Primary structure of the S peplomer gene of bovine coronavirus and surface expression in insect cells

M. D. Parker,[1]* D. Yoo,[1] G. J. Cox[1] and L. A. Babiuk[1,2]

[1]*Veterinary Infectious Disease Organization, 124 Veterinary Road, Saskatoon, Saskatchewan S7N 0W0 and*
[2]*Department of Veterinary Microbiology, University of Saskatchewan, Saskatoon, Saskatchewan S7N 0W0, Canada*

The nucleotide sequence of the S peplomer gene of bovine coronavirus (BCV) has been determined. A single open reading frame of 4089 nucleotides encodes a polypeptide of 150K with 20 potential sites for addition of *N*-linked oligosaccharides. Expression of the cloned BCV S gene by a recombinant of *Autographa californica* nuclear polyhedrosis virus resulted in production of a 180K glycosylated polypeptide which was transported to the surface of the cell. Comparison of the BCV S gene with the analogous genes of murine hepatitis viruses shows that the BCV S polypeptide contains a unique domain of 138 amino acids not present in murine hepatitis virus strain JHM, but which has a partially homologous counterpart in strain A59. This domain accounts for most of the differences in size of the S gene products of these coronaviruses.

Introduction

The family Coronaviridae is composed of a single genus of large, enveloped viruses which infect a variety of mammalian and avian species. The genome of these viruses is a single-stranded, polyadenylated RNA of 25000 to 30000 nucleotides associated with the nucleocapsid protein in a helical configuration (MacNaughton *et al.*, 1978). The ribonucleoprotein is surrounded by an envelope which, depending on the virus strain, contains two or three viral glycoproteins. The distinctive peplomers on the surface of coronaviruses are composed of multimers of the S or E2 gene product, either as cleaved subunits or an uncleaved polypeptide, and form the 'corona' which serves as a structural basis for the classification of these viruses. The small envelope glycoprotein, E1 or M, determines the intracellular site of viral morphogenesis (Tooze *et al.*, 1984) and may play a role in pathogenesis (Fleming *et al.*, 1989). Bovine coronavirus (BCV), a causative agent of severe neonatal diarrhoea (Dea, 1980; Mebus, 1978), human respiratory coronavirus OC-43 (Hogue & Brian, 1986) and haemagglutinating encephalomyelitis virus of swine (Callebaut & Pensaert, 1980) contain an additional surface glycoprotein, HE or H, which is the viral haemagglutinin (King & Brian, 1982; King *et al.*, 1985; Parker *et al.*, 1989).

Previous reports have demonstrated that coronavirus S glycoproteins carry determinants for serum neutralization (Cavanagh *et al.*, 1986*a*; Collins *et al.*, 1982; Luytjes *et al.*, 1989), cell surface binding (Cavanagh & Davis, 1986), tissue tropism, virulence (Wege *et al.*, 1988) and, in avian infectious bronchitis virus, the S polypeptide also exhibits haemagglutinating activity (Mockett *et al.*, 1984). Studies with a series of monoclonal antibodies have shown that both the S and HE proteins of BCV contain neutralizing epitopes (Deregt & Babiuk, 1987), indicating that both play a critical role in infection. In order to analyse the biochemical and immunological properties of the individual viral glycoproteins of BCV, we have constructed cDNA clones representing the S gene of BCV. We report here the sequence of the BCV S gene and compare it with the analogous gene of murine hepatitis virus (MHV) strains JHM and A59. Expression of the BCV S gene by an *Autographa californica* nuclear polyhedrosis virus (AcNPV) recombinant yields a 180K product with properties similar to the authentic polypeptide produced in BCV-infected cells.

Methods

Cells and virus. The Quebec strain of bovine coronavirus (Dea *et al.*, 1980) was propagated in Madin-Darby bovine kidney (MDBK) cells, obtained from the American Type Cell Culture collection, grown in Dulbecco's modified minimal essential medium containing 10% foetal bovine serum and 50 μg/ml gentamicin. Spinner cultures of *Spodoptera frugiperda* cells (Sf9) were propagated in TNM-FH medium containing 10% foetal bovine serum (Summers & Smith, 1987).

Molecular cloning of the BCV genome. The preparation of a cDNA library representing the genome of the Quebec strain of bovine coronavirus has been described previously (Parker *et al.*, 1989).

DNA sequencing. A series of overlapping cDNA clones representing the entire S gene of BCV were identified by colony hybridization

(Grunstein & Hogness, 1975) with a probe prepared from pCVA12H (Parker *et al.*, 1989), and sequenced by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977) after generation of an extensive series of overlapping deletions by the method of Henikoff (1984). Sequence analysis and comparisons were conducted using the IBI DNA/Protein sequence analysis system (International Biotechnologies).

Subcloning and expression of the BCV S gene in Sf9 cells. In order to produce a cassette containing the molecularly cloned S gene, the 5′ portion of the S gene was prepared by sequential *Eco*RV and exonuclease III digestion of pCVA12H to remove 1338 nucleotides upstream of the S initiation codon. A *Bam*HI linker was added and the DNA was cleaved with *Bam*HI and *Pst*I, yielding a fragment of 1565 bp. Clone pCVA12I was digested with *Taq*I and a *Bam*HI linker was added. After digestion with *Pst*I and *Bam*HI, a fragment representing the 3′ 2622 bp was ligated to the 5′ 1558 bp fragment in the *Bam*HI site of pTZ18R and used to transform *Escherichia coli* strain JM105. The termini of the cassette and the nucleotide sequences surrounding the *Pst*I site were determined to confirm the absence of cloning artefacts. The cassette, which extends from the A residue at position 7 in Fig. 2 to a *Taq*I site 77 nucleotides downstream of the S termination codon, was ligated into the *Bam*HI site of the transfer vector pAcYM1 (Matsuura *et al.*, 1987) and inserted into the genome of the AcNPV by homologous recombination as outlined by Summers & Smith (1987). Recombinants were identified by plaque hybridization and plaque purification. Six independently isolated recombinants were isolated. One recombinant, AcSD, was utilized in the following experiments.

Approximately 1 h before infection, 2×10^6 Sf9 cells were plated into 25 cm^2 flasks and allowed to attach at 28 °C. The monolayer was infected with the AcNPV recombinant, AcSD, at an m.o.i. of 5 p.f.u. per cell and incubated for 40 h at 28 °C. At 48 h after infection, the medium was replaced with methionine-free Grace's medium (Gibco) containing 1% foetal bovine serum and 100 μCi/ml [^{35}S]methionine (Amersham, 800 Ci/mmol) and incubated an additional 1·5 h. The cells were then scraped into phosphate-buffered saline (PBS), pelleted at 1000 *g* for 1 min, and lysed by vigorous vortexing in RIPA buffer (50 mM-Tris–HCl pH 7·4, 150 mM-NaCl, 1% sodium deoxycholate, 1% Triton X-100) containing 0·1% SDS. The lysate was clarified at 15000 *g* for 5 min, and radiolabelled polypeptides were immunoprecipitated with S-specific monoclonal antibodies HF8-8 and BB7-14 (Deregt & Babiuk, 1987) and analysed by SDS–gel electrophoresis on 7·5% polyacrylamide gels (acrylamide : bisacrylamide, 29·2 : 0·8) (Laemmli, 1970). The gel was treated with En^{3}Hance (New England Nuclear), dried under vacuum and autoradiographed.

Immunofluorescence analysis. Sf9 cells were infected with the baculovirus recombinant AcSD at an m.o.i. of 10 p.f.u. per cell and incubated at 28 °C for 40 h. MDBK cells were infected with BCV at an m.o.i. of 10 p.f.u. per cell and incubated at 37 °C for 24 h. In order to examine antigen distribution in fixed cells, approximately 2×10^5 cells were adsorbed to microscope slides by Cytospin centrifugation (Shandon Southern Instruments). The cells were then fixed with cold methanol for 2 min and incubated for 1 h at 37 °C with a 1 : 300 dilution of S-specific monoclonal antibodies HF8-8 and BB7-14 (Deregt & Babiuk, 1987). The cells were washed with PBS and reacted for 1 h at 37 °C with a 1 : 100 dilution of fluorescein isothiocyanate (FITC)-conjugated goat anti-mouse IgG (Boehringer Mannheim). The slides were again washed with PBS and examined for fluorescence. To measure surface distribution of antigen, washed, unfixed AcSD-infected cells were suspended in cold PBS containing the appropriate primary antibody for 1 h at 0 °C, washed with ice-cold PBS and reacted with FITC-conjugated goat anti-mouse IgG on ice for 1 h. The cells were then washed with cold PBS and adsorbed to microscope slides by centrifugation.

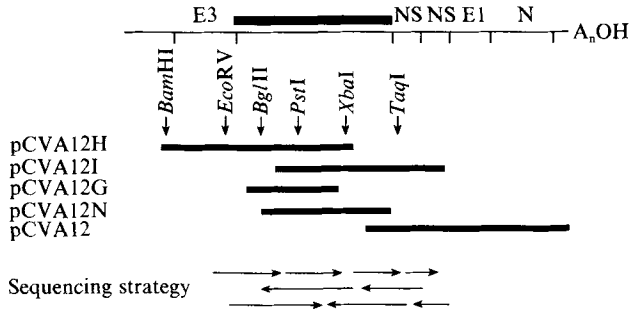


Fig. 1. Physical map of the BCV genome RNA. A series of overlapping cDNA clones of the viral genome are aligned with the genome RNA. The stippled region indicates the region characterized as the S gene.

Results

Isolation of cDNA sequences representing the BCV S gene

In a previous report, we presented the nucleotide sequence of the BCV HE haemagglutinin gene (Parker *et al.*, 1989), which extends from nucleotides 7403 to 8677 distal to the 3′ end of the BCV genome. Immediately downstream from the HE gene in cDNA clone pCVA12H (Parker *et al.*, 1989), an open reading frame was found which exhibited significant similarity to the S genes of MHV strains A59 and JHM (Luytjes *et al.*, 1987; Schmidt *et al.*, 1987). Additional clones representing this region of the viral genome were identified and mapped as shown in Fig. 1. The sequence of the S gene of BCV was obtained by the complete sequence determination of five cDNA clones by the strategy shown in Fig. 1.

A single open reading frame of 4089 nucleotides was identified which could encode a polypeptide of 1363 amino acids of 150K. The nucleotide sequence of the BCV S gene beginning with the first nucleotide adjacent to the termination codon of the HE gene and amino acid sequence of the predicted polypeptide product are shown in Fig. 2.

Several observations indicate that the sequence in Fig. 2 represents the entire S gene. First, the sequence TCTAAAC at nucleotides 8 to 14 is identical to the sequence upstream of the initiation codon of the BCV N gene (Lapps *et al.*, 1987) and closely resembles the sequence immediately preceding the initiation codons for the BCV M (E1) (Lapps *et al.*, 1987) and HE (Parker *et al.*, 1989) genes. Second, the first initiation codon at nucleotides 15 to 17 is in a favourable sequence context (Kozak, 1987) and indicated that the HE–S intergenic sequence is 14 nucleotides in length. Third, the carboxy-terminal amino acid sequence of the predicted product is very similar to the sequences of the S gene of MHV JHM and A59 (see below). The termination codon, TAA, at

Fig. 2. The nucleotide sequence and predicted amino acid sequence of the S gene of BCV. The sequence begins with the first nucleotide adjacent to the termination codon of the HE gene. Asterisks indicate the conserved intergenic sequence. Bold underline indicates probable signal sequence. Underlined Asn residues indicate potential *N*-linked glycosylation sites. Bold letters indicate possible cleavage site.

position 4103 indicates that the S gene of BCV, including the upstream intergenic sequence, extends from 3301 to 7389 nucleotides from the 3′ terminus of the viral genome excluding the poly(A) tail, very similar to the location of the S gene on the MHV A59 genome (Luytjes *et al.*, 1987).

The hydrophobicity/hydrophilicity profile of the predicted S gene product (not shown) is very similar to the analogous proteins of MHV strains JHM and A59 (Luytjes *et al.*, 1987; Schmidt *et al.*, 1987). Immediately following the initiation codon is a possible signal sequence of 16 amino acids (bold underline in Fig. 2), which, according to the −3, −1 rule (von Heijne, 1986), may be cleaved between the glycine and arginine residues at positions 17 and 18. A second hydrophobic domain, located near the carboxyl terminus of the polypeptide probably serves as an anchor for securing the polypeptide in intracellular membranes and subsequently in the virion envelope. The amino acid sequence RRSRR at positions 764 to 768 resembles the probable cleavage site of MHV proteins and is located in an extremely hydrophilic region of the molecule. Cleavage of the S 150K polypeptide at this point would yield two subunits of 85K and 65K, S1 and S2 respectively. The S1 and S2 subunits obtained from purified virus migrate on SDS–polyacrylamide gels with M_r values of 100K to 120K (Deregt *et al.*, 1987; King & Brian, 1982) indicating that both S1 and S2 contain significant levels of glycosylation. Accordingly, the predicted amino acid sequence of the S gene contains 11 possible sites for the addition of *N*-linked oligosaccharides on the S1 and nine sites in the S2 subunit.

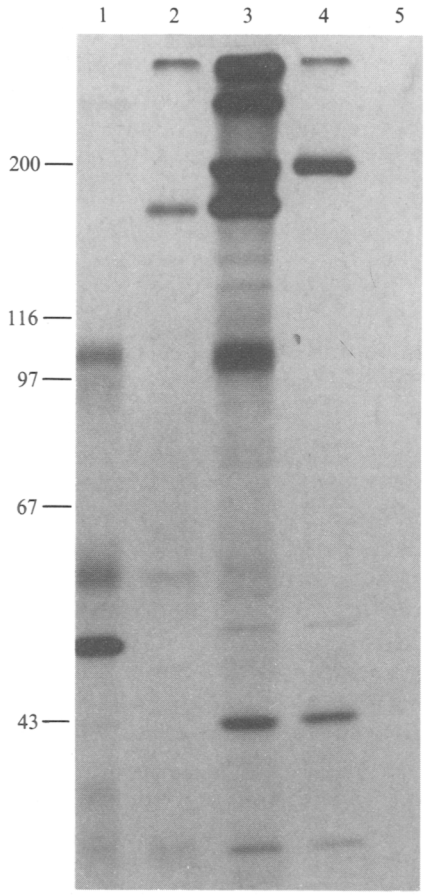


Fig. 3. Synthesis of the BCV S polypeptide by recombinant baculovirus AcSD in *S. frugiperda* cells. Cells were infected with AcSD and radiolabelled at 40 h after infection. The recombinant S polypeptides were immunoprecipitated with S-specific monoclonal antibodies and analysed by SDS–PAGE on 7·5% gels. Lane 1, ^{35}S-labelled BCV; lanes 2 to 5: immunoprecipitates from (lane 2) AcSD-infected Sf9 cells, (lane 3) BCV-infected MDBK cells, (lane 4) uninfected MDBK cells and (lane 5) uninfected Sf9 cells. Positions of M_r markers are indicated at the left, $\times 10^{-3}$.

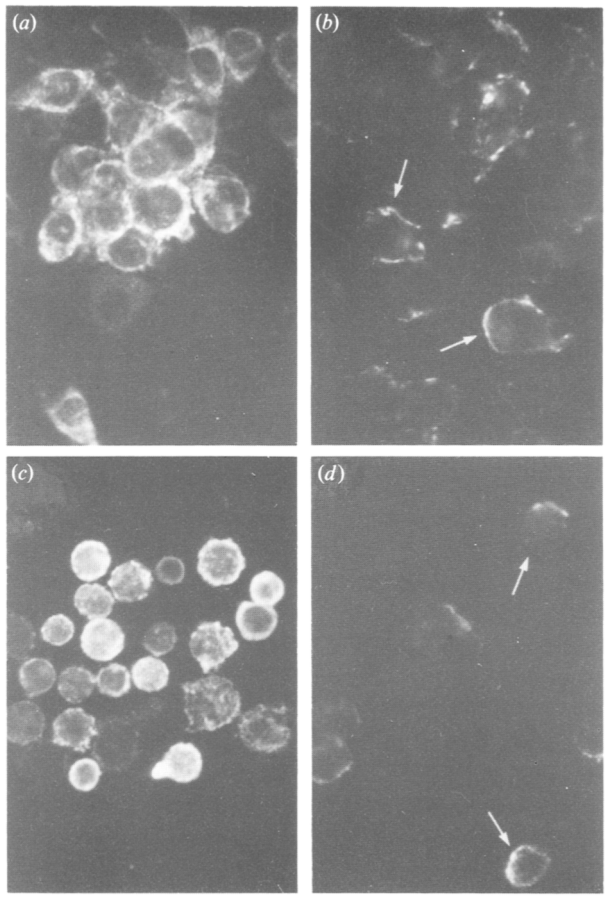


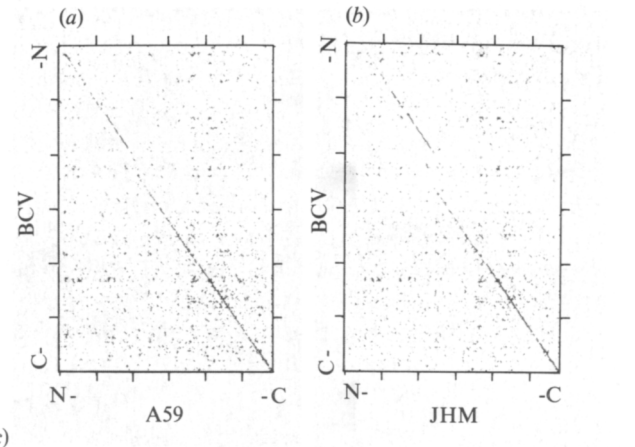
Fig. 4. Immunofluorescence of BCV-infected MDBK cells and AcNPV recombinant AcSD-infected *S. frugiperda* cells. Cells were incubated with S-specific monoclonal antibodies, washed in cold PBS and incubated with FITC-labelled goat anti-mouse IgG. (*a*) Methanol-fixed BCV-infected MDBK cells; (*b*) BCV-infected MDBK cells without methanol fixation; (*c*) methanol-fixed AcSD-infected *S. frugiperda* cells; (*d*) AcSD-infected *S. frugiperda* cells without methanol fixation. Arrows indicate patches of surface fluorescence on unfixed cells.

Expression of cloned S cDNA

In order to characterize the product of the recombinant S gene and produce sufficient quantities of the BCV S protein for analysis of its immunogenic properties in animals, an expression cassette for the S gene was assembled from two overlapping clones, pCVA12H and pCVA12I (Fig. 1) as described in Methods and inserted into the genome of AcNPV. Infection of Sf9 cells with the recombinant, AcSD, resulted in the production of a product of 180K, which is specifically immunoprecipitated by S-specific monoclonal antibodies (Fig. 3, lane 2). The 180K product is slightly smaller than the uncleaved S gene product produced in BCV-infected MDBK cells (Fig. 3, lane 3). The slightly decreased size is presumably due to the differences in oligosaccharide processing typical of insect cells (Butters & Hughes, 1981). No cleavage of the recombinant S polypeptide was detected, indicating that production of recombinant S1 and S2 subunits occurred, if at all, at a very low level. However, the S1 and S2 subunits, which comigrate as 100K polypeptides, are abundant in BCV-infected cells (lane 3).

Intracellular localization of the S gene product

In order to examine the intracellular distribution of the recombinant S polypeptide, BCV-infected MDBK cells and AcSD-infected *S. frugiperda* cells were examined by immunofluorescence. At 24 h after infection, the S gene product in BCV-infected MDBK cells was distributed throughout the cell with evidence of reticular concentrations of antigen in the cytoplasm (Fig. 4*a*). In unfixed



(c)

```
BCV 413 LQSFNYRIDTTATSCQLYYNLPAANVSVSRFNPSTWNRRFGFTEQFVFKPQPVGVFTHHDVVYAQ
JHM 411 LQtaNYKIDTaATSCQLYYsLPknNVtinnyNPSsWNRRYGFkvn--------------------
A59 411 LQtaNYKIDTaATSCQLhYtLPknNVtinnhNPSsWNRRYGFndagVFgKnQ------HDVVYAQ

BCV 478 HCFKAPKNFCPCKL-D-GSLCVGNGPGIDAGYKNSGIGTCPAGTNYLTCHNAAQCDCLCTPDPIT
JHM     ----------------------------------------------------------------
A59 470 qCFtvrssyCPCaqpDlvSpCt----------------------------------------------

BCV 541 SKSTGPYKCPQTKYL---VGIGEHCSGLAIKSDYCGG-NP---CTCQPQAFLGWSVDSCLQGDRC
JHM 456 ----------------------------------------------------------DRC
A59 492 ---------tQTKpksafVnvGdHCeGLgvleDnCGnadPhkgCiCannsFiGWShDtCLvnDRC

BCV 600 NIFANFIFHDVNSGTTCSTDLQKSNTDIILGVCVNYDLYGITGQGIFVEVNATYYNSWQNLLYDS
JHM 459 qIFANillngiNSGTTCSTDLQlpNTevatGVCVrYDLYGITGQGvFkEVkAdYYNSWQaLLYDv
A59 549 qIFANillngiNSGTTCSTDLQlpNTevvtGiCVkYDLYGITGQGvFkEVkAdYYNSWQtLLYDv
```

Fig. 5. (*a*, *b*) Amino acid sequence comparison between the predicted amino acid sequences of the S polypeptides of BCV and MHV strains A59 and JHM. The search was conducted with a window of six amino acids and a requirement of five matches. N- and -C indicate amino and carboxy termini, respectively. (*c*) Partial amino acid sequence comparison between the S polypeptides of BCV, and MHV A59 and JHM. Upper case letters indicate conserved amino acids; - indicates insertions to maximize alignment. The position in the sequence is indicated at the left.

MDBK cells, surface fluorescence was indicative of antigen expression on the extracellular surface of the plasma membrane (*b*). In AcSD-infected insect cells, the distribution of the recombinant polypeptide was virtually identical to that seen in the BCV-infected MDBK cells (Fig. 4*c* and *d*). The increased fluorescence in the fixed insect cells was most likely due to the higher levels of antigen produced by the recombinant virus compared to BCV-infected MDBK cells (data not shown). The presence of surface fluorescence on AcSD-infected *S. frugiperda* cells indicated that the S gene product was transported to the cell surface and that cleavage of S was not necessary for intracellular processing and transport of the recombinant polypeptide.

Comparison between the S glycoprotein genes of bovine coronavirus and murine hepatitis virus

The sequences of the S genes of MHV strains A59 and JHM, which are antigenically related to BCV (Hogue *et al.*, 1984), have been determined (Luytjes *et al.*, 1987; Schmidt *et al.*, 1987). Comparison of the amino acid sequence of BCV with both strains of MHV (Fig. 5) graphically demonstrates the significant degree of similarity between BCV and MHV-JHM and MHV-A59, 75·5% and 73·9%, respectively. The highest similarity between the S polypeptides of BCV and both strains of MHV is found in the carboxy-terminal half, the S2 subunits, as previously noted in the comparison of MHV-A59 and MHV-JHM (Luytjes *et al.*, 1989). It was also reported that the S genes MHV-A59 and MHV-JHM exhibited 93% similarity, with the exception of a stretch of 89 amino acids unique to MHV-A59. The same region is also indicated by the discontinuities in the plots shown in Fig. 6(*a*) and (*b*) corresponding to amino acids 457 to 596 of the BCV S amino acid sequence. Direct examination of the BCV S amino acid sequence over this region (Fig. 5*c*) indicated that the BCV has an even more extensive unique amino acid sequence when compared with MHV-JHM than was found previously with MHV-A59. Beginning at amino acid 444 of the BCV S sequence is an 11 amino acid sequence, NP(S/T)WNRRYGF, which is conserved in the three viruses. After an additional three amino acids unique to each virus, the deletion in the MHV-JHM sequence is evident and extends for the next 138 amino acids of the BCV S sequence. In contrast, there is intermittent similarity between BCV and MHV-A59 through this same region. At amino acid 596, the similarity in the sequences resumes with the conserved sequence DRC(Q/N)IFAN. The most notable feature of this region is its extremely high content of cysteine residues, 15 of the 138 amino acids in this region. The difference in amino acid sequence over this region accounts for most of the differences in the lengths of the S reading frames among the three genes.

The sequence RRSRR at positions 764 to 768 is similar to the postulated cleavage site of MHV strains JHM and A59 (Luytjes *et al.*, 1989; Schmidt *et al.*, 1987) and avian infectious bronchitis virus (IBV) (Cavanagh *et al.*, 1986*b*) except that the serine residue at position 767 in the BCV sequence is replaced by an alanine residue in both strains of MHV and phenylalanine in IBV. Luytjes *et al.* (1987) have identified the amino terminus of the mature 90A of MHV-A59 as SVSTGYRLTTFE. A similar sequence, AITTGYRFTNFE, exists in the BCV S amino acid sequence beginning at amino acid 769 further suggesting that the initial cleavage of the BCV S in virus-infected cells occurs at some point between residues 763 and 769.

The sequence KWPWYVW and the cysteine-rich sequence extending from residues 1329 to 1346 is thought to serve as part of the membrane anchor and is a common feature of coronavirus S polypeptides (Binns *et al.*, 1985; Cavanagh *et al.*, 1986*b*; Luytjes *et al.*, 1987; Schmidt *et al.*, 1987). This region may also contain the site of fatty acylation (Ricard & Sturman, 1985).

Discussion

The characterization of the S gene of BCV completes the sequence determination of the structural protein genes of BCV (Lapps *et al.*, 1987; Parker *et al.*, 1989). The open reading frame of the BCV S gene extends from nucleotides 3301 to 7389 distal to the 3′ end of the virion genome RNA and its location conforms to the consensus gene order 5′ HE–S–NS–NS–E1–N 3′, determined for all other coronaviruses yet reported, although the presence of a functional HE gene is unique to only a few coronaviruses.

Of the numerous coronavirus S genes characterized, the BCV S gene is the most closely related to the S genes of MHV strains A59 and JHM and contains the largest S gene of the three viruses. The hydropathy analysis of the S gene product indicates it has a structure typical of membrane glycoproteins. The amino-terminal hydrophobic domain, presumably a signal sequence, would probably direct the transit of the nascent polypeptide across the membranes of the rough endoplasmic reticulum. It is not known at present whether the signal is removed during maturation of the S gene product. The hydrophobic domain near the carboxyl terminus is likely to be an anchor sequence which secures the S1–S2 complex of the peplomer in the virion envelope. The sequence RRSRR, identified as the probable cleavage site, lies in an extremely hydrophilic region of the precursor polypeptide and, as demonstrated above, cleavage is not necessary prior to glycosylation and intracellular transport of the recombinant protein produced in insect cells. Fully glycosylated, uncleaved forms of the S gene product have been detected in BCV-infected MDBK cells (Deregt *et al.*, 1987) and in cells infected by a vaccinia virus recombinant expressing the infectious bronchitis virus S gene (Tomley *et al.*, 1987). It will be of interest to determine whether cell fusion can be induced by the expression of the recombinant S protein or its subunits in order to understand better the processes involved in the initiation of infection.

Numerous functional and structural properties of the S proteins of coronaviruses have been identified, including its identification as a target for neutralizing antibody (Collins *et al.*, 1982) and its involvement in cell binding (Cavanagh & Davis, 1986), cell fusion (Frana *et al.*, 1985; Sturman *et al.*, 1985), virulence (Dalziel *et al.*, 1986; Makino *et al.*, 1987), and haemagglutination (Mockett *et al.*, 1984). The ability of monoclonal antibodies specific for the haemagglutinin glycoprotein of BCV to neutralize infectivity (Deregt & Babiuk, 1987) suggests that BCV has acquired additional neutralizing epitopes not found in other coronaviruses and that some of the properties of S in other coronaviruses may have been either duplicated or shifted from the S of BCV to the HE polypeptide. The recently determined similarity between the HE haemagglutinin of BCV and the HA1 of type C influenza viruses (Parker *et al.*, 1989; Vlasak *et al.*, 1988*a*, *b*) suggests that cell surface binding by BCV is mediated by HE and that the role of S in initiation of infection by BCV occurs at a stage after initial cell surface binding, possibly by promoting fusion of the viral envelope with the membranes of acidic endosomal compartments to facilitate penetration of the target cell. Such a role is common to viral glycoproteins which require proteolytic activation and display the ability to direct cell fusion (Gething *et al.*, 1978; Richardson *et al.*, 1980). Alternatively, if the BCV S also possesses cell binding activity as in other coronaviruses, BCV would then contain two surface components capable of binding the surface of target cells. This possibility is under investigation.

Makino *et al.* (1987) have determined that the neutralizing epitope(s) of MHV-A59 are on the carboxyl portion of the S gene product, analogous to S2 of BCV, apparently in close proximity to the envelope of the virus. More recently, Luytjes *et al.* (1989) have identified a conserved amino acid sequence, LLGCIGSTC, in the S polypeptide of MHV-A59 which constitutes a neutralizing epitope, and may be involved in a vital function of the MHV-A59 S polypeptide. A search of the sequence of the BCV S gene product indicates that a similar sequence, VLGCLGSAC, extends from positions 899 to 907. Preliminary experiments (D. Yoo, unpublished) indicate that this region may constitute a portion of the fusogenic domain of BCV S2 subunit. A second neutralizing epitope, VKSQTTRIN, on the predicted peplomer stalk of MHV-JHM is very similar in sequence to VKSQSSRIN, which extends from positions 1127 to 1135 on the S2 subunit of the BCV S polypeptide. Whether or not this is also an important epitope in the case of BCV remains to be determined.

Analysis of the S-specific neutralizing monoclonal antibodies prepared by Deregt *et al.* (1987) indicates that they all bind to sites on the S1 subunit of the BCV S protein (D. Yoo, unpublished). A similar antigenic characterization was previously reported in that only the amino-terminal portion of the avian infectious bronchitis virus S polypeptide, analogous to the S1 of BCV, induces neutralizing antibody in chickens (Cavanagh *et al.*, 1986*a*). Based upon proteolytic digestion of antigen–antibody complexes, Deregt *et al.* (1989) suggested that two neutralizing epitopes on the S of BCV are located on a 37K fragment near the carboxy terminus of the S1 subunit of the S glycoprotein, just upstream from the trypsin cleavage site. Further characterization of these sites is under way.

The comparison of the amino acid sequences of BCV and MHV strains A59 and JHM S polypeptides indicates that these viruses arose from a common progenitor which

had diverged from the ancestor(s) of the remaining coronaviruses. The most notable difference between the MHV S genes was a stretch of 89 amino acids which was absent in the S gene of MHV-JHM possibly due to a recombination event mediated by a repeated nucleotide sequence immediately flanking the deletion (Luytjes *et al.*, 1987). As shown in Fig. 5(*c*), it appears that a deletion from the common progenitor virus or the acquisition of an additional amino acid domain has occurred during the divergence of BCV and the murine hepatitis viruses. Such a hypothesis concerning the evolutionary relationship between BCV and the MHV strains is complicated due to the additional variation in amino acid sequence over the entire S polypeptide, especially near the amino terminus of the S1 subunit as shown in Fig. 5. The determination of the S nucleotide sequences from other related coronaviruses such as the human virus OC-43 and haemagglutinating encephalomyelitis virus of swine should provide additional information concerning the evolution of this group of coronaviruses.

We thank Dr Dirk Deregt for his valuable suggestions and discussions and Dr C. Schmaljohn for her assistance with the AcNPV expression system. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada and is published with the permission of the Director of the Veterinary Infectious Disease Organization as article no. 81.

References

BINNS, M. M., BOURSNELL, M. E. G., CAVANAGH, D., PAPPIN, D. J. C. & BROWN, T. D. K. (1985). Cloning and sequencing of the gene encoding the spike protein of the coronavirus IBV. *Journal of General Virology* **66**, 719–726.

BUTTERS, T. C. & HUGHES, R. C. (1981). Isolation and characterization of mosquito cell membrane glycoproteins. *Biochimica et biophysica acta* **640**, 655–671.

CALLEBAUT, P. E. & PENSAERT, M. B. (1980). Characterization and isolation of structural polypeptides in haemagglutinating encephalomyelitis virus. *Journal of General Virology* **48**, 193–204.

CAVANAGH, D. & DAVIS, P. J. (1986). Coronavirus IBV: removal of spike glycopolypeptide S1 by urea abolishes infectivity and haemagglutination but not attachment to cells. *Journal of General Virology* **67**, 1443–1448.

CAVANAGH, D., DAVIS, P. J., DARBYSHIRE, J. H. & PETERS, R. W. (1986*a*). Coronavirus IBV: virus retaining spike glycopolypeptide S2 but not S1 is unable to induce virus neutralization or haemagglutination-inhibiting antibody, or induce chicken tracheal protection. *Journal of General Virology* **67**, 1435–1442.

CAVANAGH, D., DAVIS, P. J., PAPPIN, D. J. C., BINNS, M. M., BOURSNELL, M. E. G. & BROWN, T. D. K. (1986*b*). Coronavirus IBV: partial amino terminal sequencing of spike polypeptide S2 identifies the sequence Arg-Arg-Phe-Arg-Arg at the cleavage site of the spike precursor propolypeptide of IBV strains Beaudette and M41. *Virus Research* **4**, 133–143.

COLLINS, A. R., KNOBLER, R. L., POWELL, H. & BUCHMEIER, M. J. (1982). Monoclonal antibodies to murine hepatitis virus 4 (strain JHM) define the viral glycoprotein responsible for attachment and cell-cell fusion. *Virology* **119**, 358–371.

DALZIEL, R. G., LAMBERT, P. W., TALBOT, P. J. & BUCHMEIER, M. J. (1986). Site specific alteration of murine hepatitis virus type 4 peplomer glycoprotein E2 results in reduced neurovirulence. *Journal of Virology* **59**, 463–471.

DEA, S., ROY, R. S. & BEGIN, M. E. (1980). Bovine coronavirus: isolation and cultivation in continuous cell lines. *American Journal of Veterinary Research* **41**, 30–38.

DEREGT, D. & BABIUK, L. A. (1987). Monoclonal antibodies to bovine coronavirus: characteristics and topographical mapping of neutralizing epitopes on the E2 and E3 glycoproteins. *Virology* **161**, 410–420.

DEREGT, D., SABARA, M. & BABIUK, L. A. (1987). Structural proteins of bovine coronavirus and their intracellular processing. *Journal of General Virology* **68**, 2863–2877.

DEREGT, D., PARKER, M. D., COX, G. C. & BABIUK, L. A. (1989). Mapping of neutralizing epitopes to fragments of the bovine coronavirus E2 protein by proteolysis of antigen–antibody complexes. *Journal of General Virology* **70**, 647–658.

FLEMING, J. O., SHUBIN, R. A., SUSSMAN, M. A., CASTEEL, N. & STOHLMAN, S. A. (1989). Monoclonal antibodies to the matrix (E1) glycoprotein of mouse hepatitis virus protect mice from encephalitis. *Virology* **168**, 162–167.

FRANA, M. F., BEHNKE, J. N., STURMAN, L. A. & HOLMES, K. V. (1985). Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: host-dependent differences in proteolytic cleavage and cell fusion. *Journal of Virology* **56**, 912–920.

GETHING, M. J., WHITE, J. M. & WATERFIELD, M. D. (1978). Purification of the fusion protein of Sendai virus: analysis of the NH2-terminal sequence generated during precursor activation. *Proceedings of the National Academy of Sciences, U.S.A.* **75**, 2737–2740.

GRUNSTEIN, M. & HOGNESS, D. (1975). Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proceedings of the National Academy of Sciences, U.S.A.* **72**, 3961–3965.

HENIKOFF, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28**, 351–359.

HOGUE, B. G. & BRIAN, D. A. (1986). Structural proteins of human respiratory coronavirus OC43. *Virus Research* **5**, 131–144.

HOGUE, B. G., KING, B. & BRIAN, D. A. (1984). Antigenic relationships among proteins of bovine coronavirus, human respiratory coronavirus OC43, and mouse hepatitis coronavirus A59. *Journal of Virology* **51**, 384–388.

KING, B. & BRIAN, D. A. (1982). Bovine coronavirus structural proteins. *Journal of Virology* **42**, 700–707.

KING, B., POTTS, B. J. & BRIAN, D. A. (1985). Bovine coronavirus hemagglutinin protein. *Virus Research* **2**, 53–59.

KOZAK, M. (1987). Point mutations define a sequence flanking the AUG initiation codon that moderate translation by eucaryotic ribosomes. *Cell* **44**, 283–292.

LAEMMLI, U. K. (1970). Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature, London* **227**, 680–685.

LAPPS, W., HOGUE, B. G. & BRIAN, D. A. (1987). Sequence analysis of the bovine coronavirus nucleocapsid and matrix protein genes. *Virology* **157**, 47–57.

LUYTJES, W., STURMAN, L. S., BREDENBEEK, P. J., CHARITE, J., VAN DER ZEIJST, B. A. A., HORZINEK, M. C. & SPAAN, W. J. M. (1987). Primary structure of the glycoprotein E2 of coronavirus MHV-A59 and identification of the trypsin cleavage site. *Virology* **161**, 479–487.

LUYTJES, W., GEERTS, D., POSTHUMUS, W., MELOEN, R. & SPAAN, W. (1989). Amino acid sequence of a conserved neutralizing epitope of murine coronaviruses. *Journal of Virology* **63**, 1408–1412.

MACNAUGHTON, M. R., DAVIES, H. A. & NERMUT, M. V. (1978). Ribonucleoprotein-like structures from coronavirus particles. *Journal of General Virology* **39**, 545–549.

MAKINO, S., FLEMING, J. O., KECK, J. G., STOHLMAN, S. A. & LAI, M. M. C. (1987). RNA recombination of coronaviruses: localization of neutralization epitopes and neuropathogenic determinants on the carboxyl-terminus of peplomers. *Proceedings of the National Academy of Sciences, U.S.A.* **84**, 6567–6571.

MATSUURA, Y., POSSEE, R. D., OVERTON, H. A. & BISHOP, D. H. L. (1987). Baculovirus expression vectors: the requirements for high level expression of proteins, including glycoproteins. *Journal of General Virology* **68**, 1233–1250.

MEBUS, C. A. (1978). Pathogenesis of coronaviral infection in calves. *Journal of the American Veterinary Medical Association* **173**, 631–632.

MOCKETT, A. P. A., CAVANAGH, D. & BROWN, T. D. K. (1984). Monoclonal antibodies to the S1 spike and membrane proteins of avian infectious bronchitis coronavirus strain Massachusetts M41. *Journal of General Virology* **65**, 2281–2286.

PARKER, M. D., COX, G. J., DEREGT, D., FITZPATRICK, D. R. & BABIUK, L. A. (1989). Cloning and *in vitro* expression of the gene for the E3 haemagglutinin glycoprotein of bovine coronavirus. *Journal of General Virology* **70**, 155–164.

RICARD, C. S. & STURMAN, L. S. (1985). Isolation of the subunits of the coronavirus glycoprotein E2 by hydroxyapatite high-performance liquid chromatography. *Journal of Chromatography* **326**, 191–197.

RICHARDSON, C. D., SCHEID, A. & CHOPPIN, P. W. (1980). Specific inhibition of paramyxovirus and myxovirus replication by oligopeptides with amino acid sequences similar to those at the N-termini of the F1 or HA2 viral polypeptides. *Virology* **105**, 205–222.

SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, U.S.A.* **74**, 5463–5467.

SCHMIDT, I., SKINNER, M. & SIDDELL, S. (1987). Nucleotide sequence of the gene encoding the surface projection glycoprotein of coronavirus MHV-JHM. *Journal of General Virology* **68**, 47–56.

STURMAN, L. S., RICARD, C. S. & HOLMES, K. V. (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. *Journal of Virology* **56**, 904–911.

SUMMERS, M. D. & SMITH, G. E. (1987). *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures.* Texas Agricultural Experiment Station Bulletin No. 1555.

TOMLEY, F. M., MOCKETT, A. P. A., BOURSNELL, M. E. G., BINNS, M. M., COOK, J. K. A., BROWN, T. D. K. & SMITH, G. L. (1987). Expression of the infectious bronchitis virus spike protein by recombinant vaccinia virus and induction of neutralizing antibodies in vaccinated mice. *Journal of General Virology* **68**, 2291–2298.

TOOZE, J., TOOZE, S. & WARREN, G. (1984). Replication of coronavirus MHV-A59 in sac-cells: determination of the first site of budding of progeny virions. *European Journal of Cell Biology* **33**, 281–293.

VLASAK, R., LUYTJES, W., LEIDER, J., SPAAN, W. & PALESE, P. (1988*a*). The E3 protein of bovine coronavirus is a receptor-destroying enzyme with acetylesterase activity. *Journal of Virology* **62**, 4686–4690.

VLASAK, R., LUYTJES, W., SPAAN, W. & PALESE, P. (1988*b*). Human and bovine coronaviruses recognize sialic-acid containing receptors similar to those of influenza C viruses. *Proceedings of the National Academy of Sciences, U.S.A.* **85**, 4526–4529.

VON HEIJNE, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Research* **14**, 4683–4690.

WEGE, H., WINTER, J. & MEYERMANN, R. (1988). The peplomer protein E2 of coronavirus JHM as a determinant of neurovirulence: definition of critical epitopes by variant analysis. *Journal of General Virology* **69**, 87–98.

(*Received 26 May 1989; Accepted 17 October 1989*)