

1-1-1991

## Comparison of the nucleotide and deduced amino acid sequences of the S genes specified by virulent and avirulent strains of bovine coronaviruses

Xuming Zhang  
*School of Veterinary Medicine*

Konstantin G. Kousoulas  
*School of Veterinary Medicine*

Johannes Storz  
*School of Veterinary Medicine*

Follow this and additional works at: [https://digitalcommons.lsu.edu/animalsciences\\_pubs](https://digitalcommons.lsu.edu/animalsciences_pubs)

---

### Recommended Citation

Zhang, X., Kousoulas, K., & Storz, J. (1991). Comparison of the nucleotide and deduced amino acid sequences of the S genes specified by virulent and avirulent strains of bovine coronaviruses. *Virology*, 183 (1), 397-404. [https://doi.org/10.1016/0042-6822\(91\)90154-4](https://doi.org/10.1016/0042-6822(91)90154-4)

This Article is brought to you for free and open access by the School of Animal Sciences at LSU Digital Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of LSU Digital Commons. For more information, please contact [ir@lsu.edu](mailto:ir@lsu.edu).

## Comparison of the Nucleotide and Deduced Amino Acid Sequences of the S Genes Specified by Virulent and Avirulent Strains of Bovine Coronaviruses

XUMING ZHANG, KONSTANTIN G. KOUSOULAS, AND JOHANNES STORZ<sup>1</sup>

*Department of Veterinary Microbiology and Parasitology, School of Veterinary Medicine,  
Louisiana State University, Baton Rouge, Louisiana 70803*

*Received February 5, 1991; accepted April 1, 1991*

The entire nucleotide sequences of the spike glycoprotein (S) genes of the highly virulent bovine coronavirus (BCV) strain BCV-LY138, the avirulent BCV-L9 and related Norden Vaccine (BCV-Vaccine) strains were determined using the polymerase chain reaction (PCR) to amplify cDNAs obtained by reverse transcription of viral RNA, and to produce single strand cDNAs for DNA sequencing. The S gene sequences of these viral strains were compared with those of recently published strains BCV-Mebus, BCV-Quebec, and BCV-F15. An open reading frame of 4092 nucleotides, encoding a protein of 1363 amino acid residues, was found in all six strains. Frameshifts and insertions or deletions were not observed except for the BCV-F15. The S gene sequences were more than 98% conserved overall in spite of different origins of the six viruses. There were 45 to 56 nt differences between the virulent and avirulent groups while there were 6 to 14 nt differences among four avirulent strains. Comparison of the deduced amino acid sequences indicated that the S proteins had typical properties of membrane glycoproteins. Nineteen N-linked glycosylation sites were predicted in five strains, and 18 of them were conserved in the avirulent strain BCV-L9. The sequence KRRSRR at the predicted proteolytic cleavage site was identified in five strains while the sequence KRRSVR was found in BCV-F15. Substitutions of few amino acids in the putative fusogenic domains and two prolines at 507 and 567 in the antigenic domains may cause altered immunogenic and other functional properties of the S proteins specified by the virulent and avirulent BCV strains. Nine amino acid substitutions between the virulent and avirulent groups may correlate with BCV virulence. © 1991 Academic Press, Inc.

Bovine coronavirus (BCV) is a member of coronaviridae, causing severe diarrhea in newborn calves and winter dysentery in adult cattle (1, 2). It possesses a single-stranded, nonsegmented RNA genome with positive polarity (3). The virion contains four major structural proteins: the nucleocapsid protein (N), the transmembrane protein (M), the hemagglutinin/esterase protein (HE) and the spike protein (S) (4). The S glycoprotein is a predominant peplomeric structure forming the typical coronavirus morphology. It is synthesized as a high molecular weight (mol. wt.) precursor (gp190) which is cleaved to yield two comigrating subunit polypeptides: the N-terminal half (S1) and the C-terminal half (S2) with an approximate mol. wt. of 100 kDa (3, 5). The S glycoprotein functions in virus attachment to permissive cells, virus-induced cell fusion, elicitation of neutralizing antibodies and cell-mediated immunity (6). Two antigenic domains responsible for neutralization of BCV-Quebec and BCV-L9 have been identified on the S protein (Ref. (7); Hussain *et al.*, submitted). However, the location of the neutralizing epitopes on the BCV S protein has not yet been determined. Comparison of the S proteins with S-spe-

cific monoclonal antibodies (MAbs) revealed that different neutralizing and non-neutralizing epitopes as well as conformational epitopes were present in the S glycoproteins specified by the virulent and avirulent strains (Hussain *et al.*, submitted). In order to understand the molecular basis of the observed antigenic diversity of the S glycoproteins we successfully amplified, directly sequenced, and cloned the entire S genes of different BCV strains using the polymerase chain reaction (PCR). We report here the sequence characteristics of one virulent and two avirulent BCV strains and their comparison with the recently published sequences of three other BCV strains.

The strain BCV-L9 was derived from BCV-Mebus and passaged through different nonpolarized and highly polarized cells over 80 passages (8). The Norden Vaccine strain (BCV-Vaccine) was also derived from BCV-Mebus and used as vaccine by Norden Laboratories (Omaha, NE) (1, 9). The highly virulent wild-type strain BCV-LY138 was isolated from diarrheal fluid of a diseased calf in 1965 in Utah/USA (10), and maintained in calves through oral inoculation since then (8, 11). This strain replicates only in HRT-18 cells but not in numerous bovine cells (8). All strains were propagated in HRT-18 cells as described previously (8). After 24 h.p.i., cells were washed twice with phos-

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

phate buffered saline (PBS), and RNA was isolated using isothiocyanate/cesium chloride gradients according to the method described previously (12). As control, RNA was isolated from uninfected cells.

The first-strand cDNA synthesis was carried out in a volume of 25  $\mu$ l containing 50 mM Tris-HCl (pH 8.3), 50 mM KCl, 3 mM MgCl<sub>2</sub>, 10 mM DTT, 1 mM each of the four dNTPs, 25 U RNasin (Bethesda Research Laboratories (BRL)), 2'  $\mu$ M 3' primer, 1  $\mu$ g total RNA, 200 U Moloney murine leukemia virus reverse transcriptase (M-MLV RT) (BRL) for 1 hr at 42°, and the reaction was stopped by heating samples to 95° for 10 min and then cooled on ice. The RNA was denatured after adding 20 mM methylmercuric hydroxide (MeHgOH) and 70 mM 2-mercaptoethanol for 7 and 5 min at room temperature, respectively, incubated with oligonucleotide primer at 65° for 2 min, and then chilled on ice before adding other reagents. Two S gene specific primers were designed both for cDNA synthesis and PCR amplification according to the published sequence (13). The A3' primer is antisense representing the sequence at the positions 2165–2187 (5'-TTTGGATCCAGGTTGCAGCTGTCGTGAAAGA-3'). The B3' primer corresponds to the sequence downstream of the S gene (5'-AACGGATCCAATATATCGTCAGGAGCC-AATA-3'). Both primers contain an extra *Bam*HI linker in the 5' end for cloning purposes. The primers were synthesized in the Gene Assembler (Pharmacia-LKB, Piscataway, NJ) in our laboratory according to the manufacturer's instructions, and purified using Poly-PAK (Glen Research, Herndon, VA) or Oligo-PAK column (Millipore, Burlington, MA).

The reaction of second strand cDNA synthesis and amplification was carried out in PCR (MicroCycler, Eppendorf Inc., Fremont, CA) with 5  $\mu$ l of the reverse transcription reaction in a total volume of 100  $\mu$ l containing 20 mM Tris-HCl (pH 8.3), 25 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.1% Tween 20, 10  $\mu$ g of bovine serum albumin (BSA), 2.5 U of *Taq* DNA polymerase (BRL), 200  $\mu$ M each of the four dNTPs, and 20 pmol each of the primers. The sense A5' primer corresponds to the sequence upstream of the S gene (5'-GCTGAATTCGATAATGGT-ACTAGGCTGCATGAT-3'), and the B5' primer represents the sequence at the positions 2165–2187 (5'-GCTGAATTCCTCTTTCACGACAGCTGCAACCT-3'). These primers contain an extra *Eco*RI linker on the 5' end as indicated. Each cycle consisted of a 1-sec denaturation at 96°, followed by 30-sec annealing (at 59°) and 2-min extension (at 72°) steps. After 30 to 35 cycles, the final products were extended for 7 min at 72°. Single-strand cDNA fragments were generated in a second PCR for sequencing. Briefly, one-tenth of the double stranded PCR product and a single primer with 40 pmol were used. Other reagents were the same as

in the previous PCR. The cycle profile of 20 to 25 cycles was: 1-sec at 96°, 30-sec at 50°, 1-sec at 60°, 2-min at 72° (Fig. 1).

We sequenced the entire S genes of one virulent and two avirulent BCV strains which display distinct biological properties and epitopes as defined by MAbs to the reference strain BCV-L9. With the exception of 40 nt on the 5'-end and 50 nt on the 3'-end, the S gene sequences were determined in both directions at least once, as illustrated in Fig. 1. These sequences were compared with the recently published sequences for the strains BCV-Mebus (14), BCV-Quebec (15), and BCV-F15 (13), and their alignment is presented in Fig. 2 (origins of these strains are described in the legend). All of the S genes contained an ORF of 4092 nucleotides. The only variations among these sequences consisted of nucleotide substitutions. Frameshift, deletion, or insertion, and nonsense mutations were not observed except for BCV-F15. The few nucleotide substitutions (91 nt) represented approximately 2% of the sequence, and they seemed to be distributed randomly. By comparison with other strains, the following differences were detected in two regions in BCV-F15: The deletion of a base (T) at nucleotide 2316 resulted in a frameshift for 2 aa followed by the insertion of an A at nucleotide 2320 which returned the amino acid sequence to a homologous alignment (at aa position 772 within the cleavage site; also see Fig. 3); at the positions 3271–3272, the sequence GC was converted to CG. The intergenic and internal consensus sequence CYAAAC was conserved in all BCV strains. The two

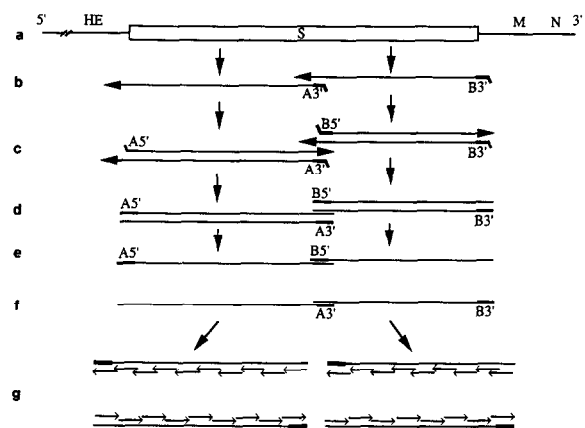


FIG. 1. Strategies of cDNA synthesis and DNA sequencing. The relative locations of S, M and N genes in the BCV genome (a); primer design and first-strand cDNA synthesis by reverse transcription (b); second-strand, PCR-mediated DNA synthesis (c); double-stranded DNA amplification by PCR (d); single-stranded DNA synthesis by PCR using an excess of 5' primer (e) and 3' primer (f); DNA sequencing strategies (g). A5' and B5' indicate two 5' primers containing an extra *Eco*RI linker; A3' and B3' indicate two 3' primers containing an extra *Bam*HI linker.

BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TTAGACCATA ATCTAACAT <u>GTTTITGATA</u> CTTTAAATT CCTTACCAAT GGCTTTTGCT GTTATAGGAG ATTTAAAGTG TACTACGGTT TCCATTAATG	100
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	ATGTTGACAC CGGTGCTCCT TCTATTAGCA CTGATATTGT CGATGTTACT AATGGTTTAG GTACTTATTA TGTTTTAGAT CGTGTGTATT TAAATACTAC	200
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	GTTGTTGCTT AATGGTTACT ACCCTACTTC AGGTTCTACA TATCGTAATA TGGCACTGAA GGAACCTTTA CTATTGAGCA GACTATGGTT TAAACCACCT	300
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TTTCTTCTG ATTTTATTAA TGGTATTTTT GCTAAGGTCA AAAATACCAA GGTATTATTA AAGGGTGTA TGTATAGTGA GTTTCCTGCT ATAACATAG	400
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	GTAGTACTTT TGTAAATACA TCCTATAGTG TGGTAGTACA ACCACATACT ACCAATTGGG ATAATAAATT ACAAGGTCTC TTAGAGATCT CTGTTTGCCA	500
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	GTATACTATG TGCAGTACC CACATACGAT TTGTCTCCT AATCTGGGTA ATAAACGCGT AGAATCTATG CATTGGGATA CAGGTGTGTG TTCTCTTTA	600
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TATAAGCGTA ATTTACATA TGATGTGAAT GCTGATTACT TGTATTTCCTA TTTTATCAA GAAGGTGGTA CTTTTATGC ATATTTTACA GACACTGGTG	700
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TTGTACTAA GTTCTGTTT AATGTTTATT TAGGCACGGT GCTTTCACAT TATTATGTCC TGCCITTTGAC TTGTAATAGT GCTATGACTT TAGAATATTG	800
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	GGTTACACCT CTCACCTCTA <u>AACAATATTT</u> ACTAGCTTTC AATCAAGATG GTGTTATTTT TAATGCTGTT GATTGTAAGA GTGATTTTAT GAGTGAGATT	900
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	AAGTGTAATA CACTATCTAT AGCACCATCT ACTGGTGTIT ATGAATTAAA CGGTACACT GTTCAGCCAA TTGCAGATGT TTACCGACGT ATACCTAATC	1000
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TTCCCGATTG TAATATAGAG GCTTGGCTTA ATGATAAGTC GGTGCCCTCT CCATTAAATT GGAACGTAA GACCTTTTCA AATTGTAATT TTAATATGAG	1100
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	CTGCCTGATG TCTTTTATTC AGGCAGACTC ATTTACTTGT AATAATATTG ATGCTGCTAA GATATATGGT ATGTGTTTTT CCAGCATAAC TATAGATAAG	1200
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TTTGCTATAC CCAATGGTAG GAAGGTTGAC CTACAATTGG GCAATTGGG CTATTGCGAG TCTTTTAACT ATAGAATTGA TACTACTGCT ACAAGTTGTC	1300
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	AGTTGTATTG TAATTTACCT GCTGCTAATG TTTCTGTTAG CAGGTTTAACT CCTTCTACTT GGAATAGGAG ATTTGGTTTT ACAGAACAAT CTGTTTTTAA	1400
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	GCCTCAACCT GTAGGTGTTT TTAATCATCA TGATGTTGTT TATGCACAAC ATTGTTTAA AGCTCCACA AATTCTGTC CGTGTAATTT GGATGGGTCT	1500

Fig. 2. Nucleotide sequence comparisons of the S genes of six BCV strains. DNA sequencing was carried out with the modified dideoxynucleotide chain termination procedure (24) using Sequenase (U.S.B., Cleveland, OH). Sequences were analyzed with the aid of the Sequence Analysis Software Package of the Genetics Computer Group of the University of Wisconsin and the MacVector Software (IBI, New Haven, CT). The S gene sequences of BCV-Mebus, BCV-Quebec, and BCV-F15 are obtained from Refs. (14), (15), and (13), respectively. The origins and properties of these strains are as follows: The strain BCV-Mebus was isolated from a calf with enteritis, and adapted to bovine fetal kidney cells described previously (25). The strain BCV-Quebec was isolated from a calf with enteritis by cultivation in Vero and MDBK cells in Quebec in 1979. Its biological properties were similar to BCV-Mebus (26). The wild-type strain BCV-F15 was isolated in HRT cells from a calf with enteritis in France in 1979 (27). The consensus sequences are underlined, and the start and stop codes are marked by asterisks.

BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TTGTGTGTAG GTAATGGTCC TGGTATAGAT GCTGGTTATA AAAATAGTGG TATAGGCACT TGTCTGCAG GTACTAATTA TTAACTTGC CATAATGCTG	1600
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	CCCAATGTGA TTGTTTGTGC ACTCCCGACC CCATTACATC TAAATCTACA GGGCCTTACA AGTGCCCCCA AACTAAATAC TTAGTTGGCA TAGGTGAGCA	1700
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	CTGTTGCGGT CTGCTATTA AAAGTGATTA TTGTGGAGGT AATCCTTGTA CTGCGCAACC ACAAGCATTT TTGGGTTGGT CTGTTGACTC TTGTTTACAA	1800
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	GGGGATAGGT GTAATATTTT TGCTAATTTT ATTTTGCATG ATGTTAATAG TGGTACTACT TGTCTACTG ATTTACAAAA ATCAAACACA GACATAATTC	1900
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TTGGTGTGTTG TGTTAATTAT GATCTTTATG GTATTACAGG CCAAGGTATT TTTGTTGAGG TTAATGCGCC TTATTATAAT AGTTGGCAGA ACCTTTTATA	2000
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TGATTCTAAT GGTAACTCTT ATGGTTTTAG AGACTACTTA ACAAACAGAA CTTTTATGAT TCGTAGTTGC TATAGCGGTC GTGTTTCAGC GGCCTTTCAT	2100
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	GCTAACTCTT CCGAACCAGC ATTGCTATTT CGGAATATTA AATGCAATTA CGTTTTTAAT AATACTCTTT CACGACAGCT GCAACCTATT AACTATTTTG	2200
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	ATAGTTATCT TGGTTGTGTT GTCAATGCTG ATAATAGTAC TTCTAGTGT GTTCAAACAT GTGATCTCAC AGTAGGTAGT GGTACTGTG TGGATTACTC	2300
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TACAAAAAGA CGAAGTCGTA GAGCGATTAC CACTGGTTAT CGGTTTACTA ATTTTGAGCC ATTTACTGTT AATTCAGTAA ATGATAGTTT AGAACCTGTA	2400
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	GGTGGTTTGT ATGAAATCA AATACCTTCA GAGTTTACTA TAGGTAATAT GGAGGAGTTT ATTCAAACAA GCTCTCTTAA AGTTACTATT GATTGTTCTG	2500
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	CTTTTGTCTG TGGTGATTAT GCAGCATGTA AATCACAGTT GGTGAATAT GGTAGCTTCT GTGACAATAT TAATGCTATA CTCACAGAAG TAAATGAAC	2600
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	ACTTGACACT ACACAGTTGC AAGTAGCTAA TAGTTTAATG AATGGTGTCA CTCTTAGCAC TAAGCTTAA GATGGCGTTA ATTTCAATGT AGACGACATC	2700
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	AATTTTCCC CTGTATTAGG TTGTTTAGGA AGCGATTGTA ATAAAGTTTC CAGCAGATCT GCTATAGAGG ATTTACTTTT TTCTAAAGTA AAGTTATCTG	2800
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	ATGTCGGITT CGTTGAGGCT TATAATAATT GTACTGGAGG TGCCGAAAT AGGGACCTCA TTTGTGTGCA AAGTTATAAT GGTATCAAAG TGTGCTCTC	2900
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	ACTGCTCTCA GTAAATCAGA TCAGTGGATA CACTTGGCT GCCACCTCTG CTAGTCTGTT TCCTCCTTGG TCAGCAGCAG CAGGTGTACC ATTTTATTTA	3000

Fig. 2—Continued

BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	AATGTTCACT ATCGTATTAA TGGGATTGGT GTTACCATGG ATGTGTTAAG TCAAAATCAA AAGCTTATTG CTAATGCATT TAACAATGCT CTTGATGCTA	3100
	TC C	
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TTGAGGAAGG GTTGTATGCT ACCAATTCTG CTTTAGTTAA AATTCAGCT GTTGTTAATG CAAATGCTGA AGCTCTTAAT AACTTATTGC AACAACTCTC	3200
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TAATAGATTT GGTGCTATAA GTTCTTCTTT ACAAGAAATT CTATCTAGAC TGGATGCTCT TGAAGCGCAA GCTCAGATAG ACAGACTTAT TAATGGGCGT	3300
	C T CG	
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	CTTACCGCTC TTAATGCTTA TGTCTTCTCA CAGCTTAGTG ATTCTACACT AGTAAATTT AGTGCAGCAC AAGCTATGGA GAAGTTAAT GAATGTGTCA	3400
	T T	
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	AAAGCCAATC ATCTAGGATA AATTTTGTG GTAATGGTAA TCATATTATA TCATTAGTGC AGAATGCTCC ATATGGTTTG TATTTTATCC ACTTTAGCTA	3500
	T	
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TGTCCCTACT AAGTATGTCA CTGCGAAGGT TAGTCCCGGT CTGTGCATTG CTGGTGATAG AGGTATAGCC CCTAAGAGTG GTTATTTTGT TAATGTAAAT	3600
	G	
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	AATACTTGGA TGTTCCTGG TAGTGGTTAT TACTACCTTG AACCCATAAC TGGAAATAAT GTTGTGTGTA TGAGTACCTG TGCTGCTAAC TATACTAAAG	3700
	C T T T T T T T T T T C	
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	CGCCGGATGT AATGCTGAAC ATTCAACAC CCAACCTCCA TGATTTTAAG GAAGAGTTGG ATCAATGGTT TAAAAACCA ACATCAGTGG CACCAGATT	3800
	A C C	
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	GTCACCTGAT TATATAAATG TTACATTCTT GGACCTACAA GATGAAATGA ATAGGTTACA GGAGGCAATA AAAGTTTTAA ATCAGAGCTA CATCAATCTC	3900
	C	
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	AAGGACATTG GTACATATGA GTATTATGTA AAATGGCCTT GGTATGTATG GCTTTTAATT GGCTTTGCTG GTGTAGCTAT GCTTGTTTTA CTATTCTTCA	4000
	C	
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TATGCTGTTG TACAGGATGT GGGACTAGTT GTTTAAGAT ATGTGGTGGT TGTGTGTATG ATTATACTGG ACACCAGGAG TTAGTAATTA AAACATCACA	4100
	A A	
BCV-L9 BCV-VACC BCV-MEB BCV-QUEB BCV-LY BCV-F15	TGACGACTAA *** 4110	

FIG. 2—Continued

most similar sequences were those of BCV-L9 and BCV-Vaccine (6 nt differences) in paired comparisons, and the most distant were those of BCV-Quebec and BCV-F15 (74 nt differences). The virulent wild-type strain BCV-LY138 was evolutionally distant from the avirulent strains (total number of differences between 45 to 56 nt). The French strain BCV-F15 also differed from the avirulent strains (total number of differences between 55 to 74 nt). Although there were 55 nt differences between BCV-LY138 and BCV-F15, the same nucleotide substitutions occurred in 29 nt.

An alignment of the deduced amino acid sequences

of the S proteins of the six BCV strains is shown in Fig. 3. All six S genes encoded a predicted protein of 1363 aa residues, having a mol. wt. of approximately 150 kDa. Nineteen potential N-linked glycosylation sites were predicted in 5 strains while 18 of them were conserved in BCV-L9. Fifty-seven cysteine residues were found in BCV-L9 while 56 of them were conserved in the other 5 strains. The BCV-F15 had an additional cysteine at position 840. The glycosylated protein had an estimated mol. wt. of approximately 190 kDa. All six S proteins possessed at least 98% sequence identity in paired comparisons. As with the nucleotide se-

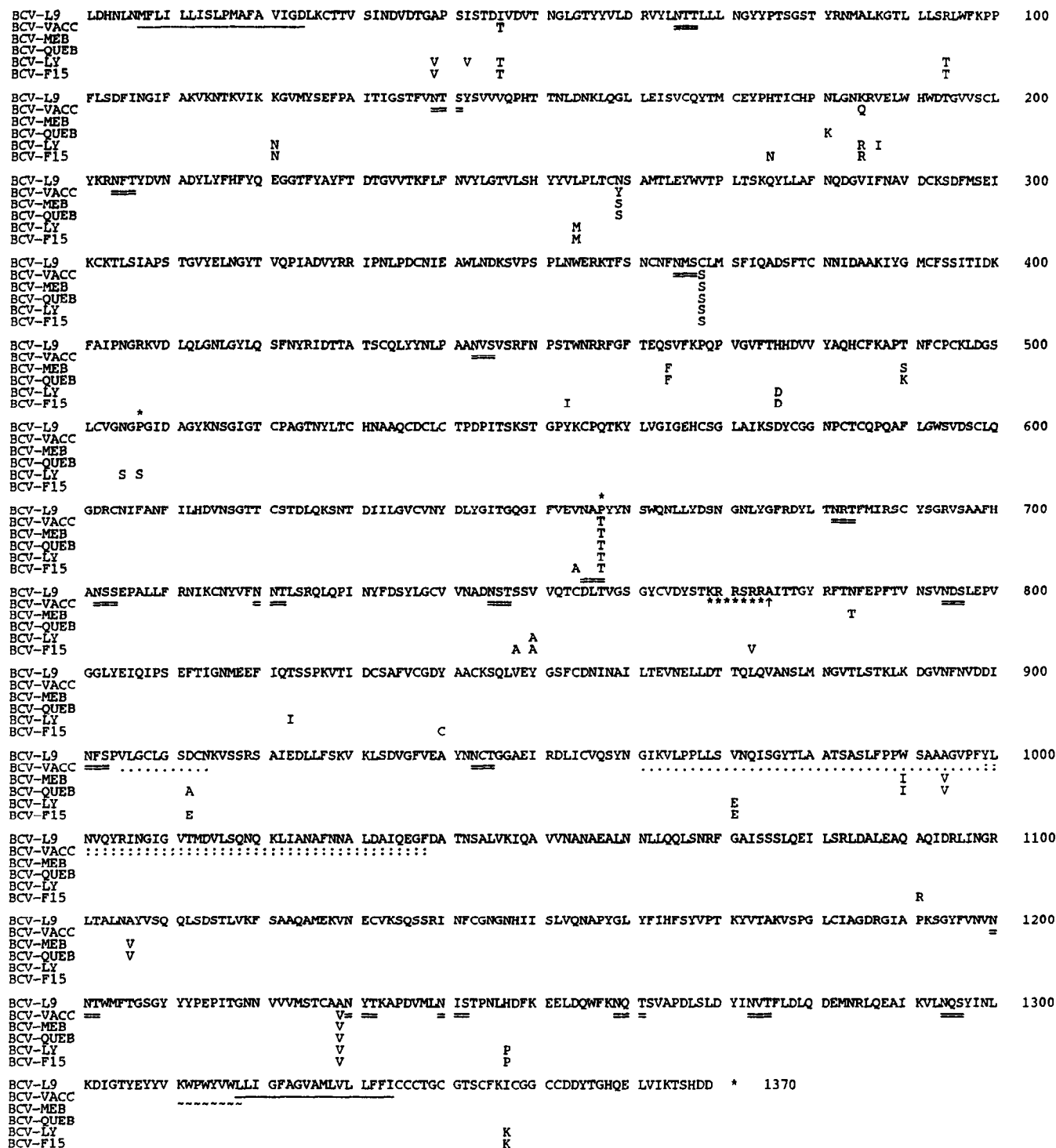


Fig. 3. Amino acid sequence comparisons of the predicted S proteins of six BCV strains. The predicted signal peptide and intramembrane anchoring sequences are underlined. The predicted cleavage site is indicated by an arrow, and its sequence is marked by asterisks underneath. The 19 predicted N-linked glycosylation sites are double underlined. The conserved sequence is marked by (~). The putative hydrophobic domains are indicated by (.), and the adjacent heptad repeat sequences by (:). Two proline substitutions are indicated by asterisks on the top.

quences, the most related proteins were those of BCV-Mebus and BCV-Quebec (4 aa differences) or BCV-L9 and BCV-Vaccine (6 aa differences); the most diver-

gent proteins were those of BCV-Mebus and BCV-F15 or BCV-Quebec and BCV-F15 (26 aa differences, respectively). The alignment reveals that the strain BCV-

Quebec is closely related to the avirulent strains BCV-Mebus, BCV-L9, and BCV-Vaccine (6 to 14 nt and 4 to 11 aa differences, respectively), suggesting that this strain is a variant of the prototype BCV-Mebus. In contrast to the nucleotide sequences, the amino acid sequences of BCV-LY138 and BCV-F15 were more homologous (13 aa differences). As shown in Fig. 3, 9 amino acids in the avirulent strains were substituted in both BCV-LY138 and BCV-F15: A to V (38), R to T (94), K to N (121), L to M (254), H to D (476), V to A (749), V to E (980), A to V (1106), H to P (1247), and I to K (1347). This suggests that the wild-type strain BCV-F15 is possibly a virulent strain, and that the 9 amino acids, especially the proline (at 1247) and lysine (at 1347) on the carboxy-terminal, may relate to BCV virulence. Amino acid differences between the virulent and avirulent groups varied from 16 to 26 aa. Apparently, most of these substitutions occurred in the S1 peptide (Fig. 3).

Comparisons of the amino acid sequences in the antigenic region (14, 16) suggested that two proline substitutions may cause the antigenic differences: at amino acid position 507, a proline in the avirulent strain BCV-L9 was substituted by a serine in the virulent strain BCV-LY138, and another proline (at 567) was substituted by a threonine in all other strains. Proline is considered a helix-breaking residue, which may influence significantly the conformation and secondary structure of peptide molecules (17). The proline substitutions resulted in decreased hydrophilicity, surface probability, and antigenicity, but increased flexibility of the peptide in this region (data not shown). Our previous study indicated that one neutralizing epitope of BCV-L9 was present in BCV-Vaccine but absent in the virulent strain BCV-LY138, and one non-neutralizing epitope was present only in the avirulent strain BCV-L9 (Hussain *et al.*, submitted). The number of proline substitutions in this region coincides with the number of conformational epitopes absent in BCV-LY138 and BCV-Vaccine lead us to conclude that substitutions of these prolines to other amino acids may alter the conformation of the S protein, and cause the loss of conformation-dependent epitopes in BCV-LY138 and BVC-Vaccine. This hypothesis is currently under investigation. It is unclear whether substitutions of these prolines and other amino acids in this region also correlate with BCV virulence.

Our previous studies revealed that the biological properties of fusion, plaque formation and host cell range were evidently different between the virulent and avirulent groups of BCV strains (8). The fusion activity of BCV is believed to be associated with the S polypeptide and the cleavage of the S into S1 and S2 subunits (5, 18). The sequence KRRSRR at the predicted proteo-

lytic cleavage site was conserved in the virulent strain BCV-LY138 and in the avirulent strains while the sequence KRRSVR occurred in the wild-type BCV-F15 (Fig. 3). The cleavage site of BCV S protein is located in a hydrophilic area, in contrast to paramyxoviruses and myxoviruses, in which the cleavage site is located in hydrophobic domains (20–30 aa) of the fusion proteins (19, 20). These observations imply that there might be a different mechanism involved in BCV-induced cell fusion. Parker *et al.* (14) indicated that a sequence VLGCLGSAC (905–913) on the S protein of BCV-Quebec may constitute a portion of the fusogenic domain within the BCV S2 subunit. This stretch is similar to the sequence LLGCIGSTC of MHV-A59, which contains a neutralizing epitope (25). By comparing with paramyxovirus and retrovirus, Chambers *et al.* (26) proposed that the hydrophobic region adjacent to heptad repeat sequences may be the potentially fusion-related domain in the S proteins of mouse hepatitis virus, infectious bronchitis virus, and transmissible gastroenteritis virus. After multiple alignments it is found that the heptad repeat sequences are located in the S2 at positions 999–1038 which are conserved in all strains, and the hydrophobic regions are located at 961–998 (see Fig. 3). However, a direct involvement of these sequences in BCV fusion has yet to be demonstrated. Similarly, the E1 protein of Semliki Forest virus (SFV) does not contain a hydrophobic N-terminal region. It has been proposed, however, that in E1 an internal uncharged stretch of 17 residues, located about 80 aa from the N-terminal, might act as the putative fusogenic domain (27). Interestingly, we found three patterns of amino acid substitution in the putative fusogenic domain: at position 971, V in the avirulent strains was substituted by E in the virulent strains; W (898) and A (993) were substituted by I and V in BCV-Mebus and BCV-Quebec, respectively. It is unclear, however, whether these amino acid changes in the predicted cleavage sites and the putative fusogenic domains reflect any functional differences, such as fusion activities among different BCV strains.

## ACKNOWLEDGMENTS

This work was supported by Grants 86-CRSR-2-2871 and 89-34116-4675 from the United States Department of Agriculture, and Louisiana Education Quality Support Fund of the Board of Regents, State of Louisiana, to J.S. and K.G.K., and Grant AI87886 to K.G.K. We thank Mamie Burrell and Li-Ju T. Huang for excellent technical assistance, Susan Newman for computer assistance, and Dr. B. Kaltenboeck for valuable discussion on PCR technique.

## REFERENCES

1. MEBUS, C. A., STAIR, E. L., RHODES, M. B., and TWIEHAUS, M. J., *Am. J. Vet. Res.* **34**, 145–150 (1973).



2. SAIF, L. J., REDMAN, D. R., BROCK, K. V., KOHLER, E. M., and HECKERT, R. A., *Vet. Rec.* **123**, 300–301 (1988).
3. SPAAN, W., CAVANAGH, D., and HORZINEK, M. C., *J. Gen. Virol.* **69**, 2939–2952 (1988).
4. CAVANAGH, D., BRIAN, D. A., ENJUANES, L., HOLMES, K. V., LAI, M. M. C., LAUDE, H., SIDDELL, S. G., SPAAN, W., TAGUCHI, F., and TALBOT, P. J., *Virology* **176**, 306–307 (1990).
5. ST. CYR-COATS, K., STORZ, J., HUSSAIN, K. A., and SCHNORR, K. L., *Arch. Virol.* **103**, 35–45 (1988).
6. STURMAN, L. S., RICHARD, C. S., and HOLMES, K. V., *J. Virol.* **56**, 904–911 (1985).
7. DEREGT, D., and BABIUK, L. A., *Virology* **68**, 410–420 (1987).
8. ST. CYR-COATS, K., and STORZ, J., *J. Vet. Med. B* **35**, 48–56 (1988).
9. SHAPEE, R. L., MEBUS, C. A., and BASS, E. P., *Am. J. Vet. Res.* **37**, 1031–1041 (1976).
10. HAJER, I., and STORZ, J., *Am. J. Vet. Res.* **39**, 441–444 (1978).
11. HAJER, I., and STORZ, J., *Arch. Virol.* **59**, 47–57 (1979).
12. KINGSTON, R. E., In "Current Protocols in Molecular Biology" (F. M. Ausubel, et al., Eds). pp. 4.2.1.–4.2.5. Greene Wiley-Interscience, New York, 1989.
13. ABRAHAM, S., KIENZLE, T. E., LAPPS, W., and BRIAN, D. A. *Virology* **176**, 296–301 (1990).
14. PARKER, M. D., YOO, D., COX, G. J., and BABIUK, L. A., *J. Gen. Virol.* **71**, 263–270 (1990).
15. BOIREAU, P., CRUCIERE, C., and LAPORTE, J., *J. Gen. Virol.* **71**, 487–492 (1990).
16. DEREGT, D., PARKER, M. D., COX, G. C., and BABIUK, L. A., *J. Gen. Virol.* **70**, 647–658 (1989).
17. CREIGHTON, T. E., In "Proteins," pp. 159–197. W. H. Freeman, New York, 1984.
18. STORZ, J., ROTT, R., and KALUZA, G., *Infect. Immun.* **31**, 1214–1222 (1981).
19. BLUMBERG, B. M., GIORGI, C., ROSE, K., and KOLAKOSKY, D., *J. Gen. Virol.* **66**, 317–331 (1985).
20. GETTING, M., BYE, J., SKEHEL, J., and WATERFIELD, M., *Nature (London)* **287**, 301–306 (1980).
21. LUYTJES, W., GEERTS, D., POSTHUMUS, W., MELOEN, R., and SPAAN, W., *J. Virol.* **63**, 1408–1412 (1989).
22. CHAMBERS, P., PRINGLE, C. R., and EASTON, A. J., *J. Gen. Virol.* **71**, 3075–3080 (1990).
23. GAROFF, H., FRISCHAUF, A., SIMONS, K., LEHRACH, H., and DELIUS, H., *Nature (London)* **288**, 236–241 (1980).
24. SANGER, F., NICKLEN, S., and COUSON, A. R. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467 (1977).
25. MEBUS, C. A., STAIR, E. L., RHODES, M. B., and TWIEHAUS, M. J., *Vet. Pathol.* **10**, 45–64 (1973).
26. DEA, S., ROY, R. S., and BEGIN, M. E., *Am. J. Vet. Res.* **41**, 30–38 (1980).
27. LAPORTE, J., HARIDON, R. L., and BOBULESCO, P. In "Enterites Virales/Viral Enteritis in Human and Animals" (F. Bricout and R. Scherrer, Eds.), pp. 99–102. INSERM, Paris, 1980.