TGEV Corona Virus ORF4 Encodes a Membrane Protein That Is Incorporated into Virions

MURIELLE GODET, RENE L'HARIDON, JEAN-FRANCOIS VAUTHEROT, AND HUBERT LAUDE[1]

Unité de Virologie et Immunologie Moléculaires, Institut National de la Recherche Agronomique, 78350 Jouy-en-Josas, France

The coding potential of the open reading frame ORF4 (82 amino acids) of transmissible gastroenteritis virus (TGEV) has been confirmed by expression using a baculovirus vector. Five monoclonal antibodies (MAbs) raised against the 10K recombinant product immunoprecipitated a polypeptide of a similar size in TGEV-infected cells. Immunofluorescence assays performed both on insect and mammalian cells revealed that ORF4 was a membrane-associated protein, a finding consistent with the prediction of a membrane-spanning segment in ORF4 sequence. Two epitopes were localized within the last 21 C-terminal residues of the sequence through peptide scanning and analysis of the reactivity of a truncated ORF4 recombinant protein. Since the relevant MAbs were found to induce a cell surface fluorescence, these data suggest that ORF4 may be an integral membrane protein having a Cexo-Nendo orientation. Anti-ORF4 MAbs were also used to show that ORF4 polypeptide may be detected in TGEV virion preparations, with an estimated number of 20 molecules incorporated per particle. Comparison of amino acid sequence data provided strong evidence that other coronaviruses encode a polypeptide homologous to TGEV ORF4. Our results led us to propose that ORF4 represents a novel minor structural polypeptide, tentatively designated SM (small membrane protein). © 1992 Academic Press, Inc.

INTRODUCTION

Transmissible gastroenteritis virus (TGEV), an important pathogen of swine, is a member of the coronaviridae, a family of enveloped viruses with a large (~30 kb) continuous, positive RNA genome. Sequencing data have led to the identification of a number of large open reading frames (ORFs). ORF1a and b, which account for the 5' two-thirds of the coronavirus genome, are assumed to encode nonstructural proteins including the viral replicase/transcriptase. The seven to eight remaining ORFs are expressed through a set of 3' coterminal, subgenomic size mRNAs of which only the unique region is translationally active. These include the ORFs coding for the virion structural proteins, i.e., the nucleocapsid (N) and two or three envelope glycoproteins: the spike (S) and the membrane (M) proteins, and the hemagglutinin-esterase (HE) present in a coronavirus subset. These ORFs are distributed following the consensus gene order 5' Pol-(HE)-S-M-N 3'. The other ORFs are interspersed within the genome and their number and position differ among coronavirus members (reviewed by Spaan *et al.*, 1988; and Lai, 1990). They have been shown to be expressed by functionally mono-, di-, or tricistronic mRNAs and were generally assumed to encode nonstructural proteins, the function of which is still unknown or conjectural.

In TGEV genome, four such ORFs have been deduced. Two of them are expressed by the same mRNA (mRNA 3) in two out of the three virus strains sequenced (Rasschaert *et al.*, 1987; Kapke *et al.*, 1988; Britton *et al.*, 1989; Wesley *et al.*, 1989). The predicted product of ORF3a is 61 to 71 codon long, with a variable C-terminal end. It appears to be dispensable for virus replication since it was found to be absent in a TGEV variant strain SP (Wesley *et al.*, 1990) as well as in the closely related porcine respiratory coronavirus (PRCV) (Rasschaert *et al.*, 1990). ORF3b (expressed by a separate mRNA species numbered 3-1 in Miller strain) has a constant length (244 residues); however, one clone of Purdue-115 strain was reported to have an ORF3b which is shortened by 79 codons at the 5' end and in several cDNA clones by 67 codons at the 3' end (Rasschaert *et al.*, 1987). ORF4 was predicted to encode a 82 amino acid long hydrophobic polypeptide (Rasschaert *et al.*, 1987; Kapke *et al.*, 1988; Britton *et al.*, 1989; Wesley *et al.*, 1989). So far, the putative products of the three above-mentioned ORFs have not been identified in infected cells. The last ORF, ORF7, is located downstream of the N gene (Kapke and Brian, 1986; Rasschaert *et al.*, 1987; Britton *et al.*, 1988), an unusual feature among coronaviruses. A polypeptide of M_r 14K, reacting with antibodies produced against an ORF7 synthetic peptide, has been characterized in TGEV-infected cells (Garwes *et al.*, 1989).

In this study we report the identification of a product of one of these ORFs (ORF4) in infected cells and its

---

preliminary characterization. In particular, we show evidence that ORF4 represents a novel virion-associated polypeptide with a possible counterpart in other coronaviruses.

METHODS

Cells and viruses

Autographa californica nuclear polyhedrosis virus (AcNPV) and recombinant baculoviruses were grown and assayed in confluent monolayers of *Spodoptera frugiperda* (Sf9) cells in medium containing 10% (v/v) fetal bovine serum, according to the procedures described by Brown and Faulkner (1977). Propagation of the high cell passage Purdue-115 strain of TGEV in swine cell lines PD5 or ST was done as previously described (Laude *et al.,* 1986).

Construction of recombinant baculovirus transfer plasmids pVLORF4 and pVLORF4Δ

Manipulations of plasmid DNA were performed according to the procedures described by Sambrook *et al.* (1989). Restriction enzymes, T4 DNA ligase, and calf intestine alkaline phosphatase (CIP) were purchased from Boehringer-Mannheim. The baculovirus transfer plasmid containing the full-length cDNA copy of TGEV ORF4 coding sequence was constructed following the general scheme outlined in Fig. 1. The *Nde*I–*Ssp*I fragment (0.7 kbp) derived from plasmid pTG2-15 (Rasschaert *et al.,* 1987) was digested with the *Dde*I restriction enzyme. *Dde*I–*Ssp*I DNA fragment was repaired with the Klenow large fragment of DNA polymerase I and cloned into the *Bam*HI site of the pVL941 vector (Luckow and Summers, 1989). The resulting plasmid, named pVLORF4, contained a 0.33 kbp insert. A second plasmid, pVLORF4Δ was constructed by inserting an ORF4 gene in which the 3′ last 63 nt were deleted through PCR mutagenesis (Scharf *et al.,* 1986) on pTG2-15 using the oligonucleotides 5′ G AAGAAGGGATCCATACCTATGAC and 5′ CTTA-TAGGGATCCTAAGCATG as 5′ and 3′ amplimers, respectively. These amplimers were designed to introduce an additional stop codon TAG at the 3′ end of the gene and a *Bam*HI cloning site at each end. The amplification product was digested by *Bam*HI and ligated into the pVL941 cloning site. The orientation and sequence of the ORF4 and ORF4Δ inserts relative to the AcNPV polyhedrin leader were determined by restriction analysis and partial DNA sequencing. In these constructs, the initiation codon of the ORF4 and ORF4Δ sequences were positioned at 54 and 7 bp from the *Bam*HI cloning site, respectively.

Transfection and selection of recombinant viruses

Transfer of the TGEV ORF4 gene into the AcNPV genome was accomplished by transfection of Sf9 cells using the calcium phosphate precipitation technique as described by Summers and Smith (1987). Recombinant baculoviruses were screened by dot blot hybridization using an ORF4-specific [^{32}P]-labeled DNA fragment as a probe. Four polyhedrin-negative clones were tested for ORF4 expression. TGEV ORF4Δ gene was introduced into a linear form of AcNPV DNA (Kitts *et al.,* 1990). Circular AcRP6-SC DNA was linearized by digestion at the unique *Bsu*36I site. Two hundred nanograms of *Bsu*36I or mock digest viral DNA was mixed with 1 μg of pVLORF4Δ DNA and transfected into Sf9 cells using the lipofectin method (Kitts *et al.,* personal communication) according to the procedure of the manufacturer (GIBCO-BRL). After a 2-day incubation the culture supernatants were harvested and plated. A dozen well-isolated plaques were picked out and screened for ORF4Δ expression using [^{35}S]methionine-labeled cultures.

Production of hybridoma antibodies

Three Balb/c mice were injected threefold intraperitoneally at a 1 month interval with 1 × 10^7 AcORF4-infected cells (disrupted in Freund complete adjuvant for the first injection). Three days before fusion, the mice were boosted both intraperitoneally and subcutaneously with ORF4 protein purified from 4 × 10^7 infected cells by 15% SDS–PAGE. Splenocytes from one mouse that tested positive by immunoprecipitation assay were fused with SP$_2$O myeloma cells. Supernatants of hybrid clones were tested in a comparative immunofluorescence assay (see below) using AcORF4- or AcNPV-infected Sf9 monolayers. Subcloning of ORF4-specific antibody-producing hybridomas and Ig isotyping were done as described elsewhere (L'Haridon *et al.,* 1991). IgGs purified from ascites fluids by ammonium sulfate precipitation and gel permeation on a Sephacryl-S200 column were used in all experiments.

Indirect immunofluorescence assay

Screening of hybridoma was performed on Sf9 cell monolayers established in 96-well microplates, infected with baculovirus (m.o.i. 10 PFU) and fixed with acetone/ethanol (v/v) at 38 hr p.i. For surface fluorescence analysis, aliquots of cells in suspension were stained with MAbs at 100 μg/ml in Grace medium and then with FITC conjugate (each step 1 hr at 4°) and spotted onto glass slides. Alternatively, spotted cells were fixed with 4% paraformaldehyde and permeabil-

ized or not with 0.1% Triton X-100 before staining (1 hr at 37°). Similar experiments were performed on ST cell monolayers infected by TGEV at a m.o.i. of 0.1 PFU and fixed 15 hr p.i.

Radioisotope labeling of insect and mammalian cell polypeptides

The procedures for metabolic labeling of insect and mammalian cells were as reported previously (Godet et al., 1991; Delmas et al., 1990). Monolayers of 4×10^5 Sf9 cells or 6×10^6 PD5 cells labeled with [^{35}S]-methionine were washed with PBS and lyzed in 4 ml of PBS-Triton (Tris, 50 mM, pH 8.5, 1% Triton X-100, 10^3 kallicrein units of aprotinin per milliliter). Resulting cytosols of Sf9 cells and PD5 cells were centrifuged 30 min at 10,000 g or 1 hr at 30,000 rpm in a 50 Ti rotor (Beckman), respectively and stored in aliquots at −70°.

Immunoprecipitation assay

Aliquots of radiolabeled cytosols or virions were adjusted to 0.5 ml with PBS-Triton buffer containing the appropriate MAb (100 µg/ml) or 3 µl of ascites fluid from a feline infectious peritonitis virus-infected cat (used as a source of anti-TGEV polyclonal antibodies) and protein A–Sepharose beads (50 µl of a 50% suspension); after a 2-hr incubation at room temperature with agitation, the immune complexes were extensively washed with PBS-Triton, then with 0.5 M NaCl + 50 mM Tris (pH 8). Beads were treated for 3 min at 100° in sample buffer. The immunoprecipitated material thus released was analyzed by 15% or 15–20% SDS–PAGE.

Analysis of virion-associated polypeptides

Virions in the supernatant of cultures labeled as above were purified following a described procedure (Laude et al., 1986). The material pelleted by ultracentrifugation at 35,000 rpm in a 45 Ti rotor was resuspended in distilled water and was applied to a linear sucrose gradient (16 ml, 20 to 45% sucrose w/v in distilled water). Centrifugation was performed in an Sw27 rotor for 3 hr at 25,000 rpm and 4°. Gradients were collected from top to bottom into 500-µl fractions. Half of each fractions was run at 100,000 rpm for 30 min and the resulting pellets were analyzed by SDS–PAGE electrophoresis on a 8–20% gradient gel. The material remaining in the virus-containing fractions was pooled, pelleted as above, and split into three parts; one was analyzed directly by 15% SDS–PAGE; the two others were solubilized in PBS-Triton, immuno-precipitated by a different MAb each, and analyzed as above. In one experiment, virion-associated material was subjected to a second round of gradient purifica-
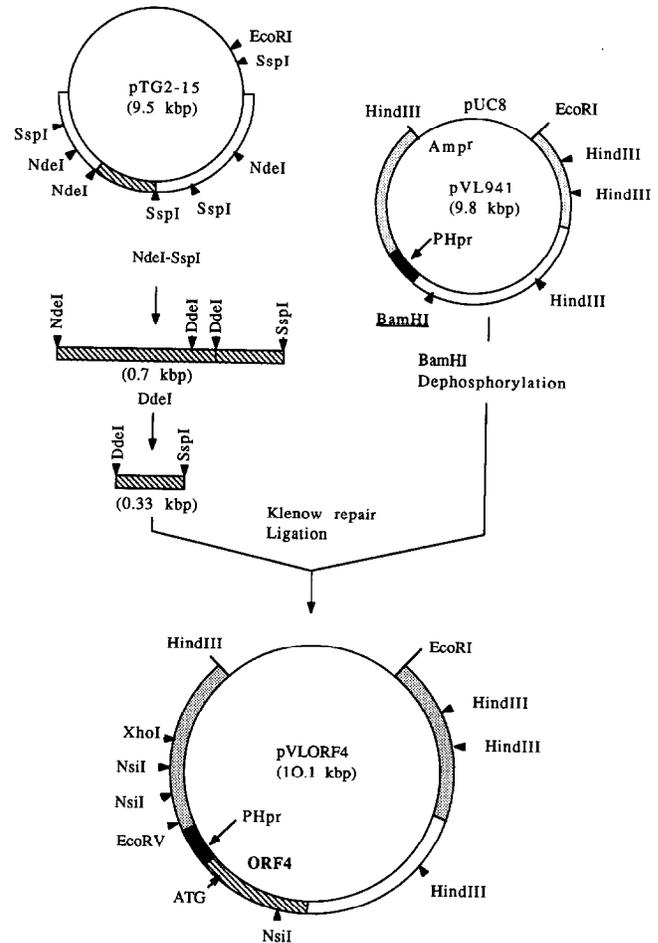
FIG. 1. Construction of the recombinant baculovirus AcORF4. See Methods for details. Abbreviations are: Ampr, ampicillin resistance gene; PHpr, polyhedrin promotor.

tion (30 to 60% concave sucrose gradient, 16 hr at 25,000 rpm; Laude et al., 1986) prior to gel analysis.

Epitope mapping

Peptides were synthesized on polyethylene pins (Geysen et al., 1984) by using a commercially available kit, according to the procedure given by the manufacturer (Cambridge Research Biochemicals). Immunoreactivity of the immobilized peptides was assayed by ELISA using anti-ORF4 MAbs (25 µg/ml) as primary antibody and an anti-mouse IgG (H + L) peroxydase conjugate (Biosys).

RESULTS

Expression of coronavirus ORF4 protein in insect cells

Insertion of the entire ORF4 coding sequence into the genome of AcNPV baculovirus was performed by using the transfer plasmid pVL941 (see Methods and Fig. 1). Clones exhibiting a positive hybridization with a
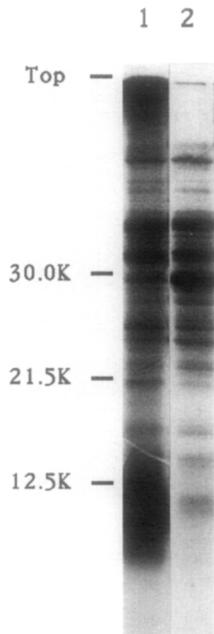
FIG. 2. Expression of TGEV ORF4 by recombinant baculovirus. Sf9 cells infected with recombinant AcORF4 (lane 1) or wild type AcNPV (lane 2) baculovirus were labeled at 28 hr p.i. with [^{35}S]methionine. Proteins from 2×10^5 cells harvested 5 hr later were resolved in a 15% SDS–PAGE and visualized by fluorography.

specific probe and a polyhedron-negative phenotype were selected, amplified, and tested for the expression of ORF4. Among four selected ORF4-expressing clones, one, designated AcORF4, was retained for subsequent studies. Analysis of [^{35}S]methionine-labeled AcORF4-infected cells revealed the presence of a major polypeptide of M_r 10K (Fig. 2), in good agreement with the predicted m.w. of ORF4 product (9.2K). The time course for its synthesis was found to be from 24 to 48 hr p.i.
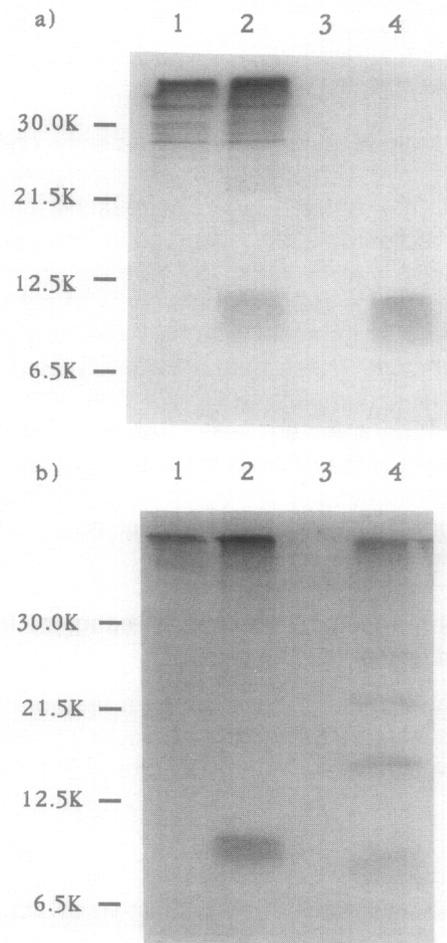
Detection of ORF4 in TGEV-infected cells using monoclonal antibodies

Screening of positive hybridoma clones was achieved by comparative indirect immunofluorescence assay on AcORF4- or AcNPV-infected cells. Among 268 hybridomas tested, 5 positive clones, all producing MAbs of IgG1 isotype, were subcloned and used for the production of ascites fluids. When assayed by immunoprecipitation of AcORF4-infected cells extracts, all 5 MAbs recognized a single species of M_r 10K, identical to the target protein (Fig. 3a, lane 4). An immunoreactive species comigrating with recombinant ORF4 protein was detected in TGEV-infected mammalian cells as well (lane 2). Analysis of the recombinant protein in nonreducing conditions revealed the

existence of oligomeric forms, which were not observed with the authentic ORF4 protein (Fig. 3b).

ORF4 product is a membrane-associated protein

Indirect immunofluorescence assays were performed to determine the subcellular location of ORF4 protein. In acetone-fixed cultures of both AcORF4-infected Sf9 and TGEV-infected cells, all anti-ORF4 MAbs induced a strong fluorescence, which were polarized in a juxtanuclear region consistent with Golgi localization (Fig. 4a). In addition, a bright fluorescence was observed on unfixed (Fig. 4b) or paraformalde-hyde-fixed (not shown) Sf9-infected cultures. In para-formaldehyde-fixed TGEV-infected cultures, plaques of

FIG. 3. Characterization of ORF4 protein in recombinant baculo-virus- and TGEV-infected cells by using a monoclonal antibody. (a) Cell extracts from 2×10^4 Sf9 or 7.5×10^5 PD5-infected cells labeled with [^{35}S]methionine were immunoprecipitated using MAb V27. Immunoreactive material was analyzed as described in Fig. 2. Mock-infected PD5 cells (lane 1), TGEV-infected PD5 cells (lane 2), AcNPV- (lane 3) or AcORF4- (lane 4) infected Sf9 cells. (b) Same materials resolved in the absence of 2-mercaptoethanol.
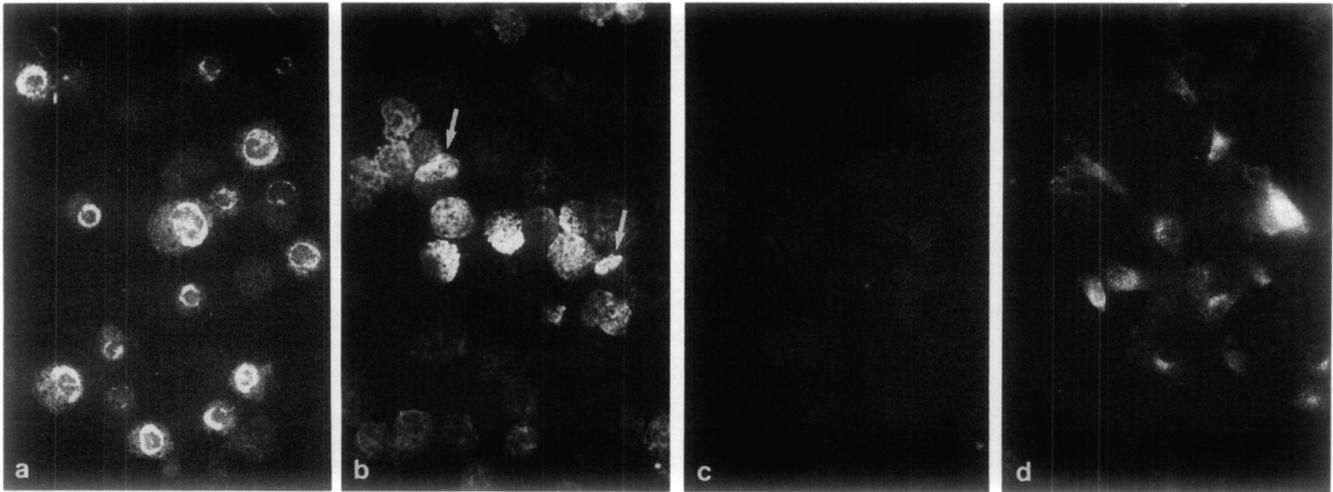
FIG. 4. Indirect immunofluorescence assay on cells expressing ORF4 protein. Sf9 cells infected at a m.o.i. of 10 PFU with AcORF4 (a, b) or AcNPV (c) baculoviruses were stained using MAb V27 (5 μg/ml) at 40 hr p.i. with (a) or without (b, c) prior fixation with ethanol–acetone. Arrows point to cells showing evidence of antigen capping, indicative of surface fluorescence. TGEV-infected ST cells (m.o.i. of 0.1 PFU) were fixed with paraformaldehyde at 15 hr p.i. then stained using the same MAb (d).
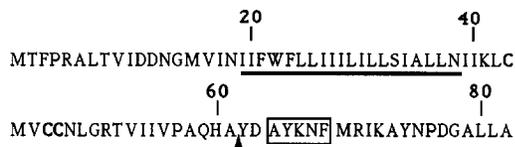
positively stained cells were consistently observed (Fig. 4d), whereas virtually all the cells showed the presence of ORF4 antigen after permeabilization (not shown). The fluorescence observed in nonpermeabilized cells was unlikely due to cell damage since only a few of them were stained positively with a MAb directed against an intracellular antigen (TGEV N protein; data not shown). These data were interpreted as reflecting a late accumulation of ORF4 at the outer membrane of TGEV-infected cells. These observations showing that ORF4 protein is found in association with cellular membranes are consistent with the prediction of a membrane-spanning domain in its amino acid sequence (Fig. 5).

Anti-ORF4 monoclonal antibodies recognize the C-terminal domain of the protein

Assuming that ORF4 was an integral membrane protein, it was of interest to determine whether exposed epitopes were located within the carboxy or amino do-

```
              20                    40
              |                     |
MTFPRALTVIDDNGMVINIIFWFLLIIILILLSIALLNIIKLC
              60                    80
              |                     |
MVCCNLGRTVIIVPAQHAYD AYKNF MRIKAYNPDGALLA
                           ▲
```

FIG. 5. Predicted amino acid sequence of TGEV ORF4 protein. The ORF4 sequence of Purdue-115 strain (Rasschaert et al., 1987) is shown. A predicted α-helix with low hydrophobic moment (program PEPLOT from UWGCG package) is underlined. Arrow marks the C-terminus of the truncated species ORF4Δ; the core sequence for the binding of 2 anti-ORF4 MAbs is boxed (see results).

main of the polypeptide chain. The peptide scanning method (Geysen et al., 1984) was used since two anti-ORF4 MAbs were shown to be reactive toward denatured and reduced protein. A set of peptides encompassing all overlapping linear nonapeptides homologous with ORF4 sequence was synthesized and tested against each of the five MAbs. As shown in Fig. 6, MAb V27 strongly recognized a linear epitope centered on residues AYKNF (positions 64 to 68; see Fig. 5). MAb S2 gave comparable results, although a lesser reactivity was observed when compared to MAb V27. No reactivity was observed with the three remaining MAbs, possibly because of a lower avidity.

To test the possibility that these three anti-ORF4 MAbs may also recognize the carboxy-subterminal region of the molecule, a second recombinant baculovirus designated AcORF4Δ, which expressed a truncated form of ORF4 lacking the 21 C-terminal amino acids, was constructed via PCR mutagenesis. Immunoprecipitation analysis revealed that a polypeptide of slightly reduced size was expressed by AcORF4Δ-infected cells and recognized by polyclonal antibodies but not by any of the anti-ORF4 MAbs (partial data in Fig. 7).

Identification of ORF4 into TGEV particles

In order to determine whether the protein is incorporated within the virion as a possible envelope protein, labeled TGEV particles were purified by centrifugation and analyzed by gel electrophoresis (Fig. 8). Five well-defined major bands were visible, which corresponds to the previously recognized virion-associated polypeptides: (i) a 220K band (S protein), (ii) a 47K band (N
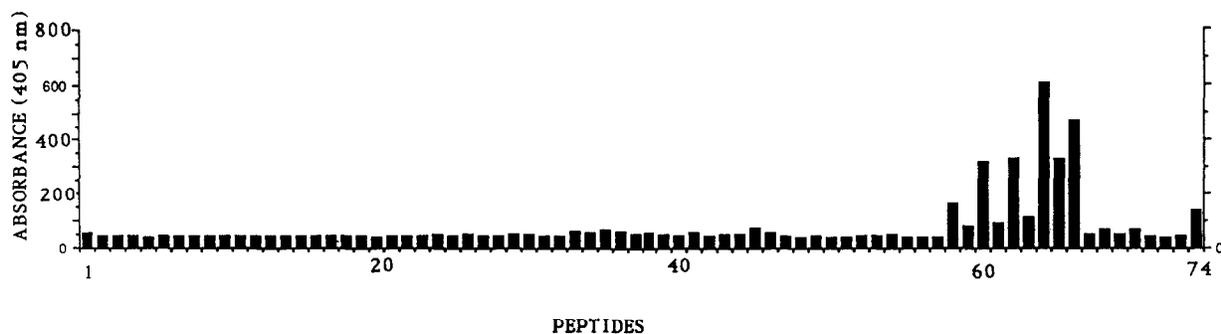
FIG. 6. Epitope mapping of anti-ORF4 MAb V27. A series of nonapeptides (overlapping 1) spanning the length of the entire ORF4 sequence was tested for ELISA reactivity toward MAb V27. N-terminus is on the left.

protein), and (iii) three bands corresponding to different species of M protein; 30–36K identified as complex type glycosylated forms, 29K (high mannose form) and 26K (unglycosylated form) (Delmas and Laude, 1991). A single additional minor band of M_r 10K, similar to that of ORF4 polypeptide, was detected (Fig. 8a). The fact that very few, if any, other polypeptides copurified renders unlikely the possibility that the 10K polypeptide remained nonspecifically associated to the sedimenting virus particles. True association of the 10K band with virions was confirmed by showing that the observed polypeptide pattern remained unchanged after an additional round of purification by isopycnic centrifugation. Immunoprecipitation with MAb V27 of detergent-dissociated material from virus-containing fractions confirmed the identity between the 10K virion-associated and ORF4 polypeptides (Fig. 8b). Densitometric tracing of autoradiograms at different times of exposure was performed to evaluate the relative amounts of the virion-associated polypeptides. The re-

sulting values were corrected according to the predicted number of Met residues in the respective sequences. This led to an estimated molar ratio of 1:20:300 for ORF4, S, and M polypeptides.

A protein homologous to TGEV ORF4 is predicted in other coronaviruses

A common feature of coronavirus genomes appears to be the existence, immediately upstream from the membrane protein M gene, of an ORF 250 to 330 nucleotide long, which is predicted to encode a hydrophobic polypeptide. Furthermore, recent studies on mouse hepatitis virus (MHV), infectious bronchitis virus (IBV), and bovine (BCV) coronaviruses have shown that the products translated from these ORFs were associated with the membrane of infected cells (Leibowitz et al., 1988; Abraham et al., 1990; Smith et al., 1990), as evidenced now for TGEV ORF4. These observations prompted us to reexamine the relevant amino acid sequences for possible similarities that earlier comparative analysis by us and others failed to detect (except for the closely related MHV and BCV viruses). Figure 9 shows a tentative alignment of the ORF4-like sequences from 5 coronavirus members, which was outlined using the program MULTALIN (Corpet, 1988) and refined manually. On the basis of the deduced consensus sequence, in which 37 residues are conserved in at least 3 out of the 5 sequences aligned (87 positions), we conclude that these proteins share significant similarities. As a striking feature, a 19–20 residue-long hydrophobic stretch, followed by a cluster of 2 or 3 cysteines is present 15 to 20 residues from the N terminus. The C-terminal part of the polypeptides seems to be more distantly related; in particular, IBV protein extended the consensus sequence by 14 to 28 residues, depending of the strain (see Liu et al., 1991 for IBV ORF3C sequence data).
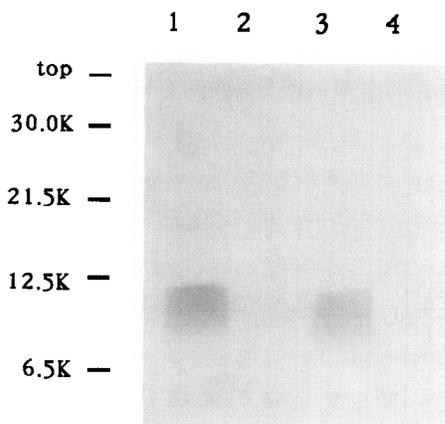
FIG. 7. Immunoreactivity of truncated ORF4 product. Labeled Sf9 cell extracts from 2 × 10⁴ AcORF4- (lane 1) or AcORF4Δ- (lane 2) infected cells were immunoprecipitated with MAb R16 at 100 μg/ml. AcORF4Δ- (lane 3) or wild type AcNPV- (lane 4) infected cell extracts were immunoprecipitated with anti-FIPV polyclonal antibodies.

DISCUSSION

In this study we have confirmed the coding potential of the ORF encoded by TGEV mRNA 4 and character-

FIG. 8. Analysis of TGEV virion-associated polypeptides. [³⁵S]methionine-labeled virions were purified by rate zonal centrifugation. Sedimentable particles in numbered fractions (Frn) were pelleted and analyzed as follows. (a) Half of the material was dissociated in sample buffer and resolved on a discontinuous 8–20% SDS–PAGE. (b) The remaining material from Frn 11 to 18 was pooled, split into three parts (1.5 × 10⁶ cpm each), and then analyzed on a 15% SDS–PAGE directly (lane 1) or after solubilization in PBS-Triton and immunoprecipitation using MAb V27 (lane 2) or anti-N MAb 5.1 (lane 3). The position of ORF4 polypeptide is indicated.

ized several properties of its translation product. Baculovirus-vectored expression of ORF4 resulted in the synthesis of a 10K polypeptide. The same species was identified in TGEV-infected cells by immunoprecipitation using monoclonal antibodies (MAbs) raised against the recombinant polypeptide.

Examination of its subcellular localization by immunostaining showed that ORF4 translation product is a membrane-associated polypeptide. This finding is consistent with the prediction in the second N-terminal quarter of a stretch of uncharged residues with proper-

ties of a membrane-spanning domain (Fig. 5). Immunofluorescence data also suggested that ORF4 may enter the exocytic pathway. In TGEV-infected cells, however, ORF4 could be detected in association with the cell surface only at a late stage of the infection.

Assuming that the observed surface fluorescence was related to externally exposed determinants, it was of interest to map the relevant epitopes on the amino acid sequence of the molecule. Peptide scanning led to the identification of residues 64-AYKNF-68 as the core sequence of the binding site of 2 of the MAbs

FIG. 9. Identification of amino acid sequences homologous to TGEV ORF4 in the genome of other coronaviruses. The amino acid sequences encoded by the indicated ORF of human coronavirus (HCV 229E), MHV-JHM, BCV, and IBV have been aligned with that of TGEV ORF4, as described under Results. Designation of the ORFs is according to the recommendations of the coronavirus study group (Cavanagh *et al.*, 1990). Pairwise homologies are marked by dots and bold letters; gaps are indicated by dashes. Bottom line, consensus sequence showing residues identical in at least three of the five sequences. Sequence data from Rasschaert *et al.*, 1987; Raabe and Siddel, 1989; Skinner *et al.*, 1985; Woloszyn *et al.*, 1990; and Boursnell *et al.*, 1985. The 22 last N-terminal residues of IBV ORF3c (Beaudette strain) are not shown.

studied. Furthermore, a recombinant ORF4 protein truncated of its last 21 C-terminal amino acids was no longer recognized by any of the 5 MAbs. These results support the view that an antigenic, possibly immunodominant site is expressed in the C-subterminal part of ORF4 protein. In addition, they led us to speculate that the region of the molecule which is translocated across the membrane would correspond to its carboxy domain. Recently, a striking correlation has been reported between the transmembrane orientation of eukaryotic proteins and the disposition of charged residues surrounding the most N-terminal membrane-spanning sequence. It has been proposed that the difference in the charge of the 15 residues flanking the presumed anchor segment determines its orientation with the more positive portion facing the cytosol (Hartmann *et al.*, 1989). Applying such a rule to TGEV ORF4 sequence gives charges of −1 and +2 for the N and C flanking segments, respectively, which predicts a Nexo-Cendo orientation, in contrast to the available experimental evidence. Thus further experiments, including the production of antibodies directed to the N-terminal part of the protein, are needed for a definite assignment of the transmembrane orientation of ORF4.

The possible role of the cysteine cluster immediately downstream the hydrophobic segment was also examined. Gel analysis of recombinant ORF4 protein under nonreducing conditions revealed the presence of multimeric, predominantly dimeric forms. In contrast, ORF4 synthesized in TGEV-infected cells could be detected in a monomeric form only, suggesting that the formation of disulfide-bridged species is an artifact, presumably linked to the high level of expression of ORF4 in the insect cells (Kiefhaber *et al.*, 1991). The possibility was tested that the cystein residues could serve as an acylation site, as demonstrated for coronavirus S protein (Schmidt, 1982). However, no incorporation of palmitic acid chains could be detected in recombinant ORF4 protein (data not shown), as would be expected if the target residues belong to the ORF4 ectodomain.

An important implication of the above findings was that ORF4 could represent a structural polypeptide present in the virus envelope. Indeed, purified preparations of labeled TGEV virions revealed the presence of a previously unrecognized 10K polypeptide which was specifically immunoprecipitated by anti-ORF4 MAbs. Based on the estimated molar ratio and assuming that coronavirions bear 100 (Roseto *et al.*, 1982) to 200 spikes, each composed of 3 S molecules (Delmas and Laude, 1990), it can be inferred that approximately 15–30 copies of ORF4 protein are incorporated into TGEV virions (Purdue strain). Such a small number of molecules in virus particles does not seem to reflect a selective exclusion since S and ORF4 accumulated in infected cells at a ratio comparable to that found in virions (data not shown). These results lead us to conclude that ORF4 may represent a minor structural polypeptide, which we propose to designate by the tentative acronym SM, standing for "small membrane" protein.

Several lines of evidence lend support to the view that a gene encoding an SM-like protein is a common feature of the coronavirus genomes: (i) an ORF predicting a polypeptide with striking similarities to TGEV ORF4 was identified in the genome sequence of each of the 5 coronaviruses examined (Fig. 9) and the fact that TGEV SM was recognized by anti-FIPV antibodies argues for the presence of a related gene also in feline infectious peritonitis virus genome; (ii) the product expressed from the relevant MHV, BCV, and IBV ORFs was reported to have properties of a transmembrane polypeptide (Leibowitz *et al.*, 1988; Smith *et al.*, 1990; Abraham *et al.*, 1990); and (iii) although expressed through a mono-, di-, or tricistronic mRNA (Abraham *et al.*, 1990; Budzilowicz and Weiss, 1987; Liu *et al.*, 1991), the assumed SM-encoding genes are all located upstream and adjacent to the M protein gene. Therefore, not only the sequences show significant similarities but the gene order 5′ . . . S . . . SM/M . . . 3′ is conserved, as would be expected for a structural protein. Finally, preliminary experiments in this laboratory allowed us to detect ORF5-encoded polypeptide in association with BCV particles (N. Woloszyn, P. Boireau, and J. F. Vautherot, unpublished results).

Small integral membrane proteins have been described in several other enveloped RNA viruses, including Sindbis and Semliki Forest togaviruses, influenza A and B viruses, Simian virus 5, and respiratory syncitial paramyxoviruses (Garoff *et al.*, 1980; Welch and Sefton, 1980; Lamb *et al.*, 1985; Hiebert, 1985; Olmsted and Collins, 1989). The influenza virus M2 protein (15K) and the alphavirus 6K protein are both acylated, Nexo-Cendo transmembrane polypeptides. M2 and 6K have been shown to represent minor structural polypeptides, with an estimated number of 40 ± 25 and 24 ± 4 molecules per virion, respectively (Zebedee and Lamb, 1988; Gaedigk-Nitschko and Schlesinger, 1990). M2 has been reported to form tetrameric channels within the membrane and to be a target of the antiviral drug amantadine and of the CTL response to influenza virus infection (Hay *et al.*, 1985; Lamb *et al.*, 1985; Surgrue and Hay, 1991). Site-directed mutagenesis studies on Sindbis and Semliki forest viruses have demonstrated that 6K protein is dispensable for virus production but exerts a role late in the assembly, possibly during virus budding (Gaedigk-Nitschko and Schlesinger, 1990; Liljestrom *et al.*, 1991). The SH protein of SV5 has been reported to be orientated in the mem-

brane with its N-terminus domain exposed at the cytoplasmic face, as it might be the case for TGEV SM. Whether SH is incorporated into virions is still questioned, as well as its potential role (Hiebert *et al.,* 1988). The apparent conservation of SM gene in the coronavirus genome strongly implies that its product is essential for an efficient replication of the virus. Based on its location and its low copy number in particles, we speculate that SM would more likely play a role in modulating assembly and/or release of the virion. Thus, elucidating the function of the coronavirus SM protein might contribute to a better understanding of an important aspect of the biology of enveloped viruses. Finally, SM may be a potent surface antigen since murine antibodies recognized a domain of the protein possibly exposed on live infected cells. Preliminary experiments indicated that anti-SM antibodies are readily detected in the serum from infected swines. Therefore, the role of SM protein in humoral and cellular immune response to TGEV infection should be worth investigating in the future.

ACKNOWLEDGMENTS

Note added in proof. During the submission process of this article, a communication by (D. X. Liu and S. C. Inglis (1991, *Virology* **185,** 911–917) reported the association of IBV ORF3c protein with the virion envelope. This strengthens the view that SM-like proteins are a general feature of coronavirus.

REFERENCES

ABRAHAM, S., KIENZLE, T. E., LAPPS, W. E., and BRIAN, D. A. (1990). Sequence and expression analysis of potential nonstructural proteins of 4.9, 4.8, 12.7, and 9.5 kDa encoded between the spike and membrane protein genes of the bovine coronavirus. *Virology* **177,** 488–495.

BOURSNELL, M. E. G., BINNS, M. M., and BROWN, D. K. (1985). Sequencing of coronavirus IBV genomic RNA: Three open reading frames in the 5' "unique" region of mRNA D. *J. Gen. Virol.* **66,** 2253–2258.

BRITTON, P., CARMENES, R. S., PAGE, K. W., GARWES, D. J., and PARRA, F. (1988). Sequence of the nucleoprotein gene from a virulent British field isolate of transmissible gastroenteritis virus and its expression in *Saccharomyces cerevisiae. Mol. Microbiol.* 2(1), 89–99.

BRITTON, P., OTIN, C. L., ALONSO, J. M. M., and PARRA, F. (1989). Sequence of the coding regions from the 3.0 kb and 3.9 kb mRNA subgenomic species from a virulent isolate of transmissible gastroenteritis virus. *Arch. Virol.* **105,** 165–178.

BROWN, M., and FAULKNER, P. (1977). A plaque assay for nuclear polyhedrosis viruses using a solid overlay. *J. Gen. Virol.* **36,** 361–364.

BUDZILOWICZ, C. J., and WEISS, S. R. (1987). *In vitro* synthesis of two polypeptides from a nonstructural gene of coronavirus mouse hepatitis virus strain A59. *Virology* **157,** 509–515.

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.

CORPET, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucl. Acids Res.* **16,** 10,881–10,891.

DELMAS, B., and LAUDE, H. (1990). Assembly of coronavirus spike protein into trimers and its role in epitope expression. *J. Virol.* **64,** 5367–5375.

DELMAS, B., RASSCHAERT, D., GODET, M., GELFI, J., and LAUDE, H. (1990). Four major antigenic sites of the coronavirus transmissible gastroenteritis virus are located on the amino-terminal half of spike glycoprotein S. *J. Gen. Virol.* **71,** 1313–1323.

DELMAS, B., and LAUDE, H. (1991). Carbohydrate-induced conformational changes strongly modulate the antigenicity of coronavirus TGEV glycoproteins S and M. *Virus Res.* **20,** 107–120.

GAEDIGK-NITSCHKO, K., and SCHLESINGER, M. J. (1990). The Sindbis virus 6K protein can be detected in virions and is acylated with fatty acids. *Virology* **175,** 274–281.

GAROFF, H., FRISCHAUF, A-M., SIMONS, K., LEHRACH, H., and DELIUS, H. (1980). Nucleotide sequence of cDNA coding for Semliki Forest virus membrane glycoproteins. *Nature* **288,** 236–241.

GARWES, D. J., STEWART, F., and BRITTON, P. (1989). The polypeptide of M_r 14000 of porcine transmissible gastroenteritis virus: Gene assignment and intracellular location. *J. Gen. Virol.* **70,** 2495–2499.

GEYSEN, H. M., MELOEN, R. H., and BARTELING, S. J. (1984). Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proc. Natl. Acad. Sci. USA* **81,** 3998–4002.

GODET, M., RASSCHAERT, D., and LAUDE, H. (1991). Processing and antigenicity of entire and anchor-free spike glycoprotein S of coronavirus TGEV expressed by recombinant baculovirus. *Virology* **185,** 732–740.

HARTMANN, E., RAPOPORT, T. A., and LODISH, H. F. (1989). Predicting the orientation of eukaryotic membrane-spanning proteins. *Proc. Natl. Acad. Sci. USA* **86,** 5786–5790.

HAY, A. J., WOLSTENHOLME, A. J., SKEHEL, J. J., and SMITH, M. H. (1985). The molecular basis of the specific anti-influenza action of amantadine. *EMBO J.* **4,** 3021–3024.

HIEBERT, S. W., PATERSON, R. G., and LAMB, R. A. (1985). Identification and predicted sequence of a previously unrecognized small hydrophobic protein, SH, of the paramyxovirus simian virus 5. *J. Virol.* **55,** 744–751.

HIEBERT, S. W., RICHARDSON, C. D., and LAMB, R. A. (1988). Cell surface expression and orientation in membranes of the 44-amino-acid SH protein of simian virus. 5. *J. Gen. Virol.* **62,** 2347–2357.

KAPKE, P. A., and BRIAN, D. A. (1986). Sequence analysis of the porcine transmissible gastroenteritis coronavirus nucleocapsid protein gene. *Virology* **151,** 41–49.

KAPKE, P. A., TUNG, F. Y. T., and BRIAN, D. A. (1988). Nucleotide sequence between the peplomer and matrix protein genes of the porcine transmissible gastroenteritis coronavirus identifies three large open reading frames. *Virus Genes* 2: 3, 293–294.

KIEFHABER, T., RUDOLPH, R., KOHLER, H. H., and BUCHNER, J. (1991). Protein aggregation *in vitro* and *in vivo:* A quantitative model of the kinetic competition between folding and aggregation. *Bio/Technology* **9,** 825–829.

KITTS, P. A., AYRES, M. D., and POSSEE, R. D. (1990). Linearization of baculovirus DNA enhances the recovery of recombinant virus expression vectors. *Nucl. Acids Res.* **18,** 5667–5672.

LAI, M. M. C. (1990). Coronavirus: Organization, replication and expression of genome. *Annu. Rev. Microbiol.* **44,** 303–333.

LAMB, R. A., ZEBEDEE, S. L., and RICHARDSON, C. D. (1985). Influenza

virus M_2 protein is an integral membrane protein expressed on the infected-cell surface. *Cell* **40**, 627–633.

LAUDE, H., CHAPSAL, J. M., GELFI, J., LABIAU, S., and GROSCLAUDE, J. (1986). Antigenic structure of transmissible gastroenteritis virus. I. Properties of monoclonal antibodies directed against virion proteins. *J. Gen. Virol.* **67**, 119–130.

L'HARIDON, R. M., BOURGET, P., LEFEVRE, F., and LA BONNARDIERE, C. (1991). Production of an hybridoma library to recombinant porcine alpha I interferon: A very sensitive assay (ISBA) allows the detection of a large number of clones. *Hybridoma* **10**(1), 35–47.

LEIBOWITZ, J. L., PERLMAN, S., WEINSTOCK, G., DEVRIES, J. R., BUDZILO-WICZ, C., WEISSEMANN, J. M., and WEISS, S. R. (1988). Detection of a murine coronavirus nonstructural protein encoded in a downstream open reading frame. *Virology* **164**, 156–164.

LILJESTRÖM, P., LUSA, S., HUYLEBROECK, D., and GAROFF, H. (1991). In vitro mutagenesis of a full-length cDNA clone of Semliki Forest virus: The small 6,000-molecular-weight membrane protein modulates virus release. *J. Virol.* **65**, 4107–4113.

LIU, D. X., CAVANAGH, D., GREEN, P., and INGLIS, S. C. (1991). A polycistronic mRNA specified by the coronavirus infectious bronchitis virus. *Virology* **184**, 531–544.

LUCKOW, V. A., and SUMMERS, M. D. (1989). High level expression of nonfused foreign genes with *Autographa californica* nuclear polyhedrosis virus expression vectors. *Virology* **170**, 31–39.

OLMSTED, R. A., and COLLINS, P. L. (1989). The 1A protein of respiratory syncytial virus is an integral membrane protein present as multiple, structurally distinct species. *J. Virol.* **63**, 2019–2029.

RAABE, T., and SIDDELL, S. (1989). Nucleotide sequence of the human coronavirus HCV 229E mRNA 4 and mRNA 5 unique regions. *Nucl. Acids Res.* **17**(15), 6387.

RASSCHAERT, D., DUARTE, M., and LAUDE, H. (1990). Porcine respiratory coronavirus differs from transmissible gastroenteritis virus by a few genomic deletions. *J. Gen. Virol.* **71**, 2599–2607.

RASSCHAERT, D., GELFI, J., and LAUDE, H. (1987). Enteric coronavirus TGEV: Partial sequence of the genomic RNA, its organization and expression. *Biochimie* **69**, 591–600.

ROSETO, A., BOBULESCO, P., LAPORTE, J., ESCAIG, J., GACHES, D., and PERIES, J. (1982). Bovine enteric coronavirus structure as studied by a freeze-drying technique. *J. Gen. Virol.* **63**, 241–245.

SAMBROOK, J., FRITSCH, E. F., and MANIATIS, T. (1989). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

SCHARF, S. J., HORN, G. T., and ERLICH, H. A. (1986). Direct cloning and sequence analysis of enzymatically amplified genomic sequences. *Science* **233**, 1076–1078.

SCHMIDT, M. F. G. (1982). Acylation of viral spike glycoproteins: A feature of enveloped RNA viruses. *Virology* **116**, 327–338.

SKINNER, M. A., EBNER, D., and SIDDELL, S. G. (1985). Coronavirus MHV-JHM mRNA 5 has a sequence arrangement which potentially allows translation of a second, downstream open reading frame. *J. Gen. Virol.* **66**, 581–592.

SMITH, A. R., BOURSNELL, M. E. G., BINNS, M. M., BROWN, T. D. K., and INGLIS, S. C. (1990). Identification of a new membrane-associated polypeptide specified by the coronavirus infectious bronchitis virus. *J. Gen. Virol.* **71**, 3–11.

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

SUGRUE, R. J., and HAY, A. J. (1991). Structural characteristics of the M_2 protein of influenza A viruses: Evidence that it forms a tetrameric channel. *Virology* **180**, 617–624.

SUMMERS, M. D., and SMITH, G. E. (1987). A manual of methods for baculovirus vectors and insect cell culture procedures. Texas Agricultural Experiment Station Bulletin No. 1555 Publishers College Station, Texas.

WELCH, W. J., and SEFTON, B. M. (1980). Characterization of a small, nonstructural viral polypeptide present late during infection of BHK cells by semliki forest virus. *J. Virol.* **33**, 230–237.

WESLEY, R. D., CHEUNG, A. K., MICHAEL, D. D., and WOODS, R. D. (1989). Nucleotide sequence of coronavirus TGEV genomic RNA: Evidence for 3 mRNA species between the peplomer and matrix protein genes. *Virus Res.* **13**, 87–100.

WESLEY, R. D., WOODS, R. D., and CHEUNG, A. K. (1990). Genetic basis for the pathogenesis of transmissible gastroenteritis virus. *J. Virol.* **64**, 4761–4766.

WOLOSZYN, N., BOIREAU, P., and LAPORTE, J. (1990). Nucleotide sequence of the bovine enteric coronavirus BECV F15 mRNA 5 and mRNA 6 unique regions. *Nucl. Acids Res.* **11**, 18(5), 1303.

ZEBEDEE, S. L., and LAMB, R. A. (1988). Influenza A virus M_2 protein: Monoclonal antibody restriction of virus growth and detection of M_2 in virions. *J. Virol.* **62**, 2762–2772.