

# Expression in Animal Cells and Characterization of the Hepatitis E Virus Structural Proteins

SHAHID JAMEEL,<sup>1\*</sup> MOHAMMAD ZAFRULLAH,<sup>1</sup> MEHMET HAKAN OZDENER,<sup>1</sup>  
AND SUBRAT KUMAR PANDA<sup>2</sup>

*Virology Group, International Centre for Genetic Engineering and Biotechnology,<sup>1</sup> and Department of Pathology,  
All India Institute of Medical Sciences,<sup>2</sup> New Delhi, India*

Received 12 July 1995/Accepted 10 October 1995

**Hepatitis E virus (HEV) is a major human pathogen in much of the developing world. It is a positive-strand RNA virus with a 7.5-kb polyadenylated genome consisting of three open reading frames (ORFs). In the absence of an in vitro culture system, the replication and expression strategy of HEV and the nature of its encoded polypeptides are not well understood. We have expressed the two ORFs constituting the structural portion of the HEV genome in COS-1 cells by using simian virus 40-based expression vectors and in vitro by using a coupled transcription-translation system. We show here that the major capsid protein, encoded by ORF2, is an 88-kDa glycoprotein which is expressed intracellularly as well as on the cell surface and has the potential to form noncovalent homodimers. It is synthesized as a precursor (ppORF2) which is processed through signal sequence cleavage into the mature protein (pORF2), which is then glycosylated (gpORF2). The minor protein, pORF3, encoded by ORF3 is a 13.5-kDa nonglycosylated protein expressed intracellularly and does not show any major processing. pORF3 interacts with a cellular protein of about 18 kDa which we call 3IP, the pORF3-interacting protein. The significance of these findings are discussed in light of an existing model of HEV genome replication and expression.**

Hepatitis E virus (HEV) is responsible for large epidemics and rampant sporadic cases of acute viral hepatitis in much of the developing world, where it is endemic (3, 11, 20, 28). In developed countries, this disease is seen primarily in travellers to areas where it is endemic. Though largely a self-limited infection, it results in significant morbidity and mortality, especially among pregnant women (12) and in situations in which coinfection with other hepatic viruses may occur (17).

The viral genome has been cloned and sequenced from a number of geographically distinct HEV strains and shows a high degree of nucleotide and amino acid sequence conservation (1, 2, 6, 24, 26). The genome is a positive-stranded RNA of about 7.5 kb with short 5' and 3' noncoding regions spanning a coding region that includes three open reading frames (ORFs) (24) (Fig. 1). Of these, the N-terminal ORF1 of about 5 kb is predicted to code for the putative nonstructural proteins, that include a methyltransferase, a papain-like cysteine protease, a replicase, and an RNA-dependent RNA polymerase (13). The C-terminal region of about 2.4 kb codes for two putative structural proteins, pORF2 and pORF3, the products of ORF2 and ORF3, respectively (Fig. 1). The fact that both of these structural-region ORFs are expressed during viral infection is demonstrated by the presence of antibodies in infected humans directed against epitopes present on pORF2 as well as on pORF3 (9, 10, 16, 18, 29).

So far, HEV has not been classified conclusively into any virus family. Its provisional classification into the *Caliciviridae* family was based primarily on the presence of morphological features similar to those of other agents in this family (3, 14). However, the genome organization shows a major difference in that the small ORF, ORF3, is located mostly within ORF2 in HEV, whereas it is C terminal in calciviruses like the Norwalk

agent (8). There has also been a suggestion that HEV is a nonenveloped "alpha-like" virus (13, 21). This is based on the presence of homologous regions across the genome (13), the detection of subgenomic HEV transcripts in the livers of experimentally infected monkeys (24), and the presence of a nucleotide sequence stretch in the HEV genome that is homologous to alphaviral junction sequences (21). A conclusive classification of HEV awaits further knowledge of its expression and replication strategy and of the nature, processing, and properties of its component proteins.

The inability to grow HEV in culture has so far precluded any such studies. No information on the nature and properties of the viral antigens is available. In this work, we expressed pORF2 and pORF3 in cultured animal cells and used this expression system to study the properties and interactions of these viral proteins.

## MATERIALS AND METHODS

**Construction of expression vectors.** The cloning of ORF2 and ORF3 from an Indian strain of HEV has been described elsewhere (18). This sequence has been deposited in the GenBank database under accession number U22532. The expression plasmid used in this study, pSG1, is a modification of plasmid pSG5 (Stratagene) in which a synthetic sequence was inserted between the *EcoRI* and *BamHI* sites. This resulted in a vector with a number of unique cloning sites, including *EcoRI*-*SmaI*-*SacI*-*EcoRV*-*KpnI*-*HindIII*-*PstI*-*XhoI*-*NaeI*-*NotI*-*BamHI*. Expression from this vector in animal cells is dependent on the simian virus 40 early promoter-enhancer region, and expression in vitro is dependent on the bacteriophage T7 promoter. For the construction of expression vector pSG-ORF2, a 2-kb *NcoI*-*BamHI* fragment encompassing the entire HEV ORF2 region was isolated, end filled with the Klenow fragment of DNA polymerase, and cloned into the *EcoRV* site within the polylinker region of plasmid pSG1. For the construction of expression vector pSG-ORF3, a 700-bp *BamHI*-*EcoRI* fragment encompassing the complete HEV ORF3 was isolated and similarly cloned into plasmid pSG1. The schematics and details of vector construction are presented in Fig. 1.

**Transfection and labeling of cultured cells.** COS-1, HepG2, and Huh-7 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum and 20  $\mu$ g of gentamicin per ml. Cells were transfected at about 50% confluency with plasmid DNA by using Lipofectin (GIBCO-BRL) according to the manufacturer's guidelines. For each 60-mm-diameter culture dish, 2.5  $\mu$ g of DNA and 10  $\mu$ l of Lipofectin were used in 1.2 ml of DMEM

\* Corresponding author. Mailing address: Virology Group, ICGBE, NII Campus, Aruna Asaf Ali Marg, New Delhi 110067, India. Phone: 91-11-6865007. Fax: 91-11-6862316 or -6862317.

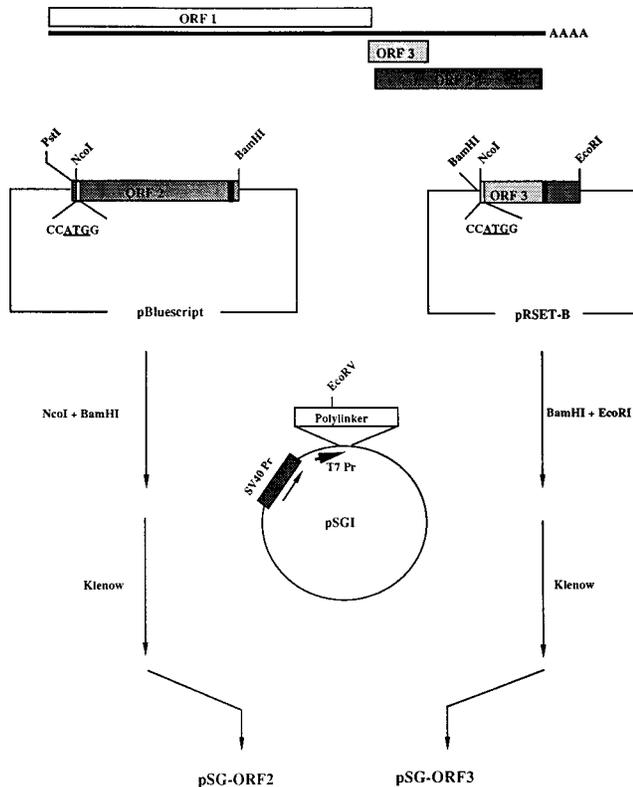


FIG. 1. HEV ORFs and cloning strategy. The three ORFs and their relative positions along the HEV genomic polyadenylated RNA are indicated. The expression vectors were constructed by subcloning the *NcoI*-*BamHI* (for ORF2) or *BamHI*-*EcoRI* (for ORF3) fragment at the *EcoRV* site within the polylinker of plasmid pSGI. The vector backgrounds for ORF2 and ORF3 fragments (shaded regions) are shown. ATG initiator codons (open bars), part of the *NcoI* site, are indicated. Closed bars represent translation termination sequences. SV40 Pr, simian virus 40 promoter-enhancer and origin of replication sequences; T7 Pr, bacteriophage T7 promoter.

without serum or antibiotics, and DNA uptake allowed to proceed for 6 h at 37°C in a CO<sub>2</sub> incubator. Forty hours posttransfection, cells were washed with 3 ml of methionine-deficient DMEM (GIBCO-BRL) and metabolically labeled with [<sup>35</sup>S]methionine (Amersham), with each 60-mm-diameter plate receiving 150  $\mu$ Ci of label in 1 ml of methionine-deficient DMEM. After a 4-h labeling period, cells were washed with ice-cold phosphate-buffered saline (PBS) and harvested for further analysis. Besides HEV ORF-containing expression plasmids, each experiment also included a control (or mock) transfection in which the same amount of the parent vector, pSGI, was used.

**Immunofluorescence.** At about 44 h posttransfection, cells were fed with 1 ml of fresh medium, kept on ice for 30 min, and scraped off with a disposable scraper. Staining was done with a 1:100 dilution of polyclonal antibodies raised in rabbits against purified pORF2 and pORF3 polypeptides expressed in *Escherichia coli* (18). For surface immunofluorescence, cells in suspension were incubated at 4°C for 1 h with diluted antibody in DMEM containing 10% fetal bovine serum. Cells were then washed three times with the above medium by centrifugation in a cold centrifuge (Hermle GmbH) at 1,000 rpm. Washed cells were incubated with a 1:100 dilution of anti-rabbit immunoglobulin G-fluorescein isothiocyanate conjugate at 4°C for 1 h and subsequently washed as described above. For intracellular localization, 10<sup>5</sup> cells in 0.5 ml were centrifuged onto glass slides at 2,000 rpm in a cytocentrifuge (Shandon). Cells were fixed in acetone and stained as described above, except that antibody dilutions and washings were carried out in PBS (pH 7.2). Stained cells were observed with an epifluorescence microscope (Nikon).

**Immunoprecipitation.** Transfected, PBS-washed COS-1 cells were harvested directly in 0.5 ml of RIPA buffer (10 mM Tris-HCl [pH 8.0], 140 mM NaCl, 5 mM iodoacetamide, 0.5% Triton X-100, 1% sodium deoxycholate, 0.1% sodium dodecyl sulfate [SDS], 2 mM phenylmethylsulfonyl fluoride) after incubation on ice for 15 to 30 min. Lysates were clarified at 10,000  $\times$  g for 10 min, and the supernatant was incubated on ice for 1 h with 5  $\mu$ l of rabbit antiserum. For immunoprecipitation with HEV immune serum, 10  $\mu$ l of pooled serum from patients with hepatitis E was used. The mix was centrifuged again as described above, and the supernatant was removed to a fresh tube. To this was added 100  $\mu$ l of a 10% suspension of RIPA buffer-washed protein A-Sepharose beads (Pharmacia, Uppsala, Sweden), and the mixture was incubated with constant shaking at 4°C for 1 h. The beads were washed five times, each time with 0.5 ml of RIPA buffer, after being centrifuged in a Costar microcentrifuge at 10,000 rpm for 10 s. Washed beads were resuspended in 50  $\mu$ l of SDS gel loading buffer (50 mM Tris-HCl [pH 6.8], 5% 2-mercaptoethanol, 2% SDS, 0.1% bromophenol blue, 10% glycerol), heated at 100°C for 2 min, and centrifuged, and the supernatants were subjected to SDS-polyacrylamide gel electrophoresis (PAGE). After electrophoresis, the gels were soaked in 0.5 M sodium salicylate for 1 h, dried, and exposed to X-ray film.

For the immunoprecipitation of antigens expressed in HepG2 and Huh-7 cells, transfected and PBS-washed cells were harvested in 1 ml of RIPA buffer. Clarified lysates were incubated with 10  $\mu$ l of preimmune rabbit serum on ice for 1

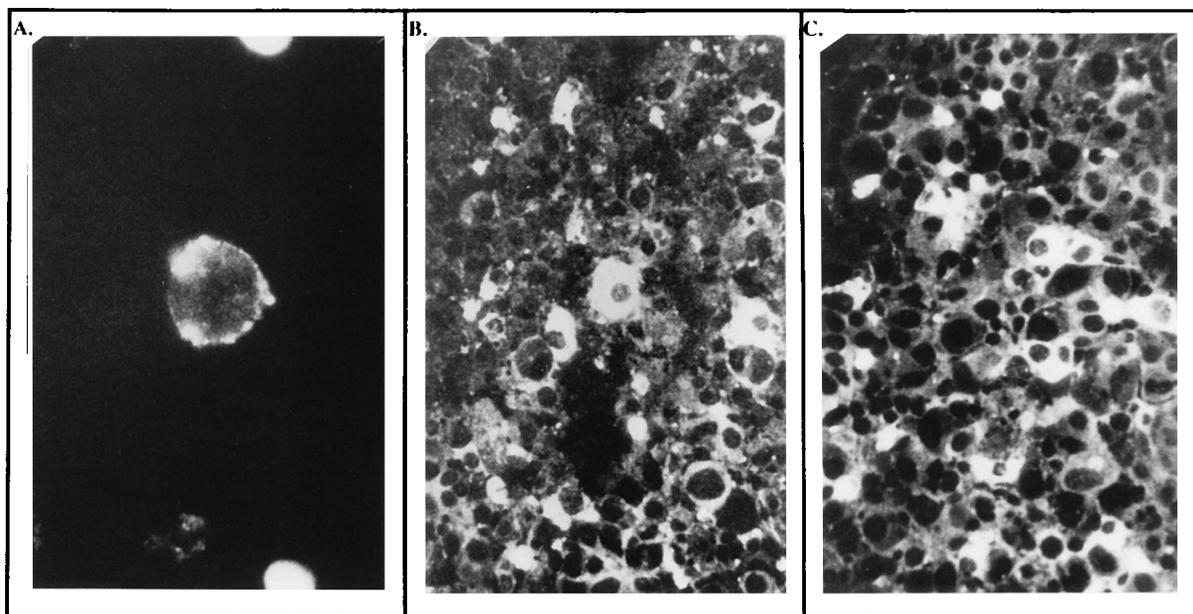


FIG. 2. Indirect immunofluorescence of HEV antigens. Immunofluorescence analysis was carried out on COS-1 cells transfected with either plasmid pSG-ORF2 (A and B) or plasmid pSG-ORF3 (C), as described in the text. Transfected cells were stained in solution for surface immunofluorescence (A) or after cytospin and fixation for intracellular immunofluorescence (B and C).

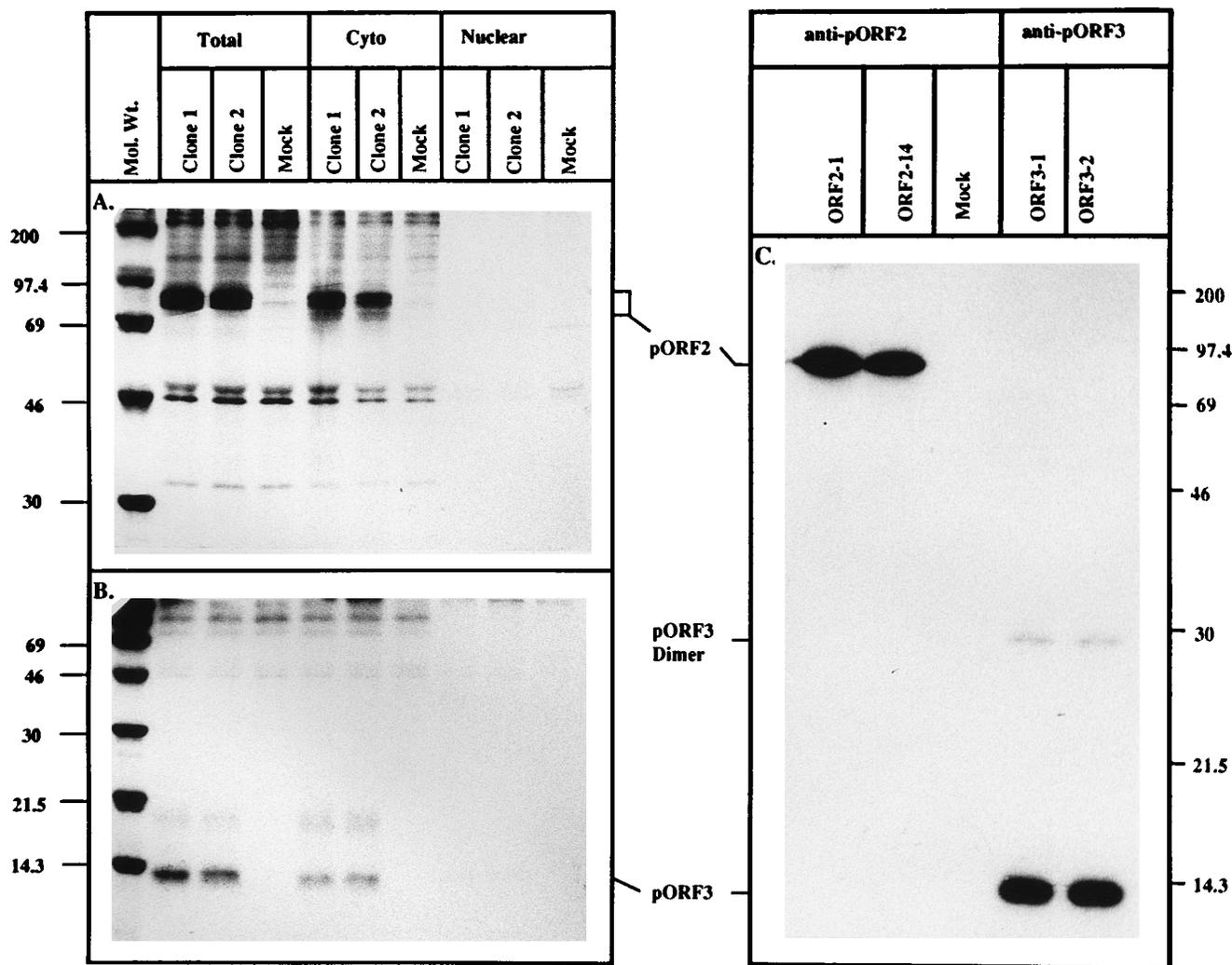


FIG. 3. Expression of HEV antigens. (A and B) Two independent clones of pSG-ORF2 (A), pSG-ORF3 (B), or the parent vector pSGI (Mock [A and B]) were transfected into COS-1 cells. Transfected cells were labeled with [<sup>35</sup>S]methionine, lysed, and fractionated into cytoplasmic (Cyto) and nuclear fractions, and the total lysates or fractions were immunoprecipitated with anti-pORF2 antiserum (A) or anti-pORF3 antiserum (B), as described in the text. Washed immunoprecipitates were separated by either SDS-10% PAGE (A) or SDS-15% PAGE (B) and visualized by fluorography. Lanes Mol. Wt., molecular size markers (in kilodaltons). (C) Two independent clones of plasmid pSG-ORF2 (ORF2-1 and ORF2-14), plasmid pSG-ORF3 (ORF3-1 and ORF3-2), and the parent vector (Mock) were expressed in a coupled *in vitro* transcription-translation system. The reaction mix was immunoprecipitated with anti-pORF2 or anti-pORF3 antiserum, separated by SDS-12% PAGE, and visualized by fluorography. The positions of pORF2, pORF3, the pORF3 dimer, and size markers (in kilodaltons) are indicated.

h with the subsequent addition of 50  $\mu$ l of a 50% protein A-Sepharose suspension. The mixture was then incubated at 4°C with shaking for 1 h. The beads were centrifuged down, and precleared lysates were subjected to immunoprecipitation with 10  $\mu$ l of the specific antiserum as described above for COS-1 cells.

**Cell fractionation.** COS-1 cells that had been transfected and labeled with [<sup>35</sup>S]methionine as described above were washed once with PBS and then scraped off the plate into 1 ml of PBS. After centrifugation, cell pellets equivalent to each 60-mm-diameter plate were resuspended in 0.5 ml of lysis buffer (10 mM Tris-HCl [pH 8.0], 140 mM NaCl, 5 mM iodoacetamide, 0.5% Triton X-100, 2 mM phenylmethylsulfonyl fluoride) and kept on ice for 1 h. Lysates were centrifuged in a Biofuge RS microcentrifuge (Heraeus Sepratech, GmbH) at 13,000 rpm for 30 min. The supernatant (cytoplasmic fraction) was removed to a fresh tube, and 50  $\mu$ l of a 10 $\times$  DOC-SDS solution (10% sodium deoxycholate, 1% SDS) was added to it. The pellets were washed once with 0.5 ml of lysis buffer as described above. Washed pellets (nuclear fraction) were resuspended in 0.5 ml of RIPA buffer. Both fractions were immunoprecipitated as described above.

**Tunicamycin treatment.** At 40 h posttransfection, cells were shifted to 1 ml of methionine-deficient DMEM without or with 10  $\mu$ g of tunicamycin (Boehringer Mannheim GmbH) for 1 h. Cells were then labeled with [<sup>35</sup>S]methionine for 4 h, as described above, in the absence or presence of 10  $\mu$ g of tunicamycin per ml. Total cell lysates in RIPA buffer were immunoprecipitated as described above.

**Pulse-chase analysis.** At 40 h posttransfection, cells were shifted to 3 ml of

methionine-deficient DMEM for 1 h. Labeling was performed as described above, except that 300  $\mu$ Ci of [<sup>35</sup>S]methionine per 60-mm-diameter plate was used and the labeling time was 20 min. After the removal of the labeling mix, 3 ml of culture medium was added and cells were harvested either immediately or after a chase of 30 min or 4 h. Total lysates in RIPA buffer were immunoprecipitated as described above.

**In vitro translation.** A coupled transcription-translation system (TNT; Promega) with bacteriophage T7 RNA polymerase was used for *in vitro* syntheses of polypeptides from pSG-ORF2 and pSG-ORF3 plasmid templates according to the supplier's guidelines. For cotranslational processing, 2  $\mu$ l of canine pancreatic membranes (Promega) was also included in the 25- $\mu$ l *in vitro* transcription-translation reaction. [<sup>35</sup>S]methionine-labeled polypeptides synthesized *in vitro* were separated by SDS-PAGE either directly or after immunoprecipitation and visualized by fluorography.

**Endoglycosidase treatment.** After the immunoprecipitation described above, the protein A-Sepharose beads containing bound antigen were resuspended in 20  $\mu$ l of 0.5% SDS-1% 2-mercaptoethanol and heated at 100°C for 10 min. To this was added 2.5  $\mu$ l of 0.5 M sodium citrate (pH 5.5) and 2  $\mu$ l of endoglycosidase H (500 U/ $\mu$ l; New England Biolabs, Beverly, Mass.). No enzyme was added to control samples. After digestion at 37°C for 2 h, 25  $\mu$ l of 2 $\times$  SDS gel loading buffer was added, and the samples were boiled for 5 min and analyzed by separation on SDS-7.5% PAGE gels and fluorography.

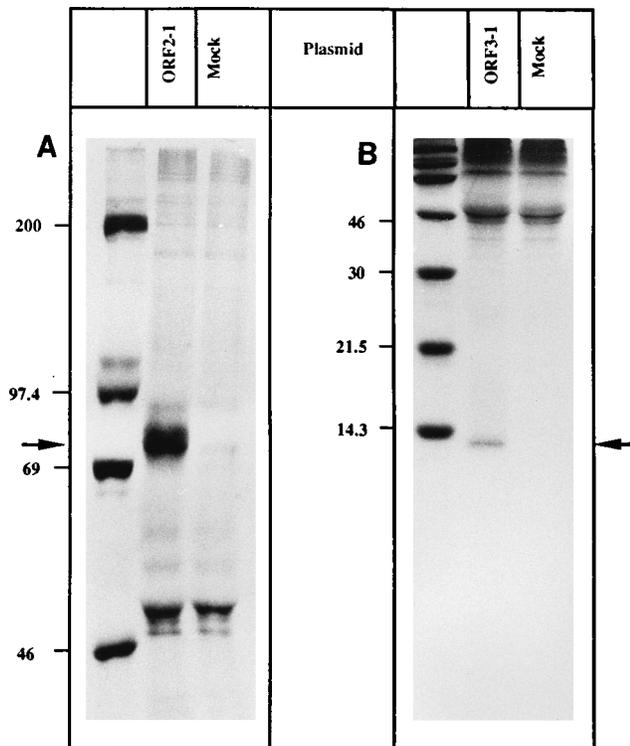


FIG. 4. Immunoprecipitation with HEV immune serum. COS-1 cells were transfected with the indicated plasmids and labeled with [ $^{35}$ S]methionine, and lysates were immunoprecipitated with pooled HEV immune serum from patients with hepatitis E. Immunoprecipitates were separated by SDS-7.5% PAGE (A) or SDS-15% PAGE (B) and visualized by fluorography. The positions of pORF2 (A) and pORF3 (B) are indicated by arrows. The positions of size markers (in kilodaltons) are also indicated.

For an analysis of glycosylation *in vitro*, to 10  $\mu$ l of the transcription-translation mixture was added 2  $\mu$ l of denaturing buffer (3% SDS, 6% 2-mercaptoethanol) and the mixture was heated at 100°C for 10 min. To this was added 20  $\mu$ l of reaction buffer (75 mM citrate-phosphate [pH 5.0], 100 mM EDTA, 5% Triton X-100, 1% 2-mercaptoethanol) and 0.4 U of endoglycosidase F (Boehringer Mannheim GmbH), and the reaction was incubated at 37°C. The control reaction included everything but the enzyme. Aliquots (10  $\mu$ l) were removed at various times, inactivated by being boiled in 40  $\mu$ l of SDS gel loading buffer, and kept frozen at -70°C. Samples were analyzed by separation on SDS-7.5% PAGE gels and fluorography.

## RESULTS

**Expression of HEV proteins.** The expression vectors pSG-ORF2 and pSG-ORF3 contain the HEV ORFs driven by the simian virus 40 control elements, including the *ori* sequences. These vectors are capable of replication to high copy numbers in T-antigen-producing monkey kidney COS-1 cells and should express high levels of the proteins encoded by HEV ORFs. Transfected cells were scored for antigen expression by immunofluorescence analysis with specific polyclonal antibodies. The results presented in Fig. 2 show that both pORF2 and pORF3 are expressed in transfected cells. pORF2 was found on the cell surface (Fig. 2A) as well as in the cytoplasm (Fig. 2B). On the other hand, pORF3 was found only in the cytoplasm (Fig. 2C). Controls with preimmune sera did not show any staining on transfected COS-1 cells (data not shown). It is evident that antisera do not react with cellular components from the staining of only a fraction of the transfected cells in a given field, in agreement with the expected transfection efficiency of about 40 to 50%.

Total lysates from COS-1 cells transfected with the appropriate vectors and metabolically labeled with [ $^{35}$ S]methionine were also subjected to immunoprecipitation with polyclonal antisera. Cells transfected with two independent clones of plasmid pSG-ORF2 showed the specific immunoprecipitation of the 74- to 88-kDa proteins absent in cells transfected with the parent vector pSG1 (Fig. 3A). Similarly, a 13.5-kDa protein was found in cells transfected with two independent clones of plasmid pSG-ORF3 but was not found in cells transfected with the parent vector (Fig. 3B). Both pORF2 and pORF3 were also found in Huh-7 hepatoma cells transfected with the expression vectors, albeit at a level much reduced in comparison with that in COS-1 cells (data not shown).

Subcellular fractionation of transfected COS-1 cells showed that both pORF2 (Fig. 3A) and pORF3 (Fig. 3B) were present in the cytoplasmic fraction, supporting the observations of immunofluorescence studies (Fig. 2). A trace amount of pORF3 was also found in the nuclear fraction. Neither protein was found to be secreted into the culture medium (data not shown).

The expression of both proteins was also carried out in a coupled *in vitro* transcription-translation system. Again, both independent clones of pSG-ORF2 as well as those of pSG-ORF3 expressed the respective proteins, as judged by immunoprecipitation with specific antisera (Fig. 3C). In total translation reactions without immunoprecipitation, pORF2 and pORF3 were the major protein bands, accounting for >80% of the synthesized protein (data not shown). The size of pORF2 expressed *in vitro* was found to be 74 kDa, indicating that the protein expressed in cultured cells may be subjected to post-translational modifications (Fig. 3A). The *in vitro*-expressed pORF3, like its cell-expressed counterpart (Fig. 3B), was 13.5 kDa. In the *in vitro* system, however, a pORF3 species of about 28 kDa was evident (Fig. 3C); it was reproducibly absent when pORF3 was expressed in transfected cells. The fraction of this 28-kDa form of pORF3 varied between different *in vitro* expression experiments. Furthermore, it was observed that rabbit polyclonal anti-pORF3 antiserum immunoprecipitated the 28-kDa form with a lower efficiency compared with that of the 13.5-kDa form of pORF3.

Both HEV antigens from lysates of pSG-ORF2- or pSG-ORF3-transfected COS-1 cells also immunoprecipitated with pooled HEV immune serum obtained from patients with hepatitis E (Fig. 4). No such precipitation was observed from lysates of mock-transfected COS-1 cells. These results further authenticate the natures of the expressed proteins.

**pORF2 is a glycoprotein.** The glycoprotein status of pORF2 and pORF3 was evaluated in experiments in which tunicamycin was used to inhibit glycosylation in transfected cells. The expression of higher-molecular-weight forms of pORF2 was quantitatively inhibited by tunicamycin (Fig. 5A), suggesting that the 74-kDa form represents the nonglycosylated protein and that the two distinctly larger forms, about 82 and 88 kDa, represent the glycosylated protein, perhaps with different extents of glycosylation. In HepG2 cells as well, treatment with tunicamycin led to a shift from the higher-molecular-weight form of gpORF2 to the faster-moving form of pORF2 (Fig. 5B). Tunicamycin treatment had no effect on pORF3 expressed either in COS-1 cells (Fig. 5C) or in HepG2 cells (Fig. 5D), suggesting that this protein is not glycosylated.

To further confirm the glycoprotein nature of pORF2, the protein expressed in COS-1 or HepG2 cells in the absence or presence of tunicamycin was subjected to endoglycosidase digestion. As shown in Fig. 6, gpORF2 expressed in both COS-1 and HepG2 cells was completely reduced to the nonglycosylated pORF2 form with endoglycosidase H, an effect similar to

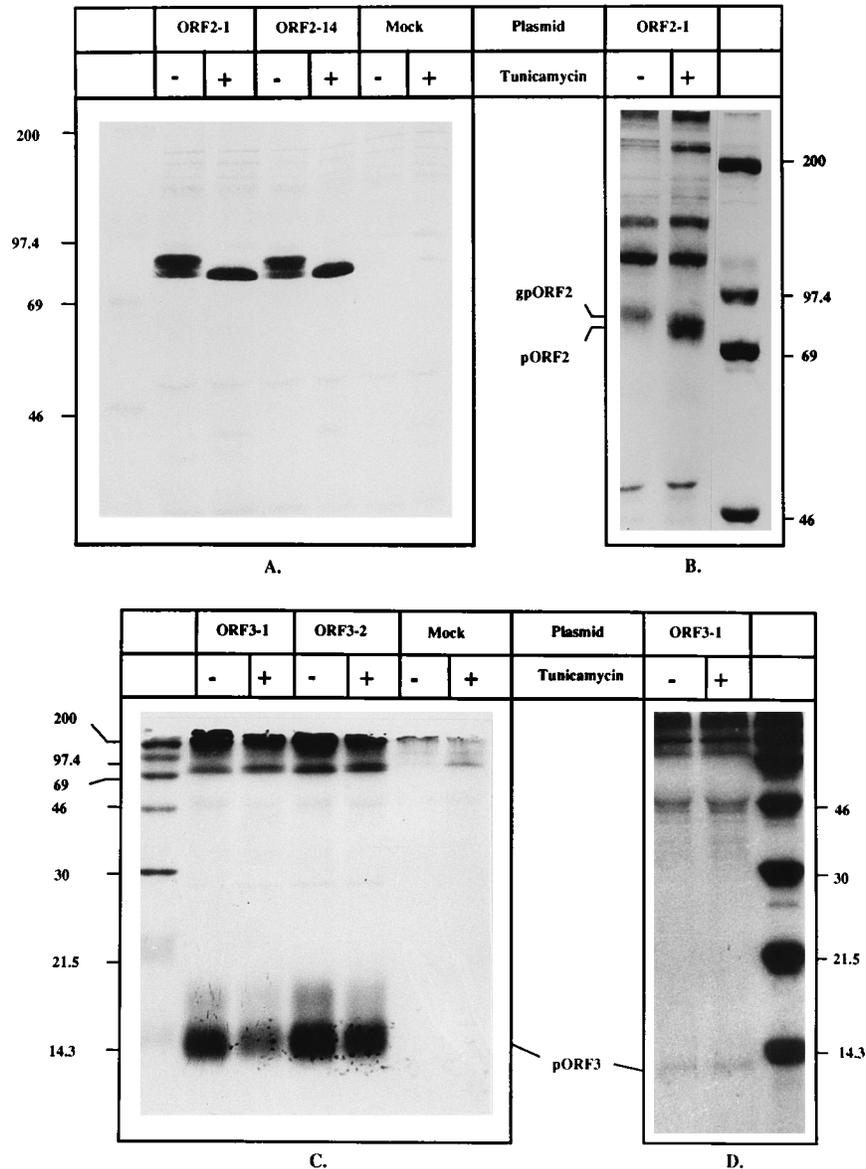


FIG. 5. Effects of tunicamycin on the expression of HEV proteins. (A and C) COS-1 cells were transfected with two independent clones of plasmid pSG-ORF2 (A) and plasmid pSG-ORF3 (C), with plasmid pSGI (Mock; A and C) as a control. (B and D) HepG2 cells were transfected with pSG-ORF2 (B) and pSG-ORF3 (D), with plasmid pSGI (Mock; B and D) as a control. The transfected cells were either treated (+) or not treated (-) with tunicamycin, as described in the text. Total cell lysates were immunoprecipitated, separated by SDS-7.5% PAGE (A and B) or SDS-15% PAGE (C and D) and visualized by fluorography. The positions of size markers (in kilodaltons), gpORF2, pORF2, and pORF3 are indicated.

that observed with tunicamycin treatment. These results unequivocally demonstrate that pORF2 is a glycoprotein and that its glycosylation is not an artifact of overexpression in COS-1 cells, which are not the natural target cells for HEV infection.

**Processing of HEV proteins.** The processing of pORF2 and pORF3 was studied in transfected cells in culture, as well as in vitro. In a pulse-chase experiment after a 20-min pulse of [<sup>35</sup>S]methionine, a single 82-kDa form of pORF2 was predominant. After a 30-min chase with unlabeled methionine, three forms of pORF2 were found, while after a 4-h chase, only the 74- and 88-kDa forms were apparent. At this stage, the 88-kDa form of pORF2 was predominant. These results (Fig. 7A) suggest that the ORF2-encoded protein is made as a precursor (ppORF2), which is first processed into the mature form (pORF2) and then glycosylated (gpORF2). Similar experi-

ments with pORF3 showed only a single form of the protein throughout the pulse-chase period (Fig. 7B), suggesting that this protein does not undergo any major processing. Further, by comparing the behaviors of these two proteins during the pulse-chase, the turnover rate of pORF3 appeared to be higher than that of pORF2.

Cotranslational processing of these two proteins was studied in vitro in the presence of canine pancreatic membranes (Fig. 8). Apart from pORF2, a slightly larger form was also observed for in vitro translations carried out in the presence of membranes (Fig. 8A). The fact that the larger form represented gpORF2 was established by endoglycosidase F digestion, which resulted in the reduction of gpORF2 into pORF2 (Fig. 8B). The translation of pORF3 was inhibited by the addition of membranes. This was a nonspecific effect, as is apparent from

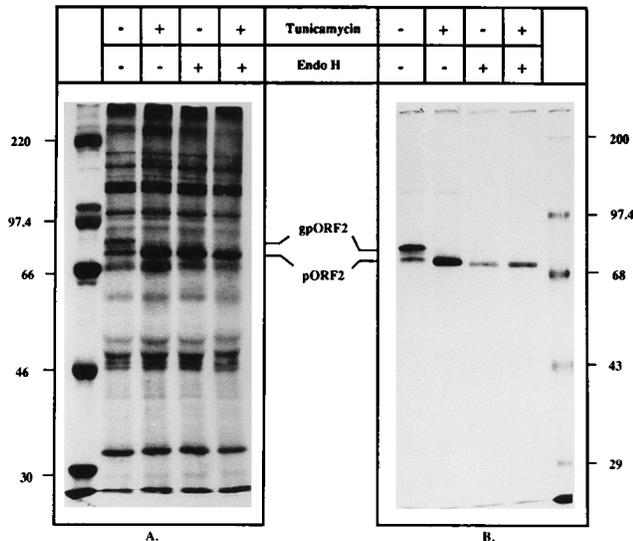


FIG. 6. Endoglycosidase digestion of gpORF2. HepG2 (A) and COS-1 (B) cells were transfected with pSG-ORF2, and cells were either treated (+) or not treated (-) with tunicamycin, as described in the text. [<sup>35</sup>S]methionine-labeled cell lysates were immunoprecipitated with rabbit anti-pORF2, and washed immunoprecipitates were treated with endoglycosidase H (Endo H) (+) or buffer alone (-), as described in the text. Digests were then subjected to separation on SDS-7.5% PAGE gels and to fluorography. The positions of the glycosylated (gpORF2) and nonglycosylated (pORF2) forms are shown. The positions of size markers (in kilodaltons) are also indicated.

the inhibition of pORF2 translation when translated either alone or along with pORF3 (Fig. 8A). In spite of the inhibitory effects of membranes, no cotranslational processing of pORF3 was observed (Fig. 8A). These *in vitro* results support the observations made for these two proteins expressed in transfected cells in culture.

**Interactions between HEV proteins.** Homologous interactions between pORF2 and pORF3 subunits were studied in the *in vitro* system. Proteins translated without the addition of membranes were subjected to SDS-PAGE in the absence or presence of 2-mercaptoethanol (5%) and/or heat (100°C, 2 min). As shown in Fig. 9A, pORF2 is unaffected by 2-mercaptoethanol treatment, but the result is a dimeric species if the sample is not heated prior to electrophoresis. This suggests that pORF2 forms a noncovalent homodimeric complex. A distinct doublet observed in the monomeric pORF2 may be the result of translation from alternate in-frame AUG codons present at amino acid positions 1, 12, and 16 in the primary sequence. As observed earlier, *in vitro*-translated pORF3 showed a predominant 28-kDa form, which was unaffected by reduction and heating (Fig. 9B).

Heteromeric interactions between subunits of pORF2 and pORF3 were studied by transfection in COS-1 cells. Cells were transfected with either pSG-ORF2, pSG-ORF3, or a combination of two independent clones of each plasmid. [<sup>35</sup>S]methionine-labeled cell lysates were then immunoprecipitated with polyclonal antibodies to one protein or the other. The results (Fig. 9C) show that while anti-pORF2 and anti-pORF3 were able to precipitate the respective proteins from cotransfected cells, there was no coimmunoprecipitation of the other protein. The specificities of the antisera were further substantiated by the inability of anti-pORF2 to precipitate pORF3 and the inability of anti-pORF3 to precipitate pORF2. The lack of coimmunoprecipitation of pORF2 and pORF3 from cotransfected cells suggests that either these two proteins do not

interact with each other *in vivo* or that the interactions are not strong enough to survive these immunoprecipitation conditions. Similar results were obtained when *in vitro*-translated pORF2 and pORF3 were mixed and cross-immunoprecipitated with these two antisera (data not shown).

**pORF3 interacts with a cellular protein.** Whenever pORF3 was immunoprecipitated from transfected COS-1 cells, another protein of about 18 kDa coprecipitated with it (Fig. 10A) (previously observed in Fig. 3B, 5C, and 9C). This protein was also observed in pORF3 immunoprecipitates from Huh-7 hepatoma cells (data not shown), but not those from HepG2 cells (Fig. 5D). It was designated 3IP, the pORF3-interacting protein. No such protein was precipitable from mock-transfected cells (Fig. 10A), showing that it was not the result of antiserum cross-reactivity. Furthermore, a Western blot (immunoblot) of lysates from pSG-ORF3-transfected cells was positive for pORF3 but negative for 3IP (data not shown). By subcellular fractionation, like pORF3 3IP, was found in the cytoplasmic fraction (Fig. 3B). It was unaffected by tunicamycin treatment of cells (Fig. 5C), suggesting that 3IP is not a glycoprotein.

To further establish that 3IP is a cellular protein, *in vitro*-translated pORF3 was mixed with [<sup>35</sup>S]methionine-labeled lysates of untransfected COS-1 cells and the mixture was precipitated with anti-pORF3 antiserum. A very weak band was seen in the presence of COS-1 lysates, but it was absent when no lysates were added (Fig. 10B). As was the pORF3-3IP interaction in transfected COS-1 cells, the *in vitro* interaction was stable for RIPA buffer washing. In fact, identical results were obtained under washing conditions that were less stringent (10 mM Tris-HCl [pH 8.0], 140 mM NaCl, 0.1% Triton X-100 [TST buffer]). Quantitatively, however, far less pORF3-3IP interaction was observed *in vitro* than in transfected cells.

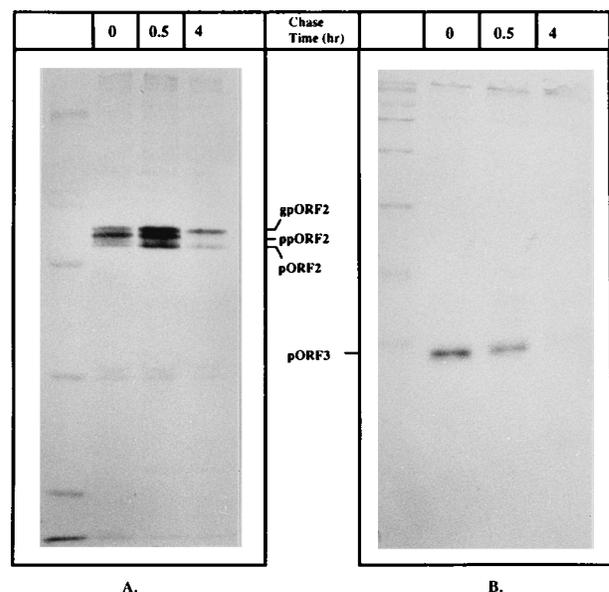


FIG. 7. Pulse-chase analysis of HEV proteins. COS-1 cells transfected with either plasmid pSG-ORF2 (A) or plasmid pSG-ORF3 (B) were labeled with [<sup>35</sup>S]methionine for 20 min and chased with unlabeled methionine for the times indicated. Total cell lysates were immunoprecipitated, separated by either SDS-7.5% PAGE (A) or SDS-15% PAGE (B), and visualized by fluorography. The size markers (from top to bottom) are 200, 97.4, 69, 46, 30, 21.5, and 14.3 kDa. The positions of multiple forms of pORF2 and that of pORF3 are indicated.

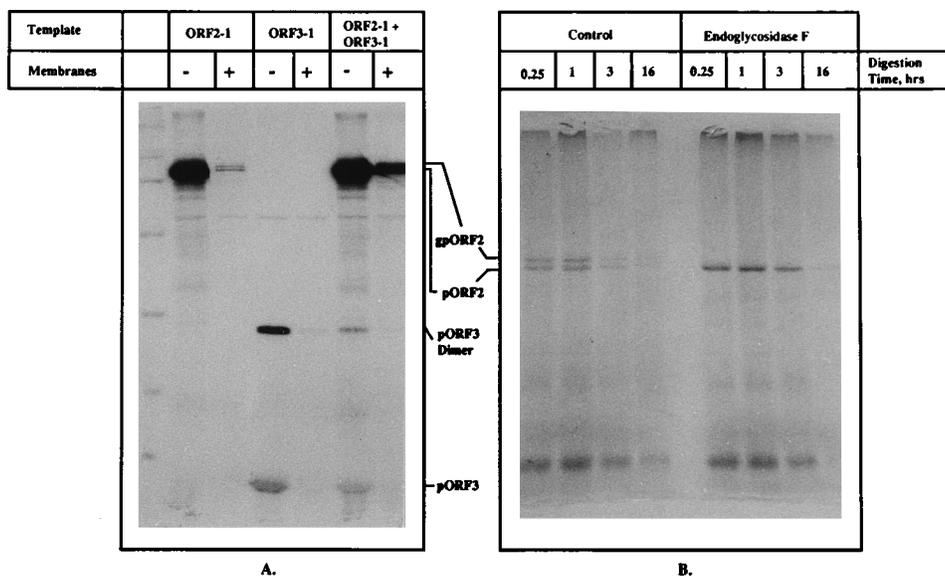


FIG. 8. In vitro processing of HEV proteins. (A) pSG-ORF2, pSG-ORF3, or both plasmids were used as templates in the in vitro coupled transcription-translation system with (+) or without (-) canine pancreatic membranes. Synthesized proteins were separated by SDS-12% PAGE and visualized by fluorography. The size markers are the same as those described in the legend to Fig. 7. (B) pORF2 synthesized in vitro in the presence of membranes was incubated in the absence (control) or presence of endoglycosidase F, as described in the text. Samples were withdrawn after digestion for the times indicated, separated by SDS-7.5% PAGE, and visualized by fluorography. The positions of pORF2, gpORF2, pORF3, and pORF3 dimer are indicated.

## DISCUSSION

The genome of HEV has been cloned from multiple geographically distinct isolates, but the inability to culture this virus has precluded studies into its molecular nature. Here we have used a subgenomic fragment expression strategy to gain an insight into the putative structural proteins of HEV. The structural region of the HEV genome consists of two ORFs. These were expressed in COS-1 and HepG2 cells. The former cell line formed the basis of subsequent analyses as larger amounts of these proteins were expressed in this cell line because of replication of the expression vector to high copy numbers (5, 15). The latter cell line was used to ascertain that HEV proteins are also expressed in liver cells, the only cells known to support viral replication.

The major structural protein, pORF2, was found in multiple forms as a 74- to 88-kDa protein by SDS-PAGE. In vitro expression showed that the 74-kDa form corresponded to unmodified pORF2. This is in agreement with the predicted size of about 72 kDa, based on the 660-amino-acid ORF. On the basis of tunicamycin inhibition experiments, cotranslational processing with glycosylation-proficient membranes, and endoglycosidase digestion experiments, we have shown that the approximately 82- and 88-kDa forms are glycosylated versions of pORF2. This is not an artifact of expression in COS-1 cells, as is borne out by the glycosylation of pORF2 expressed in HepG2 cells. Since this is a hepatoma line, it more closely mimics the natural target cell for HEV infection. In its primary amino acid sequence, pORF2 contains three potential N-linked glycosylation sites (Asn-X-Thr/Ser) at residues 137, 310, and 561. These are conserved in all of the HEV strains sequenced so far (1, 6, 18, 24, 26).

It has been suggested that pORF2 contains a signal sequence at its N-terminal end (21, 24). The putative signal sequence is a 22-amino-acid stretch consisting of positively charged residues (Arg) at the N-terminal end, a 14-residue hydrophobic core, and a turn-inducing stretch of proline residues. However, direct proof of the presence of this signal

sequence is lacking. We have shown by means of pulse-chase experiments that pORF2 is indeed synthesized as a larger precursor, ppORF2, of about 82 kDa. This is subsequently processed into the 74-kDa form and glycosylated to the mature form, gpORF2, of 88 kDa. The approximately 82-kDa species observed in tunicamycin and pulse-chase experiments are most likely distinct from each other. While one may be an intermediate in the glycosylation pathway leading to gpORF2 (88 kDa), the other is the precursor which is processed to yield pORF2. In Fig. 7A, both forms may be present but would not resolve on the gel because of very similar sizes.

The various forms of pORF2 appear to run anomalously on SDS-PAGE gels. On the basis of the primary sequence, ppORF2 should be about 72 kDa but runs as an 82-kDa protein; its signal-cleaved form, which should be smaller, runs as a 74-kDa protein. Such anomalous mobility due to conformational and charge distribution effects has previously been observed for several classes of proteins, including glycoproteins, proline-rich proteins, maleylated proteins, calcium-binding proteins, and histones (23). Apart from being a glycoprotein, pORF2 is also rich in proline residues in its N-terminal region. This is especially true for the signal sequence, which is about one-third proline (7 of 22 residues). In fact, there appears to be a far greater anomaly in the size of ppORF2 (82 versus 72 kDa) than in the size of its signal-cleaved form, pORF2 (74 versus 70 kDa).

The glycosylation pattern of a glycoprotein reflects the cellular compartment through which it has passed during its synthesis and processing. Our results with gpORF2 show that it is completely sensitive to endoglycosidase H and therefore contains only high-mannose residues. These structures are synthesized in the endoplasmic reticulum and the *cis* Golgi compartment (25). Thus, it appears that the processing of pORF2 occurs at the endoplasmic reticulum and that the protein is subsequently transported either directly or through the *cis* Golgi compartment to the cell surface. Consistent with this, immunofluorescence localization showed pORF2 to be ex-

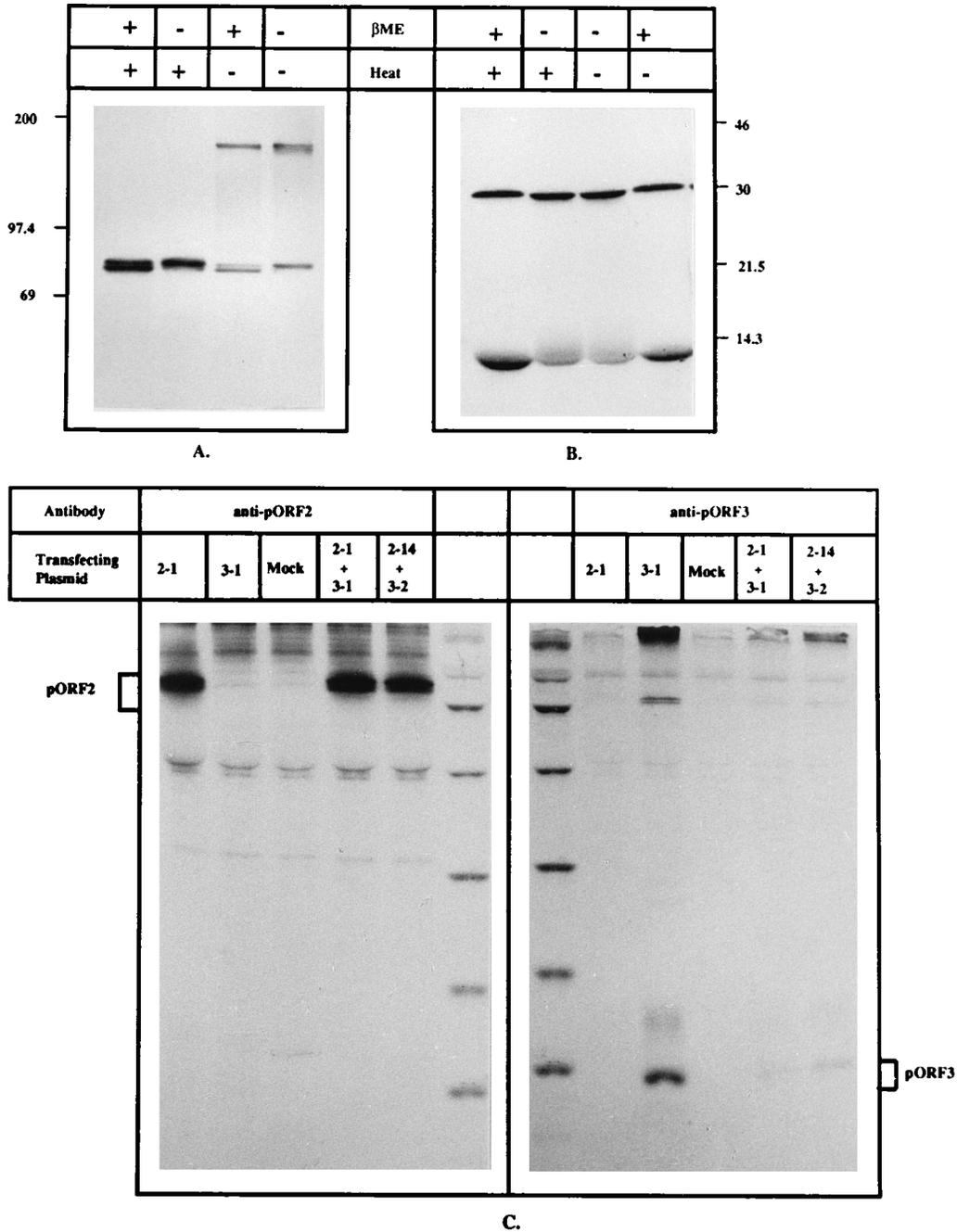


FIG. 9. Interactions between HEV proteins. (A and B) In vitro-synthesized pORF2 (A) and pORF3 (B) were treated with 5% 2-mercaptoethanol ( $\beta$ ME) and/or heated at 100°C for 2 min. The total mixture was separated by SDS-7.5% PAGE (A) or SDS-15% PAGE (B) and visualized by fluorography. The positions of size markers (in kilodaltons) are indicated. (C) COS-1 cells were transfected with the indicated plasmids and labeled with [<sup>35</sup>S]methionine, and total lysates were immunoprecipitated with the indicated antibodies. The proteins in washed immunoprecipitates were separated by SDS-12% PAGE and visualized by fluorography. The positions of pORF2 and pORF3 are indicated. The size markers are the same as those described in the legend to Fig. 7.

pressed on the cell surface as well as in the cytoplasm. On the cell surface, the protein appeared to be concentrated in some regions, suggestive of an active process of association of protein subunits, perhaps into some higher-order forms. It has been shown here that pORF2 subunits can homodimerize through noncovalent interactions. Many positive-strand RNA viruses replicate and assemble on membrane surfaces and utilize a virally encoded RNA-binding capsid protein (4). pORF2, quite basic (pI, ~10.3) in its amino-terminal half, may well

serve this purpose (21, 22). Though the replication strategy of HEV has not been worked out, such a proposal would fit well into the model proposed by Reyes et al. (22). It is not yet clear what bearing the glycosylation of pORF2 has on its cell surface localization and on the assembly of the HEV nucleocapsid. This is currently being investigated with pORF2 mutants lacking individual or all N-linked glycosylation sites.

The small ORF, ORF3, can code for a protein of 123 amino acids (24). In accordance with this, we find a protein of 13.5

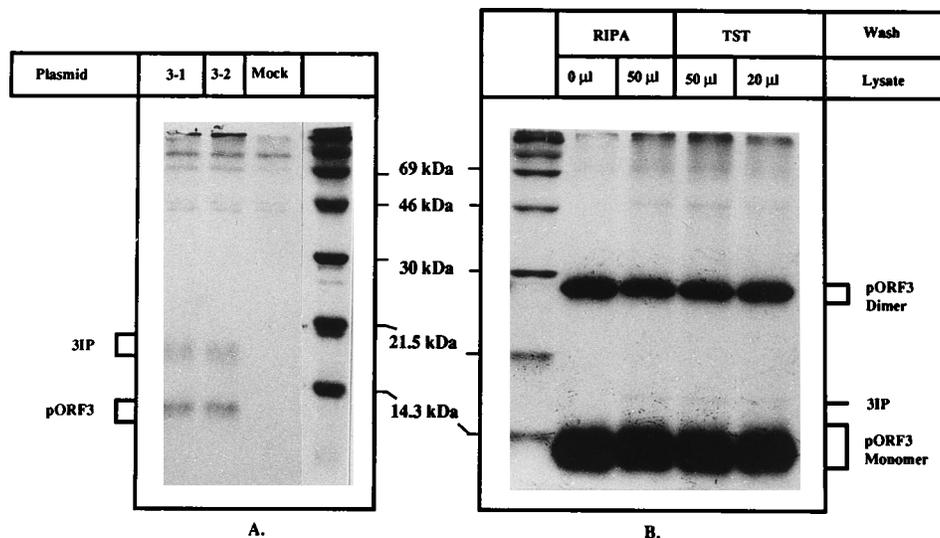


FIG. 10. Interaction between pORF3 and a cellular protein. (A) COS-1 cells were transfected with two independent clones of plasmid pSG-ORF3 or with plasmid pSG1 (Mock) and labeled with [ $^{35}$ S]methionine, and total lysates were immunoprecipitated with anti-pORF3 antiserum. (B) In vitro-synthesized pORF3 was mixed with the indicated amounts of [ $^{35}$ S]methionine-labeled, untransfected COS-1 cell lysates. After incubation, immunoprecipitation with anti-pORF3 antiserum, and washing with either RIPA or TST buffer, as described in the text, proteins were separated by SDS-15% PAGE and visualized by fluorography. The positions of the pORF3 monomer, the pORF3 dimer, pORF3, 3IP, and size markers are indicated.

kDa expressed in cells transfected with ORF3 expression vectors. This protein is localized in the cytoplasm and nonglycosylated; it does not appear to undergo any modification that would significantly alter its size. In its N-terminal half, pORF3 appears to contain two hydrophobic domains which have been proposed to constitute either transmembrane segments or a signal sequence followed by a transmembrane region (22). We have provided clear evidence that pORF3 does not undergo any significant processing to remove the proposed signal sequence.

Two forms of pORF3, 13.5 and 28 kDa, were routinely observed in vitro, but only the 13.5-kDa form was seen in transfected cells. One possibility is that the 28-kDa species represents a highly stable dimeric form of pORF3. In support of this, a similar species is also observed when pORF3 is expressed in *E. coli* (18). Alternatively, this could be an artifact of the in vitro system, whereby the ORF3 stop codon is skipped and translation terminates downstream in the vector sequences. When a hexahistidine-pORF3 fusion protein is synthesized in vitro, two forms are evident. One is the expected monomeric (6×His)pORF3 species of ~17 kDa, and the other is a species of ~28 kDa, significantly smaller than the expected homodimer (7). This supports the second alternative. In animal cells, however, a cellular protein, which we have called 3IP, was found to interact with pORF3. The fact that the in vitro-expressed pORF3 monomer interacts very weakly with 3IP suggests that this interaction takes place at the nascent chain level inside the cell. The role of 3IP, if any, in pORF3 expression or function is not clear, considering that it coimmunoprecipitates with pORF3 from COS-1 and Huh-7 hepatoma cells, not with pORF3 from HepG2 hepatoblastoma cells. However, the differences between the two liver cell lines may be explained by their different stages of differentiation. We are presently studying these homologous and heterologous interactions of pORF3 in greater detail.

It is unknown whether pORF3 is a part of the virion. Here, we tried to address this by looking for interactions between pORF2 (the major capsid protein) and pORF3. Expression in transfected cells and cross-immunoprecipitation with the re-

spective antibodies did not provide any evidence of an interaction between these two proteins. It is possible that such interactions take place at the nascent chain level for these two proteins or that the interactions are weak. Under both of these circumstances, our experimental conditions are likely to overlook such interactions. In the HEV genome, ORFs 2 and 3 overlap the entire stretch of ORF3, except for a 9-amino-acid region at the N terminus of ORF3. Though at least two subgenomic RNAs coterminal with the 3' end of the HEV genome have been reported (24), the subgenomic mRNA used for the translation of either or both of these ORFs has not been characterized. It is also possible that a single subgenomic mRNA is utilized for the synthesis of both proteins, perhaps by alternate initiator codon usage. In such a situation, the dynamics of translation and protein-protein interactions at the nascent chain level would be very distinct from those expected in our experimental system in which these two proteins are expressed from different plasmids. Work currently in progress is aimed in this direction.

How relevant are our findings to the infectious virus? Contentious issues are the glycoprotein nature of pORF2 and the fact that most viral glycoproteins are components of the virion envelope. So far, there is no evidence to suggest that HEV contains a lipid membrane. We propose that pORF2 is the major constituent of the HEV nucleocapsid, which is assembled at the cytoplasmic membrane. pORF2 is cotranslationally translocated via its N-terminal signal sequence into the endoplasmic reticulum, where the signal is processed. The protein is glycosylated and transported to the cell surface by a bulk flow mechanism in the absence of any signals for retention in the endoplasmic reticulum (19), e.g., the C-terminal KDEL peptide found in endoplasmic reticulum luminal proteins. The glycosylation of pORF2 may be coincidental, considering its passage through the endoplasmic reticulum (and perhaps the *cis* Golgi compartment as well) and its three highly conserved N-linked glycosylation sites. Some part of capsid assembly may also occur in association with the endoplasmic reticulum, as proposed by Reyes et al. (22), but final assembly and/or maturation has to be cytoplasmic to account for the encapsidation

of HEV genomic RNA. At the cytoplasmic membrane, the nucleocapsid self-assembles along with HEV positive-stranded genomic RNA. It is not clear whether viral RNA is an absolute requirement for nucleocapsid assembly, since Tsarev et al. (27) have shown that pORF2 expressed in insect cells from a baculoviral vector is able to form virus-like particles. It is also not clear what role, if any, pORF3 plays in the assembly of the viral nucleocapsid. In the absence of either a tissue culture or a genomic RNA transfection system for HEV, we are trying to answer some of these questions by using the COS-1 cell expression system described above.

#### ACKNOWLEDGMENTS

We thank V. Kumar (ICGEB, New Delhi, India) for the expression vector pSGI. We are grateful to Y. Vaishnav, D. Salunke, and S. E. Hasnain for fruitful discussions and critical readings of the manuscript. This work was supported by internal funds from the ICGEB.

#### REFERENCES

- Aye, T. T., T. Uchida, X.-Z. Ma, F. Iida, T. Shikata, H. Zhuang, and K. M. Win. 1992. Complete nucleotide sequence of a hepatitis E virus isolated from the Xinjiang epidemic (1986-1988) of China. *Nucleic Acids Res.* **20**:3512.
- Bi, S. L., M. A. Purdy, K. A. McCaustland, H. S. Margolis, and D. W. Bradley. 1993. The sequence of hepatitis E virus isolated directly from a single source during an outbreak in China. *Virus Res.* **28**:233-247.
- Bradley, D. W. 1990. Enterically-transmitted non-A, non-B hepatitis. *Br. Med. Bull.* **46**:442-461.
- Brinton, M. A., and F. X. Heinz (ed.). 1990. New aspects of positive-strand RNA viruses. American Society for Microbiology, Washington, D.C.
- Gluzman, Y. 1981. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23**:175-182.
- Huang, C. C., D. Nguyen, J. Fernandez, K. Y. Yun, K. E. Fry, D. W. Bradley, A. W. Tam, and G. R. Reyes. 1992. Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HEV). *Virology* **191**:550-558.
- Jameel, S. Unpublished data.
- Jiang, X., D. Y. Graham, K. Wang, and M. K. Estes. 1990. Norwalk virus genome cloning and characterization. *Science* **250**:1580-1583.
- Khudyakov, Y. E., M. O. Favorov, D. L. Jue, T. K. Hine, and H. A. Fields. 1994. Immunodominant antigenic regions in a structural protein of the hepatitis E virus. *Virology* **198**:390-393.
- Khudyakov, Y. E., N. S. Khudyakova, H. A. Fields, D. Jue, C. Starling, M. O. Favorov, K. Krawczynski, L. Polish, E. Mast, and H. Margolis. 1993. Epitope mapping in proteins of hepatitis E virus. *Virology* **194**:89-96.
- Khuroo, M. S. 1980. Study of an epidemic of non-A, non-B hepatitis: possibility of another human hepatitis virus distinct from post-transfusion non-A, non-B type. *Am. J. Med.* **68**:818-823.
- Khuroo, M. S., M. R. Teli, S. Skidmore, M. A. Sofi, and M. Khuroo. 1981. Incidence and severity of viral hepatitis in pregnancy. *Am. J. Med.* **70**:252-255.
- Koonin, E. V., A. E. Gorbalenya, M. A. Purdy, M. N. Rozanov, G. R. Reyes, and D. W. Bradley. 1992. Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: delineation of an additional group of positive-stranded RNA plant and animal viruses. *Proc. Natl. Acad. Sci. USA* **89**:8259-8263.
- Krawczynski, K. 1993. Hepatitis E. *Hepatology* **17**:932-941.
- Learned, R. M., R. M. Myers, and R. Tjian. 1981. Replication in monkey cells of plasmid DNA containing the minimal SV40 origin. *ICN-UCLA Symp. Mol. Cell. Biol.* **22**:555-566.
- Li, F., H. Zhuang, S. Kolivas, S. A. Locarnini, and D. A. Anderson. 1994. Persistent and transient antibody responses to hepatitis E virus detected by Western immunoblot using open reading frame 2 and 3 and glutathione S-transferase fusion proteins. *J. Clin. Microbiol.* **32**:2060-2066.
- Panda, S. K. Unpublished results.
- Panda, S. K., S. K. Nanda, M. Zafrullah, I.-H. Ansari, M. H. Ozdener, and S. Jameel. 1995. An Indian strain of hepatitis E virus (HEV): cloning, sequence, and expression of structural region and antibody responses in sera from individuals from an area of high-level HEV endemicity. *J. Clin. Microbiol.* **33**:2653-2659.
- Pelham, H. R. B., and S. Munro. 1993. Sorting of membrane proteins in the secretory pathway. *Cell* **75**:603-605.
- Purcell, R. H., and J. R. Ticehurst. 1988. Enterically transmitted non-A, non-B hepatitis: epidemiology and clinical characteristics, p. 131-137. *In* A. J. Zuckerman (ed.), *Viral hepatitis and liver disease*. Alan R. Liss, Inc., New York.
- Purdy, M. A., A. W. Tam, C. C. Huang, P. O. Yarbough, and G. R. Reyes. 1993. Hepatitis E virus: a non-enveloped member of the 'alpha-like' RNA virus supergroup? *Semin. Virol.* **4**:319-326.
- Reyes, G. R., C. C. Huang, A. W. Tam, and M. A. Purdy. 1993. Molecular organization and replication of hepatitis E virus (HEV). *Arch. Virol.* **7**:15-25.
- See, Y. P., and G. Jackowski. 1989. Estimating molecular weights of polypeptides by SDS gel electrophoresis, p. 1-21. *In* T. E. Creighton (ed.), *Protein structure: a practical approach*. IRL Press, Oxford.
- Tam, A. W., M. M. Smith, M. E. Guerra, C. C. Huang, D. W. Bradley, K. E. Fry, and G. R. Reyes. 1991. Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length viral genome. *Virology* **185**:120-131.
- Tarentino, A. L., R. B. Trimble, and T. H. Plummer. 1989. Enzymatic approaches for studying the structure, synthesis, and processing of glycoproteins. *Methods Cell Biol.* **32**:111-139.
- Tsarev, S. A., S. U. Emerson, G. R. Reyes, T. S. Tsareva, L. J. Letgers, I. A. Malik, M. Iqbal, and R. H. Purcell. 1992. Characterization of a prototype strain of hepatitis E virus. *Proc. Natl. Acad. Sci. USA* **89**:559-563.
- Tsarev, S. A., T. S. Tsareva, S. U. Emerson, A. Z. Kapikian, J. Ticehurst, W. London, and R. H. Purcell. 1993. ELISA for antibody to hepatitis E virus (HEV) based on complete open-reading frame-2 protein expressed in insect cells: identification of HEV infection in primates. *J. Infect. Dis.* **168**:369-378.
- Wong, D. C., R. H. Purcell, M. A. Sreenivasan, S. R. Prasad, and K. M. Pavri. 1980. Epidemic and endemic hepatitis in India: evidence for a non-A, non-B virus etiology. *Lancet* **ii**:876-879.
- Yarbough, P. O., A. W. Tam, K. E. Fry, K. Krawczynski, K. A. McCaustland, D. W. Bradley, and G. R. Reyes. 1991. Hepatitis E virus: identification of type-common epitopes. *J. Virol.* **65**:5790-5797.