A novel internal open reading frame product expressed from a polycistronic mRNA of porcine epidemic diarrhoea virus may not contribute to virus attenuation

Mahender Singh

Institute of Virology, Faculty of Veterinary Medicine, University of Zurich, Winterthurerstrasse 266a, 8057 Zurich, Switzerland

Cell-culture-adapted (ca) porcine epidemic diarrhoea virus (PEDV) contains three internal open reading frames (I ORF) within the nucleocapsid protein gene and lacks the downstream counterpart of porcine transmissible gastroenteritis virus ORF7 or feline infectious peritonitis virus ORF6a. To confirm whether such features also exist in wild-type (wt) PEDV, the 3’ 1800 nucleotides of its genome were sequenced and were found to be identical to those of ca virus. The coding potential of I-1 ORF was ascertained by transient expression in Vero cells followed by immunofluorescence using antipeptide sera. The I-1 protein was synthesized as a 12 kDa non-phosphorylated PEDV-specific protein that was not present in detectable amounts in virions. However, a low copy number of I-1 in the virion would suggest it is a structural component. Nevertheless, identical nucleotide sequences and gene expression strategies of attenuated ca virus and its virulent parent, wt PEDV, demonstrate that the 3’ 1800 nucleotides or the genes and gene products encoded therein may not contribute to virus attenuation.

Nearly 7400 bases at the 3’ end of the RNA genome of two different cell-culture-adapted (ca) isolates of porcine epidemic diarrhoea virus (PEDV) have been sequenced (Bridgen et al., 1993; Duarte & Laude, 1994; Duarte et al., 1994). Analysis of these sequences reveals a characteristic coronavirus gene arrangement (Fig. 1) and suggests that PEDV belongs to the group I coronaviruses. However, ca PEDV differs from group I viruses in that its genome lacks additional open reading frames (ORFs) downstream of the nucleocapsid (N) gene, whereas porcine transmissible gastroenteritis virus carries ORF7 (Kapke & Brian, 1986) and feline infectious peritonitis virus, feline enteric coronavirus and canine coronavirus contain ORF 6a and 6b in the same location (Venema et al., 1992a, b). It is not yet known whether the wild-type (wt) isolates, never propagated in cell culture, also lack such ORFs since their genomes have not yet been sequenced.

Additionally, the N gene of ca PEDV contains three internal (I) ORFs designated as I-1 (113 codons), I-2 (63 codons) and I-3 (72 codons) in +1 reading frame (Bridgen et al., 1993). Analysis of reported N gene mRNA sequences of other coronaviruses also revealed the presence of different sized I ORFs within the N gene. Only little is known about the proteins encoded by these ORFs and their biological significance. In bovine coronavirus (BCV), for example, the I ORF (208 codons in +1 frame of the N gene) is expressed as a 23 kDa membrane-associated protein during virus replication (Senanayake et al., 1992). Its functional significance, however, remains to be determined. The I ORF (207 codons) in mouse hepatitis virus (MHV) encodes a structural protein that is not essential for virus replication (Fischer et al., 1997).

The I-1 ORF of PEDV is much smaller than those of BCV and MHV, whereas the combined coding capacity of all three I ORFs of PEDV equals approximately the size of BCV or MHV I ORF. It is, therefore, interesting to determine whether the small sized I ORFs of PEDV are the result of mutations during adaptation and propagation of the wt virus in cell culture. To unravel these possibilities, the 3’ 1800 nucleotide (nt) sequence encompassing the entire N gene and the 3’ non-translated region of wt PEDV isolate CV777 were determined. Total RNA from gut specimens collected from PEDV (CV777)-infected pigs (Bernasconi, 1996) was isolated using the total RNA isolation reagent TRIzol (Gibco BRL). First strand cDNA, synthesized using oligo d(T)-tailed P25, was amplified using previously described primers (P23, P24, P25, P32 and P35) (Fig. 1) and reaction conditions (Bridgen et al., 1993). P70 was 5’ GGAGCTCGTCAATAGTAACCCCTA 3’, P85 was 5’ GGTGCCCATATCTCTCTATGC 3’ and P132 was 5’ ACA- GCTGTTGATGGTGGTGATACG 3’. The blunt-ended PCR products were cloned (Sambrook et al., 1989) and sequenced (Sanger et al., 1977). Sequences were analysed using the GCG program (Devereux et al., 1984).
Fig. 1. Gene arrangement, primer location, subcloning strategy and proteins specified by N and I ORFs. (a) The genomic RNA of PEDV (not to scale) is shown with various viral genes (Pol, polymerase; S, spike; E, small membrane; M, membrane; N, nucleocapsid; and I-1, I-2 and I-3, internal ORFs). The difference in the reading frames of the genes is shown by either upward or downward shift of the rectangles. (b) The 3’ 2 kb region shows M (partial), N and I ORFs. The arrows and numbers show the locations of different primers together with their designation. (c) N gene mRNA along with leader sequences (hatched rectangle) and 20 As was cloned in pBluescript (pMSN) or in the pSCT vector (pMCLN), and used to generate other plasmids using the indicated fragments. N gene mRNA without leader was cloned in pBluescript (pKned) and the pSCT vector (pKCN). The I-1 ORF lacking the N gene start codon (pMC1) was cloned in pSCT vector. (d) The putative proteins encoded by the N gene and I ORFs and the location of peptides used for immunizing rabbits are shown by rectangles.

Fig. 2. The indirect immunofluorescent analysis of I-1 ORF product using antipeptide serum. (a) The I-1 ORF subcloned under the CMV promoter was transiently expressed in Vero cells. (Note that only 10% of transfected cells showed positive signals). (b) Vero cells were transfected with vector plasmid only. (c) Vero cells infected with PEDV and fixed with paraformaldehyde were permeabilized with Triton X-100 for internal fluorescence. A virus syncytium with some of the cells detached is seen.

The wt virus sequences were found to be identical to those of the ca virus. No additional genes were found downstream of the N gene. The three I ORFs were absolutely conserved, with the I-1 ORF starting 55 nt downstream of the N gene AUG. Thus, the propagation of the virus in cell culture did not result in any nucleotide change in this region of the genome. However, previous experiments show that the ca virus is markedly attenuated in virulence for newborn piglets as compared to the wt parent virus (Bernasconi et al., 1995). Taken together, these observations demonstrate that the genes and gene products encoded within the 3’ 1800 nt may not contribute to PEDV attenuation.

As a first step towards identification of the putative gene product of I-1 ORF, I set out to ascertain its coding potential by expressing it under control of the cytomegalovirus (CMV) immediate early promoter. For this purpose, firstly, the 1750 nt of the 3’ end of the PEDV genome, including the poly(A) tail (20 bases), were assembled from two overlapping PCR products [P25/24 and P23/oligo d(T) P25] (Bridgen et al., 1993) to obtain plasmid pKned (Fig. 1). Secondly, cDNA representing N gene mRNA was constructed by inserting a 125 bp SacI–Eco64I fragment of plasmid pKned (Bridgen et al., 1993) to obtain plasmid pKned. (Tobler & Ackermann, 1995) containing the leader sequences in pKNL. (Tobler & Ackermann, 1995) containing the leader sequences in pKned to obtain plasmid pMSN, which served as the parent to construct
various plasmids. Subsequently, the I-1 ORF was cloned under the CMV promoter by inserting an Eco64I (blunt)–BamH I fragment of pMSN containing sequences from 39 nt downstream of the N gene start codon to the poly(A) tail in the HindIII (blunt)–BglII site of the pSCT vector (Rusconi et al., 1990) to obtain pMC1. Plasmid pMC1 was transfected into Vero cells using the calcium phosphate method and the putative product was examined by indirect immunofluorescence using rabbit antipeptide sera developed against the peptide NH$_2$ YERRILKQSFQNSLNSPV COOH, representing the C-terminal 18 residues of I-1. The I-1-specific fluorescent signals were only seen in the cytoplasm of the cells transfected with plasmid pMC1 (10% efficiency of transfection) (Fig. 2a). Experiments with vector plasmid were negative (Fig. 2b).

Further, I investigated the synthesis of I-1 during virus replication by processing PEDV-infected Vero cells (Hofmann & Wyler, 1988) for surface or internal fluorescence. The I-1 antipeptide serum showed positive fluorescence in the cytoplasm of infected cells (Fig. 2c) and no surface fluorescence was observed in non-permeabilized cells (not shown). This clearly established the expression of I-1 as a cytoplasmic protein in PEDV-infected cells.

To characterize the I-1 protein, total cellular proteins from PEDV-infected Vero cells were separated in SDS–20% PAGE (Hofmann et al., 1993) and immunoblotted. The I-1 antipeptide sera reacted with a protein of 12 kDa (Fig. 3a, lane 2) that was absent from the mock-infected cells (lane 1). The genuineness of this protein as a product of I-1 ORF was confirmed by expressing the ORF in baculovirus. For this purpose the recombinant baculoviruses containing cDNA of the N gene mRNA with leader or without leader or lacking the N gene start codon were generated using previously described procedures (Singh et al., 1996). All these viruses synthesized 12 kDa recombinant I-1 protein that migrated with the same mobility as the one made in PEDV-infected cells (Fig. 3a, lanes 6, 7 and 8). No such product was observed in parental BacPAK6 baculovirus (Clontech) or mock-infected Sf9 cells (Fig. 3a, lanes 4 and 5). The immunoblots restained with N protein-specific monoclonal antibodies (weak reaction) revealed a 56 kDa PEDV-specific N protein in lysates of cells infected with recombinant baculoviruses carrying N gene mRNA with or without leader (Fig. 3a, lanes 6 and 7), showing that the protein products were specified by the PEDV sequences.

To seek evidence for the incorporation of I-1 protein in the PEDV virion, the purified virions (Utiger et al., 1995) were disrupted in sample buffer and the proteins were analysed in Western blots. The N protein (reference viral structural protein), not the I-1 protein, was detected in such virion preparations (Fig. 3a, lane 3). This shows that the I-1 protein may not be a major constituent of the virion.

The deduced amino acid sequence of I-1 shows phosphorylation sites at residues 29, 58, 60 and 94. This prompted me to check for a possible phosphorylated I-1 protein in PEDV-infected Vero cells grown in the presence of carrier-free ($^{32}$P)orthophosphate. Total cellular proteins were separated in SDS–PAGE and PEDV-specific proteins were identified by immunostaining the blots first with I-1 antibodies and then with N protein-specific antibodies. Phosphorylated proteins were detected by exposing the blots to X-ray film. Comparison of X-ray film with immunoblots revealed no $^{32}$P-labelled protein band corresponding to the size of the I-1 protein (Fig. 3c, lane 1), while the N protein was detected as a 56 kDa phosphorylated product. Mock-infected controls did not show any of the viral proteins (Fig. 3c, lane 2).

N gene mRNA (1900 bases) has been identified as the smallest subgenomic mRNA of PEDV (Bridgen et al., 1993). To test whether N gene mRNA serves as a template for synthesis of the I-1 protein, plasmids carrying N gene mRNA sequences with (pMCLN) or without (pKCN) leader or carrying only I-1 ORF (pMC1) (Fig. 1) under control of the T7 promoter were individually used for in vitro translation. Besides the 56 kDa N protein, the 12 kDa I-1 protein was also made from the plasmids pMCLN (Fig. 3b, lane 1) and pKCN (Fig. 3b, lane 2). The 12 kDa protein synthesized from these plasmids migrated with the same mobility as did the product from plasmid pMC1 that contained the I-1 AUG as the first available start codon (Fig. 3b, lane 3). The reaction of I-1 antipeptide sera with in vitro-translated I-1 protein confirmed that this protein was synthesized from the I-1 ORF in N gene mRNA.

Further, the strategy of I-1 expression from N gene mRNA was investigated. It is possible that N gene mRNA undergoes editing in vivo and the I-1 ORF is brought in-frame with the N gene initiator codon or there is a ribosomal frameshifting during translation of the N gene mRNA. Synthesis of I-1 in such a manner would result in a fusion protein with amino-terminal residues from the N protein. This possibility was thought unlikely since there are no sequence motifs typical of frameshifting regions. In addition, the rabbit serum raised against a peptide covering the amino-terminal 16 residues of the N protein failed to detect I-1 in Western blots (not shown). The other possibilities could be that the N gene start codon is modified or removed, leaving the I-1 initiator as the first AUG on the mRNA or there may be a separate subgenomic mRNA for I-1. To detect such an edited or I-1-specific mRNA, total RNA isolated from the PEDV-infected cells was subjected to RT–PCR. First strand cDNA was synthesized using message complementary primers P70 [located 7 nt down from the I-1 initiator] or P32 [located 92 nt upstream of the poly(A) tail] or oligo d(T)-tailed P25. PCR amplification products from combinations of primer P117 (located in the leader sequences) and P70 or P32 or P25 were directly sequenced. No changes in the sequences of PCR products were found as compared to that of N gene mRNA, indicating the absence of any type of editing in the N gene mRNA. In addition, the sequencing of PCR products of primer pairs P117/32, P117/25 and P132/25 demonstrated that N gene mRNA was the only message transcribed from the 3’ end of the PEDV genome.
Thus, the successful expression of I ORFs would require either reinitiation of translation or internal entry of ribosomes in N gene mRNA. The expression pattern of I-1 from unedited N gene mRNA suggests that PEDV, like BCV (Senanayake & Brian, 1997), may use the leaky scanning model for the synthesis of I-1 protein. Since the N gene start codon is in suboptimal context (UUUAUGG), some of the 40S ribosomal subunits would bypass it to start translation at a downstream AUG (Kozak, 1989). In that case, the expression of I-1 from an initiator in the weakest primary sequence context (UCUAUGC) would require an appropriately positioned downstream secondary structure in the N gene mRNA. A stem–loop structure seems to be formed 20 nt downstream of the I-1 start codon. Such a stem–loop structure located 14 nt downstream of start codon has been shown to facilitate translation initiation from an AUG in unfavourable contexts and even from non-AUG codons (Kozak, 1990). No IRES-specific sequences, essential for direct internal entry of 40S ribosomes, were observed in the N gene mRNA, thus, obviating I-1 synthesis by internal initiation of translation. Alternatively, I-1 may also be expressed from N gene mRNA by ribosome shunting (Yueh & Schneider, 1996). However, experiments were not performed to prove or disprove these possibilities.

The exact function of coronavirus I ORFs and their proteins is not known except for MHV I protein, which is a constituent of the virion (Fischer et al., 1997). The BCV I protein is only suggested to be part of the virion (Senanayake et al., 1992). In contrast, the PEDV I-1 protein was not present in detectable amounts in virions. However, presence of I-1 in a low copy number in virions cannot be ruled out, suggesting that I-1 could be a structural protein. Although I-1 sequences carry various protease signature patterns and potential phosphorylation sites, in my experiments I-1 was not detected as a phosphorylated protein. The I-2 and I-3 ORF products were not detected in virus-infected cells (not shown). Further experiments involving disruption of I ORF initiation and stop codons or early termination of I ORFs without disturbing the N gene function are envisaged to show whether the I ORFs and their products are a requirement for virus viability.

The author is highly thankful to Dr Maria Foti for her help in immunization of rabbits and Professor Mathias Ackermann for his moral support.

These studies were supported by Swiss National Science Foundation grants no. 31-37418.93 and 31-43503.95.

References


Internal protein of PEDV


Received 23 February 1999; Accepted 29 April 1999