



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

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

## Rearrangement of Viral Sequences in Cytopathogenic Pestiviruses

GREGOR MEYERS, NORBERT TAUTZ, ROBERT STARK, JOE BROWNLIE,\*  
EDWARD J. DUBOVI,† MARC S. COLLETT,‡ AND HEINZ-JÜRGEN THIEL<sup>1</sup>

Federal Research Centre for Virus Diseases of Animals, P.O. Box 1149, 7400 Tübingen, Federal Republic of Germany; †New York State College of Veterinary Medicine, Cornell University, Ithaca, New York 14853; \*AFRC Institute for Animal Health, Compton Laboratory, Compton, Newbury, Berkshire RG 16 0NN, United Kingdom; and ‡MedImmune Inc., 19 Firstfield Road, Gaithersburg, Maryland 20878

Two cytopathogenic isolates of bovine viral diarrhea virus (cpBVDV) have been analyzed. For both viruses two regions of their genomic RNAs were found to be duplicated and rearranged. The viral genomes contain a small duplicated element (SD) derived from the genomic 5' end far downstream of its original context. This sequence is followed by a larger duplication which encompasses the region coding for the protein p80(LD), a molecular marker for cpBVDV. The SD element codes for the viral protease p20. In the case of the viruses analyzed here the aminoterminal of p80 is generated by autoproteolytic removal of the preceding SD-encoded protease. For one of the cpBVDV isolates a specific fusion protein (p28) could be identified which is composed of p20 and part of p10, another viral protein. Molecular characterization of the respective noncytopathogenic counterpart revealed that duplication and rearrangement of sequences as well as the expression of p28 and p80 are specific for the cytopathogenic virus. © 1992 Academic Press, Inc.

### INTRODUCTION

Bovine viral diarrhea virus (BVDV) represents a small enveloped RNA virus that belongs to the genus pestivirus within the family Flaviviridae (Wengler, 1991). Other members of the genus are hog cholera virus (HCV) and border disease virus (BDV) of sheep. The single-stranded RNA genome of pestiviruses has a size of about 12.5 kb and possesses positive polarity (Renard *et al.*, 1987; Collett *et al.*, 1988a; Meyers *et al.*, 1989a). Viral gene expression occurs via synthesis of a polyprotein and subsequent proteolytic processing (Collett *et al.*, 1988b,c; Meyers *et al.*, 1989a).

BVDV is the causative agent of a variety of syndromes in cattle among which the most severe is mucosal disease (MD) with almost 100% lethality (for review: Baker, 1987). Interestingly, a pair of viruses belonging to two different biotypes of BVDV can always be isolated from animals which come down with MD. These two biotypes are distinguished by their ability to replicate in tissue culture cells with or without cytopathic effect (cpBVDV or noncpBVDV, respectively). Elaborate studies revealed that a prerequisite for development of MD is an intrauterine infection with noncpBVDV (Brownlie *et al.*, 1984; Bolin *et al.*, 1985). In a second step generation of cpBVDV occurs, which

was assumed to be due to a mutation in the viral genome (Pocock *et al.*, 1987; Corapi *et al.*, 1988).

At the molecular level the only difference between cpBVDV and noncpBVDV known until recently was the presence of an 80-kDa protein (p80) in cells infected with cytopathogenic viruses. This protein is colinear with the carboxyterminal part of a polypeptide of 125 kDa (p125) detectable in cells infected with either cpBVDV or noncpBVDV (Purchio *et al.*, 1984; Pocock *et al.*, 1987; Corapi *et al.*, 1988; Collett *et al.*, 1988b). Surprisingly, for two cpBVDV strains host cell-derived insertions were detected within the genomic region coding for p125 (Meyers *et al.*, 1989b, 1990). One of the inserted sequences codes for a complete ubiquitin-like protein while the function of the cellular homologue of the other insertion is still unknown (Meyers *et al.*, 1989b, 1990). We proposed that recombination between cellular and viral RNA led to the formation of these cpBVDV genomes and that the insertion of cellular elements is linked to generation of p80 within the infected cells (Meyers *et al.*, 1990).

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. For CP1 a ubiquitin-coding element which is embedded in a large duplication of viral sequences encompassing the p80-coding region was identified. In contrast, the genome of NCP1 does not contain either insertion or duplication (Meyers *et al.*, 1991). According to the results of these studies, one possible mutation leading to cytopathogenic BVDV is a recombina-

<sup>1</sup> To whom correspondence and reprint requests should be addressed.

Sequence data from this article have been deposited with GenBank/EMBL Data Libraries under accession Nos. M96640 (Fig. 3A), M96641 (Fig. 3B), M96638 (Fig. 3C), and M96639 (Fig. 4).

tion process between cellular and viral RNA. In this report molecular characterization of two additional pairs of cpBVDV and noncpBVDV is reported which adds a new aspect to our understanding of generation of cytopathogenic BVDV.

## MATERIALS AND METHODS

### Materials

Restriction enzymes and modifying enzymes were from Pharmacia-LKB (Freiburg, FRG), New England Biolabs (Schwalbach, FRG), Boehringer-Mannheim (Mannheim, FRG), and Life Sciences, Inc. (USA). AmpliTaq DNA polymerase was obtained from Cetus-Perkin-Elmer (Langen, FRG). Radioactive compounds were purchased from Amersham-Buchler (Braunschweig, FRG). Kodak XAR5 X-ray films were used for autoradiography.

### Cells and viruses

MDBK cells and the BVDV strain NADL (Gutekunst and Malmquist, 1963) were obtained from the American Type Culture Collection (Rockville, MD). The BVDV strain Osloss (Renard *et al.*, 1987) was kindly provided by Dr. Liess (Veterinary School, Hannover, FRG). Isolation and serological characterization of BVDV strains NCP1, CP1, NCP6, and CP6 have been described by Corapi *et al.* (1988); for strains Pe515CP and Pe515NCP, see Brownlie *et al.* (1984). Cells were grown in DMEM supplemented with 10% fetal calf serum and were infected with BVDV as layer cells at  $1.5 \times 10^6$  per 3.5-cm dish. Cells and virus stocks were tested every 6 to 12 months for the absence of mycoplasma contamination.

### RNA preparation, gel electrophoresis and Northern hybridization

Preparation of RNA was done as described (Rümenapf *et al.*, 1989). Five microgramms of glyoxylated RNA (Maniatis *et al.*, 1989) was separated in phosphate-buffered 1% agarose gels containing 5.5% formaldehyde. An RNA ladder (Bethesda Research Laboratories, FRG) served as a size standard. For Northern hybridization the RNA was transferred to Duralon-UV membranes (Stratagene, Heidelberg, FRG). Membranes were hybridized to the NCP1 cDNA probe labeled with  $^{32}\text{P}$  by nick translation (Rigby *et al.*, 1977) (nick translation kit, Amersham-Buchler, Braunschweig, FRG) in 0.5 M sodium phosphate, pH 6.8, 1 mM EDTA, and 7% SDS at 54°C. Posthybridization washes were carried out with 40 mM sodium phosphate, pH 6.8, 1 mM EDTA and 5% SDS, and 40 mM sodium phosphate, pH 6.8, 1 mM EDTA, and 1% SDS

two times each for 30 min at hybridization temperature.

### Oligonucleotides

Oligonucleotides were synthesized on a Biosearch 8700 DNA synthesizer (New Brunswick Scientific, FRG) using the phosphoamidite method (Beaucage and Caruthers, 1981). Oligonucleotides BVDV 13, BVDV 14, and PES 9 are described elsewhere (Meyers *et al.*, 1991).

OIBVD32: AAATCTCTGCTGTACATCGCACATG

OIBVD33: GCATCCATCATNCCRTGRAT

### cDNA synthesis, cloning, and library screening

Synthesis of double-stranded cDNA, cloning in  $\lambda$  ZAPII (Stratagene, Heidelberg, FRG), and establishment and screening of the phage library using the NCP1 cDNA probe was done as described before (Meyers *et al.*, 1991). Subcloning of cDNA fragments into pBluescript plasmids by *in vivo* excision was performed as recommended by the supplier (Stratagene, Heidelberg, FRG).

### Polymerase chain reaction

First strand cDNA was synthesized as described (Rümenapf *et al.*, 1989) using oligonucleotide OI-BVDV33 as primer. For RNA hydrolysis 1 volume of 0.3 M NaOH and  $\frac{3}{4}$  volumes of 0.03 M EDTA were added and the mixture incubated at 95°C for 5 minutes. After neutralization with  $\frac{1}{4}$  volume of 1 M Tris/HCl (pH 8.0) the volume was adjusted to 100  $\mu\text{l}$  and the solution was passed through a Sephadex G50 spun column (Maniatis *et al.*, 1989). 1% of the single stranded cDNA served as template for the polymerase chain reaction (PCR). Buffer conditions for PCR were as described by Jeffreys *et al.* (1988) except that the 50  $\mu\text{l}$  reaction mixture contained 10% dimethylsulfoxide and 50 pMol of primers OI-BVDV32 and OI-BVDV33. Samples were overlaid with 75  $\mu\text{l}$  of paraffin and cycled 30 times for 100 s at 94°C, 100 s at 50°C and 200 s at 72°C in a thermocycler TPS (Landgraf, Hannover, FRG). After purification by chloroform extraction and preparative agarose gel electrophoresis the PCR product was cloned blunt end into the SmaI site of pBluescript SK-(Stratagene, Heidelberg, FRG) according to standard procedures (Maniatis *et al.*, 1989).

### Nucleotide sequencing

Exonuclease III and S1 were used to establish deletion libraries of cDNA clones (Hennikoff, 1987). Dideoxy sequencing (Sanger *et al.*, 1977) of double-stranded DNA templates was carried out using the T7

polymerase sequencing kit (Pharmacia-LKB). All sequences shown in the figures have been determined from two complementary strands. Computer analysis of sequence data was performed on a Digital Microvax II using the UWGCG software (Devereux *et al.*, 1984).

### Preparation of antisera against bacterial fusion proteins or synthetic peptides

For preparation of the anti D1, anti-K1, anti-L1, and antiP1 sera HCV-derived cDNA fragments coding for amino acids 102–292 (anti-D1), 2555–2766 (anti-K1), 2766–2985 (anti-L1), and 2290–2555 (anti-P1) of the HCV polyprotein (Meyers *et al.*, 1989a) were subcloned into the expression vector pEX34 using standard procedures (Maniatis *et al.*, 1989). pEX34 is identical to the previously described plasmid pEX31 (Strebel *et al.*, 1986), except for a deleted PstI site in the ampicillin resistance gene; expression and enrichment of bacterial fusion proteins were performed basically as described previously (Strebel *et al.*, 1986). Fusion proteins were further purified by preparative SDS-PAGE and, after electroelution, injected subcutaneously into rabbits with Freund adjuvant (complete for basic immunization, incomplete for booster injections). Preparation of anti-G1 (Rümenapf *et al.*, 1991) and anti-A3 (Thiel *et al.*, 1991) has been reported before.

Peptide HCV-Pep14 was synthesized on a MilliGen 9050 PepSynthesizer (Millipore) by using the Fmoc-polyamide method. For generation of antibodies the peptide was coupled to Key Hole Limpet hemocyanin using glutaraldehyde as crosslinker (Harlow and Lane, 1988). Immunization of rabbits was carried out as described above with an equivalent of 1 mg of peptide per injection.

The antiserum against the peptide corresponding to amino acids 1335–1351 of the BVDV NADL ORF has been described before (Collett *et al.*, 1991).

### Radioimmunoprecipitation and SDS-PAGE

BVDV-infected MDBK cells ( $1.5 \times 10^6$  per 3.5-cm dish) were labeled for 6 hr with 0.5 mCi/ml [ $^{35}$ S]methionine/[ $^{35}$ S]cysteine. Labeling medium contained no cysteine and  $\frac{1}{20}$  of normal methionine content. Cell extracts were prepared under denaturing conditions (Harlow and Lane, 1988). Extracts were incubated with 5  $\mu$ l of undiluted serum. Precipitates were formed with crosslinked *Staphylococcus aureus* (Kessler, 1981), analyzed by SDS-PAGE, and processed for fluorography using Enhance (New England Nuclear, Boston). Competition for anti-peptide antisera was done with 100  $\mu$ g of peptide per RIP sample.

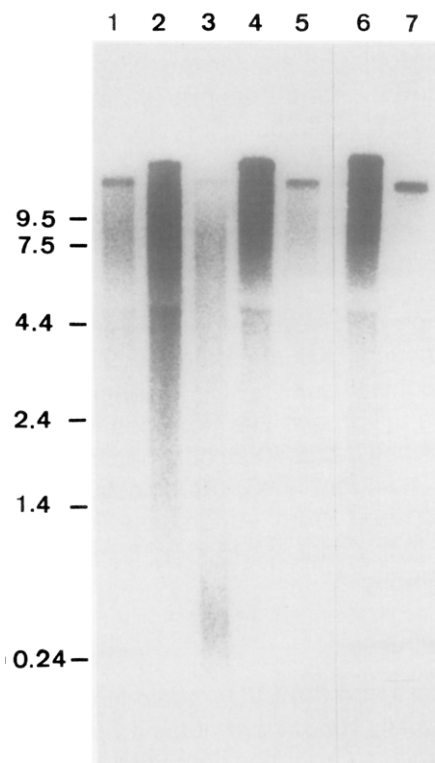


Fig. 1. Northern blot analysis of total RNA from MDBK cells infected with BVDV strains Osloss (lane 1), CP1 (lane 2), NCP1 (lane 3), Pe515CP (lane 4), Pe515NCP (lane 5), CP6 (lane 6), and NCP6 (lane 7), respectively. The blot was hybridized with the cDNA fragment from clone pNCII.1 (Meyers *et al.*, 1991). Osloss represents a cpBVDV strain with a genome containing a ubiquitin-coding insertion of 228 nucleotides (Meyers *et al.*, 1989b).

## RESULTS

### Hybridization studies

Characterization of the first BVDV pair isolated from one animal with MD revealed a striking difference between the genome sizes of the cytopathogenic virus CP1 and the noncytopathogenic virus NCP1 which was clearly detectable on Northern blots (Meyers *et al.*, 1991). The analysis of the viruses belonging to pairs Pe515 (Pe515CP and Pe515NCP) and No. 6 (CP6 and NCP6) was also started with Northern hybridization experiments (Fig. 1). The cDNA insert from clone pNCII.1 (Meyers *et al.*, 1991) was used as a probe which was hybridized against total RNA from bovine kidney (MDBK) cells infected with different virus strains. The migration rate determined for the genome of the cytopathogenic member of each virus pair was considerably slower than that of the noncytopathogenic counterpart (Fig. 1). Among the cpBVDV strains analyzed here CP6 apparently possesses the largest genome while the RNA from Pe515CP has about the same size as that of the CP1 genome (Fig. 1).

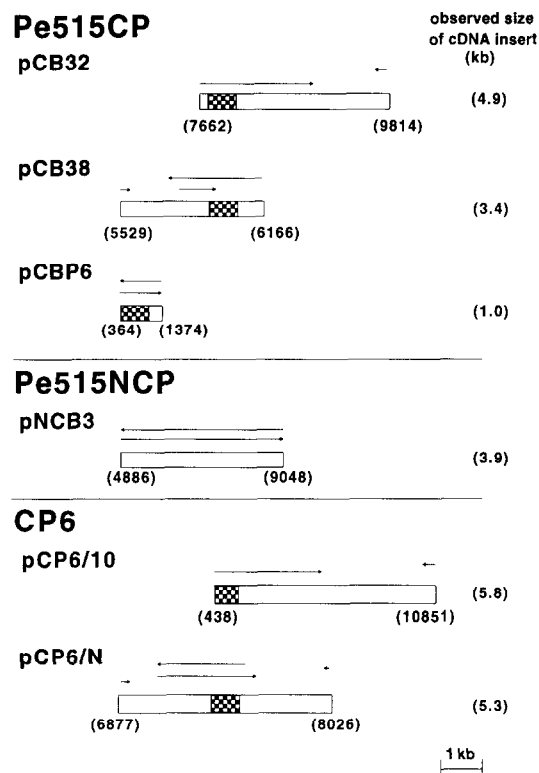


FIG. 2. Schematic drawing indicating the sequencing strategies for the BVDV Pe515CP, Pe515NCP, and CP6 cDNA clones. For each cDNA fragment the regions which have been sequenced are marked by arrows. The SD regions in the sequences from Pe515CP and CP6 are shown as chess board bars. Numbers refer to the corresponding positions of the BVDV NADL genome. The size of the different cDNA inserts determined after agarose gel electrophoresis is shown in the right panel.

According to additional hybridization experiments with a ubiquitin probe the viruses from pairs No. 6 and Pe515 did not contain ubiquitin-coding sequences in their genomes while those from BVDV Osloss and CP1 were clearly recognized (data not shown). Therefore the putative recombination leading to the genomes of BVDV CP6 and Pe515CP did not imply integration of ubiquitin-coding sequences.

#### Analysis of BVDV strains Pe515CP and Pe515NCP

For elucidation of the genome structure of BVDV Pe515CP and Pe515NCP cDNA cloning and sequencing was employed. Since all insertions identified so far in BVDV genomic RNAs were located around nucleotide 5400, construction of the cDNA library was performed with primers specific for the regions at 6 and 7 kb (Meyers *et al.*, 1991). The library screening was done with the same probe as used for the blot in Fig. 1. In order to investigate the genomic localization of the respective cDNA fragment, BVDV-specific cDNA clones were first analyzed by determination of the terminal sequences (Fig. 2). With respect to the published

BVDV genomic sequences the 5' ends of clones pCB32 and pCB38 corresponded to positions around 7662 and 5529, while the 3' ends were homologous to sequences upstream of nucleotides 9814 and 6166 (pCB32 and pCB38, respectively). Accordingly, the expected sizes of the cDNA inserts in clones pCB32 and pCB38 are 2.15 and 0.64 kb, respectively. This finding already indicated the presence of sequence rearrangements in the genome of Pe515CP since the observed length of the cloned cDNA fragments was about 4.9 kb (pCB32) and 3.4 kb (pCB38) (Fig. 2). To determine the genome organization of Pe515CP additional nucleotide sequencing of both pCB32 and pCB38 was employed (indicated by arrows in Fig. 2). The 5' end of the sequence shown in Fig. 3 corresponds to position 7089 of the published BVDV NADL genomic sequence (Collett *et al.*, 1988a). Up to nucleotide 636 the determined Pe515CP sequence is colinear with the published sequences. Then 462 residues follow which do not show any homology to the respective BVDV genomic region. The third part of the sequence corresponds to the region following position 5423 of the BVDV genome (strain NADL) (Fig. 3A). In general this genome organization is reminiscent of the one described for BVDV CP1 (Meyers *et al.*, 1991) with an insertion flanked by a duplicated viral sequence encompassing about 2.5 kb downstream of position 5423 (termed LD) (Fig. 5). For BVDV CP1 the inserted element of 366 residues was a host cell-derived sequence coding for ubiquitin. Surprisingly, the 462 nucleotides present in the genome of Pe515CP (termed SD) are not of cellular origin but are homologous to the sequence located at position 428–889 of the BVDV NADL genome (Figs. 3A and 5). Accordingly, the genome of BVDV Pe515CP represents the product of a recombination process which probably involved only viral sequences.

The differences in genome sizes observed in the Northern blot experiments (Fig. 1) suggested that the duplication and rearrangement of sequences identified for Pe515CP are not present in the RNA of Pe515NCP. To investigate this directly a cDNA library of Pe515NCP was established using a primer complementary to a sequence located at about 9 kb in a standard BVDV genome (Pes9). Priming in this region should provide cDNA clones encompassing the two recombination positions identified for Pe515CP (for nucleotide positions see Fig. 3A). After screening the library with the pNCPII.1 probe clone pNCB3 was isolated which had a size of about 3.9 kb (Fig. 2). Determination of the terminal sequences allowed the positioning of the cDNA insert between nucleotides 4886 and 9048 with respect to the BVDV NADL genome. Fig. 3B shows the complete nucleotide sequence of clone pNCB3. Insertions or discontinuances

<b>A</b>	1	TGGAGTCTATACGAGGAGGACAGCCTGTTAATAACCCAGTTAGAAATATCAAATAATTTA	60
	1	W S L Y E E D S L L I T Q L E I S N N L	21
	61	CTCATATCAGAAGACTTGCCAGCCGCTGTCAAGAACATAATGGCCAGGACTGATCACCCA	120
	21	L I S E D L P A A V K N I M A R T D H P	40
	121	GAGCCGATCCAACTTGCTTACAACAGTTATGAGGTCCAGGTCCAGTCTGTTCCTCAAAA	180
	41	E P I Q L A Y N S Y E V Q V P V L F P K	60
	181	ATAAGGAACGGAGAGGTTACAGACACCTACGAAACTACTCATTCTAAACGCTAGGAAG	240
	61	I R N G E V T D T Y E N Y S F L N A R K	80
	241	TTAGGAGAGGATGTGCCCGTATACATCTATGCCACTGAAGATGAGGACCTGGCAGTTGAC	300
	81	L G E D V P V Y I Y A T E D E D L A V D	100
	301	CTCCTAGGGCTGGACTGGCCCGACCTGGGAATCAGCAGGTCGTGGAGACCAGCAAAGCA	360
	101	L L G L D W P D P G N Q Q V V E T S K A	120
	361	CTGAAGCAAGTGGCCGGATTGTCTTCAGCTGAGAATGCCTTGCTCATAGCTTTATTTGGG	420
	121	L K Q V A G L S S A E N A L L I A L F G	140
	421	TATGTAGGTTACCAAGCTCTATCAAAGAGGCATGTCCCAATGGTCACAGATATATACACT	480
	141	Y V G Y Q A L S K R H V P M V T D I Y T	160
	481	ATAGAGGACCAGAGACTGGAAGACACCCACCTCCAGTATGCACCCAACGCCATAAAA	540
	161	I E D Q R L E D T T H L Q Y A P N A I K	180
		(BVDV-NADL) 7592 GACACCACCCACCTCCAGTATGCACCCAACGCCATAAAA	
	541	ACCGAAGGGACAGAGACTGAATTGAAGGAATTGCCGGCGGGTGACGTGGAGAAAATCATG	600
	181	T E G T E T E L K E L P A G D V E K I M	200
		(BVDV-NADL) 364 AATCTCTGCTGTACATGGCACATGGAG	
		ACCGATGGGACAGAGACTGAACTGAAAGAACTGGCGTGGGTGACGTGGAATAATCATG	
		<b>C' D'</b>	
	601	GGAGCCATCTCGGATTATGCAGCTGGAGGACTGGAA CAAAAACCCGTCGGGGTGGAGGAA	660
	201	G A I S D Y A A G G L E Q K P V G V E E	220
		TTGATCACAATGAACCTTTATACAAAACATACAAA CAAAAACCCGTCGGGGTGGAGGAA	
		GGAGCCATTTAGATTATGCAGCTGGGGACTGGAG TTTGTTAAATCCCAAGCAGAAAAG	
	661	CCTGTTTACGATCAGGCAGGTAATCCTTTATTTGGTGAAAGGGGAGCAATCCACCCTCAA	720
	221	P V Y D Q A G N P L F G E R G A I H P Q	240
		CCTGTTTATGATCAGGCAGGTGATCCCTTATTTGGTGAAAGGGGAGCAGTCCACCCTCAA	
		ATAAAAACAGCTCCTTTGTTTAAAGAAGAA 7777 (BVDV-NADL)	
	721	TCGACGCTAAAGCTCCACACAAGAGAGGGGAACGTGACGTACCCACCGATTGGCTTCT	780
	241	S T L K L P H K R G E R D V P T D L A S	260
		TCGACGCTAAAGCTCCACACAAGAGAGGGGAACGCGATGTTCCAACCAACTTGGCATCC	
	781	TTACCAAAAAGAGGTGATTGCAGGTTTGGTAACAGCAAAGGGCCTGTGAGTGGAATCTAC	840
	261	L P K R G D C R F G N S K G P V S G I Y	280
		TTACCAAAAAGAGGTGACTGCAGGTGCGGTAAATAGCAGAGGACCTGTGAGCGGGATCTAC	
	841	CTGAAGCCAGGGCCACTATTCTTTCAAGACTACAAGGTCCCGTCTATCAGAGCCCCA	900
	281	L K P G P L F F Q D Y K G P V Y H R A P	300
		CTGAAGCCAGGGCCACTATTTTACCAGGACTATAAAGGTCCCGTCTATCAGAGGCCCCG	
	901	CTGGAGCTTTTGGAGAGGGTCTATGTGTGATACGACTAAATGGATAGGGAGAGTAAT	960
	301	L E L F E E G S M C D T T K W I G R V T	320
		CTGGAGCTCTTTGGAGAGGGATCCATGTGTGAAACGACTAAACGGATAGGGAGAGTAAT	
	961	GGCAGTGACGAAAGCTGTACCACATTATGTGTGTATAGATGGATGTATAACAGTGAAG	1020
	321	G S D G K L Y H I Y V C I D G C I T V K	340
		GGAAGTGACGAAAGCTGTACCACATTATGTGTGTATAGATGGATGTATAATAATAAAA	
	1021	AGTGCCACAAGAAGTACCAAAGGGTACTTAGGTGGGTCCACAATAAGCTCAACTGCCCT	1080
	341	S A T R S H Q R V L R W V H N K L N C P	360
		AGTGCCACGAGAAGTTACCAAAGGGTGTTCAGGTGGGTCCATAATAGGCTTGACTGCCCT	
		(BVDV-NADL) 5392 TAATCTGGAACAT	

**FIG. 3.** Nucleotide and deduced amino acid sequences determined from the BVDV Pe515CP cDNA clones pCB32 and pCB38 (A), the Pe515NCP cDNA clone pNCB3 (B), and the Pe515CP PCR clone pCBP6 (C). The residues flanking the Pe515 5' recombination position are marked by the letters C' and D' and those flanking the 3' recombination position by E and B. In (C) also the location of the CP6 recombination positions is indicated (D: 5' recombination position; E: 3' recombination position). To show the homology of the Pe515CP LD and SD regions to different parts of a standard BVDV genome the corresponding sequences from the BVDV NADL genome (numbers as published by Collett *et al.*, 1988a) have been integrated in (A).

		<b>E B</b>	
1081	CTATGGGTCTATGAGCTGC	GGGCCTGCCGTGCACAAGAAGATCACAGAACGTGAAAGATGC	1140
361	L W V M S C	G P A V H K K I T E R E R C	380
	CTATGGGTCACAAGTTGC	TCAGACACGAAAGAAGAGGGAGCAACA	916 (BVDV-NADL)
	CTTGGGTGGATCCTAAGG	GGGCCTGCCGTGTGTAAGAAGATCACAGAGCAC	5455 (BVDV-NADL)
1141	CATGTCAACATACTAGACAAGTTAACTGCATTTTTCGGGATCATGCCGAGGGGGACTACA		1200
381	H V N I L D K L T A F F G I M P R G T T		400
1201	CCCAGAGCCCCAGTGAGGTTCACACGAGTTTGTAAAAAGTGAGGAGGGGTCTGGAGACT		1260
401	P R A P V R F P T S L L K V R R G L E T		420
1261	GGTTGGGCTTATACGCACCAAGGTGGGATAAGTTCGGTCCGACCGTAACCGCCGGCAAA		1320
421	G W A Y T H Q G G I S S V D H V T A G K		440
1321	GATCTATTGGTCTGTGACAGCATGGGAGAACTAGAGTGGTCTGCCAGAGCAACAACAAG		1380
441	D L L V C D S M G R T R V V C Q S N N K		460
1381	TTAACTGATGAGACAGAGTATGGTGTCAAGACTGATTGAGGATGCCAGATGGTGCCAGA		1440
461	L T D E T E Y G V K T D S G C P D G A R		480
1441	TGTTATGTGCTGAATCCAGAGGCAGTCAACATATCAGGATCCAAAGGGGAGTCGTCCAC		1500
481	C Y V L N P E A V N I S G S K G A V V H		500
1501	CTCCAAAAACGGGTGGGGAATTACGTGTGTCACTGCATCAGGCACGCCGGCTTTCTTT		1560
501	L Q K T G G E F T C V T A S G T P A F F		520
1561	GATTAAAAAAGTGAAGGATGGTGGGCCTGCCCATATTTGAAGCCTCCAGTGGAAGG		1620
521	D L K N L K G W S G L P I F E A S S G R		540
1621	GTAGTTGGCAGGGTTAAAGTAGGAAGAATGAGGAGTCTAAACCTACAAAGATAATGAGC		1680
541	V V G R V K V G K N E E S K P T K I M S		560
1681	GGTATCCAGACCGTCTCAAAGAACACAGCAGATCTTACTGAAATGGTCAAGAAGATCACC		1740
561	G I Q T V S K N T A D L T E M V K K I T		580
<b>B</b>	CCAAATGCGGACGCCATGGGAAACCCATAACATGTGGGATGTCATTAGCAGATTTTGAA		60
1	P K C G R H G K P I T C G M S L A D F E		20
61	GAAAGGCATTATAAAGGATCTTTATAAGGGAAGGCAACTTTGATGGCCCTTTCAGACAA		120
21	E R H Y K R I F I R E G N F D G P F R Q		40
121	GAATATAATGGTTTATACAACTATACTGCTAGGGGGCAGTTATTTTAAAGAAACCTGCC		180
41	E Y N G F I Q Y T A R G Q L F L R N L P		60
181	ATACTGGCAACTAAAGTAAATGCTCATGGTAGGCAACCTCGGAGAAGAAATCGGGGAC		240
61	I L A T K V K M L M V G N L G E E I G D		80
241	CTAGAACACCTAGGGTGGATCCTAAGG	GGGCCTGCCGTGTGCAAGAAGATCACAGAACAT	300
81	L E H L G W I L R	G P A V C K K I T E H	100
301	GAAAGATGCCACGTCAACATACTAGACAAGTTAACTGCATTTTTCGGGATCATGCCGAGG		360
101	E R C H V N I L D K L T A F F G I M P R		120
361	GGGACTACACCCAGAGCCCCAGTGAGGTTCCCTACGAGTTTGTAAAAAGTGAGGAGGGGT		420
121	G T T P R A P V R F P T S L L K V R R G		140
421	CTGGAGACTGGTTGGGCCTATACGCACCAAGGTGGGATAAGTTCGGTCCGACCATGTAACC		480
141	L E T G W A Y T H Q G G I S S V D H V T		160
481	GCCGGCAAAGATCTATTGGTCTGTGACAGCATGGGAGAACTAGAGTGGTCTGCCAGAGC		540
161	A G K D L L V C D S M G R T R V V C Q S		180
541	AACAACAAGTTAACAGATGAGACAGAGTATGGTGTCAAGACTGATTGAGGATGCCAGAT		600
181	N N K L T D E T E Y G V K T D S G C P D		200
601	GGTGCCAGATGTTATGTGCTGAATCCAGAGGCAGTCAACATATCAGGATCCAAAGGGGCA		660
201	G A R C Y V L N P E A V N I S G S K G A		220
661	GTCGTCCACCTCCAAAAACGGGTGGGGAATTACGTGTGTCACTGCATCAGGCACGCCG		720
221	V V H L Q K T G G E F T C V T A S G T P		240

FIG. 3—Continued

721	GCTTTCTTTGATTAAAAAAGGATGGTCGGGCCTGCCCATATTTGAAGCCTCC	780
241	A F F D L K N L K G W S G L P I F E A S	260
781	AGTGAAGGGTAGTTGGCAGGGTTAAAGTAGGGAAGATGAGGAGTCTAAACCTACAAAG	840
261	S G R V V G R V K V G K N E E S K P T K	280
841	ATAATGAGCGGTATCCAGACCGTCTCAAAGAACACAGCAGATCTTACTGAAATGGTCAAG	900
281	I M S G I Q T V S K N T A D L T E M V K	300
901	AAGATCACCAGTATGAACAGGGGAGACTTCAAGCAGATAACTTTAGCAACAGGGGCAGGA	960
301	K I T S M N R G D F K Q I T L A T G A G	320
961	AAAACTACAGAACTTCCAAAAGCAGTGATAGAGGAGATAGGAAGACACAAGCGAGTGTTA	1020
321	K T T E L P K A V I E E I G R H K R V L	340
1021	GTCCTTATACCATTTAGGGCAGCGGCAGAGTCAGTCTACCAGTACATGAGATTGAAACAC	1080
341	V L I P L R A A A E S V Y Q Y M R L K H	360
1081	CCGAGTATATCCTTTAACCTAAGGATAGGGGACATGAAAGAAGGAGACATGGCAACAGGG	1140
361	P S I S F N L R I G D M K E G D M A T G	380
1141	ATAACTTATGCATCATACGGGTACTTCTGCCAGATGCCTCAACCAAAGCTTAGAGCAGCT	1200
381	I T Y A S Y G Y F C Q M P Q P K L R A A	400
1201	ATGGTGGAATACTCTTACATATTTTATGATGAATACCATTGCGCCACTCTGAACAACCTG	1260
401	M V E Y S Y I F L D E Y H C A T P E Q L	420
1261	GCAATTATCGGAAAAATCCACAGATTTTCAGAGAGCATAAGAGTGGTTGCCATGACTGCC	1320
421	A I I G K I H R F S E S I R V V A M T A	440
1321	ACCCGAGCAGGGTCAGTAACCAACCGGGACAAAAACCCCAATAGAGGAATTCATAGCC	1380
441	T P A G S V T T T G Q K H P I E E F I A	460
1381	CCAGAGGTAATGAAAGGGGAAGACCTAGGCAGCCAGTTCTCGACATAGCGGGGCTAAAA	1440
461	P E V M K G E D L G S Q F L D I A G L K	480
1441	ATCCCGGTGGATGAGATGAAAGGCAACATGCTGGTTTTGTACCCACGAGAAACATGGCG	1500
481	I P V D E M K G N M L V F V P T R N M A	500
1501	GTTGAGGTGGCGAAGAAGTTGAAAGCTAAGGGCTACAATTCTGGGTACTACTACAGTGA	1560
501	V E V A K K L K A K G Y N S G Y Y Y S G	520
1561	GAGGATCCAGCTAATCTGAGAGTTGTAACATCACAGTCCCCCTATGTAATTGTGGCAACG	1620
521	E D P A N L R V V T S Q S P Y V I V A T	540
1621	AATGCCATTGAATCAGGAGTGACATTACCAGATCTGGACACGGTCGTGGACACAGGGTTG	1680
541	N A I E S G V T L P D L D T V V D T G L	560
1681	AAATGTGAAAAGAGGGTGAGGGTATCATCAAGATACCTTCATTGTAACAGGCCCTTAAG	1740
561	K C E K R V R V S S K I P F I V T G L K	580
1741	AGGATGGCGCTCACTGTGGCGAACAGGCTCAACGTAGAGGCAGAGTAGGTAGAGTGAAG	1800
581	R M A V T V G E Q A Q R R G R V G R V K	600
1801	CCCGGAGATATTACAGGAGCCAGGAAACAGCAACCGGGTCAAAGGACTACCACTATGAC	1860
601	P G R Y Y R S Q E T A T G S K D Y H Y D	620
1861	CTCTGACGGCACAAAGATACGGCATTGAGGACGGAATCAATGTAACAAAATCCTTTAGG	1920
621	L L Q A Q R Y G I E D G I N V T K S F R	640
1921	GAAATGAATTATGATTGGAGTCTATACGAGGAGGACAGCCTGTTAATAACCCAGTTAGAA	1980
641	E M N Y D W S L Y E E D S L L I T Q L E	660
1981	ATATTAATAATTACTCATATCAGAAGACTTGCCAGCGCTGTCAAGAACATAATGGCC	2040
661	I L N N L L I S E D L P A A V K N I M A	680
2041	AGGACTGATACCCAGAGCCGATCCAACCTTGCTTACAACAGTTATGAGGTCCAGGTCCCA	2100
681	R T D H P E P I Q L A Y N S Y E V Q V P	700
2101	GTCCTGTTCCCAAAAATAAGGAACGGAGAGGTCACAGACACCTACGAAAACCTACTCATT	2160
701	V L F P K I R N G E V T D T Y E N Y S F	720

Fig. 3—Continued



2161	CTAAACGCTAGGAAGTTAGGAGAGGATGTGCCCGTATACATCTATGCCACTGAAGATGAG	2220
721	L N A R K L G E D V P V Y I Y A T E D E	740
2221	GACCTGGCAGTTGACCTCCTAGGGCTGGACTGGCCCCGACCCTGGGAATCAGCAGGTCGTG	2280
741	D L A V D L L G L D W P D P G N Q Q V V	760
2281	GAGACCAGCAAGCACTGAAGCAAGTGGCCGGATTGTCTTCAGCTGAGAATGCCTTGCTC	2340
761	E T S K A L K Q V A G L S S A E N A L L	780
2341	ATAGCTTTATTTGGGTATGTAGGTTACCAAGCTCTATCAAAGAGGCATGTCCCAATGATC	2400
781	I A L F G Y V G Y Q A L S K R H V P M I	800
2401	ACAGATATATACACTATAGAGGACCAGAGACTAGAAGACACCACCACCTCCAGTATGCA	2460
801	T D I Y T I E D Q R L E D T T H L Q Y A	820
2461	CCCAACGCCATAAAAAACCGAAGGGACAGAGACTGAATTGAAGGAATTGGCAGCGGGTGAC	2520
821	P N A I K T E G T E T E L K E L A A G D	840
2521	GTGGAGAAAATCATGGGAGCCATCTCGGATTATGTAGCTGGAGGACTGGAA TTCGTAAAA	2580
841	V E K I M G A I S D Y V A G G L E F V K	860
2581	TCTCAAGCAGAAAAAGTGAAGACAGCCCTTTGTTTAAAGGAAAACGTAGAAGCTGCAAGG	2640
861	S Q A E K V K T A P L F K E N V E A A R	880
2641	GGGTACGTCCAAAACTAATTGACTCATTAATTGAAGAAAAAGATGTAATAATCAGATAT	2700
881	G Y V Q K L I D S L I E E K D V I I R Y	900
2701	GGCTTGTGGGAACACATACAGCACTCTATAAAGCATAGCTGCAAGACTGGGGCATGAA	2760
901	G L W G T H T A L Y K S I A A R L G H E	920
2761	ACAGCATTTGCTACACTAGTAATAAAATGGCTAGCCTTCGGAGGGGAGTCGGTGTGAGAC	2820
921	T A F A T L V I K W L A F G G E S V S D	940
2821	CATGTCAGACAGGCGGCTGTTGATTAGTGGTCTATTATGTGATGAACAAGCCTTCCTTC	2880
941	H V R Q A A V D L V V Y Y V M N K P S F	960
2881	CCAGGTGATTCTGAAACCCAGCAAGAAGGGAGGCGGTTTGTGCAAGCCTATTTATCTCC	2940
961	P G D S E T Q Q E G R R F V A S L F I S	980
2941	GCATTGGCAACTTACACATACAAAACCTGGAATTATCACATCTCTCAAAAGTGGTGAA	3000
981	A L A T Y T Y K T W N Y H N L S K V V E	1000
3001	CCAGCCCTGGCATACCTCCCTTATGCCACCAGCGCATTAATAATGTTCAACCCCAACGAGA	3060
1001	P A L A Y L P Y A T S A L K M F T P T R	1020
3061	CTGGAAAGCGTGGTGATACTGAGCACTACAATATACAAAACATACCTCTCTATAAGGAAA	3120
1021	L E S V V I L S T T I Y K T Y L S I R K	1040
3121	GGGAAGAGTGATGGATTGCTGGGTACGGGGATCAGTGCAGCCATGGAGATTCTGTCACAA	3180
1041	G K S D G L L G T G I S A A M E I L S Q	1060
3181	AACCCGTATCGGTAGGTATATCCGTGATGCTGGGGGTAGGTGCTATCGCTGCGCATAAC	3240
1061	N P V S V G I S V M L G V G A I A A H N	1080
3241	GCCATTGAATCCAGCGAGCAGAAAAGGACCCTACTTATGAAGGTATTTGTAAAGAATTTTC	3300
1081	A I E S S E Q K R T L L M K V F V K N F	1100
3301	TTGGATCAGGCTGCAACGGATGAGCTGGTTAAAGAAAAATCCAGAAAAGATCATAATGGCT	3360
1101	L D Q A A T D E L V K E N P E K I I M A	1120
3361	TTATTTGAAGCAGTCCAGACAATTGGCAACCCCTTAAGGTTGATATACCATCTATATGGA	3420
1121	L F E A V Q T I G N P L R L I Y H L Y G	1140
3421	GTTTACTACAAGGGCTGGGAGGCCAAGGAGCTGTCTGAGAGAACGGCAGGTAGAACTTA	3480
1141	V Y Y K G W E A K E L S E R T A G R N L	1160
3481	TTCACACTGATAATGTTTGGGCTTCGAATTGTTAGGATGGATTGAGAAGGAAAAATA	3540
1161	F T L I M F E A F E L L G M D S E G K I	1180
3541	AGGAACCTATCTGGAACTACATCTTAGATTGATCTATGGATTACACAGGCAGATCAAC	3600
1181	R N L S G N Y I L D L I Y G L H R Q I N	1200

FIG. 3—Continued

3601 AGAGGGCTAAAGAAGATAGTACTAGGGTGGGCCCCGCGCCCTTCAGCTGTGACTGGACT 3660  
 1201 R G L K K I V L G W A P A P F S C D W T 1220  
 3661 CCTGGCGATGAGAGAATCAGGTTGCCAACGGACAACCTATTTAAGAGTGGAAACCAGGTGC 3720  
 1221 P G D E R I R L P T D N Y L R V E T R C 1240  
 3721 CCGTGTGGCTACCAGATGAAAGCTATCAAAAATGTAGGTGGTGAGCTTACCAAGGTGGAG 3780  
 1241 P C G Y Q M K A I K N V G G E L T K V E 1260  
 3781 GAGAGAGGACCTTTCTATGTAGAAACAGACCCGGTAGGGGACCACTCAACTACAGAGTT 3840  
 1261 E R G P F L C R N R P G R G P V N Y R V 1280  
 3841 ACCAAGTACTATGATGACAACTCAATGAGATAAAGCCAGTGGCAAAGTTAGA 3893  
 1281 T K Y Y D D K L N E I K P V A K L 1297  
 D  
 C 1 AAATCTCTGCTGTACATGG CACATGGAGTTGATTGCAATGAACCTTTATACAAAACATA 60  
 1 M E L I A N E L L Y K T Y 13  
 P  
 61 CAAA CAAAACCCGTCGGGGTGGAGGAACCTGTTTACGATCAGGCAGGTAATCCTTTATT 120  
 14 K Q K P V G V E E P V Y D Q A G N P L F 33  
 121 TGGTGAAGGGGAGCAATCCACCCTCAATCGACGCTAAAGCTCCACACAAGAGAGGGGA 180  
 34 G E R G A I H P Q S T L K L P H K R G E 53  
 181 GCGTGACGTACCCACCAATTTGGCTTCTTTACCAAAAAGAGGTGACTGCAGGTCTGGTAA 240  
 54 R D V P T N L A S L P K R G D C R S G N 73  
 241 CAGCAAAGGGCCTGTGAGTGAATCTACCTGAAGCCAGGGCCACTATTCTTTCAAGACTA 300  
 74 S K G P V S G I Y L K P G P L F F Q D Y 93  
 301 CAAGGGTCCCGTCTACCCACAGAGCCCACTGGAGCTTTTTGAGAGGGGTTTATGTGTGA 360  
 94 K G P V Y H R A P L E L F E E G F M C E 113  
 361 AACAACTAAACGGATAGGGAGAGTAAGTGGCAGTGACGGAAGCTGTACCACATTTATGT 420  
 114 T T K R I G R V T G S D G K L Y H I Y V 133  
 421 GTGTATAGATGGATGTATAACAGTGAAGAGTGCCACAAGAAGTCACCAAGGGTACTTAG 480  
 134 C I D G C I T V K S A T R S H Q R V L R 153  
 E  
 481 GTGGGTCCACAATAGGCTCAACTGCCCTCTATGGGTCACAAGCTGC TCAGATACAAAAGA 540  
 154 W V H N R L N C P L W V T S C S D T K E 173  
 541 AGAAGGGGCAACAAAAAGAAACAACAAAACCCGACAGACTGGAGAAGGGGAGGATGAA 600  
 174 E G A T K K K Q Q K P D R L E K G R M K 193  
 601 AATAGTGCCCAAGAGTCTGAGAAAGACAGCAGGACTAAACCCCGGATGCTACGATAGT 660  
 194 I V P K E S E K D S R T K P P D A T I V 213  
 661 GGTAGATGGAGTGAATACCAGGTAAAGAAGAAGGGGAAAGTCAAGAGCAAAAGCACACA 720  
 214 V D G V K Y Q V K K K G K V K S K S T Q 233  
 721 GGACGGTTTATATCACAAATAAAATAAGCCGCCAGAATCAGCAAGAACTTGAGAAAGC 780  
 234 D G L Y H N K N K P P E S R K K L E K A 253  
 781 ATTATTGGCATGGGCAATATTGGCTATGGTCTTGATTCAAGTTACAATGGGCGAAAATAT 840  
 254 L L A W A I L A M V L I Q V T M G E N I 273  
 841 AACACAGTGGAACTTACAAGATAATGGGACGGAAGGGATACAACGGGCAATGTTCCAAAG 900  
 274 T Q W N L Q D N G T E G I Q R A M F Q R 293  
 901 GGGGGTGAACAGAAGTCTACACGGCATCTGGCCGAGAAAGTCTGTACAGGTGTCCCTTC 960  
 294 G V N R S L H G I W P E K V C T G V P S 313  
 961 CCATCTAGCCACCGATGTGGAACATAAAACAATACACGGTATGATGGATGC 1011

FIG. 3—Continued

TABLE 1

HOMOLOGY BETWEEN THE REGIONS FLANKING THE SD ELEMENT IN THE BVDV Pe515CP NUCLEOTIDE SEQUENCE (NUCLEOTIDES 1 TO 636 AND 1099 TO 1740 IN FIG. 3A) AND THE CORRESPONDING SEQUENCES FROM OTHER BVDV GENOMES

	Pe515NCP	Osloss	NADL	NCP1	CP1	CP6
Pe515CP	99%	82%	89%	92%	91%	83%

with respect to a standard pestivirus genome were not observed. Thus the cytopathogenic and the noncytopathogenic viruses of BVDV pair Pe515 differ dramatically in their genome organization (Fig. 5).

Comparison of the sequences determined for the regions upstream of position 636 and downstream of nucleotide 1098 in the Pe515CP cDNA (numbers refer to Fig. 3A) with the respective parts of the noncpBVDV sequence revealed 99% identity (Table 1). Since this value is much higher than the homology to any of the other analyzed BVDV strains this result supports the hypothesis that the cytopathogenic viruses represent mutants of noncp viruses generated by a recombination process.

To investigate whether a translocation or duplication is responsible for the presence of the SD element in the analyzed part of the Pe515CP genome the 5' region of the viral RNA was looked at. We performed cDNA synthesis with an oligonucleotide primer which is complementary to nucleotides 1355 to 1374 of the published BVDV NADL sequence (OIBVD33). The single-stranded cDNA was used as a template for PCR amplification; OIBVD33 and OIBVD32, the latter corresponding to position 364 to 388 of the NADL genome, served as primers. After cloning the resulting PCR fragment into pBluescript the nucleotide sequence of the resulting cDNA fragment was determined (Fig. 3C) and found to be colinear with the respective region of the BVDV NADL genome. The nucleotide sequence homology between the PCR fragment derived from the 5' end of the genome and the SD element identified in the cDNA clones pCB32 and pCB38 (Fig. 3A) is 97%; the deduced amino acid sequence is 95% identical (Table 2). Accordingly, the SD element is itself present in duplicate in the BVDV Pe515CP genome. Because of the high homology between each set of duplicated sequences (Tables 1 and 2) it can be concluded that they all originated from RNA molecules derived from just one BVDV strain, namely Pe515NCP.

### The genome of BVDV CP6

The Northern blot analysis indicated that the CP6 genomic RNA is larger than that of any pestivirus ana-

lyzed so far (Fig. 1). On the account of the molecular characterization of BVDV CP1 and Pe515CP it was expected that a duplication of viral sequences was at least in part responsible for this increase in genome size. Following the strategy outlined above cDNA cloning was performed. The terminal sequences of the insert from plasmid pCP6/10 corresponded to positions 438 and 10851 of the BVDV NADL sequence and those from plasmid pCP6/N to 6877 and 8026 (Fig. 2). Since the inserts of the two cDNA clones had sizes of 5.8 and 5.3 kb (pCP6/10 and pCP6/N, respectively) this finding strongly indicated rearranged viral sequences. For further investigation, part of pCP6/10 and pCP6/N was sequenced (Figs. 2 and 4). The first 1188 determined nucleotides are colinear with the BVDV NADL sequence from 7869 to 9057. The following 507 residues again represent a sequence derived from the 5' part of a standard pestivirus genome. The 5' end of this element corresponds to position 383 of the NADL genome, thereby starting 3 nucleotides upstream of the viral translational initiation codon. Interestingly, the 3' end of this inserted element is exactly the same as that of the Pe515CP SD sequence (corresponds to nucleotide 889 in the NADL genome) (Figs. 3A, 4, and 5). For both viruses the region downstream of this 5'-derived element is homologous to the sequence starting with nucleotide 5423 of the NADL genome.

For Pe515CP the large duplicated element encompasses 2304 nucleotides thus being 81 residues shorter than that of BVDV CP1. The respective element in the CP6 genome, however, has a size of 3635 nucleotides since it corresponds to nucleotides 5423 to 9057 of the BVDV NADL genome. Because of this large duplication together with the SD element of 507 nucleotides the CP6 genome is 4142 nucleotides longer than that of a noncytopathogenic BVDV strain.

### Recombination-induced changes at the protein level

As a marker characteristic for infection with cpBVDV a protein of 80 kDa (p80) has been described (Purchio

TABLE 2

NUCLEOTIDE SEQUENCE HOMOLOGY BETWEEN THE DOWNSTREAM SD ELEMENT (CLONES pCB32/38, NUCLEOTIDES 637 TO 1098 IN FIG. 3A) AND THE P20-CODING REGION (CLONE pCBP6, FIG. 3C) IN THE GENOME OF BVDV Pe515CP WITH RESPECT TO THE P20-CODING REGIONS OF THE BVDV OSLOSS AND NADL OPEN READING FRAMES AND THE DOWNSTREAM SD ELEMENT OF BVDV CP6 (NUCLEOTIDES 1189 TO 1696 IN FIG. 4)

	Pe515CP pCBP6	Osloss	NADL	CP6
Pe515CP pCB32/38	97%	80%	89%	79%
Pe515CP pCBP6	100%	81%	90%	79%

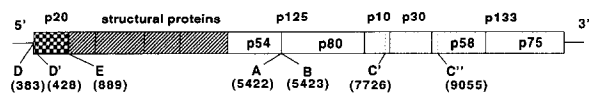
1	ACACACACGGCACTTTACAAGAGTATTGCCGCAAGATTGGGACATGAAACAGCATTTCGCT	60
1	T H T A L Y K S I A A R L G H E T A F A	20
61	ACGCTAGTGGTGAAGTGGCTAGCCTTCGGAGGTGAGTCGGTGTGATCATGTGAGACAG	120
21	T L V V K W L A F G G E S V S D H V R Q	40
121	GCGGCCGTTGACCTGGTCGTTTATTATGTGATGAATAAACCTTCCTTCCCTGGAGATTCC	180
41	A A V D L V V Y Y V M N K P S F P G D S	60
181	GAAACCCACAGGAGGGAAGGCGATTTCGTCGCCAGCTTATTATCTCCGCTCTGGCAACC	240
61	E T Q Q E G R R F V A S L F I S A L A T	80
241	TACACATACAAGACTTGAATTACCAACAACCTCTCCAAGGTAGTAGAACAGCCTTGGCA	300
81	Y T Y K T W N Y H N L S K V V E P A L A	100
301	TACCTCCCTACGCTACCACTGCACTAAAAATGTTACCCCACTAGGCTGGAGAGCGTA	360
101	Y L P Y A T S A L K M F T P T R L E S V	120
361	GTCATACTTAGCACCACAATATACAAAACCTTACCTCTCAATAAGGAAAGGAAAAAGTGAT	420
121	V I L S T T I Y K T Y L S I R K G K S D	140
421	GGACTTCTGGGTACAGGGATTAGTGCAGCAATGGAGATTCTGTACAGAAACCGGTGTCA	480
141	G L L G T G I S A A M E I L S Q N P V S	160
481	GTAGGTATATCTGTCTGCTGCTGGGGGTGGGGCAATTGCCGCTCACAATGCAATCGAGTCT	540
161	V G I S V M L G V G A I A A H N A I E S	180
541	AGTGAGCAAAAAGGACCCCTATTGATGAAAGTGTTTGTAAAAAACTTCCTGGACAGGCG	600
181	S E Q K R T L L M K V F V K N F L D Q A	200
601	GCAACAGATGAGTTGGTGAAGGAAAATCCAGAGAAAATAATAATGGCCTTATTGAAGCA	660
201	A T D E L V K E N P E K I I M A L F E A	220
661	GTACAGACAATTGGCAACCCATTGAGGCTTATATATCATCTGTATGGGGTTTACTACAAA	720
221	V Q T I G N P L R L I Y H L Y G V Y Y K	240
721	GGCTGGGAAGCGAAAGAATATCCGAGAGGACAGCTGGCAGGAACCTGTTCACTTTGATA	780
241	G W E A K E L S E R T A G R N L F T L I	260
781	ATGTTGGAAGCTTTTCGAATATTAGGGATGGACTCTGAAGGGAAGATAAGGAACCTATCT	840
261	M F E A F E L L G M D S E G K I R N L S	280
841	GGGAATTACATCTTGGATTGTATCTACAGTTTACATAAACAGATAAACAGAGGCTTGAAG	900
281	G N Y I L D L I Y S L H K Q I N R G L K	300
901	AAGATAGTCTTGGGGTGGGCTCCTGCACCCTTCAGTTGTGACTGGACTCCTAGTGATGGG	960
301	K I V L G W A P A P F S C D W T P S D G	320
961	AGAATCAGGTTACCCACAGACAATATCTGAGAGTAGAGACTAAGTGCCCATGTGGTTAT	1020
321	R I R L P T D N Y L R V E T K C P C G Y	340
1021	GAGATGAAAGCACTAAGGAATTTTCAGTGGCAGTCTCACAAGTGAAGAGAAAGGGCCT	1080
341	E M K A L R N F S G S L T K V E E K G P	360
1081	TTCTCTGTAGGAACGGGCCTGGTAGGGGGCCAGTTAACTATAGAGTCACAAAGTACTAT	1140
361	F L C R N G P G R G P V N Y R V T K Y Y	380
9004 (BVDV-NADL) TTAC		
C <sup>+</sup> D		
1141	GATGACAACCTCAGAGAGATAAAACAGTTGCAAACTAGAAGGACTG CACATGGAGTTG	1200
381	D D N L T E I K P V A K L E G L H M E L	400
360 (BVDV-NADL) CTAATAATCTCTGCTGTACATGG CACATGGAGTTG		
GATGACAACCTCAGAGAGATAAAACAGTAGCAAGTTGGAAGGACAG GTAGAGCACTAC		
1201	AACACAAATGAACTTTATACAAAACATACAAAACAAACCCGCTGGAGTGGAGGAACCA	1260
401	N T N E L L Y K T Y K Q K P A G V E E P	420
ATCACAATGAACTTTATACAAAACATACAAAACAAACCCGTCGGGGTGGAGGAACCT		
TACAAAGGGGTACAGCAAAAATTGACTACAGTAA 9103 (BVDV-NADL)		

FIG. 4. Nucleotide and deduced amino acid sequence of part of the BVDV CP6 genome. The 5' and 3' recombination positions are marked by the letters C<sup>+</sup>/D and E/B, respectively. Corresponding parts of the BVDV NADL nucleotide sequence are shown in the lines below the CP6 amino acid sequence.

1261	GTATATGACCAAGCAGGTAACCCCTTTGTTTGGCGAGAGGGGAGTGGTTCATCCGCAGGCA	1320
421	V Y D Q A G N P L F G E R G V V H P Q A	440
	GTATATGATCAGGCAGGTGATCCCTTATTTGGTGAAAGGGGAGCAGTCCACCCTCAATCG	
1321	ACGCTAAAACTGCCACATAAAAGAGGGGAGAGCGAAGTCCCTACTGATCTGGCATCCTTG	1380
441	T L K L P H K R G E S E V P T D L A S L	460
	ACGCTAAAGCTCCACACAAGAGAGGGGAACGCGATGTTCCAACCAACTTGGCATCCTTA	
1381	CCAAAAAGAGGTGACTGCATGTCAGGTAACAGCAAAGGGCCCGTGAGTGGGATCTATCTG	1440
461	P K R G D C M S G N S K G P V S G I Y L	480
	CCAAAAAGAGGTGACTGCAGGTGCGGTAATAGCAGAGGACCTGTGAGCGGGATCTACCTG	
1441	AAACCGGGGCGTTATTCTACAGGATTACAAGGGACCCGTCTATCATAGAGCTCCATTG	1500
481	K P G P L F Y Q D Y K G P V Y H R A P L	500
	AAGCCAGGGCCACTATTTTACAGGACTATAAAGGTCCCGTCTATCACAGGGCCCCGCTG	
1501	GAGTTATTTAGGAAACCTCTATGTGCGAGACAAGTAGATGGATTGGGAGAGTAACGGT	1560
501	E L F Q E T S M C E T T R W I G R V T G	520
	GAGCTCTTTGAGGAGGATCCATGTGTGAACGACTAAACGGATAGGGAGAGTAACGGA	
1561	AGTGATAGCAAATGTACCATCTATGTGTGCATAGATGGATGCATAATAATTAAGAGT	1620
521	S D S K L Y H I Y V C I D G C I I I K S	540
	AGTGACGGAAGCTGTACCACTTTATGTGTGTATAGATGGATGTATAATAATAAAAGT	
1621	GCCACAAAAGATCACAAAAGGTATTCAAATGGGTCCATAACAACTAAATTGGCCCTTTA	1680
541	A T K D H Q K V F K W V H N K L N C P L	560
	GCCACGAGAAGTTACCAAAGGGTGTTCAGGTGGGTCCATAATAGGCTTGACTGCCCTCTA	
	5392 (BVDV-NADL) TAATCTGGAACATCTT	
<b>E B</b>		
1681	TGGGTTACAAGCTGC GGGCCTGCCGTGTGCAAGAAAATCACTGAGCACGAAAAATGCCAT	1740
561	W V T S C G P A V C K K I T E H E K C H	580
	TGGGTCAAGTTGC TCAGACACGAAAGAAGAGGGAGCAACA 916 (BVDV-NADL)	
	GGGTGGATCCTAAGG GGGCCTGCCGTGTGTAAGAAGATCACAGAGCAC 5455 (BVDV-NADL)	
1741	GTCAGCATACTAGACAAACTGACTGCGTTTTTCGGGGTTATGCCAAGAGGAACACACCA	1800
581	V S I L D K L T A F F G V M P R G T T P	600
1801	AGGGCTCCGGTGAGATTCCCAACAGCATTGCTAAAGGTGAGGAGGGGACTGGAACCGGT	1860
601	R A P V R F P T A L L K V R R G L E T G	620
1861	TGGGCTTACACATCAAGGTGGTATAAGCTCAGTAGACCATGTGACCGCAGGTAAAGAT	1920
621	W A Y T H Q G G I S S V D H V T A G K D	640
1921	CTATTGGTCTGTGATAGTATGGGTAGGACTAGAGTGGTTTGCCAAAGCAACAACAAGTTA	1980
641	L L V C D S M G R T R V V C Q S N N K L	660
1981	ACTGATGAGACAGAATATGGTGTCAAGACGGACTCCGGATGTCCAGATGGAGCCAGGTGC	2040
661	T D E T E Y G V K T D S G C P D G A R C	680
2041	TATGTATTAATCCAGAGGCAGTAAATATATCAGGGTCCAAGGGAGCTGTCGTACACCTC	2100
681	Y V L N P E A V N I S G S K G A V V H L	700
2101	CAAAAAACAGGAGGGGAATTTACATGTGTCACTGCATCAGGCACTCCGGCCTTCTTTAAC	2160
701	Q K T G G E F T C V T A S G T P A F F N	720
2161	CTAAAAAATTTAAAGGGATGGTGGGTCTACCCATATATGAGGCTTCCAGCGGCAGAGTG	2220
721	L K N L K G W S G L P I Y E A S S G R V	740
2221	GTCGGCAGAGTTAAAGTAGGAAAAAATGAGGAGTCCAAACCCACAAAATTAATGAGTGGT	2280
741	V G R V K V G K N E E S K P T K L M S G	760
2281	ATACAAACCGTCTCAAAAAGCACGGCCGACTTAACAGAAATGGTCAAGAAGATAACACGC	2340
761	I Q T V S K S T A D L T E M V K K I T S	780
2341	ATGAACAGGGGAGACTTTAAGCAGATAACCCCTTGCAACAGGGGAGGAAACTACAGAA	2400
781	M N R G D F K Q I T L A T G A G K T T E	800
2401	CTCCAAAGGCAGTGACAGAGGAGATAGGACGACACAAGCGGGTGTAGTACTTATACCA	2460
801	L P K A V T E E I G R H K R V L V L I P	820
2461	TTGAGAGCAGCAGCTGAGTC	2480
821	L R A A A E	826

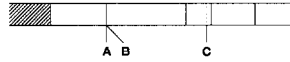
FIG. 4—Continued

### Pestivirus genome

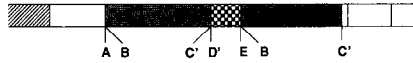


### Pair Pe515

NCP

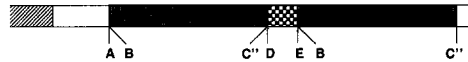


CP



### Pair #6

CP



**Fig. 5.** Schematic representation of the genome organization of the analyzed BVDV isolates. The upper bar shows a scheme of a complete pestivirus genome with indication of the various sites involved in the recombination reactions identified for the cpBVDV genomes shown below (letters below the bar). The numbers give the positions of these sites with respect to the published BVDV NADL sequence. p20-coding sequences are indicated by chess board bars, while streaked bars represent the region coding for the structural proteins. The LD regions duplicated in the genomes of Pe515CP and CP6 are shown as dotted bars. (A) Last residue preceding the position of the ubiquitin insertion in the BVDV Osloss genome. (B) First nucleotide of all LD regions and also first residue following the ubiquitin insertion in the BVDV Osloss genome. (C') Last residue of the Pe515CP LD region. (C'') Last residue of the CP6 LD region. (D) First residue of the CP6 SD region. (D') First residue of the Pe515CP SD region. (E) Last residue of both the CP6 and Pe515CP SD regions.

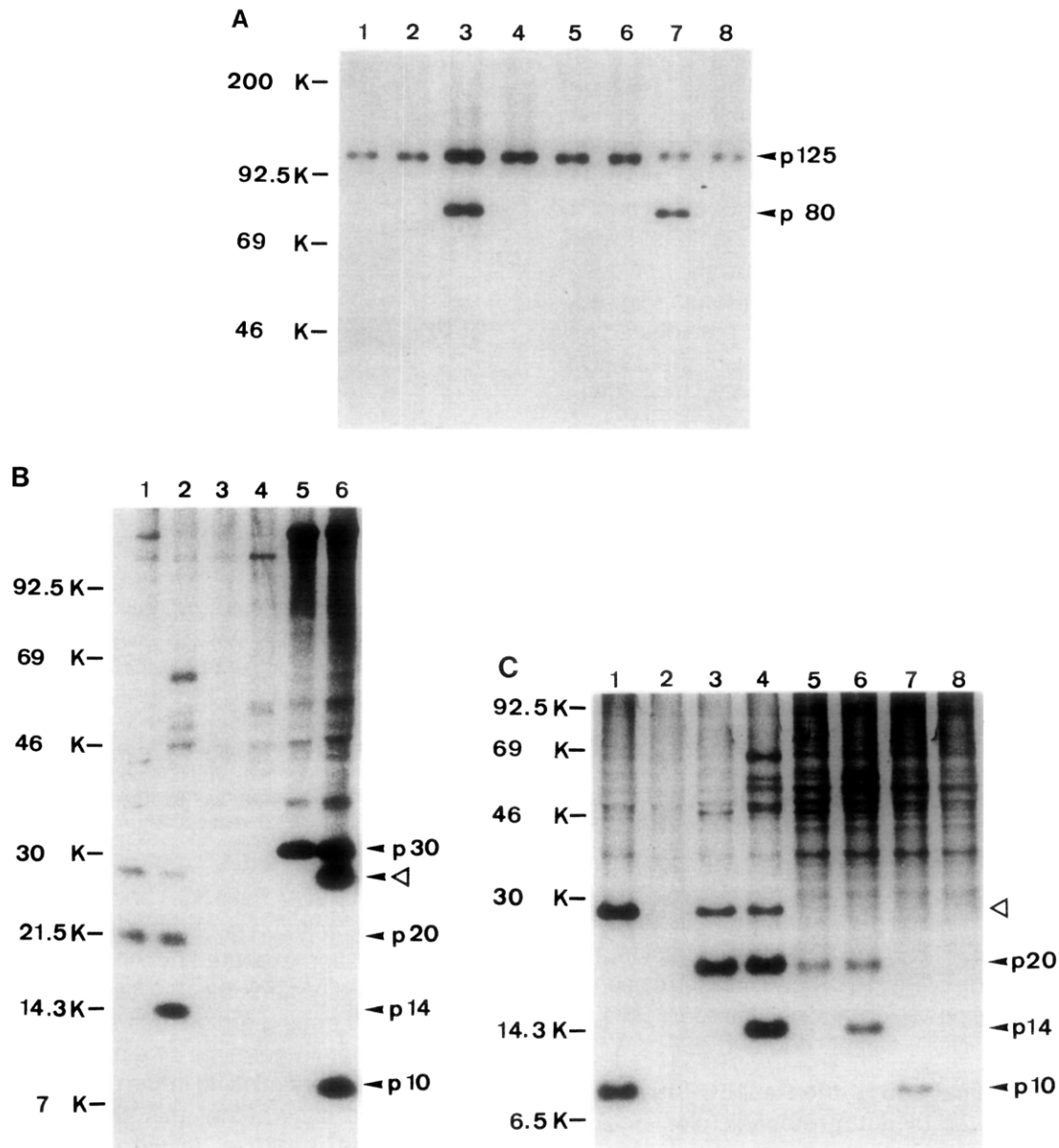
*et al.*, 1984; Donis and Dubovi, 1987a,b; Pocock *et al.*, 1987). p80 is structurally and antigenically related to a protein of 125 kDa (p125) which represents a non-structural protein present both in cells infected with cpBVDV and noncpBVDV. For the cpBVDV strains NADL and Osloss it was shown that p80 represents a carboxyterminal cleavage product of p125 (Collett *et al.*, 1988; Meyers *et al.*, 1991). The aminoterminal parts resulting from p125 processing were shown to have sizes of 54 kDa (p54) and 41 kDa (p41) (NADL and Osloss, respectively). A protein equivalent to p54 or p41 could not be demonstrated after infection with BVDV CP1 while p80 was clearly detectable. Since the duplicated region in the CP1 genome encompasses the putative p80-coding region we reasoned that p80 was expressed from the duplicated sequence and was not generated by cleavage of p125 (Meyers *et al.*, 1991). To detect the respective protein(s) after infection with BVDV Pe515CP and BVDV CP6, radioimmuno-precipitation (RIP) was performed.

Two antisera directed against different regions of p125 were employed. One of these reagents is di-

rected against a peptide corresponding to amino acids 1335 to 1351 of the BVDV NADL open reading frame (anti-1335) (Collett *et al.*, 1991). This antiserum recognizes p125 from different BVDV strains as well as p54 and p41 from BVDV NADL and Osloss, respectively. Antibodies prepared against a bacterial fusion protein served for demonstration of p80 (antiA3) (Thiel *et al.*, 1991). After infection of MDBK cells with the four BVDV strains both antisera precipitated p125 (Fig. 6A, lanes 1–8). An additional band of 80 kDa could be detected for the cytopathogenic viruses Pe515CP and CP6 in experiments with the antiA3 serum (Fig. 6A, lanes 3 and 7, respectively). Importantly, a protein corresponding to p54 of NADL or p41 of Osloss could not be precipitated by the anti-1335 serum (Fig. 6A, lanes 2, 4, 6, and 8). Since p125 is clearly recognized by the same antibodies, it can be concluded that cleavage of p125 does not occur after infection with the analyzed viruses. As already proposed for BVDV CP1, the 80-kDa protein of BVDV Pe515CP and CP6 is most likely encoded by the downstream LD regions of these cpBVDV genomes (see below).

As a consequence of recombination reactions the genomes of BVDV CP1, Pe515CP, and CP6 contain duplicated and rearranged sequences. In all these cases the genomic region coding for p125 is affected. However, the changes identified at the genome level seem not to result in expression of aberrant p125, since the respective protein from the cpBVD viruses perfectly comigrates with that from the noncp viruses (Fig. 6A and Meyers *et al.*, 1991). Accordingly, the 5' recombination point should be located downstream of the carboxyterminal processing site of p125 thereby allowing expression of this protein from the sequences upstream of the SD element. In this case aberrant protein(s) could be translated from the region downstream of the p125 gene. To investigate this hypothesis for BVDV Pe515CP RIP with different antisera, specific for the respective sequences, was performed. In addition, antibodies were included which are directed against the protein encoded by the SD-specific genomic region.

The SD region in the Pe515CP genome corresponds to codons 15–168 of the viral ORF. Therefore antisera directed against the first 55 amino acids of the pestiviral polyprotein (antiG1) and against residues 102–292 (antiD1) were employed (Thiel *et al.*, 1991). AntiG1 was shown to precipitate a protein of about 20 kDa from BVDV NADL-infected cells (p20), which represents the first protein of the BVDV ORF (Thiel *et al.*, 1991). For antiD1 precipitation of two proteins, namely p20 and the pestiviral core protein p14, has been reported (Thiel *et al.*, 1991). In addition to these proteins antiG1 and anti D1 precipitated a protein of 28 kDa



**FIG. 6.** Immunoprecipitation of proteins extracted from BVDV-infected cells after metabolic labeling with a mixture of [ $^{35}$ S]cysteine/[ $^{35}$ S]methionine. Immunoprecipitates were analyzed by SDS-PAGE (10% for A, 7.5 to 20% for B, and 12% for C). For designation of antisera see Fig. 8. (A) Extracts from MDBK cells infected with BVDV Pe515NCP (lanes 1, 2), Pe515CP (lanes 3, 4), NCP6 (lanes 5, 6), and CP6 (lanes 7, 8) were incubated with a serum directed against a peptide encompassing amino acids 1335 to 1351 of the polyprotein of BVDV NADL (Collett *et al.*, 1991) (lanes 2, 4, 6, 8) or with the antiA3 serum (Thiel *et al.*, 1991) (lanes 1, 3, 5, 7). (B) Extracts from MDBK cells infected with BVDV Pe515CP were incubated with the following rabbit antisera: anti-G1 (lane 1), anti-D1 (lane 2), rabbit preimmuneserum (lane 3), anti-L1 (lane 4), anti-K1 (lane 5), and anti-P1 (lane 6). The bands corresponding to BVDV proteins p10, p14, p20, and p30 are indicated. The fusion protein p28 is marked by an open triangle. (C) To show the specificity of p28 for the cp virus of pair Pe515 extracts from cells infected with Pe515CP (lanes 1, 2, 3, 4) or Pe515NCP (lanes 5, 6, 7, 8) were incubated with anti-HCVp14 (lanes 1, 2, 7, 8), anti-G1 (lanes 3 and 5), and anti-D1 (lanes 4 and 6). For lanes 2 and 8 100  $\mu$ g of peptide HCVp14 were added as competitor to the immunoprecipitation reaction. The position of proteins p10, p14, and p20 is indicated. p28 is marked by a triangle.

(p28) from extracts of BVDV Pe515CP-infected cells (Fig. 6B, lane 1 and 2, respectively). For further investigation three antisera specific for the region following p125 in the pestiviral polyprotein were used; they were prepared against bacterial fusion proteins containing

amino acids 2290 to 2555 (antiP1) or 2555 to 2766 (antiK1) or 2766 to 2985 (antiL1) of the HCV ORF. For BVDV NADL it has been reported that downstream of the p125 gene a protein of 10 kDa (p10) is encoded while for the following nearly 1 kb no translation prod-

uct except for a short-lived precursor of 42 kDa was found (Collett *et al.*, 1991). A stable polypeptide of 30 kDa (p30) could, however, be demonstrated for HCV and different BVDV strains with antiP1 and antiK1 (Stark and Thiel, unpublished). By using these antisera, p30 was also precipitated from extracts of Pe515CP-infected cells (Fig. 6B, lanes 5 and 6). The antiP1 serum also recognized p10 since the bacterial fusion protein employed for generation of this antiserum overlaps the respective region of the polyprotein (Fig. 6B, lane 6). Importantly, the antiP1 serum precipitated an additional antigen which comigrated with the 28-kDa protein described above and presumably represents the same polypeptide (Fig. 6B, lane 6). Since p28 reacted with antiG1 and antiD1 as well as with antiP1 but not with antiL1 or antiK1 (Fig. 6B, lanes 1, 2, 6, 4, 5, respectively) this polypeptide most likely represents a fusion protein containing sequences from p20 as well as p10.

According to the hypothesis presented above p28 should be specific for Pe515CP. To investigate this more closely, RIP with extracts from cells infected with either the cytopathogenic virus (Fig. 6C, lanes 1–4) or the noncytopathogenic virus (Fig. 6C, lanes 5–8) was performed. In addition to antiG1 (lanes 3 and 5) and antiD1 (lanes 4 and 6) a serum specific for p10 (antiHCVp14, prepared against a peptide corresponding to amino acids 2309–2325 of the HCV ORF) was employed. While p20, p14, and p10 can easily be demonstrated for both viruses p28 is only precipitated from extracts of Pe515CP-infected cells. In conclusion, the rearranged genome of Pe515CP leads to expression of at least two proteins which are not found in cells infected with Pe515NCP, namely p80 and p28.

#### The carboxyterminal end of the Pe515CP fusion protein is generated by autoproteolytic cleavage

The SD sequences integrated between the LD elements in the genomes of BVDV Pe515CP and CP6 are derived from a region of the viral RNA which has been shown to code for p20, the first protein of the BVDV ORF (Collett *et al.*, 1988b; Thiel *et al.*, 1991). Because of the rapid liberation of this polypeptide during translation which can also be observed *in vitro*, it has been proposed, that p20 represents an autoprotease (Wiskerchen *et al.*, 1991; Thiel *et al.*, 1991). Recently experimental evidence has been obtained for this theory (Wiskerchen *et al.*, 1991; Stark *et al.*, unpublished). The sequence tryptophane-valine (position 164/165) has been proposed as the cleavage recognition site of this enzyme (Wiskerchen *et al.*, 1991). Since the conserved 3' end of the SD elements maps to codon 168 of the BVDV ORF the putative cleavage signal is also duplicated and transferred to the downstream position. Ac-

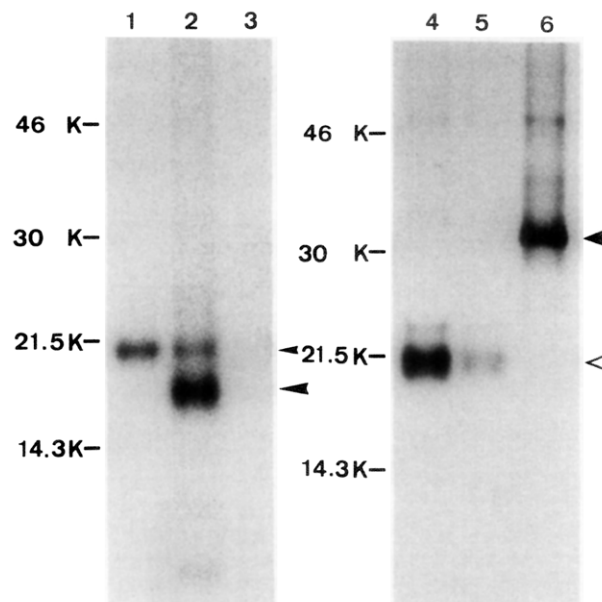
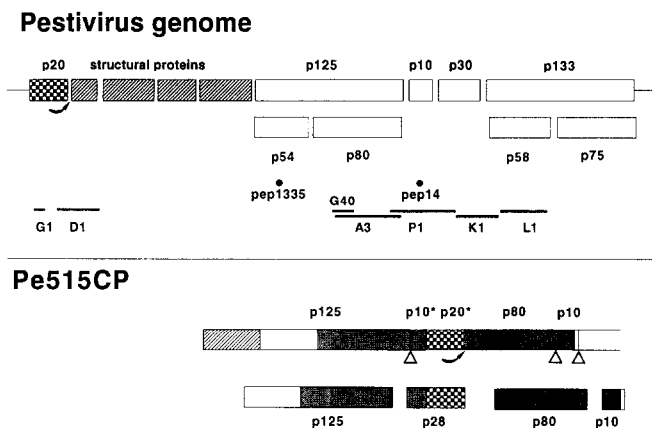


Fig. 7. Demonstration of autoproteolytic processing by the BVDV p20 protease. The products of *in vitro* translation of RNA transcribed from the linearized plasmids pCBP6 (lanes 1, 2, 3) and pCB32 (lanes 4, 5, 6) were precipitated with the sera antiG1 (lanes 1 and 4), antiD1\* (lanes 2 and 5), and antiG40 (lanes 3 and 6) (for designation of antisera see Fig. 8) and separated by 11% SDS-PAGE. The amino- and carboxyterminal products of the autoproteolytic cleavage reactions are marked by small and large arrowheads (pCBP6) or open and filled triangles (pCB32), respectively.

cordingly, for BVDV Pe515CP and CP6 the aminoterminal of p80 could be generated in an autoproteolytic cleavage reaction executed by the preceding p20 protease. For BVDV CP6 the SD sequence covers the complete p20-coding region. In the case of Pe515CP, however, the aminoterminal 14 amino acids of the protease molecule are missing in the downstream SD element. To analyze whether this aminoterminal truncated enzyme is active, *in vitro* studies were performed. Clone pCB32 was linearized with HindIII and used as template for *in vitro* transcription. The synthesized RNA contains an AUG suitable for translation initiation at position 94–96 (position 598–600 in Fig. 3A). In order to provide a control for the cleavage reaction, clone pCBP6 which contains the 1-kb PCR fragment encompassing the complete p20-coding region from the 5' part of the Pe515CP genome was also used for *in vitro* transcription. Translation of the resulting RNA should start at the original initiation codon of the viral ORF (position 23–25 in Fig. 3C). For analysis of the *in vitro* translation products, RIP with different antisera was performed. Translation of the pCBP6-derived RNA gave the following results: A full-length product of 38 kDa could not be identified (Fig. 7). Instead, an aminoterminal cleavage product of about 20 kDa was precipitated with antiG1 and antiD1 (Fig. 7, lanes 1 and 2,





**Fig. 8.** Schematic presentation of the proteins encoded by the analyzed region of the Pe515CP genome. The upper part shows the genetic map of BVDV (Collett *et al.*, 1991; Thiel *et al.*, 1991; Stark *et al.*, unpublished) and the location and designation of bacterial fusion proteins (black bars) and peptides (dots) used for production of the different antisera. For Pe515CP the respective parts of the polyprotein as well as the mature proteins are indicated. The autoproteolytic action of the p20 protease is indicated by a bent arrow. The other cleavages involved in processing of the BVDV polyprotein nonstructural region are carried out by p80 and presumable p125 (Wiskerchen and Collett, 1991) (cleavage sites indicated by triangles).

respectively). This band most likely represents the autoprotease p20. Since antiD1 is directed against amino acids 102 to 292 of the pestiviral ORF, this serum also reacts with the carboxyterminal processing product which has an expected size of about 18 kDa (Fig. 7, lane 2).

Translation of the pCB32-derived RNA again did not result in substantial amounts of the full length product with an expected size of 51 kDa (Fig. 7, lanes 4, 5, 6). AntiG1 and antiD1\* both recognize a predominant band of about 20 kDa, while a serum directed against a bacterial fusion protein encompassing part of p80 (anti-G40, Collett *et al.*, 1988b) precipitates a protein of about 32 kDa. This finding correlates with a proteolytic cleavage at the predicted site (between amino acids 362 and 363 in Fig. 3A). Additional experiments with 3' terminally shortened templates revealed that this result is not dependent on the serine protease motif present in the p80 moiety of the protein translated after *HindIII* digest of pCB32. Analogous results were obtained in reactions with a corresponding cDNA fragment from CP6 (data not shown). These data strongly suggest that the transferred element composed of protease and cleavage site is indeed functional, leading to an autoproteolytical removal of the fusion protein from the following p80 region (Fig. 8).

## DISCUSSION

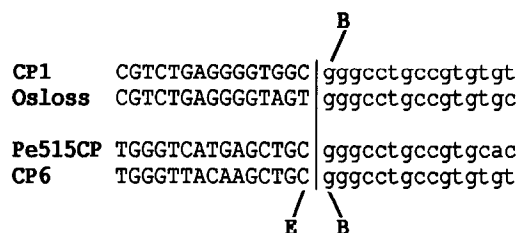
The generation of cytopathogenic BVDV within a calf persistently infected with noncytopathogenic virus re-

sults in development of mucosal disease and death of the animal (Brownlie *et al.*, 1984; Bolin *et al.*, 1985). The discovery of host cell-derived insertions in the genomes of two cpBVDV strains (Osloss and NADL) led us to propose a recombination as mechanism for development of cpBVDV (Meyers *et al.*, 1989b, 1990). Clear evidence for this model was gained by analysis of a BVDV pair consisting of the viruses CP1 and NCP1. The genome of CP1 contains a ubiquitin-coding sequence which is flanked by a large duplication of viral sequences. Neither insertion nor duplication are present in the respective part of the NCP1 genome. However, the two viruses exhibit 99.6% of sequence identity within corresponding regions of their genomes (Meyers *et al.*, 1991).

The data presented in this report broaden the proposed model. In the case of BVDV Pe515CP and CP6 no host cell-derived insertion could be identified. However, elaborate duplication and rearrangement of viral sequences were found for both genomes. The analysis of the noncp virus (Pe515NCP) 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.

Until now recombination-induced changes at the genome level have been identified for 6 cpBVDV strains. Three of these viruses contain ubiquitin-coding insertions, namely BVDV Osloss, CP1, and CP14 (manuscript in preparation for the latter), while the NADL strain has integrated another cellular sequence. The other two, Pe515CP and CP6, have been generated in a recombination process involving only viral sequences. On the basis of PCR analyses it has been found that the genomes of several cpBVDV strains including Pe515CP do not contain insertions within the p125 gene (DeMoerloose *et al.*, 1990). For Pe515CP, this finding describes only one feature of the respective genomic region. The viral RNA contains a p125 gene without insertion but nevertheless it represents the product of a recombination which affected p125-coding sequences. Since part of the p125-coding region has been duplicated the respective rearrangement in the Pe515CP RNA could not be detected by a simple PCR assay.

The situation is different for some cpBVDV laboratory strains, like Singer, Oregon, and Danmark. None of these strains has a genome size indicating duplications (G. Meyers, unpublished). The PCR analysis revealed that the region around position 5400 of the BVDV Singer genome does not contain an insertion



**Fig. 9.** Comparison of the sequences flanking the 3' recombination site (vertical line) of different cpBVDV isolates. The residues adjacent to the recombination site are marked with the letters B and E according to the nomenclature listed in legend of Fig. 5. Sequences 5' to the recombination site are shown in capital letters, nucleotides 3' of this position in lower case.

(DeMoerlooze *et al.*, 1990). After conventional cDNA cloning and sequencing of about 2 kb derived from the respective genomic region the same result was obtained for the strains Oregon and Danmark (G. Meyers, unpublished). More elaborate cloning and sequencing is necessary to identify the genetic basis for generation of p80 and the establishment of a cytopathogenic phenotype by these viruses.

Recombination at the RNA level has been observed for several RNA viruses. Template switching of the viral polymerase during replication or transcription has been proposed as a mechanism for these reactions (Lazzarini *et al.*, 1981; Lai *et al.*, 1985; Kirkegaard and Baltimore, 1986; Makino *et al.*, 1986; King *et al.*, 1987; and Khatchikian *et al.*, 1989). Based on this hypothesis a model for the generation of the cpBVDV genomic RNAs has been put forward which includes two subsequent switches (Meyers *et al.*, 1991). Since the cellular mRNAs which serve as templates for integration of, e.g., ubiquitin-coding sequences, are present in the "sense" orientation in the cell and in the viral genome, the recombination should occur during synthesis of negative strand viral RNA. Interestingly, the 3' recombination position was the same for the genomes of CP1 and Osloss and it was obvious to look for sequences responsible for this site specificity. Two different bovine polyubiquitin mRNAs were analyzed but no homology or complementarity between the sequences flanking the putative recombination positions in the viral and cellular RNAs could be identified (Meyers *et al.*, 1991). The same result was obtained when the nucleotides adjacent to the SD and LD regions of BVDV Pe515CP and CP6 were analyzed. Surprisingly, sequence comparison revealed identical 3' recombination positions regardless whether viruses with ubiquitin-coding insertions or duplication of p20- and p80-coding sequences were looked at (Fig. 9). It therefore has to be considered that the site-specific recombination observed for these viruses might be a result of functional pressure

rather than being guided by specific sequences. This functional pressure could be based on the requirement of a given aminoterminal for the cpBVDV-specific protein p80. Accordingly, development of a cytopathogenic virus in conjunction with outbreak of MD would represent the final step with regard to selection for the correct recombinant. Studies on coronaviruses tend to support this view since selection for certain recombination hotspots is preceded by random RNA recombination (Banner and Lai, 1991).

The genome rearrangements which have been identified for the different cpBVDV strains result in changes of virus-encoded proteins. One of these changes represents the expression of aberrant fusion proteins composed of (i) cellular and viral sequences (NADL p54 and Osloss p41) or (ii) sequences from two viral polypeptides derived from different regions of the genome (Pe515CP) (Fig. 8). For BVDV CP6 a fusion protein could so far not be demonstrated. This might either be due to a technical problem or to rapid degradation of the respective product which according to the genome organization of CP6 should be composed of p20 sequences and a small part of p58, the second last polypeptide encoded by the pestiviral ORF (Collett *et al.*, 1991). Further investigation is needed to clarify this point.

The other important change at the protein level distinguishing cpBVDV from noncpBVDV is the expression of p80 in addition to p125. While the above-mentioned fusion proteins show a high degree of variability concerning nature and size, not only the mere presence of p80 in cpBVDV-infected cells but also its size is strictly conserved (Purchio *et al.*, 1984; Donis and Dubovi, 1987a; Greiser-Wilke *et al.*, 1992; Meyers *et al.*, unpublished). A prerequisite for generation of p80 is the introduction of a protease 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 eucaryotic cells in form of fusion proteins which are subsequently cleaved at the carboxyterminal end of the ubiquitin moiety (Finley *et al.*, 1987, 1989; Özkaynak *et al.*, 1987; Redmann and Rechsteiner, 1989, Rechsteiner, 1987). 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. In addition a processing signal is 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 already present in the virus serves this purpose. 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 *in trans* action of the p20 protein. This would be in accordance with the hypothesis that p20 can only act as an autoprotease (Wiskerchen *et al.*, 1991).

Both alternatives outlined above would allow identical aminoterminal ends of the different p80 proteins since the insertion point for the ubiquitin gene in the Osloss RNA and the 5' end of the LD region in the LD/SD type of genomes correspond to the same genomic position. Even though further investigation is necessary to fully understand the generation of p80, one has to keep in mind that expression of this protein is strictly correlated with the cytopathogenic phenotype and the development of MD. p80 therefore represents the prime candidate for the agent responsible for killing the infected cells while the different fusion proteins might be only byproducts resulting from the recombination. According to sequence comparison studies p80 contains protease as well as helicase motifs (Gorbalenya *et al.*, 1989a,b). Experimental evidence for the protease function of p80 has recently been obtained (Wiskerchen *et al.*, 1991; Tautz *et al.*, in preparation). If p80 turns out to be required for the cytopathogenic effect it will be interesting to determine why p125, having the same enzymatic domains, does not exhibit this effect.

The BVDV/MD system offers the opportunity to study RNA virus evolution in action. Development of MD appears to be obligatorily linked to generation of cpBVDV. The required recombination process leads to expression of novel proteins, especially p80. The analysis of additional virus pairs will provide insights into the different possibilities to generate viruses with a given phenotype.

## ACKNOWLEDGMENTS

The authors thank Carmen Rein, Petra Ulrich, and Silke Mauritz for excellent technical assistance.

## REFERENCES

- BAKER, J. C. (1987). Bovine viral diarrhea virus: A review. *J. Amer. Vet. Med. Assoc.* **190**, 1449–1458.
- BANNER, L. R., and LAI, M. M. C. (1991). Random nature of coronavirus recombination in the absence of selection pressure. *Virology* **185**, 441–445.
- BEAUCAGE, S. L., and CARUTHERS, M. H. (1981). Deoxynucleoside phosphoramidites: A new class of key intermediates for deoxynucleotide synthesis. *Tetrahedron Lett.* **22**, 1859–1862.
- BOLIN, S. R., MCCLURKIN, A. W., CUTLIP, R. C., and CORIA, M. F. (1985). Severe clinical disease induced in cattle persistently infected with noncytopathic bovine viral diarrhea virus by superinfection with cytopathic bovine viral diarrhea virus. *Amer. J. Vet. Res.* **46**, 573–576.
- BROWNIE, J., CLARKE, M. C., and HOWARD, C. J. (1984). Experimental production of fatal mucosal disease in cattle. *Vet. Rec.* **114**, 535–536.
- COLLETT, M. S., LARSON, R., GOLD, C., STRINCK, D., ANDERSON, D. K., and PURCHIO, A. F. (1988a). Molecular cloning and nucleotide sequence of the pestivirus bovine viral diarrhea virus. *Virology* **165**, 191–199.
- COLLETT, M. S., LARSON, R., BELZER, S. K., and RETZEL, E. (1988b). Proteins encoded by bovine viral diarrhea virus: The genomic organization of a pestivirus. *Virology* **165**, 200–208.
- COLLETT, M. S., ANDERSON, D. K., and RETZEL, E. (1988c). Comparisons of the pestivirus bovine viral diarrhea virus with members of the flaviviridae. *J. Gen. Virol.* **69**, 2637–2643.
- COLLETT, M. S., WISKERCHEN, M. A., WELNIAK, E., and BELZER, S. K. (1991). Bovine viral diarrhea virus genomic organization. *Arch. Virol.* [Suppl 3], 19–27.
- CORAPI, W. V., DONIS, R. O., and DUBOVI, E. J. (1988). Monoclonal antibody analyses of cytopathic and noncytopathic viruses from fatal bovine viral diarrhea virus infections. *J. Virol.* **62**, 2823–2827.
- DEMOERLOOZE, L., DESPORT, M., RENARD, A., LECOMTE, C., BROWNIE, J., and MARTIAL, J. A. (1990). The coding region for the 54-kDa protein of several pestiviruses lacks host insertions but reveals a "zinc finger-like" domain. *Virology* **177**, 812–815.
- DEVEREUX, J., HAEBERLI, P., and SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**(1), 387–395.
- DONIS, R. O., and DUBOVI, E. J. (1987a). Molecular specificity of the antibody responses of cattle naturally and experimentally infected with cytopathic and noncytopathic bovine diarrhea virus biotypes. *Amer. J. Vet. Res.* **48**, 1549–1554.
- DONIS, R. O., and DUBOVI, E. J. (1987b). Characterization of bovine viral diarrhea-mucosal disease virus-specific proteins in bovine cells. *J. Gen. Virol.* **68**, 1597–1605.
- FINLEY, D., ÖZKAYNAK, E., and VARSHAVSKY, A. (1987). The yeast poly-ubiquitin gene is essential for resistance to high temperatures, starvation and other stresses. *Cell* **48**, 1035–1046.
- FINLEY, D., BARTEL, B., and VARSHAVSKY, A. (1989). The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. *Nature (London)* **338**, 394–401.
- GORBALENYA, A. E., DONCHENKO, A. P., KOONIN, E. V., and BLINO, V. M. (1989a). N-terminal domains of putative helicases of flaviviruses and pestiviruses may be serine proteases. *Nucleic Acids Res.* **17**, 3889–3897.
- GORBALENYA, A. E., KOONIN, E. V., DONCHENKO, A. P., and BLINOV, V. M. (1989b). Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res.* **17**, 4713–4730.
- GREISER-WILKE, I., DITTMAR, K. E., LIESS, B., and MOENNIG, V. (1992). Heterogeneous expression of the non-structural protein p80/p125 in cells infected with different pestiviruses. *J. Gen. Virol.* **73**, 47–52.
- GUTEKUNST, D. E., and MALMQUIST, W. A. (1963). Separation of a soluble antigen and infectious particles of bovine virus diarrhea virus and their relationship to hog cholera. *Canad. J. Comp. Med. Vet. Sci.* **27**, 121–123.
- HARLOW, E., and LANE, D. (1988). "Antibodies, A Laboratory Manual." Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- HENNIKOFF, S. (1987). Unidirectional digestion with exonuclease III in

- DNA sequence analysis. In "Methods in Enzymology" (R. Wu, Ed.) Vol. 155, pp. 156–165, Academic Press, San Diego.
- JEFFREYS, A. J., WILSON, V., NEUMANN, R., and KEYTE, J. (1988). Amplification of human minisatellites by the polymerase chain reaction: Towards DNA fingerprinting of single cells. *Nucleic Acids Res.* **16**, 10953–10971.
- KESSLER, S. W. (1981). Use of protein A-bearing staphylococci for the immunoprecipitation and isolation of antigens from cells. In "Methods in Enzymology" (J. J. Langone and H. Van Vunakis, Eds.), Vol. 73, pp. 442–459. Academic Press, New York.
- KHATCHIKIAN, D., ORLICH, M., and ROTT, R. (1989). Increased viral pathogenicity after insertion of a 28S ribosomal RNA sequence into the haemagglutinin gene of an influenza virus. *Nature* **340**, 156–157.
- KING, A. M. Q., ORTLEPP, S. A., NEWMAN, J. W. I., McCAHON, D. (1987). Genetic recombination in RNA viruses. In "The Molecular Biology of the Positive Stranded RNA Viruses" (D. J. Rowlands, M. A. Mayo, B. W. J. Mahy, Eds.), pp. 129–152. Academic Press, London.
- KIRKEGAARD, K., and BALTIMORE, D. (1986). The mechanism of RNA recombination in poliovirus. *Cell* **47**, 433–443.
- LAI, M. M. C., BARIC, R. S., MAKINO, S., KECK, J. G., EGBERT, J., LEIBOWITZ, J. L., and STOHLMANN, S. A. (1985). Recombination between nonsegmented RNA genomes of murine coronaviruses. *J. Virol.* **56**, 449–456.
- LAZZARINI, R. A., KEENE, J. D., and SCHUBERT, M. (1981). The origins of defective interfering particles of the negative-strand RNA viruses. *Cell* **26**, 145–154.
- MAKINO, S., KECK, J. G., STOHLMANN, S. A., and LAI, M. M. C. (1986). High-frequency RNA recombination of murine coronaviruses. *J. Virol.* **57**, 729–737.
- MANIATIS, T., FRITSCH, E. F., and SAMBROOKS, S. (1989). "Molecular Cloning, A Laboratory Manual." Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- MEYERS, G., RÜMENAPF, T., and THIEL, H.-J. (1989a). Molecular cloning and nucleotide sequence of the genome of hog cholera virus. *Virology* **171**, 555–567.
- MEYERS, G., RÜMENAPF, T., and THIEL, H.-J. (1989b). Ubiquitin in a Togavirus. *Nature (London)* **341**, 491.
- MEYERS, G., RÜMENAPF, T., and THIEL, H.-J. (1990). Insertion of ubiquitin-coding sequence identified in the RNA genome of a Togavirus. In "New Aspects of Positive Strand RNA Viruses" (M. A. Brinton and F. X. Heinz, Eds.) pp. 25–29. American Society for Microbiology, Washington, DC.
- MEYERS, G., TAUTZ, N., DUBOVI, E. J., and THIEL, H.-J. (1991). Viral cytopathogenicity correlated with integration of ubiquitin-coding sequences. *Virology* **180**, 602–616.
- ÖZKAYNAK, E., FINLEY, D., SOLOMON, M. S., and VARSHAVSKY, A. (1987). The yeast ubiquitin genes: A family of natural gene fusions. *EMBO J.* **6**, 1429–1439.
- POCOCK, D. H., HOWARD, C. J., CLARKE, M. C., and BROWNLIE, J. (1987). Variation in the intracellular polypeptide profiles from different isolates of bovine viral diarrhoea virus. *Arch. Virol.* **94**, 43–53.
- PURCHIO, A. F., LARSON, R., and COLLETT, M. S. (1984). Characterization of bovine viral diarrhoea viral proteins. *J. Virol.* **50**, 666–669.
- RECHSTEINER, M. (1987). Ubiquitin-mediated pathways for intracellular proteolysis. *Annu. Rev. Cell Biol.* **3**, 1–30.
- REDMANN, K. L., and RECHSTEINER, M. (1989). Identification of the long ubiquitin extension as ribosomal protein S27a. *Nature (London)* **338**, 438–440.
- RENARD, A., DINO, D., and MARTIAL, J. (1987). Vaccines and diagnostics derived from bovine diarrhoea virus. European Patent Application number 86870095. 6. Publication number 0208672, 14 January 1987.
- RIGBY, P. W. J., DIECKMANN, M., THODES, C., and BERG, P. (1977). Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.* **113**, 237–251.
- RÜMENAPF, T., MEYERS, G., STARK, R., and THIEL, H.-J. (1989). Hog cholera virus-Characterization of specific antiserum and identification of cDNA clones. *Virology* **171**, 18–27.
- RÜMENAPF, T., STARK, R., MEYERS, G., and THIEL, H.-J. (1991). Structural proteins of hog cholera virus expressed by vaccinia virus: further characterization and induction of protective immunity. *J. Virol.* **65**, 589–597.
- SANGER, F., NICKLEN, S., and COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
- STREBEL, K., BECK, E., STROHMAIER, K., and SCHALLER, H. (1986). Characterization of foot-and-mouth disease virus gene products with antisera against bacterially synthesized fusion proteins. *J. Virol.* **57**, 983–991.
- THIEL, H.-J., STARK, R., WEILAND, E., RÜMENAPF, T., and MEYERS, G. (1991). Hog cholera virus: Molecular composition of virions from a pestivirus. *J. Virol.* **65**, 4705–4712.
- WENGLER, G. (1991). Family Flaviviridae. In "Classification and Nomenclature of Viruses: Fifth Report of the International Committee on Taxonomy of Viruses" (R. I. B. Francki, C. M. Fauquet, D. L. Knudson, and F. Brown, Eds.). Springer-Verlag, Berlin.
- WISKERCHEN, M. A., BELZER, S. K., and COLLETT, M. S. (1991). Pestivirus gene expression: The first protein of the bovine viral diarrhoea virus large open reading frame, p20, possesses proteolytic activity. *J. Virol.* **65**, 4508–4514.
- WISKERCHEN, M. A. and COLLETT, M. S. (1991). Pestivirus gene expression: Protein p80 of bovine viral diarrhoea virus is a proteinase involved in polypeptide processing. *Virology* **184**, 341–350.