization may very well determine the budding site. Unfortunately we were unable to study the intracellular transport of the E protein itself due to a

lack of suitable antibodies. We therefore constructed the EM hybrid protein assuming that the replacement of the cytoplasmic tail of the E protein by the corresponding tail of a similar coronavirus protein would not affect transport. Indeed, the available data do not indicate a role in targeting for the C-terminus of the M protein. For instance, replacing the cytoplasmic tail of the vesicular stomatitis virus G protein by that of the coronavirus infectious bronchitis virus M protein did not alter the cell surface expression of the resulting mutant G protein (Puddington *et al.,* 1987). Moreover, we have recently shown that deletion of the C-terminus from the MHV M protein does not alter its membrane topogenesis or its accumulation in the Golgi system (P. J. M. Rottier and J. Krijnse Locker, unpublished observations).

The expressed EM hybrid protein appeared to stay intracellularly in the ER region and not to be transported to the cell surface to an appreciable extent. There was no indication of any accumulation in the Golgi apparatus. Apparently, the protein is not targeted to the major site of torovirus budding. Again, this is reminiscent of coronaviruses: whereas MHV is known to bud into pre-Golgi smooth membranes, the viral M protein when expressed from cloned cDNA accumulates in the Golgi complex (Rottier and Rose, 1987). We interpret these data to indicate that in infected cells the envelope proteins interact with each other and possibly also with other viral factors resulting in their accumulation at the site where budding occurs.

The BEV E and P protein (Snijder *et al.,* 1990b) have now been shown to be structurally analogous to the coronaviral M and spike proteins. In view of the recently described evolutionary relationship between toro- and coronaviruses (Snijder *et al.,* 1990b,c), we postulate that these similarities reflect common ancestry, even though significant amino acid sequence similarity can no longer be detected. The membrane proteins of toro- and coronaviruses probably contain properties which are essential for virus assembly and which have therefore been conserved during the evolution of these intracellularly budding RNA viruses.

ACKNOWLEDGMENTS

The authors thank Willy Spaan for helpful discussions and continuing support, and Harry Vennema for sharing sequence data prior to publication. This project was supported by the Division for Health Research TNO in cooperation with the Organization for Scientific Research (NWO; Project 900-502-081).

REFERENCES

ARMSTRONG, J., NIEMANN, H., SMEEKENS, S., ROTTIER, P. J. M., and WARREN, G. (1984). Sequence and topology of a model intracellular membrane protein, E1 glycoprotein, from a coronavirus. *Nature* **308,** 751–752.

BLOBEL, G., and DOBBERSTEIN, B. (1975). Transfer of proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. *J. Cell Biol.* **67,** 852–862.

BOURSNELL, M. E. G., BROWN, T. D. K., and BINNS, M. M. (1984). Sequence of the membrane protein gene from avian coronavirus IBV. *Virus Res.* **1,** 303–314.

CAVANAGH, D., BRIAN, D. A., ENJUANES, L., HOLMES, K. V., LAI, M. M. C., LAUDE, H., SIDDELL, S. G., SPAAN, W., TAGUCHI, F., and TALBOT, P. J. (1990). Recommendations of the coronavirus study group for the nomenclature of the structural proteins, mRNAs, and genes of coronaviruses. *Virology* **176,** 306–307.

CHAKRABARTI, S., BRECHLING, K., and MOSS, B. (1985). Vaccinia virus expression vector: coexpression of β-galactosidase provides visual screening of recombinant virus plaques. *Mol. Cell. Biol.* **5,** 3403–3409.

CHEN, E. Y., and SEEBURG, P. H. (1985). Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. *DNA* **4,** 165–170.

DEVEREUX, J., HAEBERLI, P., and SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12,** 387–395.

DUBOIS-DALCQ, M. E., DOLLER, E. W., HASPEL, M. V., and HOLMES, K. V. (1982). Cell tropism and expression of mouse hepatitis virus (MHV) in mouse spinal chord cultures. *Virology* **119,** 317–331.

EISENBERG, D., SCHWARZ, E., KOMAROMY, M., and WALL, R. (1984). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **179,** 125–142.

FAGERLAND, J. A., POHLENZ, J. F. L., and WOODE, G. N. (1986). A morphological study of the replication of Breda virus (proposed family *Toroviridae*) in bovine intestinal cells. *J. Gen. Virol.* **67,** 1293–1304.

FRANCKI, R. I. B. (1991). *I.C.T.V. Sixth Report of the International Committee on Taxonomy of Viruses,* in press.

HOLMES, K. V., DOLLER, E. W., and STURMAN, L. S. (1981). Tunicamycin resistant glycosylation of a coronavirus glycoprotein: Demonstration of a novel type of viral glycoprotein. *Virology* **115,** 334–344.

HORZINEK, M. C., and WEISS, M. (1984). *Toroviridae:* A taxonomic proposal. *Zentralbl. Veterinaermed. Reihe B* **31,** 649–659.

HORZINEK, M. C., EDERVEEN, J., and WEISS, M. (1985). The nucleocapsid of Berne virus. *J. Gen. Virol.* **66,** 1287–1296.

HORZINEK, M. C., EDERVEEN, J., KAEFFER, B., DE BOER, D., and WEISS, M. (1986). The peplomers of Berne virus. *J. Gen. Virol.* **67,** 2475–2483.

HORZINEK, M. C., FLEWETT, T. H., SAIF, L. J., SPAAN, W. J. M., WEISS, M., and WOODE, G. N. (1987). A new family of vertebrate viruses: *Toroviridae. Intervirology* **27,** 17–24.

KAPKE, P. A., TUNG, F. Y. T., HOGUE, B. G., BRIAN, D. A., WOODS, R. D., and WESLEY, R. (1988). The amino-terminal signal peptide on the porcine transmissible gastroenteritis coronavirus matrix protein is not an absolute requirement for membrane translocation and glycosylation. *Virology* **165,** 367–376.

LAPPS, W., HOGUE, B. G., and BRIAN, D. A. (1987). Sequence analysis of the bovine coronavirus nucleocapsid and matrix protein genes. *Virology* **157,** 47–57.

LAUDE, H., RASSCHAERT, D., and HUET, J. C. (1987). Sequence and N-terminal processing of the transmembrane protein E1 of the coronavirus transmissible gastroenteritis virus. *J. Gen. Virol.* **68,** 1687–1693.

MACKETT, M., SMITH, G. L., and MOSS, B. (1984). General method for production and selection of infectious vaccinia virus recombinants expressing foreign genes. *J. Virol.* **49,** 857–864.

MANIATIS, T., FRITSCH, E. F., and SAMBROOK, J. (1982). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

PEARSON, W. R., and LIPMAN, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85,** 2444–2448.

PUDDINGTON, L., MACHAMER, C. E., and ROSE, J. K. (1987). Cytoplasmic domains of cellular and viral integral membrane proteins substitute for the cytoplasmic domain of the VSV G protein to the cell surface. *J. Cell Biol.* **102,** 2147–2157.

RAABE, T., and SIDDELL, S. G. (1989). Nucleotide sequence of the gene encoding the membrane protein of human coronavirus 229E. *Arch. Virol.* **107,** 323–328.

ROSE, J. K., and BERGMANN, J. E. (1982). Expression from cloned cDNA of cell surface secreted forms of the glycoprotein of vesicular stomatis virus in eukaryotic cells. *Cell* **30,** 753–762.

ROTTIER, P. J. M., HORZINEK, M. C., and VAN DER ZEIJST, B. A. M. (1981). Viral protein synthesis in mouse hepatitis virus strain A59-infected cells: Effect of tunicamycin. *J. Virol.* **40,** 350–357.

ROTTIER, P. J. M., BRANDENBURG, D., ARMSTRONG, J., VAN DER ZEIJST, B. A. M., and WARREN, G. (1984). Assembly in vitro of a spanning membrane protein of the endoplasmic reticulum: The E1 glycoprotein of coronavirus mouse hepatitis virus A59. *Proc. Natl. Acad. Sci. USA* **81,** 1421–1425.

ROTTIER, P. J. M., WELLING, G. W., WELLING-WESTER, S., NIESTERS, H. G. M., LENSTRA, J. A., and VAN DER ZEIJST, B. A. M. (1986). Predicted membrane topology of the coronavirus protein E1. *Biochemistry* **25,** 1335–1339.

ROTTIER, P. J. M., and ROSE, J. K. (1987). Coronavirus E1 glycoprotein expressed from cloned cDNA localizes in the Golgi region. *J. Virol.* **61,** 2042–2045.

SNIJDER, E. J., EDERVEEN, J., SPAAN, W. J. M., WEISS, M., and HORZINEK, M. C. (1988). Characterization of Berne virus genomic and messenger RNAs. *J. Gen. Virol.* **69,** 2135–2144.

SNIJDER, E. J., DEN BOON, J. A., SPAAN, W. J. M., VERJANS, G. M. G. M., and HORZINEK, M. C. (1989). Identification and primary structure of the gene encoding the Berne virus nucleocapsid protein. *J. Gen. Virol.* **70,** 3363–3370.

SNIJDER, E. J., HORZINEK, M. C., and SPAAN, W. J. M. (1990a). A 3′-coterminal nested set of independently transcribed messenger RNAs is generated during Berne virus replication. *J. Virol.* **64,** 331–338.

SNIJDER, E. J., DEN BOON, J. A., SPAAN, W. J. M., WEISS, M., and HORZINEK, M. C. (1990b). Primary structure and post-translational processing of the Berne virus peplomer protein. *Virology* **178,** 355–363.

SNIJDER, E. J., DEN BOON, J. A., BREDENBEEK, P. J., HORZINEK, M. C., RIJNBRAND, R., and SPAAN, W. J. M. (1990c). 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.

SPAAN, W. J. M., CAVANAGH, D., and HORZINEK, M. C. (1988). Coronaviruses: Structure and genome expression. *J. Gen. Virol.* **69,** 2939–2952.

VON HEIJNE, G. (1981). On the hydrophobic nature of signal sequences. *Eur. J. Biochem.* **116,** 419–421.

WEISS, M., STECK, F., and HORZINEK, M. C. (1983). Purification and partial characterization of a new enveloped RNA virus Berne virus. *J. Gen. Virol.* **64,** 1849–1858.

WEISS, M. and HORZINEK, M. C. (1986). Morphogenesis of Berne virus (proposed family Toroviridae). *J. Gen. Virol.* **67,** 1305–1314.