

**Molecular characterization of positive-strand RNA viruses:
pestiviruses and the porcine reproductive and respiratory
syndrome virus (PRRSV)**

**H.-J. Thiel¹, G. Meyers¹, R. Stark¹, N. Tautz¹, T. Rumenapf², G. Unger¹,
and K.-K. Conzelmann**

¹Federal Research Centre for Virus Diseases of Animals, Tübingen,
Federal Republic of Germany

²California Institute of Technology, Division of Biology, Pasadena, CA, U.S.A.

Summary. Molecular characterization has become an important tool for the analysis of viruses including their classification. The manuscript focuses on the molecular analysis of two members of the genus pestivirus (hog cholera virus, HCV and bovine viral diarrhoea virus, BVDV) and of the recently discovered porcine reproductive and respiratory syndrome virus (PRRSV). The first protein encoded within the single large pestivirus ORF is a nonstructural protein with autoproteolytic activity. The cleavage site between the protease and the capsid protein p14 has been predicted previously, but recent experimental data indicate that processing occurs at a different site. The capsid protein is followed by a putative internal signal sequence and three glycoproteins which are part of the virion envelope. According to a new proposal for the nomenclature of the structural proteins of pestiviruses they are termed C, E0, E1 and E2. The genomes of BVDV pairs isolated from animals which came down with mucosal disease were analyzed. The genomes from cytopathogenic (cp) BVD viruses may contain insertions highly homologous to cellular sequences. In addition, cp BVDV may differ from its non cytopathogenic (noncp) counterpart by mere rearrangement of viral sequences. The disease PRRS, which emerged a few years ago, is caused by a single strand RNA virus; the viral genome is of positive polarity and has a size of 15 kb. Data concerning morphology, morphogenesis and virion composition suggested already that PRRSV belongs to a group of so-called arteriviruses which comprises equine arteritis virus (EAV), lactate dehydrogenase elevating virus (LDV) and simian hemorrhagic fever virus (SHFV). This conclusion has now been confirmed by analysis of genome organization, gene expression strategy and by comparison of deduced protein sequences.

Introduction

Pestiviruses are causative agents of important animal diseases such as hog cholera, bovine viral diarrhoea and border disease of sheep. The pestivirus genome represents a single-stranded RNA of about 12.5 kb, which is of positive polarity. The genomic RNA comprises a single large open reading frame [8, 15]; accordingly, membership of the genus pestivirus in the family *Togaviridae* was no longer justified. Instead, pestiviruses have recently been reclassified as members of the family *Flaviviridae* which now comprises the genera flavivirus, pestivirus and the hepatitis C virus group. These viruses have certain characteristics in common like overall genome organization and strategy of gene expression. Antisera against bacterial fusion proteins and synthetic peptides as well as monoclonal antibodies were used to identify proteins encoded by pestiviruses [5, 7, 27–30]. Interestingly, the genomes of several cytopathogenic BVDV strains contain insertions highly homologous to cellular sequences [16, 17, 19, 23].

The porcine reproductive and respiratory syndrome virus (PRRSV) has been described as a small enveloped RNA virus [2, 32] with morphological and morphogenetical similarities to members of the arterivirus group, including equine arteritis virus (EAV) and lactate dehydrogenase-elevating virus of mice (LDV). In addition, relationships between PRRSV and arteriviruses are suggested by the nature of permissive cells. The arteriviruses infect particular subpopulations of macrophages [21] and PRRSV apparently grows exclusively in alveolar lung macrophages. However, serological crossreactions could so far not be demonstrated between PRRSV and any of the arteriviruses.

Results and discussion

Pestiviruses

a. Hog cholera virus

Analysis of the autoprotease HCV p23. The nonstructural protein HCV p23 represents the first protein of the pestivirus ORF. As already shown for BVDV p20 [33], HCV p23 also possesses autoproteolytic activity. A precursor molecule consisting of p23 and p14 could never be demonstrated even when very short pulse periods were used. This result was obtained after infection with HCV as well as different Vaccinia virus/HCV recombinants. Processing was also observed after in vitro translation of HCV RNA and in vitro transcription/translation experiments

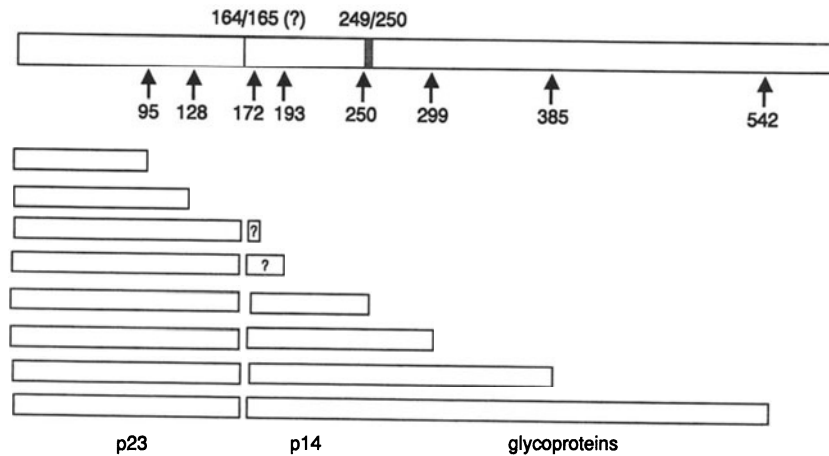


Fig. 1. Evidence that HCV p23 possesses proteolytic activity. After linearization of an HCV cDNA construct with different restriction enzymes, in vitro transcription/translation was performed. The proteins were identified after immunoprecipitation by SDS-PAGE

with cDNA constructs. The latter experiments also showed that the proteolytic activity resides in HCV p23 (Fig. 1).

For BVDV and HCV it has been proposed that cleavage occurs after Trp-164 of the ORF [28, 33] and this assumption was supported by in vitro mutagenesis studies [33]. However, there is new evidence indicating that cleavage occurs after Cys-168 (Fig. 2). Firstly, p14 could not be labeled with [³⁵S]-cysteine which is absent from p14 only if cleavage occurs after Cys-168. Secondly, the genomic region comprising HCV p23 and p14 was expressed in bacteria. In this prokaryotic expression system the release of a 14 kD protein was observed. N-terminal sequencing of this protein revealed that the protein starts with Ser-169 (manuscript in preparation). Further efforts concern characterization of the protease and elucidation of its function.

Interestingly, the analysis of two genomes of cytopathogenic (cp) BVDV strains revealed that their genomes contain duplications encompassing the autoprotease p20. The 3'-terminal codon of the inserted p20 coding region is identical in both cp-BVDV genomes and codes for a cysteine analogous to Cys-168 of the pestiviral ORF [18].

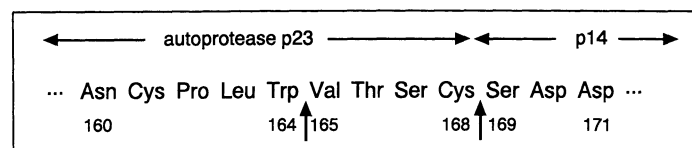


Fig. 2. Cleavage site of pestivirus autoprotease

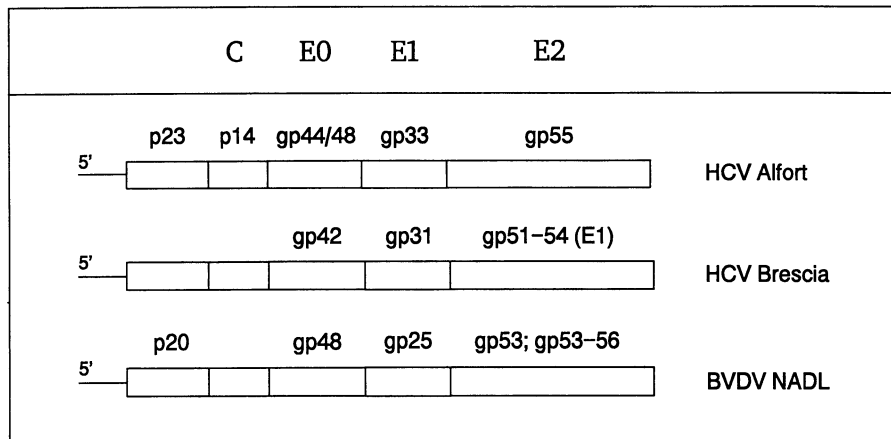


Fig. 3. Suggested nomenclature for structural proteins of pestiviruses

Structural proteins of HCV

Within the pestiviral polyprotein the autoprotease HCV p23/BVDV p20 is followed by the nucleocapsid protein p14 which has been demonstrated for HCV as well as BVDV [28]. The following internal signal sequence probably mediates translocation of the first structural glycoprotein, HCV gp44/48. Together with gp33 and gp55, the three glycoproteins constitute the virion envelope. The glycoproteins form parts of disulfide-linked dimers [28, 29]. While gp33 probably represents a transmembrane protein, the other two glycoproteins both are exposed on the surface of virions and induce virus neutralizing antibodies [30].

So far, the designation of pestivirus encoded glycoproteins varies among different laboratories because most research groups use the apparent molecular weights as a basis for nomenclature (Fig. 3). In order to obtain a common nomenclature, we suggest to use the abbreviations C, E0, E1 and E2 for nucleocapsid protein and the glycoproteins, respectively (Fig. 3). This approach allows also to have the same designations for analogous glycoproteins from the hepatitis C virus group (Fig. 4). It remains to be seen whether the (glyco)proteins from members of the genus flavivirus will be termed accordingly. At this point the designation E0 (for HCV gp44/48, BVDV gp48) is being debated because (1) it is clearly a structural protein of pestiviruses and (2) the "0" may be misleading because it is used in other virus systems for precursor molecules. It appears that E0 represents a unique glycoprotein among pestiviruses which is not only a structural protein but also secreted from infected cells (manuscript in press).

In order to determine the processing sites between C and E0 as well as between the glycoproteins, N-terminal sequencing was performed.

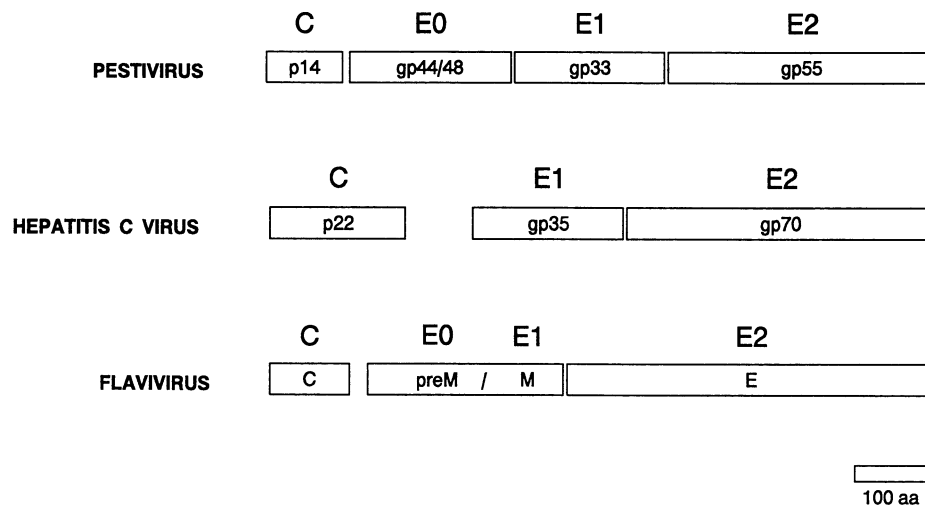


Fig. 4. Structural proteins of *Flaviviridae*

	E 0	E 1	E 2
Number of Amino Acids	227	195	373 (?)
Calculated Molecular Weight (kDa)	25.8	21.8	41.9
Apparent Molecular Weight (kDa)	48 (gp 44/48)	33 (gp33)	55 (gp55)
Number of Possible N-glycosylation Sites	9	3	6
Apparent Molecular Weight after Endoglycosidase F (kDa)	similar to the respective calculated molecular weights		

Fig. 5. Glycoproteins of hog cholera virus

For this purpose the glycoproteins were purified from virus infected cells by immunoaffinity columns using monoclonal antibodies against E0 and E2. E1 was copurified with E2 because of the linkage of the two glycoproteins by disulfide bridges. The results of N-terminal sequencing allow important conclusions concerning the three glycoproteins, namely the actual sizes of the protein backbones and the calculated contribution of carbohydrate moieties to apparent molecular weights (Fig. 5). The indicated sizes of the protein backbones are based on the assumption that only one cleavage occurs between the glycoproteins. With regard to E2 the C terminus remains to be determined.

On the basis of the data outlined above we intend to study the biosynthesis of pestivirus glycoproteins. A particularly interesting aspect of this study will be the analysis of a supposed hierarchy of cleavage

events (manuscript in press). In addition, C-terminal truncations of HCV gp55 (E2) are being performed to determine its C-terminus and to identify signals important for heterodimerization and homodimerization (manuscript in preparation).

With regard to the development of vaccines against pestiviruses the role of individual glycoproteins in the induction of protective immunity will be studied. These efforts include use of (1) recombinant vaccines based upon different vectors, primarily vaccinia virus and pseudorabies virus and (2) subunit vaccines containing pestivirus glycoprotein(s) expressed for example in insect cells after infection with recombinant baculoviruses.

b. Bovine viral diarrhea virus

Pathogenesis of mucosal disease. Comparison of the genomic sequences of two BVDV strains (BVDV Osloss [26] and BVDV NADL [8]) led to the identification of small insertions located in a region coding for a nonstructural protein [6, 15, 19]. The insertion of 228 nucleotides identified in the BVDV Osloss genome encodes a complete ubiquitin-like element with only two amino acid exchanges with respect to the ubiquitin sequence conserved in all animals [16, 19]. The sequence of 270 nucleotides which is inserted in the BVDV NADL genome shows no homology to a ubiquitin gene but is almost identical with another bovine mRNA sequence (Fig. 6) [16].

In tissue culture two BVDV biotypes, cytopathogenic BVDV (cpBVDV) and noncytopathogenic BVDV (noncpBVDV), can be distinguished [1] (Table 1). Both biotypes are involved in pathogenesis of mucosal disease (MD), the most severe clinical manifestation of BVDV infections. A prerequisite for MD is a persistent infection with

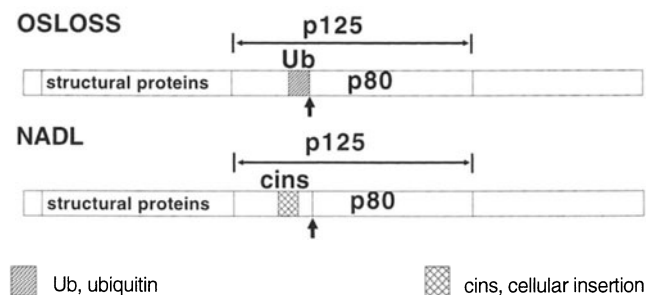
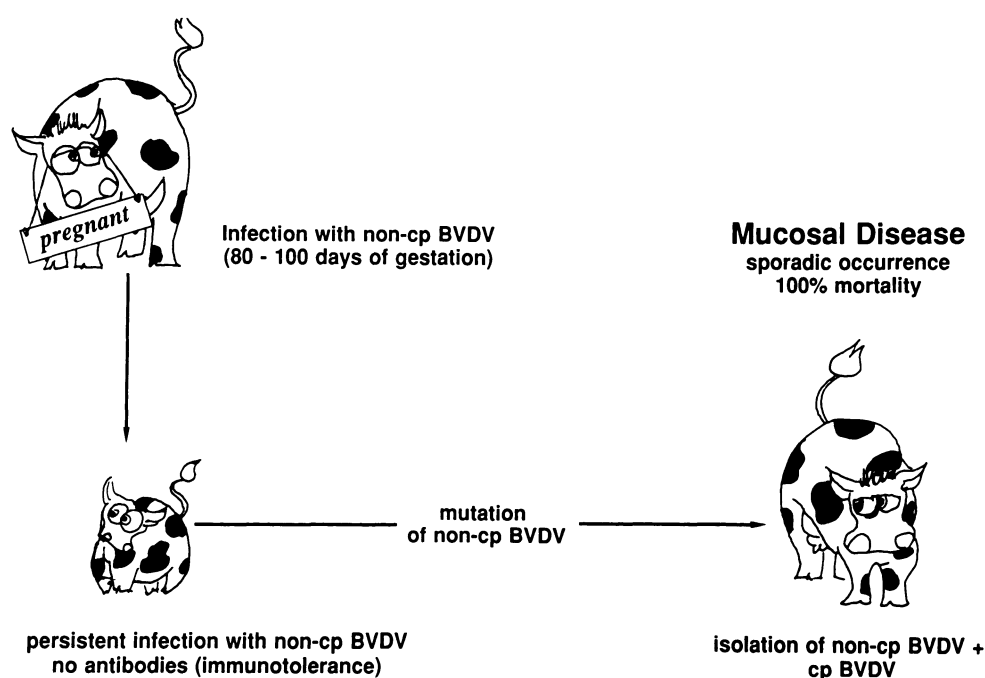


Fig. 6. Insertions identified in CP BVDV strains. *OSLOSS* 228 nucleotides; coding for a ubiquitin-like protein; equivalent to one ubiquitin monomer; 2 amino acid exchanges. *NADL* 270 nucleotides; no homology to ubiquitin; 99% identity to a bovine mRNA sequence

Table 1. Characteristics of BVDV biotypes

	Non-CP	CP
Cytopathogenicity (CP) in tissue culture cells	No	Yes
Strains (laboratory)	New York	NADL, Osloss, Singer, Oregon, Denmark
Detection of p80	No (only p125)	Yes (in addition to p125)
Pathogenicity in cattle	Yes	Yes
Occurrence of mucosal disease	Always both biotypes isolated	

**Fig. 7.** Pathogenesis of mucosal disease

noncpBVDV (Fig. 7). Surprisingly, cpBVDV can always be isolated from MD animals in addition to the persisting noncp virus [3, 4]. In contrast to the described antigenic variability of BVDV field isolates the members of such a “pair” of noncpBVDV and cpBVDV are antigenically very closely related [10, 22]. This observation led to the hypothesis that during pathogenesis of MD a cpBVD virus develops from the noncp virus by acquiring some kind of mutation [10]. We proposed as a working hypothesis that recombination between viral and cellular RNA led to the formation of these cpBVDV genomes [16, 19].

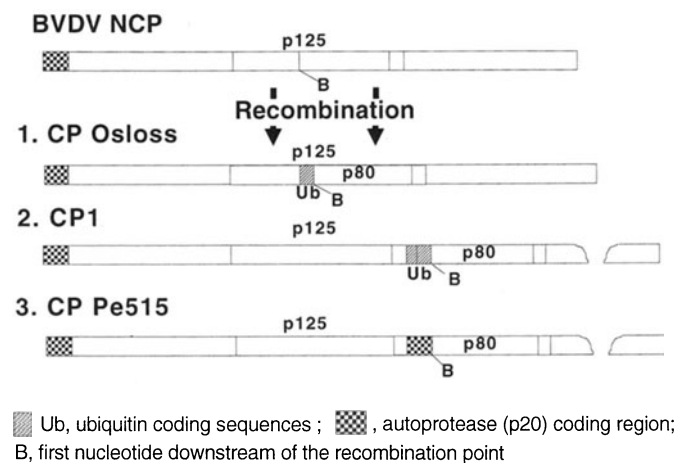


Fig. 8. Comparison of BVDV genomes

To investigate directly the difference between a cytopathogenic virus and its noncytopathogenic counterpart we analyzed the genomes of a pair of cpBVDV (CP1) and noncpBVDV (NCP1) isolated from one MD animal. The RNA of CP1 was found to contain a ubiquitin-coding element which is embedded in a large duplication of viral sequence encompassing the p80-coding region (Fig. 8). In contrast, the genome of NCP 1 does not contain either insertion or duplication [17]. According to the results of these studies, one possible mutation leading to cytopathogenic BVDV is a recombination process between cellular and viral RNA.

In the case of two other cpBVDV isolates from BVDV pairs no host cell-derived insertion could be identified. However, elaborate duplication and rearrangement of viral sequences were found in both genomes (for CP Pe515 see Fig. 8). The analysis of the noncp virus (NCP Pe515) revealed that neither duplication nor rearrangement of sequences are present, and thus supported the linkage between recombination and establishment of the cytopathogenic phenotype. Accordingly, generation of cpBVDV is not restricted to recombination between cellular and viral sequences but can also be achieved by mere rearrangement of viral sequences.

One important change at the protein level distinguishing cpBVDV from noncpBVDV is the expression of p80 in cpBVDV infected cells (Fig. 6, Fig. 8). A prerequisite for generation of p80 is the presence of a cleavage site at the aminoterminal of this protein. Based on our current knowledge two basically different ways can be distinguished. One possibility is represented by the Osloss strain where the signal for cleavage is provided by a cellular ubiquitin sequence inserted into the p125 region of

the viral polyprotein. Ubiquitin itself is synthesized within eukaryotic cells in form of fusion proteins which are subsequently cleaved at the carboxyterminal end of the ubiquitin moiety [12, 13, 20, 24, 25]. Accordingly, the insertion of ubiquitin into p125 introduces a signal for processing by a cellular protease (manuscript in preparation). The second way to generate p80 requires duplication of the respective coding region. Again a processing signal is necessary which is now placed in front of the duplicated sequence. In the case of CP1 ubiquitin provides again the signal for cleavage by a cellular protease (manuscript in preparation). For Pe515CP and CP6 a processing site encoded at the end of a second virus-derived duplicated element serves this purpose. This second duplicated element codes for the viral protein p20. Interestingly, insertion of the p20-coding region also transfers the protease responsible for the cleavage to its point of action. The integration of both protease and cleavage site might be necessary because of the lack of trans function of the p20 protease. This would be in accordance with the hypothesis that p20 can only act as an autoprotease [33].

Both alternatives outlined above would allow identical aminoterminal ends for p80. The expression of this protein is strictly correlated with the cytopathogenic phenotype and the development of MD. Therefore p80 represents the prime candidate for the agent responsible for killing the infected cells.

Porcine reproductive and respiratory syndrome virus (PRRSV)

Members of the arterivirus group are currently classified within the *Togaviridae* family [31], but the need for reclassification has become obvious after cloning and molecular analysis of the total EAV genome [11] and of parts of the LDV genome [14]. In contrast to togaviruses, arterivirus gene expression is characterized by transcription of multiple subgenomic mRNAs, each encoding one protein. Similar to coronaviruses, arteriviral mRNAs form a 3' coterminal nested set and possess common 5' terminal leader sequences which are joined to the bodies of the mRNAs during transcription. Moreover, the putative EAV polymerase gene is probably expressed by ribosomal frameshifting as in coronaviruses and possesses conserved domains also present in corona- and torovirus polymerases [11].

Using purified PRRS virions from infected macrophages as starting material molecular cDNA cloning and sequencing was performed. PRRSV specific cDNA clones spanning the 3' terminal 5kb of the genomic RNA were isolated, sequenced and used as probes for identification of PRRSV specific RNAs. The PRRSV genome is a positive

stranded polyadenylated RNA of about 15 kb. In infected cells a 3' coterminal nested set of six major subgenomic mRNAs could be demonstrated. Within the 3' terminal 3.5 kb of the PRRSV genome six overlapping reading frames (ORFs) were identified, each most likely expressed by one of the subgenomic mRNAs. Amino acid sequence comparisons revealed that the most 3' terminal ORF (ORF7) encodes the PRRSV nucleocapsid protein with a calculated molecular weight of 14 kD. It displays 44.8% amino acid identity with the capsid protein of lactate dehydrogenase-elevating virus (LDV) and 23.6% with that of equine arteritis virus (EAV). The product of ORF6, the second 3' terminal ORF, represents a putative membrane protein and exhibits 53.2% and 27.2% amino acid identity with the corresponding LDV and EAV polypeptides. Similar to EAV, ORFs 2 through 5 might encode glycosylated viral proteins. The polypeptide deduced from the most 5' ORF (ORF1b) contains two conserved domains common to EAV and coronavirus polymerases. Genome organization, strategy of gene expression and the sequence of deduced proteins show that PRRSV belongs to the arterivirus group of viruses [9].

Acknowledgements

This study was in part supported by the Deutsche Forschungsgemeinschaft (Th 298/3-1), the Bundesministerium für Forschung und Technologie and Intervet International BV (project 0319028A).

References

1. Baker JC (1987) Bovine viral diarrhea virus: A review. *J Am Vet Med Assoc* 190: 1449–1458
2. Benfield DA, Nelson E, Collins JE, Harris L, Goyal SM, Robison D, Christianson WT, Morrison RB, Gorcyca D, Chladek D (1992) Characterization of swine infertility and respiratory syndrome virus (isolate ATCC VR-2332). *J Vet Diagn Invest* 4: 127–133
3. Bolin SR, McClurkin AW, Cutlip RC, Coria MF (1985) Severe clinical disease induced in cattle persistently infected with noncytopathic bovine viral diarrhea virus by superinfection with cytopathic bovine viral diarrhea virus. *Am J Vet Res* 46: 573–576
4. Brownlie J, Clarke MC, Howard CJ (1984) Experimental production of fatal mucosal disease in cattle. *Vet Rec* 114: 535–536
5. Collett MS, Wiskerchen MA, Welniak E, Belzer SK (1991) Bovine viral diarrhea virus genomic organization. In: Liess B, Moennig V, Pohlenz J, Trautwein G (eds) *Ruminant pestivirus infections*. Springer, Wien New York, pp 19–27 (*Arch Virol [Suppl]* 3)

6. Collett MS, Moennig V, Horzinek MC (1989) Recent advances in pestivirus research. *J Gen Virol* 70: 253–266
7. Collett MS, Larson R, Belzer SK, Retzel E (1988) Proteins encoded by bovine viral diarrhea virus: the genomic organization of a pestivirus. *Virology* 165: 200–208
8. Collett MS, Larson R, Gold C, Strinck D, Anderson DK, Purchio AF (1988) Molecular cloning and nucleotide sequence of the pestivirus bovine viral diarrhea virus. *Virology* 165: 191–199
9. Conzelmann K-K, Visser N, Van Woensel P, Thiel H-J (1993) Molecular characterization of porcine reproductive and respiratory syndrome virus, a member of the arterivirus group. *Virology* 193: 329–339
10. Corapi WV, Donis RO, Dubovi EJ (1988) Monoclonal antibody analyses of cytopathic and noncytopathic viruses from fatal bovine viral diarrhea virus infections. *J Virol* 62: 2823–2827
11. Den Boon JA, Snijder EJ, Chirnside ED, De Vries AAF, Horzinek MC, Spaan WJM (1991) Equine arteritis virus is not a togavirus but belongs to the coronavirus superfamily. *J Virology* 65: 2910–2920
12. Finley D, Bartel B, Varshavsky A (1989) The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. *Nature* 338: 394–401
13. Finley D, Özkaynak E, Varshavsky A (1987) The yeast polyubiquitin gene is essential for resistance to high temperatures, starvation and other stresses. *Cell* 48: 1035–1046
14. Godeny EK, Speicher DW, Brinton MA (1990) Map location of lactate dehydrogenase-elevating virus (LDV) capsid protein (Vp1) gene. *Virology* 177: 768–771
15. Meyers G, Rümenapf T, Thiel H-J (1989a) Molecular cloning and nucleotide sequence of the genome of hog cholera virus. *Virology* 171: 555–567
16. Meyers G, Rümenapf T, Thiel H-J (1990) Insertion of ubiquitin-coding sequence identified in the RNA genome of a Togavirus. In: Brinton MA, Heinz FX (eds) *New aspects of positive strand RNA viruses*. American Society for Microbiology, Washington, D.C., pp 25–29
17. Meyers G, Tautz N, Dubovi EJ, Thiel H-J (1991) Viral cytopathogenicity correlated with integration of ubiquitin-coding sequences. *Virology* 180: 602–616
18. Meyers G, Tautz N, Stark R, Brownlie J, Dubovi EJ, Collett MS, Thiel H-J (1992) Rearrangement of viral sequences in cytopathogenic pestiviruses. *Virology* 191: 368–386
19. Meyers G, Rümenapf T, Thiel H-J (1989b) Ubiquitin in a togavirus. *Nature* 341: 491
20. Özkaynak E, Finley D, Solomon MS, Varshavsky A (1987) The yeast ubiquitin genes: A family of natural gene fusions. *EMBO J* 6: 1429–1439
21. Plagemann PGW, Moennig V (1992) Lactate dehydrogenase-elevating virus, equine arteritis virus and simian hemorrhagic fever virus. *Adv Virus Res* 41: 99–192
22. Pocock DH, Howard CJ, Clarke MC, Brownlie J (1987) Variation in the intracellular polypeptide profiles from different isolates of bovine viral diarrhea virus. *Arch Virol* 94: 43–53
23. Qi F, Ridpath JF, Lewis T, Bolin SR, Berry ES (1992) Analysis of the bovine viral diarrhea virus genome for possible cellular insertions. *Virology* 189: 285–292
24. Rechsteiner M (1987) Ubiquitin-mediated pathways for intracellular proteolysis. *Annu Rev Cell Biol* 3: 1–30

- 52 H.-J. Thiel et al.: Molecular characterization of positive-strand RNA viruses
25. Redmann KL, Rechsteiner M (1989) Identification of the long ubiquitin extension as ribosomal protein S27a. *Nature* 338: 438–440
 26. Renard A, Dino D, Martial J (1987) Vaccines and diagnostics derived from bovine diarrhea virus. European Patent Application number 86870095.6. Publication number 0208672, 14 January 1987
 27. Stark R, Rüménapf T, Meyers G, Thiel H-J (1990) Genomic localization of hog cholera virus glycoproteins. *Virology* 174: 286–289
 28. Thiel H-J, Stark R, Weiland E, Rüménapf T, Meyers G (1991) Hog cholera virus: molecular composition of virions from a pestivirus. *J Virol* 65: 4705–4712
 29. Weiland E, Stark R, Haas B, Rüménapf T, Meyers G, Thiel H-J (1990) Pestivirus glycoprotein which induces neutralizing antibodies forms part of a disulfide-linked heterodimer. *J Virol* 64: 3563–3569
 30. Weiland E, Ahl R, Stark R, Weiland F, Thiel H-J (1992) A second envelope glycoprotein mediates neutralization of a pestivirus, hog cholera virus. *J Virol* 66: 3677–3682
 31. Westaway EG, Brinton MA, Gaidamovich SYA, Horzinek MC, Igarashi A, Kääriäinen L, Lvov DK, Porterfield JS, Russel PK, Trent DW (1985) Togaviridae. *Intervirology* 24: 125–139
 32. Wensvoort G, Terpstra C, Pol JMA, Wagenaar F (1991) Lelystad virus, the cause of porcine epidemic abortion and respiratory syndrome (mystery swine disease). In: *Proc Porcine Reprod Respir Syndrome Semin*, November 4–5, Brussels, Belgium, pp 27–35. European Commission VI/BII.2
 33. Wiskerchen MA, Belzer SK, Collett MS (1991) Pestivirus gene expression: The first protein of the bovine viral diarrhea virus large open reading frame, p20, possesses proteolytic activity. *J Virol* 65: 4508–4514

Authors' address: Dr. H.-J. Thiel, Federal Research Centre for Virus Diseases of Animals, Paul-Ehrlich-Strasse 28, D-72076 Tübingen, Federal Republic of Germany.