Cloning, sequencing and expression of the S protein gene from two geographically distinct strains of canine coronavirus ☆

Brian C. Horsburgh [1], T. David K. Brown *

Division of Virology, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK

Abstract

The gene encoding the spike (S) protein from two geographically distinct strains (American and British) of canine coronavirus (CCV) was cloned and sequenced. The nucleotide sequence revealed open reading frames of 1443 or 1453 amino acids, respectively. Structural features include an N-terminal hydrophobic signal sequence, a hydrophilic cysteine-rich cluster near the C-terminus, two heptad repeats and 29 or 33 potential N-glycosylation sites. Pairwise comparisons of S amino acid sequences from these isolates with other CCV strains (Insavc1 and K378) revealed that heterogeneity, found mostly in the form of conservative substitutions, is distributed throughout the canine sequences. However, 5 variable regions could be identified. Similar analysis with feline, porcine, murine, chicken and human coronavirus sequences revealed that the canine sequences are much more closely related to the feline S protein sequence than to the porcine S protein sequences even though they are all from the same antigenic group. Moreover, the sequence similarity between CCV isolates and the feline coronavirus, feline infectious peritonitis virus (FIPV) was comparable. Expression of the CCV or the transmissible gastroenteritis virus (TGEV) S gene using the vaccinia virus system produced a protein of the expected size which could induce extensive syncytia formation in infected canine A72 cells.

Keywords: Coronavirus; S protein; Syncytium

Canine coronavirus (CCV) belongs to one of the major antigenic groups of coronaviruses (Siddell et al., 1983; Spaan et al., 1988) and is serologically and genetically related to feline infectious peritonitis virus (FIPV), feline enteric coronavirus (FeCV), transmissible gastroenteritis virus (TGEV) and porcine respiratory coronavirus (PRCV) (Sanchez et al., 1990, Horsburgh et al., 1992; Wesseling et al., 1994). This close genetic relationship indicates that these viruses may have a common ancestor (Horzinek et al., 1982).

The CCV virion is known to contain at least 4 protein species: the 50K nucleocapsid protein, N, the 32K integral membrane protein, M, the 9K small membrane protein, SM, and the spike glycoprotein, S (Garwes and Reynolds, 1981; Godet et al., 1992; Horsburgh et al., 1992). The 160K S protein, which forms the projecting spike on the surface of the virion (Tyrrell et al., 1968; Spaan et al., 1988), is synthesized on ribosomes bound to the rough endoplasmic reticulum (RER) where it is co-translationally glycosylated (Holmes et al., 1981; Niemann et al., 1982; Sturman and Holmes, 1985). After glycosylation, the majority of the S proteins are oligomerized into trimeric structures (Cavanagh 1983a,b; Delmas and Laude, 1990) and transported to the Golgi where a large number become incorporated into virions. The S glycoprotein mediates binding of virions to the host cell receptor (Williams et al., 1991), is the major target for neutralizing antibodies (Spaan et al., 1988), is recognized by T-cell lymphocytes (Korner et al., 1991) and induces polykaryocyte formation by cell fusion in cultured cells (Collins et al., 1982; Sturman and Holmes, 1985).

Clearly, to understand the biological and pathogenic properties of CCV at the molecular level, the primary structure of CCV S genes and data on their processing are essential. Indeed, the S protein gene has been cloned for two strains of CCV (Horsburgh et al., 1992; Wesseling et al., 1994). However, coronavirus genomes are dynamic, subject to recombination, insertion and deletion and it appears that deletions do occur at a higher frequency within S; the polymorphism of S found in MHV strains accounting for differences in their biological activities is well documented (Spaan et al., 1988; Parker et al., 1989; Gallagher et al., 1990). We expected that cloning and sequencing of two geographically distinct isolates of CCV, CCV-6 (American) and CCV-C54 (British), would help to build a consensus picture of this virus within a region where deletions do occur at a higher frequency with respect to background mutation.

Oligo(dT) selected CCV genomic RNA from either CCV-6 or CCV-C54 was used to generate cDNA libraries (Horsburgh et al., 1992) using oligonucleotides 10 and 1690 as primers (Fig. 1). It was hoped that using two cDNA libraries would increase the likelihood of finding clones that covered the S gene. Inserts from recombinant clones of 1 kb or greater were retained for further analyses. Sequencing the ends of these clones permitted initial alignment of the CCV-6 and CCV-54 clones with respect to the CCV-Insavc1 genome (Horsburgh et al., 1992). This approach proved fruitful in that 5 clones were identified which spanned the S coding gene of both CCV-6 and CCV-C54 (see Fig. 1). The relationships between putative overlapping clones were confirmed by Southern hybridization. The nucleotide sequence of the overlapping cDNA clones was determined using the
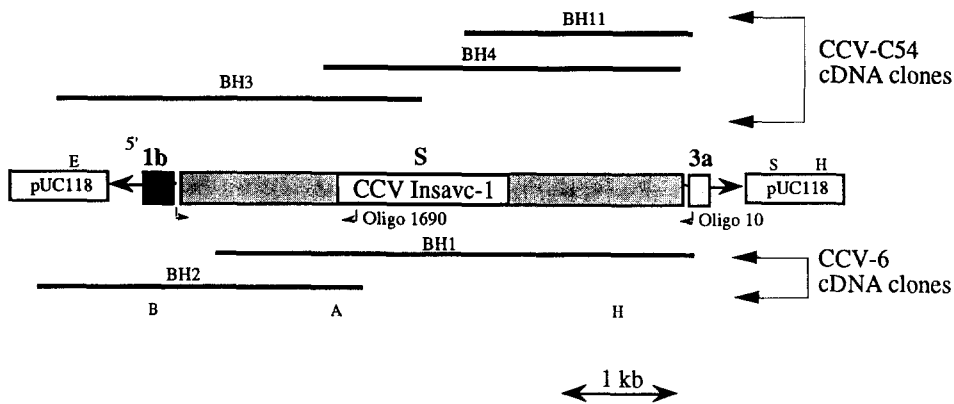
Fig. 1. Alignment of CCV-6 and CCV-C54 cDNA clones with respect to the CCV-Insavc1 genome (Horsburgh et al., 1992) using partial sequence information. Overlaps were confirmed by Southern blotting. Libraries of cDNA clones were generated from poly(A)-containing infected cell RNA using primers 10 and 1690 (GACCTGTAATGACTCAT and TAGGTAGTAACACAACA, respectively) and cloned into pUC118 as described by Horsburgh et al. (1992). Restriction enzymes sites used for cloning are indicated. B, BamHI; A, AflIII; H, HindIII (see text). The BamHI site was introduced by site-directed mutagenesis using primer 17 (TAATCACTTGGATCCTTAATGTGCC) as described by Brierley et al. (1987).

Sanger dideoxy chain termination method and analysed using the SAP programs of Staden (1986). The deduced S gene amino acid sequence from CCV-6 and CCV-C54 and the results of comparisons with other S gene protein sequences is presented in Fig. 2 and summarized in Table 1.

Thirty-two nucleotides (nts) upstream of the ATG start codon of either CCV-6 or CCV-C54 is the sequence CTAAAC. This motif, which is believed to be the signal for transcription of subgenomic messenger RNA species (Spaan et al., 1988) is found upstream of nearly every coronavirus S gene sequenced so far (DeGroot et al., 1987; Luytjes et al., 1987; Schmidt et al., 1987; Britton and Page, 1990; Parker et al., 1990; Rasschaert et al., 1990; Britton et al., 1991; Horsburgh et al., 1992; Mounir and Talbot, 1993; Wesseling et al., 1994). The exceptions are IBV-Beaudette, HCV-229E and PEDV which are CTGAAC, CTCAAC and GTAAAC, respectively (Binns et al., 1985; Raabe et al., 1990; Duartre and Laude, 1994). It would appear that the consensus sequence XTXAAC, is conserved throughout this family of viruses; however, it is likely that this motif alone is only a component of a larger transcription initiation signal as the surrounding sequences are quite different in these viruses. This may in part explain the different ratios of subgenomic mRNAs found in infected cells.

Analysis of the CCV S gene sequences revealed features common to all coronavirus S proteins; namely, signal sequence, transmembrane anchor, hydrophobic profile, the presence of 2 heptad repeats, the C-terminal cysteine rich cluster, the KWPWY(W)VWL motif and a large number of potential N-glycosylation sites (Fig. 2 and Table 1). The homology that these canine proteins share is

```
                                                          ****                    90
>CCV   6   MIVLILCLLL F-SYNSVICT SNNDCVQ-VN VTQLPGNENI IKDFLFH--- TFKEEGSLVV GGYYPT-EVW YNCSRSATTT AYKDFSNIHA
>CCV C54   T                                         Q          N                   T     T          HY
>CCV  In   T    F    L S S                           Q          N                        TTQQ       Y
>CCV K378  T    F              G                                  P V
>FIPV sp   VT       LC HT LS  T E I         A    L R    S        N                        T R       FQY N
>TGEV sp   KK FVV VV M----PLIYG D FP SKLT   RTIGNQW- L  ET  LNYSS RLPPNSDD L  D F  VQP  F  I NNSND L VTLE LK
>PRCV sp   KK FVV VV M----PLIYG DKFP------  ---------- ---------- ---------- ---------- ---------- ----------

                                  ****************                               180
>CCV   6   FYFDMEDMEK STGNARGKPL LVHVHGGPVS IIIICAR--- ----KASLKH GLLCITKNK- -IIDYNTFTS AQWSAICLGD DRKIPFSVIP
>CCV C54        A AN                     S   VYISAYRD DVQNRPL      S- -T   S    RD   T            V
>CCV  In        A  N                     N   VYISAYRD DVQFRPL      D- -TV    I N  RD             V
>CCV K378       A  N                     N    YISAYRD DVQPRPL        - -
>FIPV sp   V  A  N           F    E        --SAYRD DVQQRPL     V     R- -H N EQ    N  NST T A
>TGEV sp   L W YAT-EN I W H QR- NVVVN Y Y   TVTTTRNFS AEG------- AII  C GSP PTTTTESSLT CNWGSE RLN H-K-- PIC
>PRCV sp   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                                                                                270
>CCV   6   TDN----GTK IFGLEWNDDY VTAYISDRSH HLNINNNWFN NVTILYSRSS TATWQKSAAY V----YQGVS NFTYYKLNNT NGLKSYELCE
>CCV C54   ----      L    T         D    R    T                        ----
>CCV  In   ----      L                   E    R            L    T       H       ----            K A
>CCV K378  ----                                                S                 ----
>FIPV sp   ----      Y    QWFA E  V  LHGA Y RISPE Q SG T  L    --FGDMR  AT LEVAGTL  DLWWFNP Y DVS  RV  K          T
>TGEV sp   SNSEANC NM Y    QWFA E  V  LHGA Y RISPE Q SG T  --FGDMR AT LEVAGTL  DLWWFNP Y DVS  RV  K          T
>PRCV sp   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- --TSVVSN T

                                                                                360
>CCV   6   DYEYCTGYAT NVFAPTVGGY IPDGFSPNNW FMLTNSSTFV SGRFVTNQPL LVNCLVPVPS FGVAAQEFCF EGAQFSQCNG VSLNNTVDVI
>CCV C54               S                     L                    I  W
>CCV  In                S                                          W
>CCV K378                        H          R                      W        Q
>FIPV sp   H           S                     L                  I  W
>TGEV sp   --Q AS VA   TTQP  F   SD          L      L   KL  K              EE  ST       G D      AS
>PRCV sp   --Q AS VA   TTQP  F   SD          I      L   KL  K              EE  ST       D

                                                                                450
>CCV   6   RFNLNFTTDV QSGMGAIVFS LNTTGGVILE ISCYNDTVSE SSFYSYGEIS IGVTDGPRYC YALYNGQALK CLGTLPPSVK EIAISKWGHF
>CCV C54        N        T                                 P F      V    T
>CCV  In        N        T                                 P F      V    T
>CCV K378      NL        T                                 P F      F    T
>FIPV sp        A                          S               P F I    YV   T
>TGEV sp        N    K   T          I        --  D   S     P F V    V         Y
>PRCV sp        N    K   T          I            D   S     P F V    V    T    Y

                                                                                540
>CCV   6   YINGYNFFST FPIDCISFNL TTGDSGAFWT IAYTSYTDAL VQVENTAIKK YTYCNSHINN IKCSQLTANL QNGFYPVASS EVGLVNKSVV
>CCV C54                                      E                V
>CCV  In                          A        A    E                V
>CCV K378                         AS        D    E                V
>FIPV sp                          V         E          N V             V                         F
>TGEV sp                          DV        E          T  V    V        I    N    S
>PRCV sp                          DV        E          TN V    YV            NN   S    S

                            *****                                               630
>CCV   6   LLPSFYSHTS VNITIDLGMK RS-VMVTIAS TLSNITLPMQ DNNTDVYCIR SNQFSVYVHS TCKSSLWDDV FNSDCTDVLY ATAVIKTGTC
>CCV C54                        GYGQP                                           I          H
>CCV  In                        -V         P          I                      NN    A    D
>CCV K378                       GYGQP                           F
>FIPV sp   FTY A         L GYGQP                                            NI    Q    E
>TGEV sp   T   I     G   GYGQP                            D              A  NI  KRN    D
>PRCV sp   LT  I     G   GYGQP                      V     D              A  NV  KRN    D

                                                                                720
>CCV   6   PFSFDKLNNY LTFNKFCLSL NPVGANCKFD VAARTRTNEQ VVRSLYVIYE EGDNIAGVPS DNSGLHDLSV LHLDSCTDYN IYGRTGVGII
>CCV C54
>CCV  In                                                 M            DRP
>CCV K378                                        FG                V
>FIPV sp                          S                               V
>TGEV sp                          S                               V         V
>PRCV sp                          S                 D  F          SIV

                                                                                810
>CCV   6   RQTNSTLLSG LYYTSLSGDL LGFKNVSDGV IYSVTPCDVS VQAAVIDGAI VGAMTSINSE L-GLTHWTTT PNFYYYSIYN YTNERTRGTA
>CCV C54        IF                          V          A                    L                     SV
>CCV  In                                    V          A                    L                     VMN
>CCV K378                                               AH                   L
>FIPV sp   R                                           A                    L                     S
>TGEV sp       R  I                                    A    T    I          L                     D
>PRCV sp       R              V                        A    I         I     LA    I               DK      P
```

Fig. 2.

```
                                                                                     900
>CCV   6    IDSNDVDCEP IITYSNIGVC KNGALVFINV THSDGDVQPI CTGNVTIPTN FTISVQVEYI QVYTTPVSID CSRYVCNGNP RCNKLLTQYV
>CCV  C54                                                                             A
>CCV  In                                                                              A
>CCV  K378
>FIPV sp              V                                                    M          A
>TGEV sp              V
>PRCV sp    G         V
```

```
                                                        ******                        990
>CCV   6    SACQTIEQAL AMGARLENME IDSMLFVSEN ALKLASVEAL -IVGNLDPIY KEWPNIGGSW LGGLKDILPS HNSKRKYRSA IEDLLFDKVV
>CCV  C54                                              F NSTET
>CCV  In                        I                      F NSTE
>CCV  K378                      I                      F NSTE
>FIPV sp                        V                      FNSTE      S         I                    G
>TGEV sp                        V                      FNSSET                    E    Y
>PRCV sp                        V                      FNSSET               F    E    Y    D                    S
```

```
                                                                                    1080
>CCV   6    TSGLGTVDED YKRCTGGYDI ADLVCAQYYN GIMVLPGVAN DDKMAMYTAS LAGGITLGAL GGGAVSIPFA IAVQARLNYV ALQTDVLNKN
>CCV  C54                                                                             A         V
>CCV  In               SA           R  N                          T         S    A    V
>CCV  K378
>FIPV sp                                              T                              V
>TGEV sp                                              T                              V
>PRCV sp                                              T                              V
```

```
                                                                                    1170
>CCV   6    QQILANAFNQ AIGNITQAFG KVNDAIHQTS QGLATVAKAL AKVQDVVNTQ GQALSHLTVQ LQNNFQAISS SISDIYNRLD ELSADAQVDR
>CCV  C54
>CCV  In                        K
>CCV  K378                      Q          V
>FIPV sp
>TGEV sp    S           S       R
>PRCV sp    S           S       R                     R
```

```
                                                                                    1260
>CCV   6    LITGRLTALN AFVSQTLTRQ AEVRASRQLA KDKVNECVRS QSQRFGFCGN GTHLFSLANA APNGMIFFHT VLLPTAYETV TAWSGICASD
>CCV  C54
>CCV  In
>CCV  K378
>FIPV sp
>TGEV sp                                                                                             P
>PRCV sp                        K          H                                                         L
```

```
                                                                                    1350
>CCV   6    GDRTFGLVVK DVQLTLFRNL DDKFYLTPRT MYQPRVATSS DFVQIEGCDV LFVNATVIDL PSIIPDYIDI NQTVQDILEN FRPNWTVPEL
>CCV  C54
>CCV  In    S       E        E                                    G    E
>CCV  K378
>FIPV sp                                                                                   Y
>TGEV sp                                                              S
>PRCV sp                        N                                T    S
```

```
                                                        *****                        1440
>CCV   6    TLDIFNATYL NLTGEIKCLE FRSEKLHNTT VELAILIDNI NNTLSILM-L NRIETYVKWP WYVWLLIGLV VIFCIPILLF CCCSTGCCGC
>CCV  C54                ND                          V    VN EW
>CCV  In    P    H       ND                               VN EW
>CCV  K378  P            ND                               VN EW
>FIPV sp                 DD                               VN EW                        L    F
>TGEV sp                 DD                               VN EW                        L
>PRCV sp    M V          DD                               VN EW                        L
```

```
                                      1469
>CCV   6    IGCLGSCCHS ICSRRQFESY EPIEKVHVH
>CCV  C54            M          T
>CCV  In            G
>CCV  K378           G
>FIPV sp                       N
>TGEV sp                       N
>PRCV sp             F         N
```
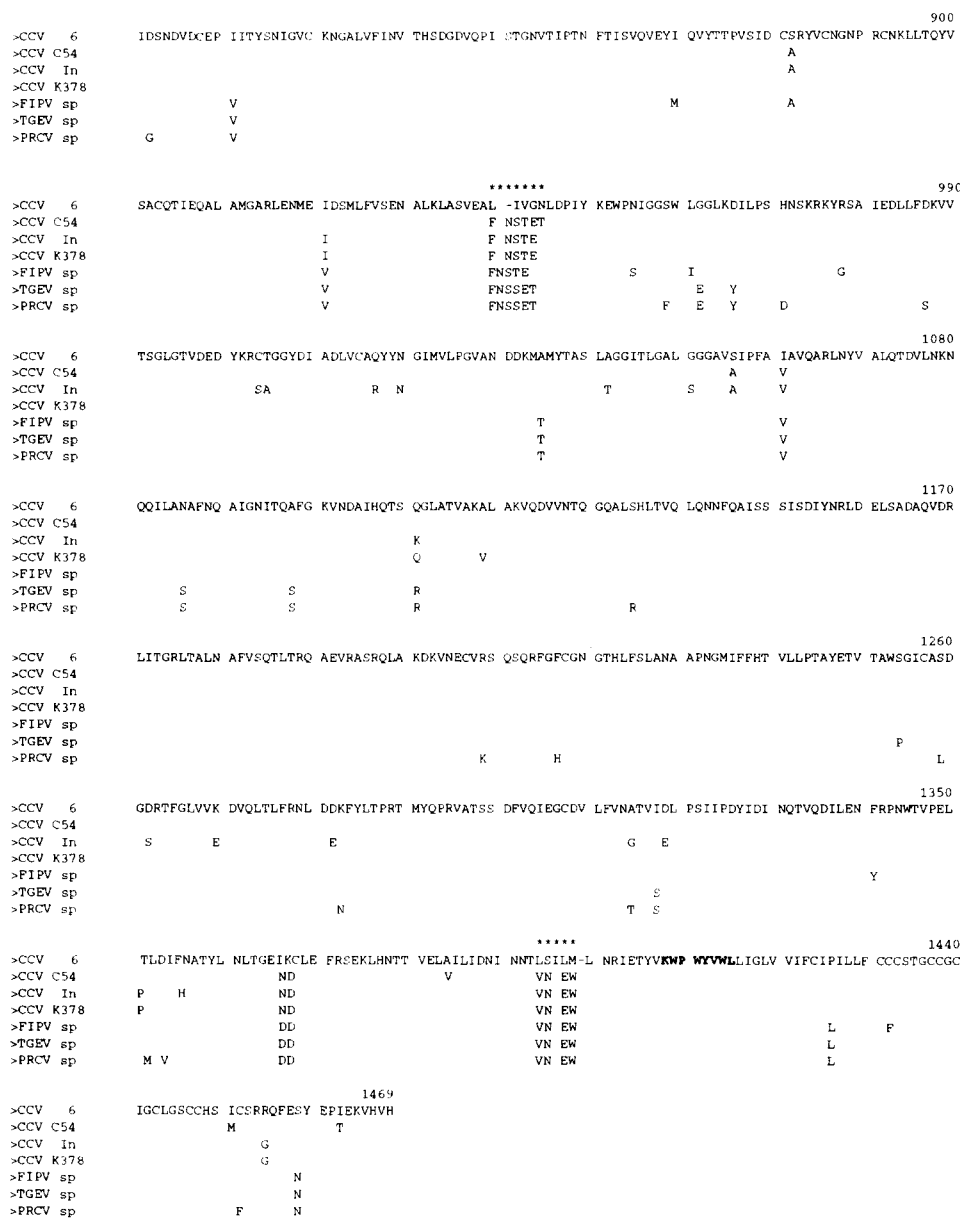
Fig. 2. Alignment of the spike protein amino acid sequences of CCV-6, CCV-C54, CCV-Insavc1 (Horsburgh et al., 1992) CCV-K378 (Wesseling et al., 1994), FIPV (DeGroot et al., 1987), TGEV (Rasschaert and Laude, 1987) and PRCV (Britton et al., 1991) using the CLUSTAL program of Higgins and Sharp (1989). Spaces indicate positions for which the amino acid is identical to that of CCV-6 and minuses represent putative deletions. The presumed signal sequence (Heijne, 1986) is underlined and the conserved motif, KWPWY(W)VWL, is shown in bold. Asterisks indicate regions where the canine sequences vary by more than 3 residues.

Table 1
Summary of results

Coronavirus [a]	Transcription signal motif	Cleavage sequence	No. of N-gly-cosylation sites [b]	Apoprotein (kDa)	Length (amino acids)	Similarity [c] (%)
CCV-Insavc1	CTAAAC	n.a. [d]	30	160	1452	100
CCV-6	CTAAAC	n.a.	29	158	1443	92.7
CCV-C54	CTAAAC	n.a.	33	160	1453	94.7
CCV-K378	CTAAAC	n.a.	31	160	1453	93.3
FIPV 79-1146	CTAAAC	n.a.	35	160	1452	91.1
TGEV FS772	CTAAAC	n.a.	32	158	1447	78.5
PRCV 86	CTAAAC	n.a.	29	135	1225	74.5
HCV-229E	CTCAAC	n.a.	30	129	1173	39.3
BCV-Que	CTAAAC	RRSRR	20	150	1363	23.9
MHV-A59	CTAAAC	RRARR	20	146	1324	23.3
MHV-JHM	CTAAAC	RRARR	21	136	1360	22.1
IBV-Bea	CTGAAC	RRFRR	28	127	1162	22.1

[a] The strains are as described in the text.
[b] The potential number of sites.
[c] Similarity relative to CCV-Insavc1.
[d] n.a., not applicable.

approximately 93% (Table 1) although the British isolates are more related to each other (Insavc1 and C54) than to the American isolates (6 and K378). The majority of these sequence changes are conservative, however there are 5 regions of variability (changes of 4 or more residues (shown in bold in Fig. 2). Interestingly, there is some heterogeneity in the length of the canine S protein genes (CCV-6 = 1443 residues, whereas CCV-C54 = 1453 residues; Table 1). This is not unusual as the plasticity of the coronavirus genome, especially within and around the S gene, is well documented (Keck et al., 1988; Kusters et al., 1989; Parker et al., 1990; Horsburgh et al., 1992). Indeed, the S protein genes from different isolates of TGEV and especially MHV may differ in length by as much as 159 amino acids (Rasschaert and Laude, 1987; Britton and Page, 1990; Parker et al., 1990).

Computer-aided analysis of the canine sequences with those of FIPV (DeGroot et al., 1987), TGEV (Britton and Page, 1990), PRCV (Britton et al., 1991), HCV-229E (Raabe et al., 1990), HCV-OC43 (Mounir and Talbot, 1993), BCV (Parker et al., 1990), PEDV (Duatre and Laude, 1994), MHV-A59 (Luytjes et al., 1987), MHV-JHM (Schmidt et al., 1987) and IBV-Beaudette (Binns et al., 1985) revealed that the canine sequences are much more closely related to FIPV than to the TGEV and PRCV sequences even though they are all from the same antigenic group (Table 1). This observation is highlighted by the finding that CCV and FIPV have near identical subgenomic mRNA patterns in infected cells and that both CCV and FIPV possess an extra ORF, 7b/6b, which is not present in TGEV or PRCV (DeGroot et al., 1988; Horsburgh et al., 1992). Furthermore, it is interesting to note that CCV-Insavc1 is almost as related to CCV-6 (92.7% identity) as it is to FIPV-791146 (91.1% identity), suggesting a very close evolutionary relationship between the canine and feline coronaviruses.

In an attempt to characterize their biological properties, the canine S glycoproteins were expressed using recombinant vaccinia viruses. Briefly, the full length coding region of the S gene of CCV-6 was reconstructed from two overlapping cDNA clones, BH1 and BH2 (Fig. 1). The 3.0-kb insert from pBH1 contains sequences from S and the polymerase sequence, 1b. In order to express S, the polymerase coding sequence had to be removed. A BamHI site was introduced immediately 5′ of the initiating methionine by site-directed mutagenesis, described by Brierley et al. (1987). Mutants were screened by restriction enzyme digestion. Positive clones were sequenced across this site. A mutant which had the introduced BamHI site was selected and designated pBH1-bam. This plasmid overlapped pBH2 by approximately 300 bp. A unique AflII site was located in this region of overlap. The proximal S coding sequence was isolated from pBH1-bam as a 1.5-kb AflII–SphI fragment and ligated into AflII–SphI digested pBH2 generating pCCV6. The full length S coding sequence was isolated as a 4.4-kb BamHI fragment then ligated into the BamHI site of the transfer vector pRK19 to form pRKCCV6. pRK19, a gift from Dr. G.L. Smith, contains vaccinia virus TK flanking sequences and the 4b late promoter. (A late promoter was utilized as the *cis* signals for termination of early transcription in vaccinia, $(T)_5NT$ are found several times within the S gene coding sequences). The C54 S gene coding sequence was assembled from the 3 overlapping clones pBH3, pBH4 and pBH11 (Fig. 1). A unique BamHI site was created 10 bp upstream of the peplomer ATG start codon by site-directed mutagenesis (in the proximal clone), pBH3, generating pBH3-bam. A 2.0-kb AflII–EcoRI fragment was isolated from this plasmid and ligated to AflII–EcoRI digested pBH4 forming pBH5′MS. This plasmid was cleaved with HindIII, phosphatased and gel eluted. The 3′ coding sequence was excised as a 1.1-kb HindIII fragment from pBH11, then ligated to the HindIII digested pBH5′MS generating pBHC54. Plasmid TGEVS, which contains the TGEV strain FS772/70 S gene coding sequence was a gift from Dr. Paul Britton, AFRC Compton. The gene was excised as a 4.5-kb BamHI fragment and subcloned into the BamHI digested transfer vector pRK19 to form pRK19TG. Correct orientation of the coronavirus S genes was confirmed by restriction enzyme digestion. Recombinant vaccinia viruses were constructed by established procedures (Mackett and Smith, 1986). Vaccinia virus infected HeLa cells were transfected with either pRKCCV6, pBHC54 or pRK19TG and recombinant viruses selected for as described (Mackett and Smith, 1986). Stocks were prepared after 3 rounds of plaque purification. The recombinant viruses were called VAc4b-C6, Vac4b-C54 and Vac4b-TG respectively. The control recombinant vaccinia virus, Vac4b-gB, expressing the CMV glycoprotein gB under the control of the 4b promoter, was a gift from Dr. Helena Browne, University of Cambridge.

In vitro characterization revealed that Vac4b-C6, Vac4b-C54 and Vac4b-TG all expressed a polypeptide which co-migrated with the CCV (Insavc1) S protein on SDS–PAGE gels (data not shown). As expected, sera from either FIPV infected cats or CCV infected dogs could immunoprecipitate the recombinant CCV and TGEV S proteins (Enjuanes et al., 1990; Sanchez et al., 1990).

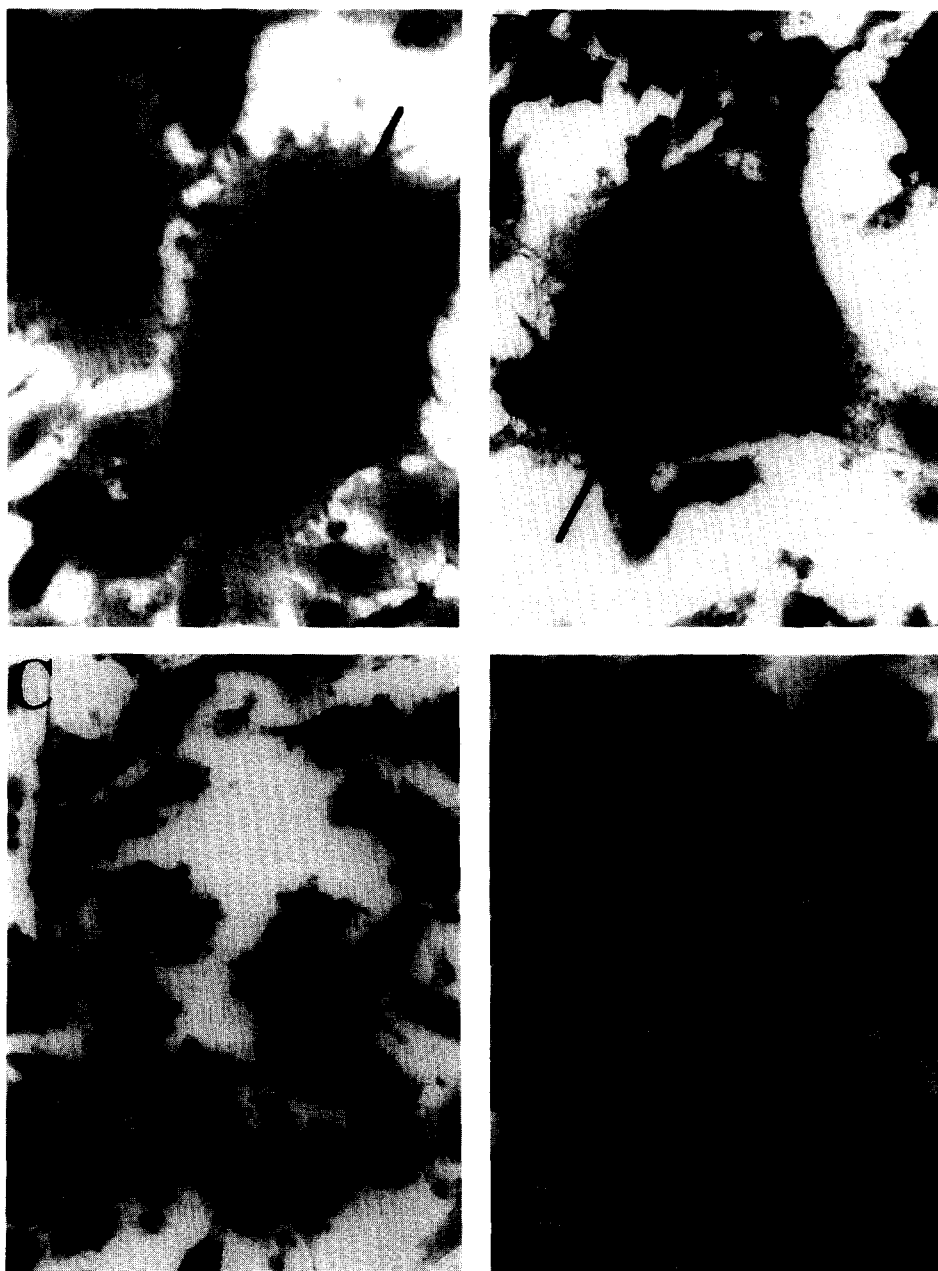As one of the prominent biological features of coronavirus S proteins in

Fig. 3. Biological activity of the recombinant S proteins. A72 cells, used for propagation of CCV, were infected (m.o.i. of 0.01) with the recombinant vaccinia viruses Vac4b-C6 (A), Vac4b-C54 (B), Vac4b-gB (C), or mock infected (D). Infected cells were stained with 0.1% crystal violet in 20% ethanol and photographed under light microscopy. Arrows indicate multinucleated cells.

cultured cells is the induction of polykaryocyte formation, canine A72 cells were infected with the recombinant vaccinia viruses or CCV, both at an m.o.i. of 0.05 or mock infected. The infection was allowed to proceed for 24 h and the cells were then fixed with 20% ethanol and stained with 0.1% crystal violet. Syncytia were evident in Vac4b-C6, Vac4b-C54 (Fig. 3), Vac4b-TG and CCV infected cells (data not shown). Multinucleated cells were not observed in Vac4b-gB and mock infected cells (Fig. 3). Syncytia were not evident upon repetition of these experiments in other cell types, e.g. BHK, RK, HeLa and TK⁻ 143 cells. These data show that the recombinant CCV-6, CCV-C54 and TGEV S proteins are expressed on the cell surface and that they retain their biological activity.

The S genes of FIPV, MHV, IBV and TGEV have been expressed using the vaccinia virus system (Tomley et al 1987; Pulford et al., 1990; Vennema et al., 1990). Induction of cell–cell fusion by these recombinant viruses was observed but was type-specific. In other words, the recombinant FIPV S protein could only induce fusion in cells of feline origin; likewise, polykaryocytosis due to the recombinant MHV S protein was restricted to murine cells (Daya et al., 1989; Vennema et al., 1990). Similarly, the recombinant CCV-6 and CCV-C54 S proteins could induce cell fusion in canine A72 cells. The observation that the recombinant porcine S glycoprotein expressed from Vac4b-TG could induce polykaryocytosis in canine A72 cells is not surprising as TGEV, as well as FIPV and CCV can replicate in this cell type (S. Chalmers; personal communication). Therefore, it is likely that S will induce syncytia only in cells that support replication of the parental coronavirus. Fusion may conceivably be mediated by the interaction of S with its respective coronavirus cell receptor as cell lines expressing the envelope glycoprotein of HIV will only induce syncytia in cells bearing the CD4 receptor (Sodroski et al., 1986). As expected, the CCV S protein, like its counterparts from FIPV, FeCV, TGEV and PRCV, is probably not cleaved as it lacks the basic amino acid motif, RRXRR, at which MHV and IBV are proteolytically cleaved. Furthermore, cleavage products have not been detected by immunoprecipitation studies from this work and those reported by Wesseling et al. (1994).

In conclusion, this work has gone some way towards the elucidation of aspects relative to phylogenetic relationships within the coronavirus genus and the reagents produced will serve as useful tools in understanding its molecular pathogenesis.

Acknowledgements

References

Binns, M.M., Boursnell, M.E.G., Cavanagh, D., Pappin, D.J.C. and Brown, T.D.K. (1985) Cloning and sequencing of the gene encoding the spike protein of coronavirus IBV. J. Gen. Virol. 66, 719–726.

Brierley, I., Boursnell, M.E., Binns, M.M., Bilimoria, B., Blok, V.C., Brown, T.D.K. and Inglis, S.C. (1987) An efficient ribosomal frameshifting signal in the polymerase encoding region of the coronavirus IBV. EMBO J. 6, 3779–3785.

Britton, P, and Page, K.W. (1990) Sequence of the S gene from a virulent British field isolate of transmissible gastroenteritis virus. Virus Res. 18, 71–80.

Britton, P., Mawditt, K.L. and Page, K.W. (1991) The cloning and sequencing of the virion protein genes from a British isolate of porcine respiratory coronavirus: comparison with transmissible gastroenteritis virus genes. Virus Res. 21, 181–198.

Cavanagh, D. (1983a) Coronavirus IBV: structural characterization of the spike protein. J. Gen. Virol. 64, 2577–2583.

Cavanagh, D. (1983b) Coronavirus IBV glycopeptides: size of their polypeptide moieties and nature of their oligosaccharides. J. Gen. Virol. 64, 1187–1191.

Collins, A.R., Knobler, R.L., Powell, H. and 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.

Daya, M., Wong, F., Cervin, M., Evans, G., Vennema, H., Spaan, W. and Anderson, R. (1989) Mutation of host cell determinants which discriminate between lytic and persistent mouse hepatitis virus infection results in a fusion-resistant phenotype. J. Virol. 70, 3335–3346.

Delmas, B. and Laude, H. (1990) Assembly of coronavirus spike protein into trimers and its role in epitope expression. J. Virol. 64, 5367–5375.

Degroot, R.J., Maduro, J., Lenstra, J.A., Horzinek, M.C., Ziejst, B.A.M., van der and Spaan, W.J.M. (1987) cDNA cloning and sequence analysis of the gene encoding the peplomer protein of feline infectious peritonitis virus. J. Gen. Virol. 68, 2639–2646.

DeGroot, R.J., Andeweg, A.C., Horzinek, M.C. and Spaan, W.J.M. (1988) Sequence analysis of the 3' end of the feline coronavirus FIPV 79-1146 genome: comparison with the genome of porcine coronavirus TGEV reveals large insertions. Virology 167, 370–376.

Duatre, M. and Laude, H. (1994) Sequence of the spike protein of the porcine epidemic diarrhoea virus. J. Gen. Virol. 75, 1195–1200.

Enjuanes, L., Gebauer, F., Correa, I., Bullido, M.J., Sune, C., Smerdou, C., Sanchez, C., Lenstra, J.A., Posthumus, W.P.A. and Meloen, R.H. (1990) Localization of antigenic sites of the S-glycoprotein of transmissible gastroenteritis virus and their conservation in coronaviruses. Adv. Exp. Med. Biol. 276, 159–172.

Gallagher, T.M., Parker, S.E and Buchmeier, M.J. (1990) Neutralization-resistant variants of a neurotropic coronavirus are generated by deletions within the amino-terminal half of the spike glycoprotein. J. Virol. 64, 731–741.

Garwes, D.J. and Reynolds, D.J. (1981) The polypeptide structures of canine coronavirus and its relationship to porcine transmissible gastroenteritis virus. J. Gen. Virol. 52, 153–157.

Godet, M., l'Haridon, R., Vautherot, J.-F. and Laude, H. (1992) TGEV coronavirus ORF4 encodes a membrane protein that is incorporated into virions. Virology 188, 666–675.

Heijne, G., von (1986) A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 14, 4683–4690.

Higgins, D.G. and Sharp, P.M. (1989) Fast and sensitive multiple sequence alignments on a microcomputer. Cabios 5, 151–153.

Holmes, K.V., Dolfer, E.W. and Behnke, J.N. (1981) Analysis of the functions of coronavirus glycoproteins by differential inhibition of synthesis with tunicamycin. Adv. Exp. Med. Biol. 142, 133–142.

Horsburgh, B.C., Brierley, I. and Brown, T.D.K. (1992) Analysis of a 9.6kb sequence from the 3'-end of canine coronavirus. J. Gen. Virol. 73, 2849–2862.

Horzinek, M.C., Lutz, H. and Pedersen, N. (1982) Antigenic relationships among coronavirus homologous structural polypeptides of porcine, feline and canine coronaviruses. Infect. Immun. 37, 1148–1155.

Keck, J.G., Matsushima, G.K., Makino, S., Fleming, J.O., Vannier, D.M., Stohlman, S.A. and Lai, M.M.C. (1988) In vivo RNA-RNA recombination of coronavirus in mouse brain. J. Virol. 62, 1810–1813.

Korner, H., Schliephake, A., Winter, J., Zimprich, F., Lassmann, H., Sedgwick, J., Siddell, S. and Wege, H. (1991) Nucleocapsid or spike protein-specific CD4 + T lymphocytes protect against coronavirus-induced encephalomyelitis in the absence of CD8 + T cells. J. Immunol. 147, 2317–2323.

Kusters, J.G., Jager, E.J. and van der Zeijst, B.A.M. (1989) Sequence evidence for in vivo RNA recombination in avian coronavirus IBV. Nucleic Acids Res. 17, 6726–6729.

Luytjes, W., Sturman, L.S., Bredenbeek, P.J., Charite, J., van der Zeijst, B.A.M., Horzinek, M.C. and 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.

Mackett, M. and Smith, G.L. (1986) Vaccinia virus expression vectors. J. Gen. Virol. 67, 2067–2082.

Mounir, S. and Talbot, P.J. (1993) Molecular characterization of the S protein gene of human coronavirux OC43. J. Gen. Virol. 74, 1981–1987.

Neimann, H., Boschek, B., Evans, D., Rosing, M., Tamura, T. and Klenk, H.D. (1982) Post-translational glycosylation of coronavirus glycoprotein E1: inhibition by monensin. EMBO J. 1, 1499–1504.

Parker, S.E., Gallagher, T.M. and Buchmeier, M.J. (1989) Sequence analysis reveals extensive polymorphism and evidence of deletions within the E2 glycoprotein gene of several strains of murine hepatitis virus. Virology 173, 664–673.

Parker, M.D., Yoo, D., Cox, G.J. and Babiuk, L.A. (1990) Primary structure of the S peplomer gene of bovine coronavirus and surface expression in insect cells. J. Gen. Virol. 71, 263–270.

Pulford, D.J., Britton, P., Page, K.W. and Garwes, D.J. (1990) Expression of TGEV structural genes in virus vectors. Adv. Exp. Med. Biol. 276, 223–231.

Raabe, T., Schelle-Prinz, B. and Siddell, S.G. (1990) Nucleotide sequence of the gene encoding the spike glycoprotein of human coronavirus HCV 229E. J. Gen. Virol. 71, 1065–1073.

Rasschaert, D. and Laude, H. (1987) The predicted primary structure of the peplomer protein E2 of the porcine coronavirus transmissible gastroenteritis virus. J. Gen. Virol. 68, 1883–1890.

Rasschaert, D., Duarte, M. and Laude, H. (1990) Porcine respiratory coronavirus differs from transmissible gastroenteritis virus by a few genomic deletions. J. Gen. Virol. 71, 2599–2607.

Sanchez, C.M., Jimenez, G., Laviada, M.D., Correa, I., Sune, C., Bullido, M.J., Gebaues, F., Smerdou, C., Callebaut, P., Escribano, J.M. and Enjuanes, L. (1990) Antigenic homology among coronaviruses related to transmissible gastroenteritis virus. Virology 174, 410–417.

Schmidt, I., Skinner, M. and Siddell, S.G. (1987) Nucleotide sequence of the gene encoding the surface projection glycoprotein of coronavirus MHV-JHM. J. Gen. Virol. 68, 47–56.

Siddell, S., Wege, H. and ter Meulen, V. (1983) The biology of coronaviruses. J. Gen. Virol. 64, 761–776.

Sodroski, J., Goh, W.C., Rosen, C., Dayton, A., Terwilligier, E. and Haseltine, W.A. (1986) A second post-transcriptional trans activator gene required for HTLV-III replication. Nature 321, 412–417.

Spaan, W., Cavanagh, D. and Horzinek, M.C. (1988) Coronaviruses: structure and genome expression. J. Gen. Virol. 69, 2939–2952.

Staden, R. (1986) The current status and portability of our sequence handling software. Nucleic Acids Res. 14, 217–231.

Stauber, R., Pfleiderer, M. and Siddell, S. (1993) Proteolytic cleavage of the murine coronavirus surface glycoprotein is not required for fusion activity. J. Gen. Virol. 74, 183–191.

Sturman, L. and Holmes, K. (1985) The novel glycoproteins of coronaviruses. Trends Biochem. Sci. 10, 17–20.

Sturman, L.S., Ricard, C.S. and 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. J. Virol. 56, 904–911.

Tomley, F.M., Binns, M.M., Boursnell, M.E.G., Mockett, A.P.A., Brown, T.D.K. and Smith, G.L. (1987) Expression of IBV spike protein by a vaccinia virus recombinant. J. Gen. Virol. 68, 2291–2298.

Tyrrell, D.A., Almeida, J.D. and Berry, D.M. (1968) Coronaviruses. Nature 220, 650.

Vennema, H., Heijnen, A., Zijderveld, A., Horzinek, M.C. and Spaan, W.J.M. (1990) Intracellular transport of recombinant coronavirus spike proteins: implications for virus assembly. J. Virol. 64, 339–346.

Wesseling, J.G., Vennema, H., Godeke, G.-J., Horzinek, M.C. and Rottier, P.J.M. (1994) Nucleotide sequence and expression of the spike (S) gene of canine coronavirus and comparison with the S proteins of feline and porcine coronaviruses. J. Gen. Virol. 75, 1789–1794.

Williams, R.K., Jiang, G.-S. and Holmes, K.V. (1991) Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins. Proc. Natl. Acad. Sci. USA 88, 5533–5536.