

## **Comparative sequence analysis of a polymorphic region of the spike glycoprotein S1 subunit of enteric bovine coronavirus isolates**

**M. R. Rekik and S. Dea**

Centre de recherche en virologie, Institut Armand-Frappier,  
Université du Québec Laval, Québec, Canada

Accepted November 26, 1993

**Summary.** Complementary oligonucleotide primers which flank a 1146-nucleotide gene fragment (S1B: nt 1185 to 2333) encompassing a polymorphic region (nt 1368 to 1776) of the S1 subunit of bovine coronavirus spike glycoprotein were used for enzymatic amplification by PCR. We chose four clinical isolates, recovered from cases of epidemic diarrhea in neonatal calves in Québec dairy herds between 1987–1990, to specifically amplify and analyze their sequences in the selected genomic area. Nucleotide sequence analysis of the four clinical isolates indicated that their S1B gene fragments were highly conserved. We also compared the S1B gene sequences of the Québec BCV isolates to the published corresponding sequences from BCV-L9 [37], BCV-MEB [1], and BCV-F15 [3] reference strains. A high degree of similarity was demonstrated for all viruses, no deletions or insertions were observed, and the only variations that were identified consisted of nucleotide substitutions. The differing nucleotides and amino acids (aa) were not distributed randomly over the entire sequence but rather were clustered in the polymorphic region. Of these, four sporadic aa changes were located in antigenic domain II (aa residues 517 to 720) of S1. This correlates with varied antigenicity observed among the BCV Québec isolates when reacting with MAbs directed against the S glycoprotein of the Mebus strain. The other mutations seem to be fixed in all Québec isolates.

### **Introduction**

Bovine coronavirus (BCV) is an important cause of acute enteritis in newborn calves [18]. Some recent reports suggest that BCV is also responsible for upper respiratory tract illness in growing calves [22, 24], winter dysentery (hemorrhagic enteritis) or chronic shedding in adult cattle [2, 25]. The viral particle is mostly spherical, enveloped with a diameter of around 100 nm, and displays two fringes of surface projections [28, 31]. The viral genome consists of a large single-stranded RNA with positive polarity, approximately 30 Kb in length [31], and encodes four major structural proteins which are the phosphonucleoprotein

(N; 52 kDa), the matrix glycoprotein (M; 24–26 kDa), the spike protein (S; 180–200 kDa) and the hemagglutinin/esterase (HE; 62–65 kDa) [11, 13, 16].

The S glycoprotein often is post-translationally cleaved by host-cell proteases into two 90–100 kDa fragments, S1 and S2, respectively corresponding to the N- and C-terminal subunits [3]. This proteolytic cleavage may have an enhancing effect on viral infectivity and fusion activity [5, 29, 30]. Both S subunits contain antigenic domains responsible for inducing neutralizing antibodies [9, 10, 15, 33], but S1 appears to elicit the production of the monoclonal antibodies (MAbs) displaying the highest neutralizing activity [34]. Competitive ELISA tests using anti-S MAbs suggest the existence of at least four independent antigenic domains (A, B, C and D) for the S1 subunit, antigenic sites S1A and S1B being associated to the neutralizing activity [34]. Alterations in the S protein resulted in changes in viral antigenicity, as demonstrated for neutralization-escape mutants selected under the pressure of MAbs specific for the S protein of BCV [36]. A highly polymorphic region of the S1 subunit of BCV (aa residues 456 to 592) and the mouse hepatitis virus type 4 (MHV-4) (aa residues 426 to 477) has been identified [20, 21, 36]. This region has been demonstrated to undergo deletions or point mutations selectable by neutralizing MAbs [21]. The mutant viruses displayed decreased virulence, suggesting that the S1 polymorphic region may be directly involved in the viral pathogenicity [12].

Recent comparison of Québec BCV isolates using polyclonal antisera and MAbs directed against the S glycoprotein of the Mebus strain confirmed their close antigenic relationship, but also revealed that they are assigned to at least three distinct antigenic subgroups [19]. In the present study, we analysed the nucleotide (nt) sequence of a 1146-nt gene fragment (S1B) encompassing the highly polymorphic region and the proteolytic cleavage site of 4 serologically distinct Québec BCV isolates and compared them to the homologous sequences of the Mebus [1] and F15 [3] reference strains.

## Materials and methods

### *Cells and viruses*

The cell culture-adapted Mebus strain [18] of BCV was obtained from the American Type Culture Collection (ATCC VR-874), Rockville, MD. This prototype BCV strain was originally isolated in bovine fetal kidney cells from diarrhea fluid of calf [18]. Viral isolates BCQ 9, BCQ 571, BCQ 20 and BCQ 2070 were recovered during winter 1989 from clinical cases of epidemic diarrhea in newborn calves affecting Québec dairy herds located in four different geographic areas [19]. No commercial BCV vaccine had been applied in these herds during the year preceeding the emergence of clinical cases. The various Québec BCV isolates could be differentiated serologically by indirect immunofluorescence, hemagglutination-inhibition, seroneutralization and competitive ELISA tests using MAbs directed against the S protein of the prototype Mebus strain [19]. For the present studies, the BCV strains were passaged not more than five times in human rectal tumor (HRT-18) cells in the presence of 10 U/ml of bovine pancreatic trypsin [8]. The extracellular virions were purified from the supernatants of infected cell cultures by differential and isopycnic ultracentrifugation on sucrose gradients as previously described [8].

### *Isolation of viral RNA*

Viral RNA was extracted following the guanidinium isothiocyanate-acid phenol method [4]. Briefly, 100 µl of purified virus was resuspended and vortexed vigorously in 500 µl of solution D (4 M guanidinium isothiocyanate, 0.75 M sodium citrate, 0.5% Sarkosyl and 0.1 M 2-mercaptoethanol). After acid phenol treatment, the aqueous phase was removed and RNA was precipitated with isopropanol. The RNA pellets were washed with 75% ethanol, dissolved in 20 µl RNase-free water, and stored at  $-70^{\circ}\text{C}$ .

### *Preparation of oligonucleotide primers*

The oligonucleotide primers used in the RT-PCR assay were chosen on the basis of the previously published sequence for the Mebus strain of BCV [1], EMBL/Gen Bank accession No. M 31053. The primers were: oligo S1s (5'TGCTATACCGAATGGTAGG 3') complementary to a sequence at positions 1185–1203 on the S gene and oligo S1as (5'GTAAACCGATAACCGAGTGG 3') corresponding to antisense of the sequence at position 2315–2333 on the S gene. This amplification primer pair flanks a 1146-nucleotide gene fragment (S1B) encoding most neutralizing epitopes and the proteolytic cleavage site [34, 36]. In order to clone the PCR products using the CloneAmp system (GIBCO BRL, Gaithersburg, MD), the following sequences were incorporated into the 5' end of each amplification primer: Sn 5'-CUA CUA CUA CUA-3', and Asn 5'-CAU CAU CAU CAU-3'. Oligonucleotide syntheses and incorporation of dTMP into 5' termini were carried out on an automated DNA synthesizer (Pharmacia Biotech Inc., Baie d'Urfé, QC, Canada).

### *First-strand cDNA synthesis*

Prior to PCR amplification, the target RNA was converted to a cDNA by use of reverse transcriptase (RT) and S1as primer. The RT reaction contained 2 µl of 10X PCR buffer (500 mM KCl, 100 mM Tris-HCl (pH 9), 15 mM MgCl<sub>2</sub>, and 1% Triton X-100), 1 µl containing 20 units of RNA guard (Pharmacia Biotech Inc.), 2 µl of 10 mM each dNTP, 2 µl of 25 mM MgCl<sub>2</sub>, 1 µl of 50 pmol antisense primer, and 2–5 µl of the extracted viral RNA. A 20-µl total reaction volume was obtained by adding sterile DEPC-treated distilled water. After heating to 94 °C for 1 min and 65 °C for 5 min, 20 units MuLV RT were added (Pharmacia Biotech Inc.). The mixture was incubated 60 min at 42 °C, and the RT reaction was stopped by heating to 95 °C for 5 min and stored at  $-20^{\circ}\text{C}$ .

### *PCR amplification of first-strand cDNA*

A volume of 80 µl of the PCR reaction mix was added to 20 µl of the cDNA mix such that the final 100 µl volume contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2.5 mM MgCl<sub>2</sub>, 0.01% gelatin, 50 pmol of S1s and S1as primers and 2.5 units of Taq DNA polymerase (BIO 101, La Jolla, CA, U.S.A.). After overlaying with 100 µl of mineral oil, amplification was carried out in a DNA thermal cycler (Perkin Elmer) with the following cycling program: 94 °C denaturation for 1 min, 60 °C annealing for 1 min, and 72 °C elongation for 2 min. A total of 35 cycles was performed, the final polymerisation step was at 74 °C for 10 min. Following amplification, 5 µl of the final product was analyzed on a 0.8% agarose gel in TBE buffer (0.089 M Tris-HCl, (pH 8.3), 0.089 M boric acid, and 2 mM EDTA) and stained with ethidium bromide.

#### *Cloning of PCR-generated fragments*

The PCR products were ligated into the pAMP 1 vector using the CloneAmp system (GIBCO BRL). The ligation mixture was used to transform competent cells *E. coli* DH5 $\alpha$  (GIBCO BRL) following the manufacturer's direction.

#### *Sequencing of cloned PCR fragment*

DNA sequencing was performed by the dideoxynucleotide chain-termination method [26], following the procedure outlined in the instruction of T7 sequencing kit (Pharmacia Biotech Inc.). Sequences of the additional primers used for sequencing the S1B fragment were: 5'-CGTGTAATTGGATG-3', 5'-GTTGGCAAGTACAAG-3', 5'-AATCCTTGTACTIONTGC-3', 5'-GTTAAGTAGTCTCTA-3', 5'-AATTTACACGGACAG-3' and 5'-TAGAGACTACTTAAC-3'. To assess the error rate of the reverse transcriptase and TaqDNA polymerase, clones from different cDNA synthesis and PCR events were sequenced. Errors generated by T7 polymerase were avoided by sequencing clones in both DNA strands and directions. Sequences were analysed by the computer programs, McVector 3.5 (Rainbow Tech. Inc., CA, U.S.A.) and GeneWorks 2.2 (IntelliGenetics Inc., Mountain View, CA, U.S.A.).

#### *Nucleotide sequence accession numbers*

The nucleotide sequence accession numbers (EMBL/GeneBank/DDBJ) are as follows: U06090 for BCQ-2070; U06093 for BCQ-571; U06092 for BCQ-20; and U06091 for BCQ-9.

## **Results**

#### *Comparison of S1B sequences of various clinical BCV isolates*

The complete S1B nucleotide (nt) and deduced amino acid (aa) sequences of the four clinical BCV isolates were aligned and compared (Figs. 1 and 2). Each S1B gene fragment is 1146 bp long encoding the antigenic domain II identified within aa residues 517 to 720 [36], the polymorphic region located between aa residues 456 and 592 [36], and the proteolytic cleavage site beginning with aa 763 [1, 36]. The degree of homology and sequence conservation in the nucleotide and deduced aa sequences of the S1B gene fragment were analysed using a microcomputer program GeneWorks. The isolates BCQ-2070, BCQ-9 and BCQ-571 possessed at least 98% nt sequence identity in paired comparisons. However, nt sequence of the S1B gene fragment of BCQ-20 was only 96% conserved and, thus, was the most distantly related. These sequences demonstrated a high degree of similarity, and the only variations observed consisted of point mutations that apparently not occurred at random. Fifty percent of the nucleotide substitutions were located in the third positions of the respective codons and therefore represent translationally silent mutations. This resulted in only thirteen aa changes among the four isolates. The most similar deduced aa sequences were those of BCV-2070 and BCV-571 or BCV-2070 and BCV-9 (5 aa differences) in paired comparisons, and the most distant were those of BCV-571 and BCV-9 (10 aa differences). Four potential N-linked glycosylation sites were predicted in cases of the S1B fragment of the four Québec BCV isolates, these glycosylation sites being located downstream the polymorphic

BCV-MEB	<b>TTTGCTATAC</b>	<b>CCAATGGTAG</b>	<b>GAAGGTTGAC</b>	<b>CTACAATTGG</b>	<b>GCAATTGGG</b>	<b>CTATTTGCAG</b>	<b>TCFTTTAACT</b>	<b>ATAGAAFTGA</b>	<b>TACTACTGCT</b>	<b>ACAAGTTGTC</b>	100
BCV-F15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100
BCQ-2070	.A.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100
BCQ-571	.A.....	.....	.....	.....	.T.....	.....	.....	.....	.....	.....	100
BCQ-20	.A.....	.....	.....	.....	.T.....	.....	.....	.T.....	.....	.....	100
BCQ-9	.A.....	.....	.....	.....	.T.....	.....	.....	.....	.....	.....	100
BCV-MEB	<b>AGTTGTATTA</b>	<b>TAATTTACCT</b>	<b>GCTGCTAATG</b>	<b>TTTGTGTTAG</b>	<b>CAGGTTTAAT</b>	<b>CCTTCTACTT</b>	<b>GGAATAGSAG</b>	<b>ATTTGGTTTT</b>	<b>ACAGAACAAT</b>	<b>TTGTTTTTAA</b>	200
BCV-F15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200
BCQ-2070	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200
BCQ-571	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200
BCQ-20	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200
BCQ-9	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200
BCV-MEB	<b>GCCTCAACCT</b>	<b>GTAGGTTGTT</b>	<b>TTACTCATCA</b>	<b>TGATGTTGTT</b>	<b>TATGCACAAC</b>	<b>ATGTTTTTAA</b>	<b>AGTCCCCTCA</b>	<b>AATTTCTGTC</b>	<b>CGGTAAAT</b>	<b>GGATGGGTCT</b>	300
BCV-F15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	300
BCQ-2070	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	300
BCQ-571	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	300
BCQ-20	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	300
BCQ-9	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	300
BCV-MEB	<b>TTGTGTGTAG</b>	<b>GTAATGGTCC</b>	<b>TGGTATAGAT</b>	<b>GCTGGTTATA</b>	<b>AAAATAGTGG</b>	<b>TATAGGCACT</b>	<b>TGTCCTGCAG</b>	<b>GTACTAATTA</b>	<b>TTTAACTTGC</b>	<b>CATAATGCTG</b>	400
BCV-F15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	400
BCQ-2070	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	400
BCQ-571	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	400
BCQ-20	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	400
BCQ-9	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	400
BCV-MEB	<b>CCCAATGTAA</b>	<b>TTGTTTGTGC</b>	<b>ACTCCCGACC</b>	<b>CCATTACATC</b>	<b>TAAATCTACA</b>	<b>GGGCCCTTACA</b>	<b>AGTGCCECCA</b>	<b>AACTAAATAC</b>	<b>TTAGTTGGCA</b>	<b>TAGGTGAGCA</b>	500
BCV-F15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	500
BCQ-2070	.T.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	500
BCQ-571	.T.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	500
BCQ-20	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	500
BCQ-9	.T.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	500
BCV-MEB	<b>CTGTTCGGGT</b>	<b>CTTGCTATTA</b>	<b>AAAGTGATTA</b>	<b>TTGTGGAGGT</b>	<b>AATCCTTGTA</b>	<b>CTTGCCAACC</b>	<b>ACAAGCATT</b>	<b>TTGGGCTGGT</b>	<b>CTGTTGACTC</b>	<b>TTGTTTACAA</b>	600
BCV-F15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	600
BCQ-2070	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	600
BCQ-571	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	600
BCQ-20	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	600
BCQ-9	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	600
BCV-MEB	<b>GGGGATAGGT</b>	<b>GTAATATTTT</b>	<b>TGCTAATTTT</b>	<b>ATTTTGCATG</b>	<b>ATGTTAATAG</b>	<b>TGGTACTACT</b>	<b>TGTTCTACTG</b>	<b>ATTTACAAAA</b>	<b>ATCAAACACA</b>	<b>GACATAATTC</b>	700
BCV-F15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	700
BCQ-2070	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	700
BCQ-571	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	700
BCQ-20	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	700
BCQ-9	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	700
BCV-MEB	<b>TTGGTGTGTT</b>	<b>TGTTAATTAT</b>	<b>GATCTTTATG</b>	<b>GTATTACAGG</b>	<b>CCAAGGTATT</b>	<b>TTTGTGAGG</b>	<b>TTAATGCGAC</b>	<b>TTATTATAAT</b>	<b>AGTTGGCAGA</b>	<b>ACCTTTTATA</b>	800
BCV-F15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	800
BCQ-2070	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	800
BCQ-571	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	800
BCQ-20	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	800
BCQ-9	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	800
BCV-MEB	<b>TGATICTAAT</b>	<b>GGTAATCTCT</b>	<b>ATGGTTTTAG</b>	<b>AGACTACTTA</b>	<b>ACAAACAGAA</b>	<b>CTTTTATGAT</b>	<b>TCGTAGTTCG</b>	<b>TATAGCGGTC</b>	<b>GTGTTTCAGC</b>	<b>GGCCTTTCAT</b>	900
BCV-F15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	900
BCQ-2070	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	900
BCQ-571	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	900
BCQ-20	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	900
BCQ-9	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	900
BCV-MEB	<b>GCTAATICTT</b>	<b>CCGAACCAGC</b>	<b>ATTGCTATTT</b>	<b>CGGAATATTA</b>	<b>AATGCAATTA</b>	<b>CGTTTTTAA</b>	<b>AATACTCTTT</b>	<b>CACGACAGCT</b>	<b>GCAAACCTATT</b>	<b>AACTATTTTG</b>	1000
BCV-F15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1000
BCQ-2070	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1000
BCQ-571	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1000
BCQ-20	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1000
BCQ-9	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1000
BCV-MEB	<b>ATAGTTATCT</b>	<b>TGGTTGTGTT</b>	<b>GTCAATGCTG</b>	<b>ATAATAGTAC</b>	<b>TTCTAGTGT</b>	<b>GTTCAAACAT</b>	<b>GTGATCTCAC</b>	<b>AGTAGGTAGT</b>	<b>GGTTACTGTG</b>	<b>TGGATTACTC</b>	1100
BCV-F15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1100
BCQ-2070	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1100
BCQ-571	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1100
BCQ-20	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1100
BCQ-9	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1100
BCV-MEB	<b>TACAAAAAGA</b>	<b>CGAAGTCTTA</b>	<b>GAGCGATTAC</b>	<b>CACTGGTTAT</b>	<b>CGGTTTA</b>	.....	.....	.....	.....	.....	1147
BCV-F15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1147
BCQ-2070	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1147
BCQ-571	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1147
BCQ-20	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1147
BCQ-9	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	1147

**Fig. 1.** Nucleotide sequence comparisons of the S1B gene fragment (nt 1185 to 2333) of four Québec BCV isolates, BCQ-2070, BCQ-571, BCQ-20 and BCQ-9, and two reference strains, BCV-MEB [1] and BCV-F15 [3]. Bold letters show the polymorphic region corresponding to aa residues 456 to 592. The underlined sequences refer to the proteolytic cleavage site. The sequence of reference strain BCV-MEB is written in full, while only changes in nucleotides are indicated for the sequences of the other strains

BCV-MEB	FAIPNGRKVD LQLGNLGYLQ SFNYRIDTTA TSCQLYLNLP AANVSVSREFN PSTWNRREFGF <b>TEQFVFKPQP</b> VGVFTHHDVV <u>YAQHCPKAFS</u> <u>NFCPCCKLDGS</u>	500
BCV-F15	.....I.....D.....	500
BCQ-2070	Y.....S.....A.....D.....T.....	500
BCQ-571	Y.....S.....A.....D.....T.....	500
BCQ-20	Y.....S.....A.....A.....T.....	500
BCQ-9	Y.....S.....A.....D.....T.....	500
<b>BCV-MEB</b>	<b>LCVGNPGLD AGYKNSGIQT CPAGTNYLTC HNAAQCNCLC TPDFITSRST GPYKCPQTKY LVGIGEHCSG LAIKSDYCGG NPCTCQPAF LGSVSDSCLQ</b>	600
BCV-F15	.....A.....S.....H.....	600
BCQ-2070	.....S.....A.....S.....A.....	600
BCQ-571	.....S.....A.....S.....A.....	600
BCQ-20	.....T.....A.....L.....H.....	600
BCQ-9	.....S.....A.....L.....H.....	600
BCV-MEB	GDRCNIFANF ILHDVNSGTT CSTDLQKSNT DIILGVCVNY DLYGITGQGI <u>FVEVNATYYN</u> SWQNLLYDSN GNLYGFRDYL <u>TNRTFMIRSC</u> YSGRVSAAFH	700
BCV-F15	.....A.....	700
BCQ-2070	.....A.....	700
BCQ-571	.....L.....	700
BCQ-20	.....L.....	700
BCQ-9	.....L.....	700
BCV-MEB	<u>ANSSEPALLF</u> <u>RNIKCNYVFN</u> <u>NTLSRQLQPI</u> NYFDSYLGCV VNADNSTSSV <u>VQTCDLTVGS</u> GYCVDYSTKR <u>RSRRAITGY</u> RF	782
BCV-F15	.....A.A.....V.....S.....	782
BCQ-2070	.....V.....S.....	782
BCQ-571	.....S.....S.....	782
BCQ-20	.....S.....S.....	782
BCQ-9	.....S.....I.....S.....	782

**Fig. 2.** Alignment of the deduced amino acid sequences of the S1B gene fragment of four Québec BCV isolates, BCQ-2070, BCQ-571, BCQ-20, and BCQ-9; and two reference strains, BCV-MEB [1], and BCV-F15 [3]. The polymorphic region (aa residues 456 to 592) is written in bold face, the proteolytic cleavage site (aa residues 663 to 767) is underlined and the potential N-linked glycosylation sites are double underlined. The sequence of reference strain BCV-MEB is written in full, while only changes in amino acid residues are indicated for the sequences of the other strains

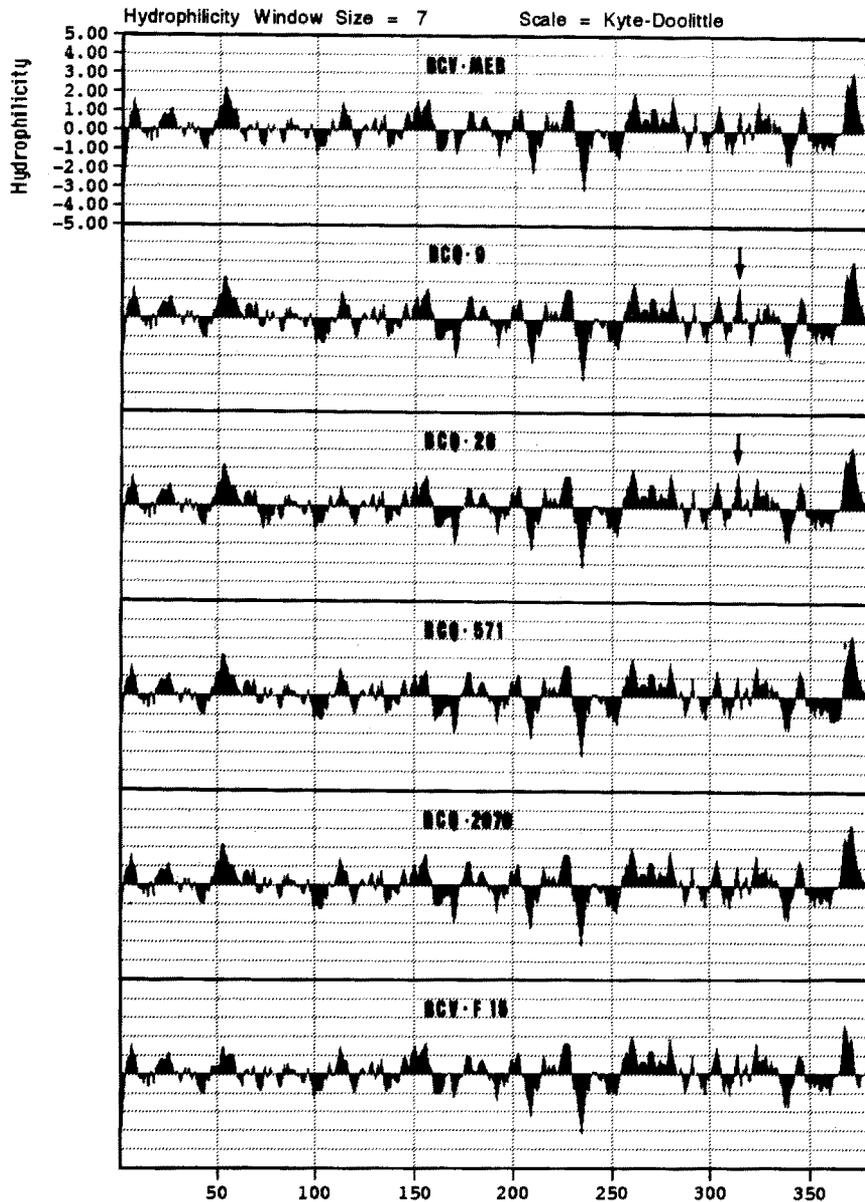
region and antigenic domain II. Twenty-six cysteine residues were conserved among all of the Québec isolates.

#### *Comparison of S1B sequences of Québec BCV isolates with other known BCV strains*

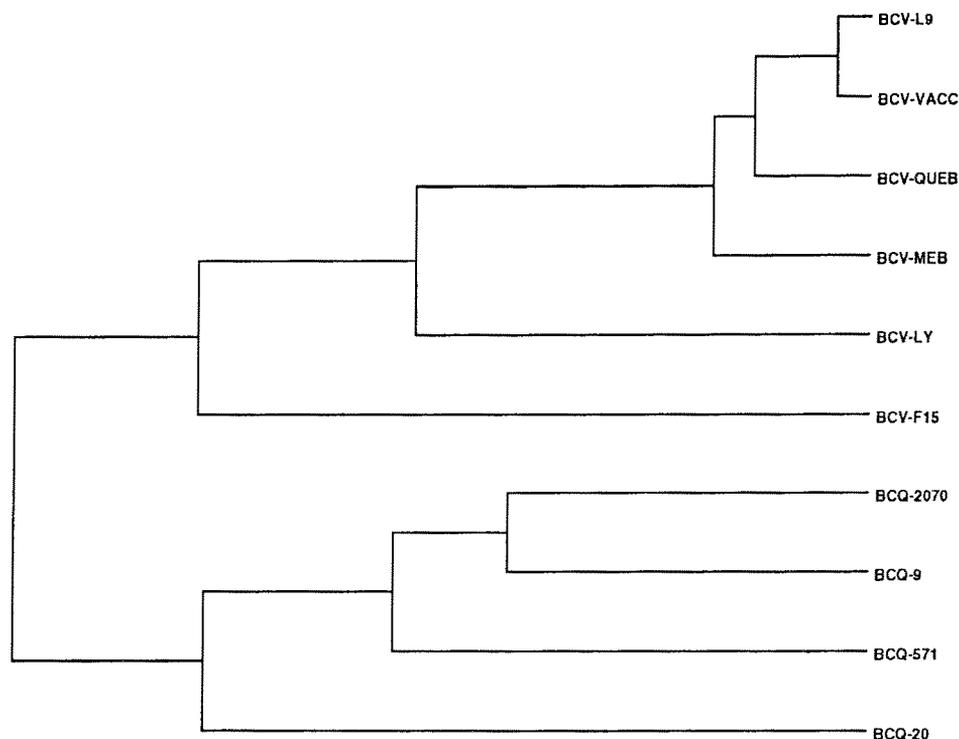
As illustrated in Fig. 1, the complete S1B nt sequences of the four Québec clinical BCV isolates were aligned and compared with the homologous sequences of the Mebus (BCV-MEB) and F15 (BCV-F15) reference strains. The level of nt sequence identity between the S1B fragments of the four Québec BCV isolates and that from BCV-MEB and BCV-F15 strains was 97% and 96%, respectively. In contrast, the polymorphic region (nt 168–576) seemed to be more conserved among the Québec BCV isolates and BCV-F15 (97%).

Frameshift, deletion, or insertion, and non sense mutations were not observed; the only variations among these sequences consisted of 40 nt substitutions, which represented 4% of the sequence. The differing nucleotides were not distributed randomly over the entire sequence but rather were clustered in the polymorphic region. This was also reflected in the deduced aa sequences (Fig. 2). The 60 aa residues immediately upstream the polymorphic region were almost completely conserved among all the isolates. However, 12 aa substitutions were found between residues 456 and 592 of the polymorphic region. The fact, that these aa changes occurred in almost identical location suggests these differences may be significant. The 29 aa residues (aa 593 to 621) immediately downstream the hypervariable region have been shown to be part of the antigenic domain II which is recognized by neutralizing Mabs BB7–14 [36].

This region was totally conserved among all BCV isolates tested. The twenty-six cysteine residues identified above were perfectly conserved among all of the Québec isolates. The five remaining differences in the 160 residues upstream the proteolytic cleavage site represented variability which seemed to be unique to some isolates. The residue changes consisted of Phe for Leu and Asp for



**Fig. 3.** Hydrophilicity profile of the S1B gene fragment of four Québec BCV isolates (BCQ-2070, BCQ-571, BCQ-20, and BCQ-9); and two BCV reference strains, BCV-MEB [1], and BCV-F15 [3]. Hydrophilicity plots (according to the Kyte-Doolittle methods, using a seven-amino acid window) starts at amino acid 401 of the mature S1 protein. Arrows in hydrophilicity plots of isolates BCQ-20 and BCQ-9 show regions of increased hydrophilicity



**Fig. 4.** Phylogenetic tree of the S1B gene fragment of four Québec BCV isolates (BCQ-2070, BCQ-571, BCQ-20, and BCQ-9); and six BCV strains; BCV-MEB [1]; BCV-F15 [3]; BCV-L9, BCV-LY, and BCV-VACC [37]; and BCV-QUEB [20]. The length of the horizontal lines connecting one sequence to another is proportional to the estimated genetic distance between the sequences

Val for BCQ-571, Ile for Ser for BCQ-20 and BCQ-9, and Thr for Ile for BCQ-9 which resulted in the loss of one of the four potential N-linked glycosylation sites identified in case of the S1B fragments of the other BCV isolates. Residue substitutions Ile for Ser for BCQ-20 and BCQ-9, and Thr for Ile for BCQ-9 also resulted in increased hydrophilicity (Fig. 3) and surface probability in this region.

The phylogenetic tree in Fig. 4 shows the relationship of the four Québec clinical BCV isolates to the published homologous sequences from BCV-F15 [3], BCV-MEB [1], BCV-LY, BCV-L9, BCV-VACC [37] based on their S1B fragment genes. The homologous sequences of BCV-QUEB [20], a cell culture-adapted enteropathogenic strain isolated in Québec in 1977, was also included in this comparison. Phylogenetic analyses by Unweighted Pair Group Method with Arithmetic Mean (GeneWorks) classify recent Québec clinical isolates in a distinct sublineage.

#### *Comparison of the proteolytic cleavage site sequences*

The third region that was examined contains the coding sequence for the proteolytic cleavage site. The prototype BCV cleavage site sequence consists

of six residues KRRSRR beginning with aa 763, and predicts a cleavage between aa 768 and 769. Cleavage at the Arg–Ala bond by a host cell trypsin-like enzyme generates the S1 and S2 subunits of the S protein [1]. The sequence of the predicted proteolytic cleavage site (KRRSRR) was conserved among all the isolates studied. However, three of the Québec clinical BCV isolates examined had modifications in aa proximal to the cleavage site. The isolates BCQ-2070, BCQ-20 and BCQ-9 had a Ala-769 to Ser change immediately upstream the cleavage site.

### Discussion

Considering the complexity of the pathogenicity and tropism of BCV and other coronaviruses, such as turkey coronavirus or human coronavirus HCoV-OC43, comparative analysis of different clinical isolates will probably not produce meaningful results. However, with accumulating data on genomic and antigenic characterizations, immunodominant and functional regions can be selected for sequence analysis. In the present study, we chose four Québec BCV isolates recovered from cases of epidemic diarrhea in neonatal calves in Québec dairy herds between 1987–1990 to analyze their sequences in a selected genomic area which codes for immunologically and functionally important domains. Nucleotide sequence analysis of the four clinical isolates indicated that their S1B gene fragments were highly conserved. When compared with the prototype BCV-MEB strain, the S1B gene fragments of these Québec isolates differed by only 2.5 to 3.5%, the highest concentration of nt divergence being located in the polymorphic region. This region has greater than 94% nt identity with that of the prototype Mebus strain, and apparently, the differing nt were not distributed randomly over the sequences. The mutations have different effects on the encoded aa sequences. Over the S1B sequence most of mutations were silent, in contrast, in the polymorphic region most of the nt substitutions resulted in changes of aa residues. The S1B deduced aa sequences were more than 97% conserved among the strains, however, the polymorphic region was only 91% conserved. Recently, comparisons of the S gene sequences of six highly virulent (including BCV-F15) or avirulent (including BCV-MEB, and BCV-QUEB) strains have revealed more than 98% sequence identity in paired comparison, and most of the substitutions occurred in the S1 subunit and were distributed randomly over the entire sequence [37]. The polymorphic regions were almost completely conserved among all strains, except for the virulent BCV-F15 strain [37]. In the present study, a closer genomic relatedness was demonstrated between BCV-F15 strain and the four Québec clinical BCV isolates. At least twelve aa substitutions have been identified between the virulent and avirulent (BCV-MEB and BCV-QUEB) groups in the highly polymorphic region of the S1 peptide (Fig. 2), suggesting that aa changes in this particular region of the S glycoprotein may be related to BCV virulence. However, the results obtained do not exclude that other regions of the BCV genome may be also involved in the virulence or tissue tropism.

It has been previously demonstrated that antigenic domain II, identified within aa residues 517 to 720, is functionally and immunologically important [36]. The epitopes that mapped in this domain are capable of inducing neutralizing antibodies [34, 36] and, therefore, represent targets of immunological pressure. This antigenic domain, together with the polymorphic region, demonstrated a variability among the Québec clinical isolates. Although the number of aa substitutions was relatively small, the fact that most substitutions appeared to cluster in an hypervariable region of the S gene, suggests these differences may be significant. The present data are in agreement with recent findings using various immunological assays that have shown that the four BCV isolates tested, reacted differently with neutralizing MABs directed against the S1 subunit of the Mebus strain [19].

Three of the Québec BCV isolates examined could be differentiated from the prototype BCV-MEB strain by the nt sequences proximal to the proteolytic cleavage site. These isolates had a Ala to Ser change at residue 769 immediately upstream the cleavage site. This change maintained the number and pairing of basic residues of the prototype sequence. Changes in amino acids proximal to the cleavage site of the HA molecule of the influenza A virus have been demonstrated to modulate its cleavability [23]. In addition, changes from Phe to Ser immediately following the proteolytic cleavage site of the F protein of Sendai virus conferred the acquisition of susceptibility to plasmin cleavage; this suggests that regions of the F protein contiguous to the cleavage site can also affect its cleavability by interfering with the efficacy of host proteases involved in the maturation of the viral envelopped glycoprotein [14]. However, in the present study, the Ala to Ser change immediately following the cleavage site of the S glycoprotein of three of the Québec BCV isolates did not appear to modulate cleavability, and no correlation was found between this substitution and the rate of viral replication and type of cytopathic changes induced in HRT18 cells (data not shown).

The genomic variations which have been identified in the present study seemed to represent characteristic features of the Québec BCV isolates. Considering the fact that these isolates were not passaged more than five times in HRT-18, and because of the finding of limited identity between Québec BCV isolates, the possibility that our sequence data contain only aa substitutions arising after cell culture passages may be excluded. However, such aa substitutions may have occurred in cases of BCV-MEB and BCV-QUEB isolates for which numerous passages were done in cell lines other than HRT-18 and from heterologous animal species [7, 27]. The human rectal adenocarcinoma cell line has been reported to be susceptible and permissive to human, canine and turkey enteric coronaviruses [6, 8, 16,17]. This human cell line grows in culture in tightly packed colonies of cells that have retain some characteristics of mature enterocytes [32]. The adaptation of field strains of enteric coronaviruses to this cell line usually requires only few passages. For RNA viruses, it is expected that numerous cell culture passages can result in genomic variations which may occur as point mutations randomly distributed throughout the entire virus

genome, rather than clustered in a specific genomic region. In the present study, the sequence heterogeneity was located mostly in the polymorphic region of the S1 glycoprotein.

In summary, our data have shown that Québec clinical BCV isolates are genetically divergent from the reference strains [1, 3, 20, 37], and that the polymorphic region demonstrates a high variability. The antigenic as well as genetic relationships among the currently circulating strains of BCV provide useful information for determining the most appropriate composition for the coronaviruses vaccines. Continued surveillance of these viruses will be necessary to detect the emergence of new lineages of BCV with altered antigenic properties and assure that the future vaccines contain the most appropriate strains of BCV.

### Acknowledgements

We thank Ms. N. Sawyer and J. Roger for technical assistance. This work was supported in part by grants received from the National Research Council of Canada (OGP0046412) and the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec (92-NC-0511).

### References

1. Abraham S, Kienzle TE, Lapps W, Brian DA (1991) Deduced sequence of the bovine coronavirus spike protein and identification of the internal proteolytic cleavage site. *Virology* 176: 296–301
2. Benfield DA, Saif L (1990) Cell culture propagation of a coronavirus isolated from cows with winter dysentery. *J Clin Microbiol* 28: 1454–1457
3. Boireau P, Cruciere C, Laporte J (1990) Nucleotide sequence of the glycoprotein S gene of bovine enteric coronavirus and comparison with the S proteins of two mouse hepatitis virus strains. *J Gen Virol* 71: 487–492
4. Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform. *Anal Biochem* 162: 156–159
5. Cyr-Coats K, Storz J, Hussain KA, Schnorr KL (1988) Structural proteins of bovine coronavirus strain L9: Effects of the host cell and trypsin treatment. *Arch Virol* 103: 35–45
6. Dea S, Verbeek AJ, Tijssen P (1990) Antigenic and genomic relationships among turkey and bovine enteric coronaviruses. *J Virol* 64: 3112–3118
7. Dea S, Roy RS, Begin ME (1980) Bovine coronavirus isolation and cultivation in continuous cell lines. *Am J Vet Res* 41: 30–38
8. Dea S, Garzon S, Tijssen P (1989) Isolation and trypsin-enhanced propagation of turkey enteric (bluecomb) coronaviruses in a continuous human rectal adenocarcinoma cell line. *Am J Vet Rec* 50: 1310–1318
9. Dereg D, Babiuk LA (1987) Monoclonal antibodies to bovine coronavirus: characteristics and topographical mapping of neutralizing epitopes on the E2 and E3 glycoproteins. *Virology* 161: 410–420
10. Dereg D, Parker MD, Cox GC, Babiuk LA (1989) Mapping of neutralizing epitopes to fragments of the bovine coronavirus E2 protein by proteolysis of antigen-antibody complexes. *J Gen Virol* 70: 647–658
11. Dereg D, Sabara M, Babiuk LA (1987) Structural proteins of bovine coronavirus and their intracellular processing. *J Gen Virol* 68: 2863–2877
12. Gallagher TM, Parker SE, Buchmeier MJ (1990) Neutralization-resistant variants of

- aneurotropic coronavirus are generated by deletions within the amino-terminal half of the spike glycoprotein. *J Virol* 64: 731–741
13. Hogue BG, King B, Brian DA (1984) Antigenic relationships among proteins of bovine coronavirus, human respiratory coronavirus OC43, and mouse hepatitis coronavirus A59. *J Virol* 51: 384–388
  14. Hsu M-C, Scheid A, Choppin PW (1987) Protease activation mutants of Sendai viruses: sequence analysis of the mRNA of the fusion protein (F) gene and direct identification of the cleavage-activation site. *Virology* 156: 84–90
  15. Hussain KA, Storz J, Kousoulas KG (1991) Comparison of bovine coronavirus (BCV) antigens: monoclonal antibodies to the spike protein distinguish between vaccine and wild-type strains. *Virology* 183: 442–445
  16. King B, Potts BJ, Brian DA (1985) Bovine coronavirus hemagglutinin protein. *Virus Res* 2: 53–59
  17. Laporte J, Bobulesco P, Rossi F (1980) Une lignée particulièrement sensible à la réplication du coronavirus entéritique bovin: les cellules HRT-18. *CR Acad Sci (III) Paris* 290D: 623–626
  18. Mebus CA, Stair EL, Rhodes MB, Twiehaus MJ (1973) Neonatal calf diarrhoea: propagation, attenuation, and characteristics of a coronavirus-like agent. *Am J Vet Res* 34: 145–150
  19. Michaud L, Dea S (1993) Characterization of monoclonal antibodies to bovine enteric coronavirus and antigenic variability among Quebec isolates. *Arch Virol* 131: 455–465
  20. Parker MD, Yoo D, Cox GJ, Babiuk LA (1990) Primary structure of the S peplomer gene of bovine coronavirus and surface expression in insect cells. *J Gen Virol* 71: 263–270
  21. Parker SE, Gallagher TM, Buchmeier MJ (1989) Sequence analysis reveals extensive polymorphism and evidence of deletion within E2 glycoprotein gene of several strains of murine hepatitis virus. *Virology* 173: 664–673
  22. Reynolds DJ, Debney TG, Hall GA, Thomas LH, Parsons KR (1985) Studies on the relationship between coronaviruses from the intestinal and respiratory tracts of calves. *Arch Virol* 85: 71–83
  23. Rott R, Orlich M, Klenk H-D, Wang ML, Skehel JJ, Wiley DC (1984) Studies on the adaptation of influenza viruses to MDCK cells. *EMBO J* 3: 3329–3332
  24. Saif LJ, Redman DR, Moorhead PD, Theil KW (1986) Experimentally-induced coronavirus infections in calves: viral replication in the respiratory and intestinal tracts. *Am J Vet Res* 47: 1426–1432
  25. Saif LJ, Brock KV, Redman DR, Kohler EM (1991) Winter dysentery in dairy herds: electron microscopic and serological evidence for an association with coronavirus infection. *Vet Rec* 128: 447–449
  26. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463–5467
  27. Sharpee RL, Mebus CA, Bass EP (1976) Characterization of a calf diarrheal coronavirus. *Am J Vet Res* 37: 1031–1041
  28. Siddell S, Wege H, Ter Meulen V (1983) The biology of coronaviruses. *J Gen Virol* 64: 761–776
  29. Storz J, Rott R, Kaluza G (1981) Enhancement of plaque formation and cell fusion of an enteropathogenic coronavirus by trypsin treatment. *Infect Immun* 31: 1214–1222
  30. Sturman LS, Ricard CS, Holmes KV (1985) Proteolytic cleavage of the E2 Glycoprotein of murine coronavirus: activation of cell-fusing activity of virions by trypsin and separation of two different 90K cleavage fragments. *J Virol* 56: 904–911
  31. Sturman LS, Holmes KV (1983) The molecular biology of coronaviruses. *Adv Virus Res* 28: 35–112

32. Tompkins WAF, Watrach AW, Schmale JD, Schultz RM, Harris JA (1974) Cultural and antigenic properties of newly established cell strains derived from adenocarcinomas of the human colon and rectum. *J Natl Cancer Inst* 52: 101–106
33. Vautherot JF, Laporte J, Madelaine MF, Bobulesco P, Roseto A (1984) Antigenic and polypeptide structure of bovine enteric coronavirus as defined by monoclonal antibodies. *Adv Exp Med Biol* 173: 117–132
34. Vautherot JF, Madelaine MF, Boireau P, Laporte J (1992) Bovine coronavirus peplomer glycoproteins: detailed antigenic analysis of S1, S2 and HE. *J Gen Virol* 73: 1725–1737
35. Vlasak R, Luytjes W, Leider J, Spaan W, Palese P (1988) The E3 protein of bovine coronavirus is a receptor-destroying enzyme with acetylsterase activity. *J Virol* 62: 4686–4690
36. Yoo D, Parker MD, Song J, Cox GJ, Deregt D, Babiuk LA (1991) Structural analysis of the conformational domains involved in neutralization of bovine coronavirus using deletion mutants of the spike glycoprotein S1 subunit expressed by recombinant baculoviruses. *Virology* 183: 91–98
37. Zhang X, Kousoulas KG, 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: 397–404

Authors' address: Dr. S. Dea, Centre de recherche en virologie, Institut Armand-Frappier, 531 boulevard des Prairies, Laval, Québec H7N 4Z3, Canada.

Received August 25, 1993