VIROLOGY 176, 296–301 (1990)

Deduced Sequence of the Bovine Coronavirus Spike Protein and Identification of the Internal Proteolytic Cleavage Site[1]

SUSHMA ABRAHAM, THOMAS E. KIENZLE,[2] WILLIAM LAPPS,[3] AND DAVID A. BRIAN[4]

Department of Microbiology, The University of Tennessee, Knoxville, Tennessee 37996-0845

Received October 13, 1989; accepted January 17, 1990

The sequence of the spike (also called peplomer or E2) protein gene of the Mebus strain of bovine coronavirus (BCV) was obtained from cDNA clones of genomic RNA. The gene sequence predicts a 150,825 mol wt apoprotein of 1363 amino acids having an N-terminal hydrophobic signal sequence of 17 amino acids, 19 potential N-linked glycosylation sites, a hydrophobic anchor sequence of approximately 17 amino acids near the C terminus, and a hydrophilic cysteine-rich C terminus of 35 amino acids. An internal Lys–Arg–Arg–Ser–Arg–Arg sequence predicts a protease cleavage site between amino acids 768 and 769 that would separate the S apoprotein into S1 and S2 segments of 85690 and 65153 mol wt, respectively. Amino terminal amino acid sequencing of the virion-derived gp100 spike subunit confirmed the location of the predicted cleavage site, and established that gp120 and gp100 are the glycosylated virion forms of the S1 and S2 subunits, respectively. Sequence comparisons between BCV and the antigenically related mouse hepatitis coronavirus revealed more sequence divergence in the putative knob region of the spike protein (S1) than in the stem region (S2).

The bovine coronavirus (BCV) is an important cause of neonatal calf diarrhea (*14, 26*) and may also be the cause of winter dysentery in adult cattle (*30*). The mechanisms by which BCV causes disease and persistent infection are not understood, nor are current vaccines universally regarded as effective. Toward these ends, we have begun a detailed study of the BCV protein and genome structure.

BCV is comprised of four major structural proteins (*17*). These are (i) a 200-kDa spike (peplomer) glycoprotein (S), that exists on the virion as cleaved subunits of approximately 120 and 100 kDa, (ii) a 140-kDa glycoprotein (HE) that has both hemagglutinating (*18*) and esterase (*37*) activities, and which is comprised of two identical, disulfide-linked 65-kDa subunits (*10, 12, 16, 28*), (iii) a 26-kDa integral membrane glycoprotein (M) (*21*), and , (iv) an internal phosphorylated nucleocapsid protein (N) (*21*). Of these, the S protein is presumed to be the major structure by which coronaviruses attach to cells and initiate infection (reviewed by Spaan *et al.* (*34*)). The HE protein, however, may also bind to cells to initiate infection, and for BCV, the relative importance of these two proteins in initiating infection is not known. Both S and HE are probably important in inducing immunity since antibodies to each are known to neutralize virus infectivity in cell culture and in calves (*8, 9*). S and HE, therefore, may both be useful in developing effective engineered vaccines against BCV.

cDNA cloning of BCV genomic RNA was accomplished essentially as previously described (*11, 21*) except that random 5-mer oligodeoxynucleotides (Pharmacia) and 17-mer oligodeoxynucleotides of specific sequences were used as primers for first-strand synthesis. Clones were mapped relative to one another and to the 3′ end of the genome using a matrix spot hybridization technique. Some clones were sequenced by the chemical method of Maxam and Gilbert (*25*) and some by the dideoxynucleotide-induced chain termination method of Sanger (*31*) as described by Kraft *et al.* (*19*) using Sequenase enzyme (United States Biochemicals). For much of the sequencing, restriction endonuclease fragments were subcloned into the pGEM4Z vector (Promega) and forward and reverse sequencing

[1] Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession No. M31053.

[2] Present address: Department of Molecular Biology and Microbiology, Case Western Reserve University, Cleveland, OH 44106.

[3] Present address: DNX Corporation, One President Street, Athens, OH 45701-2979.

[4] To whom requests for reprints should be addressed.

0042-6822/90 $3.00

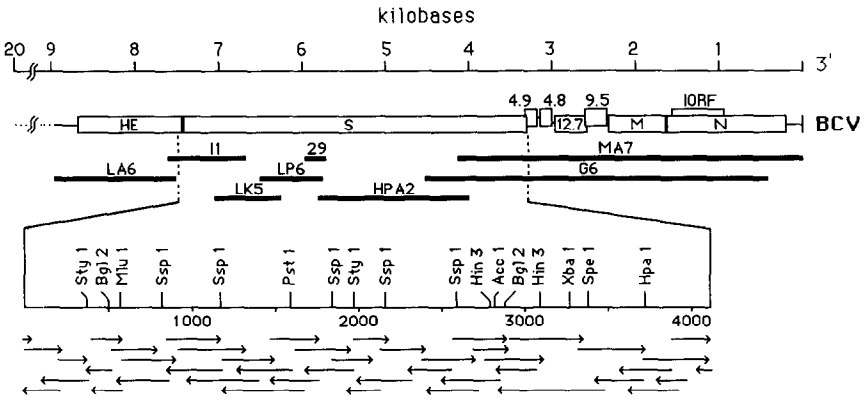

FIG. 1. Gene map of the BCV genome, cDNA clone positions, and strategy for sequencing the S gene.

primers for the pGEM vectors were used. Sequence-specific oligodeoxynucleotides were also synthesized and used for sequencing within certain regions of the large clones.

The amino-terminal ends of purified gp120 and gp100 subunits were subjected to sequencing by the method of Matsudaira (*24*). Unlabeled BCV was purified by isopycnic sedimentation in sucrose gradients and the proteins were electrophoretically separated after reduction in 2-mercaptoethanol (*17*) and electroblotted (*13*) onto polyvinylidene difluoride membrane (*24*). Proteins were visualized by staining with Coomassie brilliant blue and the gp120 and gp100 bands were excised and shipped to Dr. Matsudaira for analysis.

Complete sequencing of clone MA7 which extends 4.2 kilobases from the 3′ end of the genome (Fig. 1) revealed a continuous open reading frame located on the 5′ side of the ORF for a potential 4.9-kDa protein (Abraham *et al.*, to be published elsewhere). The deduced amino acid sequence of the extended ORF demonstrated high sequence similarity to the C-terminal end of the antigenically related MHV-A59 (*22*) and MHV-JHM (*32*) S proteins, both antigenic homologs of the BCV S protein (*13*). These data suggested that the S protein gene of BCV lies in the same relative position on the genome as does the spike protein gene of MHV. To complete the sequencing of the S gene, both strands of three clones, I1, HPA2, and G6, generated by random priming, and three clones, LK5, LP6 and 29, generated by specific priming, were sequenced (Fig. 1).

The total sequence for the putative S ORF extended to a position 7.4 kb from the 3′ end of the genome and contained 4089 bases (Fig. 2). We conclude this ORF to be the S gene since it potentially encodes a 1363 amino acid protein of 150,825 Da, the approximate size of the unglycosylated spike precursor (*10*), and because its deduced amino acid sequence shows extensive sequence similarity throughout with the S proteins of both strains of MHV. Five other open reading frames ranging in size from 34 to 66 amino acids were also found within the S gene sequence in the plus one reading frame, but their significance is not known at this time. The putative S ORF is preceded immediately upstream (beginning at base 12 in Fig. 2) by the consensus CYAAAC sequence thought to play a role in leader priming of coronavirus transcription. The sequence is also found three times within the S ORF, beginning at positions 817, 1667, and 3776, but it is not established that transcripts initiate at any of these sites.

Five features of the deduced BCV S protein reflect the properties of four other coronavirus spike proteins that have been characterized to date from nucleotide sequence data (*1, 2, 15, 20, 22, 27, 29, 32*). (i) There is an N-terminal hydrophobic stretch of amino acids which predicts a signal peptide with a cleavage site between amino acids 17 and 18 (*38*). (ii) There are 19 potential asparagine-linked glycosylation sites that could give rise to the only kind of glycosylation demonstrated for this protein (Hogue and Brian, unpublished data; *10*). (iii) There is a hydrophobic stretch of 17 amino acids near the C terminus that could serve as a stop-transfer and anchor sequence. (iv) There is a stretch of 8 amino acids on the immediate N-terminal side of the predicted anchor sequence (-K-W-P-W-Y-V-W-L-, beginning with amino acid 1305) that is identical in all coronavirus S proteins sequenced to date. (v) There is a cysteine-rich hydrophilic C-terminus of 35 amino acids that is probably the intravirion domain. In common with MHV- (*22, 32*) and IBV (*1, 2, 20, 27*), but not in common with TGEV (*15, 29*; Tung and Brian, unpublished) and FIPV (*15*), is also an internal sequence of basic amino acids that, in the case of MHV and IBV, lies on the immediate N-terminal side of the protease cleavage site (*6, 22*). In BCV the sequence is K-R-R-S-R-R beginning

```
          10        20        30        40        50        60        70        80        90       100       110       120
TAGACCATAATCTAAACATGTTTTTGATACTTTTAATTTCCTTACCAATGGCTTTTGCTGTTATAGGAGATTTAAAGTGTACTACGGTTTCCATTAATGATGTTGACACCGGTGCTCCTT
                  M  F  L  I  L  L  I  S  L  P  M  A  F  A  V  I  G  D  L  K  C  T  T  V  S  I  N  D  V  D  T  G  A  P

         130       140       150       160       170       180       190       200       210       220       230       240
CTATTAGCACTGATATTGTCGATGTTACTAATGGTTTAGGTACTTATTATGTTTTAGATCGTGTGTATTTAAATACTACGTTGTTGCTTAATGGTTACTACCCTACTTCAGGTTCTACAT
S  I  S  T  D  I  V  D  V  T  N  G  L  G  T  Y  Y  V  L  D  R  V  Y  L [N] T  T  L  L  L  N  G  Y  Y  P  T  S  G  S  T
               40

         250       260       270       280       290       300       310       320       330       340       350       360
ATCGTAATATGGCACTGAAGGGAACTTTACTATTGAGCAGACTATGGTTTAAACCACCTTTTCTTTCTGATTTTATTAATGGTATTTTTGCTAAGGTCAAAAATACCAAGGTTATTAAAA
Y  R  N  M  A  L  K  G  T  L  L  L  S  R  L  W  F  K  P  P  F  L  S  D  F  I  N  G  I  F  A  K  V  K  N  T  K  V  I  K
               80

         370       380       390       400       410       420       430       440       450       460       470       480
AGGGTGTAATGTATAGTGAGTTTCCTGCTATAACTATAGGTAGTACTTTTGTAAATACATCCTATAGTGTGGTAGTACAACCACATACTACCAATTTGGATAATAAATTACAAGGTCTCT
K  G  V  M  Y  S  E  F  P  A  I  T  I  G  S  T  F  V [N] T  S  Y  S  V  V  V  Q  P  H  T  T  N  L  D  N  K  L  Q  G  L
              120

         490       500       510       520       530       540       550       560       570       580       590       600
TAGAGATCTCTGTTTGCCAGTATACTATGTGCGAGTACCCACATACGATTTGTCATCCTAATCTGGGTAATAAACGCGTAGAACTATGGCATTGGGATACAGGTGTTGTTTCCTGTTTAT
L  E  I  S  V  C  Q  Y  T  M  C  E  Y  P  H  T  I  C  H  P  N  L  G  N  K  R  V  E  L  W  H  W  D  T  G  V  V  S  C  L
              160

         610       620       630       640       650       660       670       680       690       700       710       720
ATAAGCGTAATTTCACATATGATGTGAATGCTGATTACTTGTATTTCCATTTTTATCAAGAAGGTGGTACTTTTTATGCATATTTTACAGACACTGGTGTTGTTACTAAGTTTCTGTTTA
Y  K  R [N] F  T  Y  D  V  N  A  D  Y  L  Y  F  H  F  Y  Q  E  G  G  T  F  Y  A  Y  F  T  D  T  G  V  V  T  K  F  L  F
              200

         730       740       750       760       770       780       790       800       810       820       830       840
ATGTTTATTTAGGCACGGTGCTTTCACATTATTATGTCCTGCCTTTGACTTGTTCTAGTGCTATGACTTTAGAATATTGGGTTACACCTCTCACTTCTAAACAATATTTACTAGCTTTCA
N  V  Y  L  G  T  V  L  S  H  Y  Y  V  L  P  L  T  C  S  S  A  M  T  L  E  Y  W  V  T  P  L  T  S  K  Q  Y  L  L  A  F
              240

         850       860       870       880       890       900       910       920       930       940       950       960
ATCAAGATGGTGTTATTTTTAATGCTGTTGATTGTAAGAGTGATTTTATGAGTGAGATTAAGTGTAAAACACTATCTATAGCACCATCTACTGGTGTTTATGAATTAAACGGTTACACTG
N  Q  D  G  V  I  F  N  A  V  D  C  K  S  D  F  M  S  E  I  K  C  K  T  L  S  I  A  P  S  T  G  V  Y  E  L  N  G  Y  T
              280

         970       980       990      1000      1010      1020      1030      1040      1050      1060      1070      1080
TTCAGCCAATTGCAGATGTTTACCGACGTATACCTAATCTTCCCGATTGTAATATAGAGGCTTGGCTTAATGATAAGTCGGTGCCCTCTCCATTAAATTGGGAACGTAAGACCTTTTCAA
V  Q  P  I  A  D  V  Y  R  R  I  P  N  L  P  D  C  N  I  E  A  W  L  N  D  K  S  V  P  S  P  L  N  W  E  R  K  T  F  S
              320

        1090      1100      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
ATTGTAATTTTAATATGAGCAGCCTGATGTCTTTTATTCAGGCAGACTCATTTACTTGTAATAATATTGATGCTGCTAAGATATATGGTATGTGTTTTTCCAGCATAACTATAGATAAGT
N  C  N  F [N] M  S  S  L  M  S  F  I  Q  A  D  S  F  T  C  N  N  I  D  A  A  K  I  Y  G  M  C  F  S  S  I  T  I  D  K
              360

        1210      1220      1230      1240      1250      1260      1270      1280      1290      1300      1310      1320
TTGCTATACCCAATGGTAGGAAGGTTGACCTACAATTGGGCAATTTGGGCTATTTGCAGTCTTTTAACTATAGAATTGATACTACTGCTACAAGTTGTCAGTTGTATTATAATTTACCTG
F  A  I  P  N  G  R  K  V  D  L  Q  L  G  N  L  G  Y  L  Q  S  F  N  Y  R  I  D  T  T  A  T  S  C  Q  L  Y  Y  N  L  P
              400

        1330      1340      1350      1360      1370      1380      1390      1400      1410      1420      1430      1440
CTGCTAATGTTTCTGTTAGCAGGTTTAATCCTTCTACTTGGAATAGGAGATTTGGTTTTACAGAACAATTTGTTTTTAAGCCTCAACCTGTAGGTGTTTTTACTCATCATGATGTTGTTT
A  A [N] V  S  V  S  R  F  N  P  S  T  W  N  R  R  F  G  F  T  E  Q  F  V  F  K  P  Q  P  V  G  V  F  T  H  H  D  V  V
              440

        1450      1460      1470      1480      1490      1500      1510      1520      1530      1540      1550      1560
ATGCACAACATTGTTTTAAAGCTCCCTCAAATTTCTGTCCGTGTAAATTGGATGGGTCTTTGTGTGTAGGTAATGGTCCTGGTATAGATGCTGGTTATAAAAATAGTGGTATAGGCACTT
Y  A  Q  H  C  F  K  A  P  S  N  F  C  P  C  K  L  D  G  S  L  C  V  G  N  G  P  G  I  D  A  G  Y  K  N  S  G  I  G  T
              480

        1570      1580      1590      1600      1610      1620      1630      1640      1650      1660      1670      1680
GTCCTGCAGGTACTAATTATTTAACTTGCCATAATGCTGCCCAATGTAATTGTTTGTGCACTCCCGACCCCATTACATCTAAATCTACAGGGCCTTACAAGTGCCCCCAAACTAAATACT
C  P  A  G  T  N  Y  L  T  C  H  N  A  A  Q  C  N  C  L  C  T  P  D  P  I  T  S  K  S  T  G  P  Y  K  C  P  Q  T  K  Y
              520

        1690      1700      1710      1720      1730      1740      1750      1760      1770      1780      1790      1800
TAGTTGGCATAGGTGAGCACTGTTCGGGTCTTGCTATTAAAAGTGATTATTGTGGAGGTAATCCTTGTACTTGCCAACCACAAGCATTTTTGGGCTGGTCTGTTGACTCTTGTTTACAAG
L  V  G  I  G  E  H  C  S  G  L  A  I  K  S  D  Y  C  G  G  N  P  C  T  C  Q  P  Q  A  F  L  G  W  S  V  D  S  C  L  Q
              560

        1810      1820      1830      1840      1850      1860      1870      1880      1890      1900      1910      1920
GGGATAGGTGTAATATTTTTGCTAATTTTATTTTGCATGATGTTAATAGTGGTACTACTTGTTCTACTGATTTACAAAAATCAAACACAGACATAATTCTTGGTGTTTGTGTTAATTATG
G  D  R  C  N  I  F  A  N  F  I  L  H  D  V  N  S  G  T  T  C  S  T  D  L  Q  K  S  N  T  D  I  I  L  G  V  C  V  N  Y
              600

        1930      1940      1950      1960      1970      1980      1990      2000      2010      2020      2030      2040
ATCTTTATGGTATTACAGGCCAAGGTATTTTTGTTGAGGTTAATGCGACTTATTATAATAGTTGGCAGAACCTTTTATATGATTCTAATGGTAATCTCTATGGTTTTAGAGACTACTTAA
D  L  Y  G  I  T  G  Q  G  I  F  V  E  V [N] A  T  Y  Y  N  S  W  Q  N  L  L  Y  D  S  N  G  N  L  Y  G  F  R  D  Y  L
              640

        2050      2060      2070      2080      2090      2100      2110      2120      2130      2140      2150      2160
CAAACAGAACTTTTATGATTCGTAGTTGCTATAGCGGTCGTGTTTCAGCGGCCTTTCATGCTAACTCTTCCGAACCAGCATTGCTATTTCGGAATATTAAATGCAATTACGTTTTTAATA
T [N] R  T  F  M  I  R  S  C  Y  S  G  R  V  S  A  A  F  H  A [N] S  S  E  P  A  L  L  F  R  N  I  K  C  N  Y  V  F [N]
              680

        2170      2180      2190      2200      2210      2220      2230      2240      2250      2260      2270      2280
ATACTCTTTCACGACAGCTGCAACCTATTAACTATTTTGATAGTTATCTTGGTTGTGTTGTCAATGCTGATAATAGTACTTCTAGTGTTGTTCAAACATGTGATCTCACAGTAGGTAGTG
N  T  L  S  R  Q  L  Q  P  I  N  Y  F  D  S  Y  L  G  C  V  V  N  A  D [N] S  T  S  S  V  V  Q  T  C  D  L  T  V  G  S
              720

        2290      2300      2310      2320      2330      2340      2350      2360      2370      2380      2390      2400
GTTACTGTGTGGATTACTCTACAAAAAGACGAAGTCGTAGAGCGATTACCACTGGTTATCGGTTTACTACTTTTGAGCCATTTACTGTTAATTCAGTAAATGATAGTTTAGAACCTGTAG
G  Y  C  V  D  Y  S  T  K  R  R  S  R  R  A  I  T  T  G  Y  R  F  T  T  F  E  P  F  T  V  N  S  V [N] D  S  L  E  P  V
              760

        2410      2420      2430      2440      2450      2460      2470      2480      2490      2500      2510      2520
GTGGTTTGTATGAAATTCAAATACCTTCAGAGTTTACTATAGGTAATATGGAGGAGTTTATTCAAACAAGCTCTCCTAAAGTTACTATTGATTGTTCTGCTTTTGTCTGTGGTGATTATG
G  G  L  Y  E  I  Q  I  P  S  E  F  T  I  G  N  M  E  E  F  I  Q  T  S  S  P  K  V  T  I  D  C  S  A  F  V  C  G  D  Y
              800
```

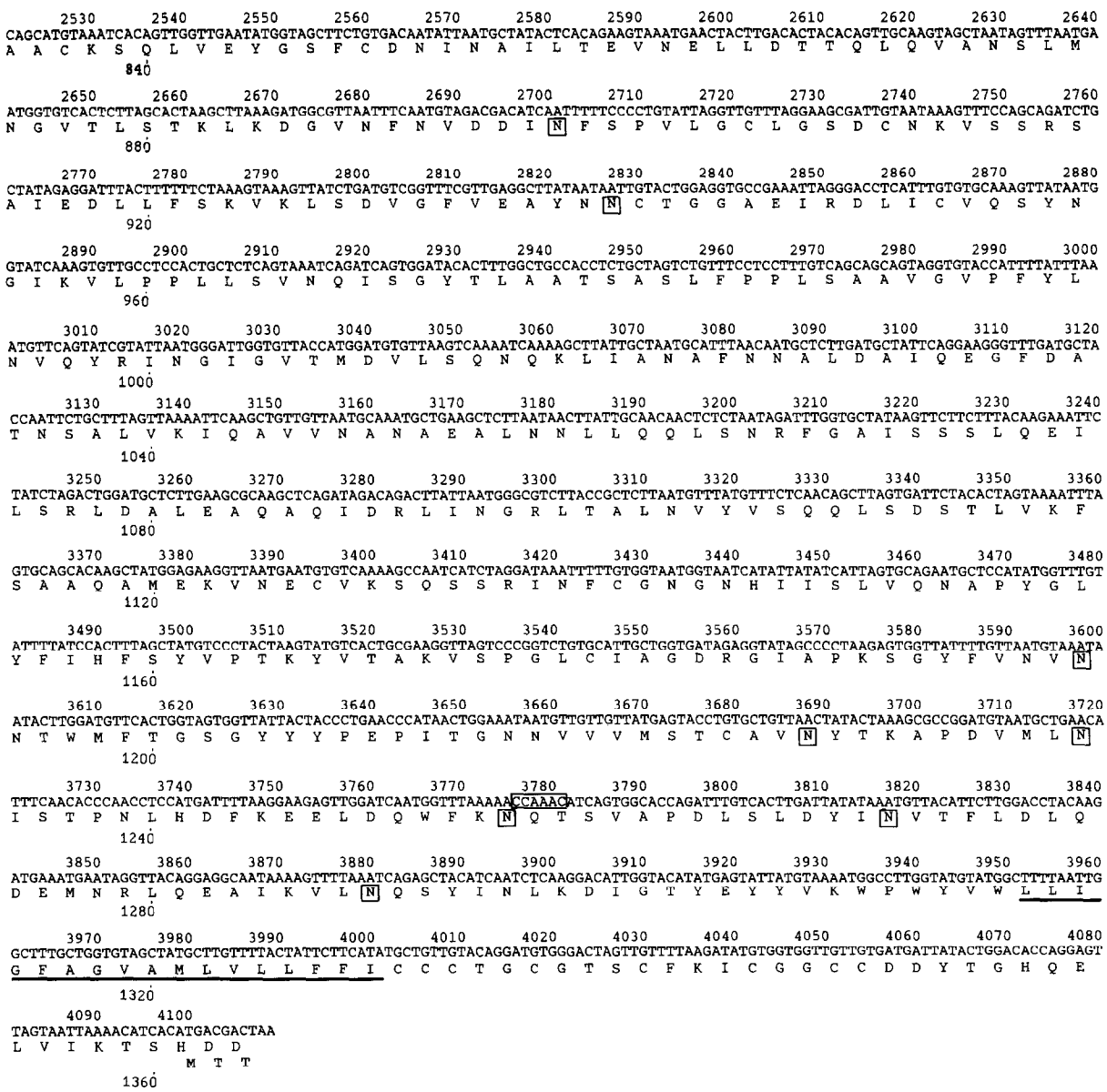

FIG. 2. Nucleotide sequence of the S gene and its deduced amino acid sequence. The nucleotide sequence shown begins with the TAG termination codon of the HE gene (underlined) 17 bases upstream of the presumed S start site (7407 bases from the poly(A) tail), and ends with the TAA termination codon of the S protein. The first three amino acids of the putative 4.9-kDa protein are shown beginning at base position 4099. Consensus CYAAAC sequences are boxed. The presumed amino-terminal signal peptide and carboxy-terminal anchor sequences are underlined. Potential N-linked glycosylation sites (NXS or NXT, where X ≠ P) are boxed. The proteolytic cleavage site separating S1 and S2 is identified with an arrow. The extended sequence of amino acids missing in MHV JHM is identified by individually underlined amino acids, and that missing in MHV A59, by asterisks.

with amino acid 763, and, on the basis of the pattern in MHV and IBV, predicts a cleavage between amino acids 768 and 769 (note arrow in Fig. 2). Cleavage at this point would divide the unglycosylated S protein into an N-terminal segment of 85,690 Da (S1) and a C-terminal segment of 65,153 Da (S2).

From amino acid sequencing studies, no N-terminal sequence could be obtained from the virion-derived 120-kDa subunit, possibly because of N-terminal blockage. The N-terminal sequence of the 100-kDa subunit could be obtained, however, and was determined to be X-I-T-T-G-Y-X-F-, identifying the first amino acids downstream from the predicted internal cleavage site. These results confirmed the predicted internal cleavage site and established that the 120-kDa subunit is S1 and the 100-kDa subunit is S2.

The BCV and MHV S proteins show remarkable sequence homology suggesting that these viruses are re-

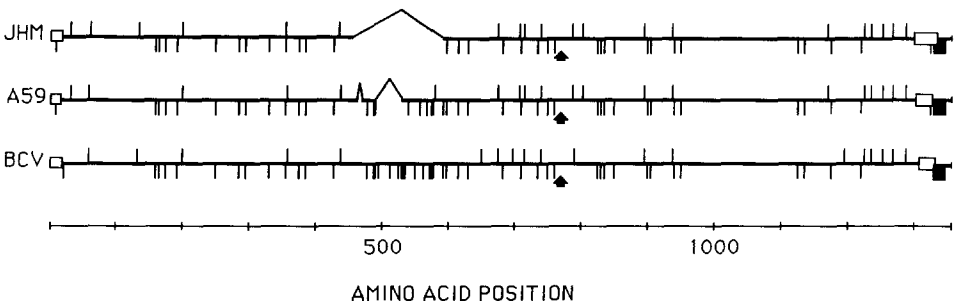

FIG. 3. Structural comparison of the S proteins of MHV-JHM, MHV-A59, and BCV. Sequences are aligned for maximum homology. A sequence found in BCV but not found in MHV-JHM or MHV-A59 is expressed as a gap (broken line) in the MHV sequences. Putative N-terminal signal peptides and C-terminal anchor sequences are boxed. Vertical lines above the sequence indicate potential asparagine-linked glycosylation positions, and below the sequence, cysteine positions. The identified (BCV, MHV-A59) and putative (MHV-JHM) proteolytic cleavage sites are identified by arrows.

cently diverged. After aligning sequences for maximal homology, the following points emerge. (i) Relative to BCV, a large deletion appears in the MHV S1 subunits. For JHM it is a contiguous gap of 138 amino acids, and for A59 it is a discontiguous gap of 50 amino acids (Figs. 2 and 3). The function of the additional sequence in the BCV S1 subunit is not known, but it is possibly a structure that interacts in some way with the HE glycoprotein, a structural protein not found on MHV (*13, 34*) except under certain rare conditions (*33*). No electron micrographic or chemical data exist, however, to suggest that S and HE do physically interact (*3, 17, 18*). It is interesting to note that the entire region in the BCV S protein corresponding to the gap region of the JHM S protein is especially rich in cysteine residues and contains 15 (26%) of the 56 total cysteines in the BCV S protein (Figs. 2 and 3). This suggests that this part of the molecule may be important for intramolecular or intermolecular disulfide linkages. (ii) Exclusive of the large gap in the MHV sequences, the S1 subunits of JHM and A59 show 62 and 60% identity, respectively, with BCV, and the S2 subunits show 75 and 74%, respectively. Throughout the S protein, 41 of 56 cysteine positions and 13 of 19 potential N-linked glycosylation sites are conserved. The internal proteolytic cleavage position (not yet confirmed for JHM) is also conserved. The pattern of greater amino acid sequence divergence in the S1 subunit is consistent with the model of Cavanagh (*4*) and De Groot *et al.* (*7*) which proposes that the S1 subunit comprises the exposed bulbous structure of the spike and probably contains most (*5*), but not all (*23, 36*), of the neutralizable antigenic sites. It is the structure most likely to undergo changes as a result of immunologic selective pressures.

Fusion of cells in culture is one biological activity associated with cleavage of the MHV S protein (*35*). Despite its extensive sequence similarity with the MHV S protein, however, the BCV S protein shows little fusion activity. In fact, fusion is a behavior we have not observed with the Mebus strain of BCV even though the S protein is primarily in the cleaved form on the virion (*13, 17*). It is not clear why BCV and MHV behave so differently in their fusogenic properties, but functional evaluation of sequence differences near the cleavage sites of these two viruses may aid in clarifying the mechanisms of fusion by MHV. This is especially interesting since hydrophobic regions, common at the cleavage sites on fusion proteins of paramyxoviruses and myxoviruses, are absent in the MHV S protein (*22*) and different mechanisms of fusion may be employed.

ACKNOWLEDGMENTS

This work was supported by Grant AI-14367 from the National Institute of Allergy and Infectious Diseases, and by Grant 82-CRSR-2-1090 from the United States Department of Agriculture. T.E.K. and W.L. were predoctoral trainees on Grant T32-AI07123 from the National Institutes of Health.

REFERENCES

1. BINNS, M. M., BOURSNELL, M. E. G., CAVANAGH, D., PAPPIN, D. J. C., and BROWN, T. D. K., *J. Gen. Virol.* **66,** 719–726 (1985).
2. BINNS, M. M., BOURSNELL, M. E. G., TOMLEY, F. M., and BROWN, T. D. K., *J. Gen. Virol.* **67,** 2825–2831 (1986).
3. BRIDGER, J. C., CAUL, E. O., and EGGLESTONE, *Arch. Virol.* **57,** 43–51 (1978).
4. CAVANAGH, D., *J. Gen. Virol.* **64,** 2577–2583 (1983).
5. CAVANAGH, D., DAVIS, P. J., and MOCKETT, A. P. A., *Virus Res.* **11,** 141–150 (1988).
6. CAVANAGH, D., DAVIS, P. J., PAPPIN, D. J. C., BINNS, M. M., BOURSNELL, M. E. G., and BROWN, T. D. K., *Virus Res.* **4,** 133–143 (1986).
7. DEGROOT, R. J., LUYTJES, W., HORZINEK, M. C., VAN DER ZEIJST, B. A. M., SPAAN, W. J. M., and LENSTRA, J. A., *J. Mol. Biol.* **196,** 963–966 (1987).
8. DEREGT, D., and BABIUK, L. A., *Virology* **161,** 410–420 (1987).
9. DEREGT, D., GIFFORD, G. A., IJAZ, M. K., WATTS, T. C., GILCHRIST, J. E., HAINES, D. M., and BABIUK, L. A., *J. Gen. Virol.* **70,** 993–998 (1989).

10. DEREGT, D., SABARA, M., and BABIUK, L. A., *J. Gen. Virol.* **68,** 2863–2877 (1987).
11. GULBER, U., and HOFFMAN, B. J., *Gene* **25,** 263–269 (1983).
12. HOGUE, B. G., KIENZLE, T. E., and BRIAN, D. A., *J. Gen. Virol.* **70,** 345–352 (1989).
13. HOGUE, B. G., KING, B., and BRIAN, D. A., *J. Virol.* **51,** 384–388 (1984).
14. HOUSE, J. A., *J. Amer. Vet. Med. Assoc.* **173,** 573–576 (1978).
15. JACOBS, L., DE GROOT, R., VAN DER ZEIJST, B. A. M., HORZINEK, M. C., and SPAAN, W., *Virus Res.* **8,** 363–371 (1987).
16. KIENZLE, T. E., ABRAHAM, S., HOGUE, B. G., and BRIAN, D. A., *J. Virol.,* **64,** in press (1990).
17. KING, B., and BRIAN, D. A. *J. Virol.* **42,** 700–707 (1982).
18. KING, B., POTTS, B. J., and BRIAN, D. A., *Virus Res.* **2,** 53–59 (1985).
19. KRAFT, R., TARDIFF, J., KRAUTER, K., and LEINWARD, L., *BioTechniques* **6,** 544–549 (1988).
20. KUSTERS, J. G., NIESTERS, H. G. M., LENSTRA, J. A., HORZINEK, M. C., and VAN DER ZEIJST, B. A. M., *Virology* **169,** 217–221 (1989).
21. LAPPS, W., HOGUE, B. G., and BRIAN, D. A., *Virology* **157,** 47–57 (1987).
22. LUYTJES, W., STURMAN, L. S., BREDENBEEK, P. J., CHARITE, J., VAN DER ZEIJST, B. A. M., HORZINEK, M. C., and SPAAN, W. J. M., *Virology* **161,** 479–487 (1987).
23. MAKINO, S., FLEMING, J. O., KECK, J. G., STOHLMAN, S. A., and LAI, M. M. C., *Proc. Natl. Acad. Sci. USA* **84,** 6567–6571 (1987).
24. MATSUDAIRA, P., *J. Biol. Chem.* **262,** 10,035–10,038 (1987).
25. MAXAM, A. M., and GILBERT, W., *In* "Methods in Enzymology" (L. Grossman and K. Moldave, Eds.), Vol. 65, pp. 499–560. Academic Press, Orlando, FL (1980).
26. MEBUS, C. A., STAIR, E. L., RHODES, M. B., and TWIEHAUS, M. J., *Amer. J. Vet. Res.* **34,** 145–150 (1973).
27. NIESTERS, H. G. M., LENSTRA, J. A., SPAAN, W. J. M., ZIJDERVELD, A. J., BLEUMINK-PLUYM, N. M. C., HONG, F., VAN SCHARRENBURG, G. J. M., HORZINEK, M. C., and VAN DER ZEIJST, B. A. M., *Virus Res.* **5,** 253–263 (1986).
28. PARKER, M. D., COX, G. J., DEREGT, D., FITZPATRICK, D. R., and BABIUK, L. A., *J. Gen. Virol.* **70,** 155–164 (1989).
29. RASSCHAERT, D., and LAUDE, H., *J. Gen. Virol.* **68,** 1883–1890 (1987).
30. SAIF, L. J., REDMAN, D. R., BROCK, K. V., KOHLER, E. M., and HECKERT, R. A., *Vet. Rec.* **123,** 300–301 (1988).
31. SANGER, F., NICKLEN, S., and COULSON, A. R., *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467 (1977).
32. SCHMIDT, I., SKINNER, M., and SIDDELL, S., *J. Gen. Virol.* **68,** 47–56 (1987).
33. SHEIH, C-K., LEE, H-J., YOKOMORI, K., MONICA, N. L., MAKINO, S., and LAI, M. M. C., *J. Virol.* **63,** 3729–3736 (1989).
34. SPAAN, W., CAVANAGH, D., and HORZINEK, M. C., *J. Gen. Virol.* **69,** 2939–2952 (1988).
35. STURMAN, L. S., RICARD, C. S., and HOLMES, K. V., *J. Virol.* **56,** 904–911 (1985).
36. TALBOT, P. J., DIONNE, G., and LACROIX, M., *J. Virol.* **62,** 3032–3036 (1988).
37. VLASAK, R., LUYTJES, W., LEIDER, J., SPAAN, W., and PALESE, P., *J. Virol.* **62,** 4686–4690 (1988).
38. VON HEIJNE, G., *J. Mol. Biol.* **184,** 99–105 (1985).