Molecular Cloning and Nucleotide Sequence of a Pestivirus Genome, Noncytopathic Bovine Viral Diarrhea Virus Strain SD-1

RUITANG DENG AND KENNY V. BROCK

Ohio Agricultural Research And Development Center, Food Animal Health Research Program, 1680 Madison Avenue, Wooster, Ohio 44691

Received July 10, 1992; accepted August 12, 1992

Genomic RNA of noncytopathic (NCP) bovine viral diarrhea virus (BVDV) strain SD-1 was extracted directly from serum obtained from a persistently infected animal. cDNA was synthesized and amplified by polymerase chain reaction (PCR) before cloning. The complete genomic nucleotide sequence was determined by sequencing at least two different clones from independent PCR reactions. The 5' and 3' end sequences of the SD-1 genome was determined from 5'-3' ligation clones. The complete genome sequence was comprised of 12,308 nucleotides containing one large open reading frame which encodes an amino acid sequence of 3898 residues with a calculated molecular weight of 438 kDa. In contrast to cytopathic (CP) BVDV strain NADL, which contains a cellular RNA insert of 270 nucleotides and CP BVDV strain Osloss, which has an inserted ubiquitin RNA sequence of 228 nucleotides, the NCP strain SD-1 had no insertion along the genome. Sequence comparison with other pestiviruses revealed that the overall nucleotide sequence homologies of SD-1 are 88.6% with NADL, 78.3% with Osloss, 67.1% with HoCV Alfort, and 67.2% with HoCV Brescia. The overall deduced amino acid sequence homologies of SD-1 are 92.7% with NADL, 86.2% with Osloss, 72.5% with HoCV Alfort, and 71.2% with HoCV Brescia. The most conserved nucleotide and amino acid sequences are located in the 5' untranslated region (5'UTR) and nonstructural protein p80 region, respectively. The viral glycoproteins, particularly gp53, and nonstructural proteins p54 and p58 have the lowest homology comparing both nucleotide and amino acid sequences between SD-1 and other pestiviruses. Extensive analyses of amino acid sequences for the viral structural proteins and nonstructural protein p54 regions from five pestiviruses led to the identification of four conserved domains (designated as C1, C2, C3, C4) and three highly variable domains (designated as V1, V2, V3) within this region. The C1, C2, and C3 domains are located in the capsid protein p14, glycoprotein gp48, and gp25, respectively. The C4 domain is located in the junction between gp53 and p54. Interestingly, out of three variable domains, two (V1, V2) are located in the same glycoprotein gp53. The third variable domain is located in the nonstructural protein p54.

INTRODUCTION

Bovine viral diarrhea virus (BVDV), a small enveloped virus, is one of the most important viral pathogens of cattle (Duffel and Harkness, 1985). BVDV infection can result in a variety of clinical diseases in cattle, such as abortion, persistent infection, and mucosal disease (MD). Its genome is a single-stranded RNA with positive polarity and consists of about 12,500 nucleotides (Renard et al., 1987; Collett et al., 1988a). Together with the other two serologically and structurally related viruses: hog cholera virus (HoCV) of swine and border diseases virus (BDV) of sheep, BVDV belongs to the Pestivirus group. Based on their similarities of genome organization and strategy of gene expression with that of the Flaviviruses, Pestiviruses were recently reclassified into the Flaviviridae family (Collett et al., 1988c; Francki et al., 1991; Hufnagel, 1991). However, differences in virion composition with the Flaviviridae (Thiel et al., 1991) and the absence of a 5' cap structure of its RNA genome (Brock et al., 1992), reflecting different mechanisms of viral RNA translation, are some objections to this reclassification. Based on the cytopathogenicity in cell culture, BVDV has been divided into two biotypes: cytopathic (CP) BVDV and noncytopathic (NCP) BVDV (Bolin et al., 1985). Only NCP BVDV is capable of establishing persistent infections in cattle following in utero infection (Brownlie et al., 1984). Furthermore, mucosal disease, a severe clinical syndrome, occurs only in persistently infected animals when they are superinfected with a second, antigenically indistinguishable CP BVDV (Brownlie et al., 1984; Bolin et al., 1985; Corapi et al., 1988). These observations led to the hypothesis that CP BVDV may originate from NCP BVDV by genomic mutation (Corapi et al., 1988).

To date, two CP strains of BVDV: NADL (Collett et al., 1988a) and Osloss (Renard et al., 1987), and two strains of HoCV: Alfort (Meyers et al., 1989a) and Brescia (Moormann et al., 1990) have been cloned and sequenced. Recently, a partial genomic sequence located in the p125 region of a pair of BVDV, CP BVDV
strain CP1 (about 5.7 kb), and NCP BVDV strain NCP1 (about 3.2 kb) were published (Meyers et al., 1991). Comparison of the genomic sequence between BVDV and HoCV led to the finding of cellular sequence inserts in a region coding for the N-part of p125 in the CP BVDV (Meyers et al., 1989b; Collett et al., 1989). Therefore, a hypothesis was proposed by Meyers et al. (1991) that the insertion of a cellular RNA sequence by RNA recombination into the NCP BVDV genome was responsible for the development of CP BVDV from NCP BVDV. However, this hypothesis was challenged by observations that some CP BVDV strains lack the insertion in their genomes (Moerlooze et al., 1990; Desport et al., 1991) which suggests that the understanding of BVDV cytopathogenicity is far from complete.

In this report, the complete nucleotide sequence of NCP BVDV, strain SD-1, is presented for the first time. In addition, analyses and comparisons of the nucleotide and amino acid sequences are made with those of other pestiviruses.

MATERIALS AND METHODS

Persistently infected animal and virus

A persistently infected heifer was maintained in an isolation facility to prevent exposure and infection with other BVDV strains. The virus isolated from this heifer was a NCP BVDV and was designated as strain SD-1. Blood was collected from the persistently infected heifer and serum recovered for virus purification and extraction of viral RNA.

Virus purification and RNA extraction

Partial purification of virus and viral RNA extraction by the guanidine thiocyanate method were performed as described previously (Brock et al., 1992).

cDNA synthesis, PCR amplification, and cloning

cDNA synthesis using random primer and genomic viral RNA extracted from partially purified viral particles was done as previously reported (Brock et al., 1992). Following the first-strand cDNA synthesis, PCR was carried out to amplify the defined viral segments. The two primers used to amplify the first SD-1 fragment were designed based on the sequence of the NADL strain in the nonstructural protein region. After determination of the nucleotide sequence of the first SD-1 cDNA clone, some of the primers were designed based on SD-1 sequence. The size of amplified segments was chosen to be 1.5 to 2.0 kb in order to obtain optimal amplification. Primer length was chosen to be 18 to 24 bases in order to maintain the T_m value above 55°C. PCR amplification was performed with Taq polymerase (Perkin-Elmer Cetus, Norwalk, CT) for 30 to 35 cycles. The working profiles were as follows: 94°C for 1 min to denature the DNA, 42 to 60°C for 1.5 min to allow primer annealing and 72°C for 2 to 5 min for DNA extension. All the primers were purchased from either Biochemical Instrument Center of The Ohio State University (Columbus, OH) or Genosys (Woodlands, TX). Purification and C-tailing of the PCR products was done as described previously (Brock et al., 1992). C-tailed PCR products were cloned into G-tailed pUC9 plasmid and used to transform competent Escherichia coli strain JM109 (Hanahan, 1983). Colony blots were done using nitrocellulose membranes and positive clones were screened by the corresponding [32P]-dCTP-labeled NADL cDNA fragments. The cloning of 5' and 3' end sequences of viral RNA was done by 5'-3' ligation and PCR as previously described (Mandl et al., 1991; Brock et al., 1992).

Sequencing of cDNA clones

Prior to nucleotide sequencing, restriction enzyme mapping was performed to determine the appropriate restriction enzyme sites in the cDNA clones for subcloning of restriction fragments into pGEM-3Z vector (Promega, Madison, WI). The alkaline lysis and PEG precipitation method was used to extract and purify plasmid DNA (Birnboim, 1983; Lis and Schleif, 1975) for double-stranded DNA sequencing (Chen and Sambrook, 1985). The nucleotide sequence of the inserts was determined by the dideoxy chain termination method (Sanger et al., 1977) using sequencing kits (United State Biochemicals, Cleveland, OH). The nucleotide sequences of the clones that lacked the appropriate restriction enzyme sites for subcloning were determined by a progressive oligonucleotide primer method (Sambrook et al., 1989). Considering the potential of sequence errors created by Taq polymerase the entire genomic sequence of NCP BVDV SD-1 was determined by completely sequencing a minimum of two clones from independent PCR reactions for each region. If a different nucleotide sequence was obtained from the two clones, the common nucleotide sequence was verified by sequencing a third or even fourth clone from other independent PCR reactions. The 5' and 3' end sequences were confirmed by sequencing nine independent 5'-3' ligation clones.

Computer analysis

Nucleotide sequence comparison and analysis were made with HIBIO DNASIS (Hitachi Software Engineering Co., Ltd., Brisbane, CA) (Lipman and Pearson, 1985; Needleman and Wunsch, 1970). The predicted amino acid sequence was analyzed and compared by
Molecular cloning of BVDV SD-1 RNA

To determine the nucleotide sequence, viral RNA was directly extracted from serum obtained from a persistently infected heifer. The virus titer in the serum was $10^3 - 10^4$ CCID$_{50}$/ml. Viral RNA extracted from 1 to 5 ml of serum was enough to carry out the PCR amplification and cDNA cloning. To optimize PCR amplification and minimize the nonspecific priming, PCR profiles varied for each set of primers. The annealing temperatures ranged from 42 to 60°C depending on the T$_{m}$ value of the primer and the homology between the primer and the sequence to be amplified. Several cloning methods were tried to clone PCR products. Compared with either blunt-end ligation or AT annealing and ligation cloning methods for PCR products, the GC-tailing method had a higher cloning efficiency. In order to eliminate the potential sequence errors created by Taq polymerase, repeat cDNA cloning of three independent PCR reactions for each viral RNA region was done. The corresponding restriction fragments of NADL cDNA clones were used as probes for identification of positive cDNA clones of SD-1 genomic RNA. To determine the extreme 5' and 3' end sequences of the genome, genomic RNA ligation was performed before PCR amplification and cloning. Based on previous results suggesting that there is no cap structure at the 5' end of BVDV genome (Brock et al., 1992), genomic RNA of SD-1 was directly ligated without the treatment of pyrophosphatase to remove a 5' cap structure. After restriction enzyme mapping, about 70% of inserts in the original pUC9 vector were subcloned into pGEM vectors for sequencing. A total of 29 cDNA clones that almost overlapped the whole genome three times were used to determine the nucleotide sequence of SD-1 genome (Fig. 1).

Nucleotide sequence of NCP BVDV strain SD-1

The complete SD-1 sequence was determined by sequencing at least two clones from independent PCR reactions. Seventy percent of the sequence was determined from both strands of cDNA clones. The remainder was obtained from multiple determinations on a single strand. A total of 36 nucleotides were different by comparison of the two independent nucleotide sequences in the cDNA sequence of about 30,000 nucleotides. In those cases, the consensus nucleotide sequence was obtained by sequencing a third or even fourth independent clone. Interestingly, one of the errors found in one of the cDNA clones has been previously described for BVDV NADL (Collett et al., unpublished data) and HoCV Brescia cDNA clones (Moor-mann et al., 1990), involving a stretch of five sequential adenosines, where six adenosines were the final correct nucleotide sequence.

The 5' and 3' end sequences of the SD-1 genome were determined from the 5'-3' ligation clones. Of nine clones sequenced, four clones have the 3' end sequence of 5' . . . CAGCCCCC 3', four clones have 5' . . . CAGCCCCC 3', and one clone has 5' . . . CAGCCC 3'. Therefore, the dominant and longest sequence, 5' . . . CAGCCCCC 3', is the authentic 3' end sequence of SD-1 genome, which is identical to the 3' end sequence of NADL. The complete nucleotide sequence consists of 12,308 nucleotides (Fig. 2), which is 270 bases less than CP BVDV strain NADL (12,578 nucleotides) and 124 bases less than Osloss genomes (12,430 nucleotides), but closer to HoCV strain Alfort and Brescia genomes (12,284 and 12,283 nucleotides, respectively). Base composition of the entire genome of SD-1 is 32.2% A, 22.0% U, 25.6% G, and 20.2% C and is similar to NADL RNA (Collett et al., 1988a).

Analysis of the SD-1 sequence revealed one transla-
tional open reading frame (ORF) in the second phase of one strand. The other two reading phases of this strand and the three reading phases of the complementary strand contain multiple stop codons throughout the sequence. No other significant ORF can be predicted in these reading phases. The large ORF starts with the AUG at position 386 to 388 and ends with a stop codon UGA at position 12,080 to 12,082. This ORF is capable of encoding a polyprotein of 3898 amino acids with the calculated molecular weight of 438 kDa (Fig. 2). The predicted amino acid sequence of the large ORF is given in Fig. 2. The 5' untranslated region (5'UTR) preceding the large ORF consists of 385 nucleotides, including six AUG start codons and sev-
signal sequence. The putative glycosylation sites in the viral glycoprotein region are underlined. The double-underlined sequence is the cysteine-rich stretch. AAA indicates the catalytic triad of His, Asp, and Ser residues of the viral p80 proteinase. 000 denotes the tripeptide sequence highly conserved in viral RNA dependent RNA polymerases.

FIG.

DENG AND BROCK
Fig. 2—Continued
eral small ORFs. Following the stop codon of the large ORF, the 3' untranslated region (3'UTR) continues for another 229 nucleotides.

Amino acid sequence of NCP BVDV strain SD-1

Hydrophobicity analysis of the entire amino acid sequence revealed two characteristic regions. The first region, located at the N-terminal of the polyprotein (the first 250 amino acid residues), is highly hydrophilic, particularly in the region from residues 165 to 250 which was reported to encode the viral capsid protein p14 (Thiel et al., 1991). Of 85 amino acid residues, 25 are positively charged residues (22 lysines, 3 arginines) and 12 are negatively charged residues (6 aspartic acids, 6 glutamic acids). The second region located at position 1036 to 1301 consists of 266 amino acid residues and is highly hydrophobic. This region contains 51 leucines, 35 valines, 29 isoleucines, and 20 threonines.

Twenty-eight potential N-linked glycosylation sequences (Asn-X–Ser or Asn-X–Thr) were predicted along the polyprotein. Fourteen sites are located in the glycoprotein region between residues 271 and 992 (Fig. 2). Of the 14 sites, 8 were conserved among all the pestiviruses and 3 other sites were conserved among the BVDV strains. A cysteine-rich stretch, reported to conform to a "zinc finger" binding domain (Moerlooze et al., 1990), was also found in the SD-1 polyprotein at amino acid residues 1484 to 1512 (Fig. 2). The tripeptide sequence, Gly–Asp–Asp, highly conserved in viral RNA dependent RNA polymerases (Kamer and Argos, 1984), was uniquely identified in the nonstructural protein p75 region at position 3626 to 3628 (Fig. 2). This finding supports the proposal that this protein is a candidate BVDV replicase (Collett et al., 1991). Sequence comparisons and structural pattern analysis predicted the BVDV p80 protein to be a trypsin-like serine proteinase (Bazan and Fletierick, 1989; Gorbalenya et al., 1989a; Moormann et al., 1990). Recently, Wiskerchen and Collett (1991) experimentally demonstrated that this protein is a viral proteinase and is responsible for the nonstructural protein processing. The catalytic triad of His, Asp, and Ser residues was also observed in the SD-1 sequence and located at positions 16b8, 16b9, and 17b2, respectively (Fig. 2).

Comparison of nucleotide and amino acid sequences

The inserts identified in CP BVDV Osloss (Meyers et al., 1988c) and NADL (Collett et al., 1989) and a deletion of 41 nucleotides in the 3'UTR of Osloss (Deng and Brock, in preparation) are responsible for the variation of genome sizes (Fig. 3). There is a 270-nucleotide cellular RNA insert in NADL between nucleotides 4992 and 4993 of SD-1 and a 228-base ubiquitin RNA insert in Osloss between nucleotides 5152 and 5153 of SD-1. Our results reveal that, in contrast to CP BVDV, NADL and Osloss, NCP BVDV SD-1 has no RNA insertions along its genome.

The overall nucleotide sequence homologies of SD-1 are 88.6% with NADL, 78.3% with Osloss, 67.1% with HoCV Alfort and 67.2% with HoCV Brescia. The most conserved region is located in the 5'UTR in which the degree of homologies of SD-1 sequence are 93% with NADL, 86% with Osloss, 74% with HoCV Alfort and Brescia (Fig. 3). A moderately high homology is found along most of the viral nonstructural protein region. The viral structural proteins (p14, gp48, gp25, and gp53) and nonstructural proteins p54 and p58 have the lowest homology between SD-1 and other pestivirus sequences (Fig. 3).

The predicted amino acid sequence is more conserved than nucleotide sequence among all the pestiviruses. The overall amino acid sequence homologies of SD-1 are 92.7% with NADL, 86.2% with Osloss, 72.5% with HoCV Alfort, and 71.2% with HoCV Brescia. The p80 is the most conserved viral protein with sequence homologies of SD-1 with NADL of 98%, with Osloss of 95%, and with HoCV Alfort and Brescia of 86% (Fig. 3). Similar to the nucleotide sequence, the amino acid sequence of the viral structural proteins and nonstructural proteins p54 and p58 is more variable than other regions. Extensive analysis of amino acid sequence led to the identification of four conserved domains (designated C1, C2, C3, C4) and three highly variable domains (designated V1, V2, V3) in the viral structural proteins and nonstructural protein p54 region. The regions that had average amino acid sequence homologies of above 90% between SD-1 and other pestiviruses were defined as conserved domains. The average percentages of homology between SD-1 and other pestiviruses were defined as conserved domains. The average percentages of homology between SD-1 and other pestiviruses in C1, C2, C3, and C4 domains are 95, 92, 94, and 90%, respectively. The regions that had average amino acid homologies of below 61% between SD-1 and other pestiviruses were defined as variable domains. The average percentages of homology between SD-1 and other pestiviruses in C1, C2, V1, V2, and V3 domains are 54, 61, and 48%, respectively. A comparison of amino acid sequence in those domains is shown in Fig. 4. The C1, C2, and C3 domains are located in the viral capsid protein P14, glycoproteins gp48 and gp25, respectively. The C4 domain is located in the junction between gp53 and p54. Of three highly variable domains, two (V1 and V2) are located in the same viral glycoprotein gp53 (Fig. 5). The hydrophobicity analysis of the amino acid sequence in this area show that the V1 and V2 domains identified in gp53 are hydrophobic (Fig. 5). The third variable domain,
Strain SD-1 is a NCP BVDV obtained from a persistently infected heifer maintained in isolation. In order to eliminate the possibility of the adaptive selection of virus during cell culture passage, which may change the dominance of virus in the population, and to eventually evaluate the naturally occurring mutation rate of the viral genome in the persistently infected animal in vivo, the persistently infected animal was the source of virus for cDNA cloning. However, the virus titer in serum obtained from the persistently infected heifer was approximately 10^3 logs lower than that in infected cell culture supernatants. Initially, direct cloning of cDNA from SD-1 RNA extracted from serum was attempted. Because of low concentrations of viral RNA in the preparations, no positive clones of SD-1 were obtained after screening of the cDNA library. To overcome this problem, PCR was used to amplify the lower viral RNA levels in the serum before cloning. Based on dot–blot hybridization results (Brock et al., 1992), there was a high homology between SD-1 and NADL nucleotide sequences. It was also reported that the highest sequence homology was located in the nonstructural protein region, particularly in the p80 region (Meyers et al., 1989a; Moormann et al., 1990). Therefore, the first two primers used to amplify SD-1 RNA were designed based on NADL sequence in the p80 region, which is located in the middle of the genome. Because of the lower homology between NADL and SD-1 nucleotide sequence in the regions of the viral glycoproteins and the nonstructural protein p58, NADL sequence primers were not able to amplify SD-1 sequence in these regions. These two gaps, therefore, were filled in following cloning and determination of the 5' and 3'UTR nucleotide sequences of SD-1 and the use of specific SD-1 sequence primers. The cDNA sequence of SD-1 reported represents the genomic sequence of the virus in vivo, which may be different from that of the virus obtained in vitro due to the adaptive selection of virus during cell culture passage.

It was reported that the error rate of Taq polymerase during the DNA polymerization is about 2 × 10^{-5} errors/nucleotide/cycle under standard conditions (Eckert and Kunkel, 1990; Lundberg et al., 1991). In this study, the actual error rate was 4 × 10^{-5} errors/nucleotide/cycle, which is two times higher than the theoretical error rate of Taq polymerase. Therefore, out of 36 “errors,” some may have been created by either reverse transcriptase or DNA polymerase. In addition, some nucleotide differences may have been due to the sequence heterogeneity in the viral RNA population.

The cDNA sequence of NCP BVDV strain SD-1 RNA was determined to be 12,308 nucleotides in length using a 5'-3' ligation strategy to determine the extreme 5' and 3' end sequence. Therefore, it is reasonable to state that the sequence of our cDNA clones represents the complete nucleotide sequence of NCP BVDV SD-1 genome. Of nine 5'-3' ligation clones sequenced, four had one nucleotide and one had two nucleotides shorter than the authentic 3' end sequence. Whether this variation represents the heterogeneity in the viral RNA population or indicates the presence of a low level of RNA exonuclease activity in the viral RNA preparations, resulting in cleavage of 1 or 2 bases at the 3' end of viral RNA prior to the 5'-3' ligation, is unknown.
Comparison of the nucleotide sequence between CP BVDV and HoCV led to the identification of a cellular RNA and a ubiquitin sequence insert within the region coding for the p54 in CP BVDV NADL and Osloss, respectively (Meyers et al., 1989c; Collett et al., 1989). Because the NADL and Osloss strains are cytopathic, it has been proposed that cellular RNA insertion into NCP BVDV genome is responsible for the development of cytopathogenicity (Meyers et al., 1989c). Recently, Meyers et al. (1991) cloned and determined the partial genomic nucleotide sequence in the p125 region of a pair of CP BVDV strain CP1 and NCP BVDV strain NCP1. A ubiquitin insert and p80 nucleotide sequence duplicate were found in the CP1, but not in the NCP1 within this region, which supports the previous hypothesis. However, whether other differences are present in the remainder of the genome is unknown. Comparison of the complete nucleotide sequence of NCP BVDV SD-1 with that of CP BVDV NADL and Osloss revealed that the most remarkable difference is the absence of an inserted cellular RNA sequence in the p125 region.

It has been reported that the p80 protein, generated by either cleavage of p125 or expression from a p80 duplicated region, is the only marker of CP BVDV (Meyers et al., 1991). Therefore, the release of the p80
protein from the viral polyprotein is the key event associated with the development of cytopathogenicity. Although the p125 proteins of the different CP BVDV strains were heterogeneous in size, the processed p80 had the identical size (Akkina, 1991; Greiser-Wilke et al., 1992). These results indicate that the size variation of p125 is determined in the p54 region and the cleavage site at the N-terminal of p80 is present in the authentic BVDV sequence rather than in the cellular inserts (Greiser-Wilke et al., 1992). Some of the CP BVDV strains had p125 almost identically sized with that of NCP BVDV (Akkina, 1991; Greiser-Wilke et al., 1992), suggesting that minor base changes in CP BVDV, such as base substitutions, small insertions, deletions, and gene duplication (Meyers et al., 1991), may be also responsible for the release of p80 from the polyprotein. Comparison of NCP BVDV SD-1 nucleotide sequence with CP BVDV NADL and Osloss at the N-terminal part of p125 revealed that, in addition to the obvious difference of cellular RNA insertion, there are many base substitutions in NADL and Osloss, which result in the change of amino acid sequence, particularly in the V3 domain. Secondary structure prediction of amino acid sequence in this region reveals that some of the amino acid changes alter the secondary structure of the N-terminal portion of p125 (data not shown). Therefore, in addition to the insertion, the base substitutions resulting in conformational change of the N-terminal part of p125 may also contribute to the release of p80 from the polyprotein. It is tempting to propose that the mutations that occurred in NCP BVDV, such as insertion, deletion, duplication, and substitution, result in conformational changes in the N-terminal portion of p125, which also may affect its electrophoretic mobility. Following the conformational change, the cleavage site at the N-terminus of p80 becomes functional and the p80 is released from the viral polyprotein, resulting in the development of CP BVDV from NCP BVDV. The construction of an infectious cDNA clone and the following experimental conversion of both biotypes will help to elucidate the mechanism for cytopathogenicity of BVDV.

The most conserved nucleotide and amino acid sequences are located in the 5'UTR and nonstructural protein p80 respectively, which reflect the functional importance of these two regions for either virus replication or viral RNA translation or viral polyprotein processing. The p80 was reported to possess protease activity which is responsible for the processing of all the viral nonstructural proteins (Wiskerchen and Collett, 1991). In addition, a helicase motif was also predicted within the p80 region (Gorbalenya et al., 1989b; and Moormann et al., 1990). The possible regulatory elements and the secondary structure of 5'UTR of BVDV RNA will be analyzed and reported elsewhere (Deng and Brock, in preparation).

The comparison of both nucleotide and amino acid sequence of SD-1 with that of other pestiviruses revealed the high heterogeneity in the region of viral structural proteins and nonstructural proteins p54 and p58. It is reasonable to assume that the conserved domains identified in this region are under highly functional selection and are critical for either viral RNA packaging or virion assembly or important for viral interaction with receptors on infected cells. However, the functions of these domains remain to be determined. It has been reported repeatedly that neutralizing monoclonal antibodies against BVDV bound to and immunoprecipitated glycoprotein gp53 (Magar et al., 1988; Donis et al., 1988; Bolin et al., 1988; Xue et al., 1990; Weiland et al., 1990). Antigenic variation of this protein among BVDV isolates was demonstrated by
the fact that none of the monoclonal antibodies neutralized all the BVDV isolates tested (Magar et al., 1988; Xue et al., 1990). The amino acid sequence analysis revealed that two hypervariable domains (V1, V2) are located in the viral glycoprotein gp53. The hydrophobicity analysis indicated that these two domains are hydrophilic, suggesting that V1 and V2 domains are present on the outer surface of virions. Therefore, it is proposed that V1 and V2 domains in gp53 may represent the protective epitopes of this protein, which are the important targets for the immune response of the host, hence under high immunological pressure. The high diversity of amino acid sequence within the protective epitopes of gp53 among BVDV and HoCV implies the difficulty in developing effective vaccine for preventing pestivirus infections by expression of this immunodominant glycoprotein. Recently, Weiland et al. (1992) reported that antibodies to a second envelope glycoprotein gp44/48 were able to mediate neutralization of HoCV. Whether the corresponding glycoprotein gp48 in BVDV has the same ability to induce neutralizing antibodies remains to be investigated. Amino acid sequence analyses indicated that glycoprotein gp48 was more conserved than gp53 among pestiviruses. Furthermore, a highly conserved domain (C2) was observed within this protein. Therefore, gp48 may be a candidate for a subunit vaccine as well. Similar to pestiviruses, the variable and hypervariable domains were also identified in the regions of Hepatitis C virus (HCV) corresponding to the gp53 of pestivirus (Weiner et al., 1991). However, no similar hypervariable regions have been reported in the corresponding envelope E protein and NS5 protein of flaviviruses (Chu et al., 1989; Deubel et al., 1986; Lobigs et al., 1988; Hahn et al., 1988, lrie et al., 1989).

The largest variable domain was identified within the first nonstructural protein p54, the N-part of precursor protein p125. The characteristic feature of this protein is that a highly hydrophobic stretch including 266 amino acids is present in this protein and it was usually not detected by immunoprecipitation (Moonnig and Plagemann, 1992). The V3 domain is located in the highly hydrophobic stretch of this protein. Although the primary amino acid sequence is hypervariable, the hydrophobic character of this domain is maintained in all the pestiviruses, indicating that the hydrophobicity rather than the primary sequence of this protein is under the functional selection. Although a membrane-associated function of this protein was proposed (Wiskerchen and Collett, 1991), further work is needed to elucidate its function.

ACKNOWLEDGMENTS

Salaries and research support were provided by state and federal funds appropriated to the Ohio Agricultural Research and Development Center, The Ohio State University. Journal article no. 200-92. This work was supported in part by Cooperative Research Agreement 88-34116-3563 from the U.S. Department of Agricultural Science and Education Administration. The authors thank Dr. M. S. Collett for kindly providing BVDV NADL cDNA clones, Sylva Riblet and Jerry Meitzler for excellent technical assistance, and Fengmin Tian for typing the manuscript.

REFERENCES


