THE HAEMAGGLUTININ OF BOVINE CORONAVIRUS EXHIBITS SIGNIFICANT
SIMILARITY TO THE HAEMAGGLUTININ OF TYPE C INFLUENZA VIRUS

Michael D. Parker, Graham J. Cox, Dongwan Yoo,
David R. Fitzpatrick and Lorne A. Babiuk

Veterinary Infectious Disease Organization
University of Saskatchewan
Saskatoon, Saskatchewan

INTRODUCTION

Coronaviruses are large enveloped RNA viruses containing a single-stranded genome complexed with a phosphorylated nucleocapsid protein which is surrounded by an envelope derived from intracellular membranes. The envelope contains two or three species of glycoproteins. The largest, E2 or S, is cleaved to S1 and S2 subunits and forms the large club-shaped peplomers on the surface of the virion which gives the virus its characteristic appearance. The E1 glycoprotein is believed to be responsible for determining the intracellular site of viral morphogenesis. A smaller group of coronaviruses contain a third membrane glycoprotein (Hogue and Brian, 1986, King et al., 1985, Pocock and Garwes, 1977), E3 or H, which, in the case of bovine coronavirus, exhibits haemagglutinating and acetylcholinesterase activities.

Bovine coronavirus is one cause of severe scours in newborn cattle and is responsible for significant morbidity and mortality. Previous work has demonstrated that monoclonal antibodies against E2 or E3 are capable of neutralization and are protective in vivo (Deregt et al., 1989). In order to examine the biochemical and immunological properties of these glycoproteins in more detail, the gene for the E3 polypeptide has been molecularly cloned (Parker et al., 1989). Expression of the cloned E3 gene has shown that the polypeptide possesses receptor-destroying activity similar to that exhibited by the haemagglutinin/esterase (HE) of type C influenza virus. In addition, the sequence of the E3 polypeptide is very similar to the HE polypeptide of type C influenza virus.

RESULTS AND DISCUSSION

A series of cDNA clones representing the 3' 10kb of the bovine coronavirus genome have been mapped and sequenced and the resulting map is shown in Figure 1. The E2 and E3 genes are immediately adjacent on the genome. The open reading frame of the E2 gene is 4089 nucleotides in length and encodes a 145K polypeptide (Parker et al., submitted). Immediately upstream is the E3 gene which encodes a polypeptide of 47K (Parker et al., 1989). Immediately downstream of the E2 gene is a sequence of 946 nucleotides thought to encode two non-structural polypeptides (Leibowitz et al., 1982). As shown in the map in Figure 1,

Figure 1. The 3' end of the genome of bovine coronavirus. The numbers below each designated gene indicate the length of the open reading frame exclusive of intergenic regions. The number in parentheses indicates the length of an uncharacterized sequence.

we have tentatively identified a 252 nucleotide open reading frame which would encode a 9.5K polypeptide (Cox and Parker, submitted). The remaining 694 nucleotides in this region adjacent to the E2 gene have consistently failed to generate an open reading frame capable of generating a polypeptide of greater than 7K. Further characterization of this region is underway. The E1 and N genes are adjacent and precede a non-coding sequence of 291 nucleotides immediately upstream of the poly-A at the 3' terminus of the genomic RNA (Lapps et al., 1987). Therefore, the characteristic gene order reported for other coronaviruses is also found in BCV.

Characterization of the Esterase Activity of Bovine Coronavirus

Recently, Vlasak et al. (1989a, 1989b) reported that the bovine coronavirus exhibited receptor binding specificity and receptor-destroying activity similar to the type C influenza viruses. Indirect evidence that the BCV haemagglutinin was responsible for the observed enzymatic activity was obtained by radiolabeling the E3 polypeptide with ³H-DFP which specifically inhibited the hydrolysis of p-nitrophenyl acetate. We previously reported that E3-specific monoclonal antibodies which inhibited haemagglutination also neutralized infectivity (Parker et al., 1989). In order to determine if the BCV E3 polypeptide was responsible for the virus-specific hydrolysis of acetate from O-acetylated sialic acids, we tested the ability of monoclonal antibodies specific for the E3-polypeptide to inhibit the release of acetate from bovine submaxillary mucin. When 2 ug of purified virus was incubated in phosphate buffered saline containing 25 mg/ml of BSM, an average of three determinations showed that 26 ug of acetate was released as determined with a commercially available kit. When the virus was incubated with a pool of E3-specific monoclonal antibodies (Deregt et al., 1988) prior to addition of substrate, background levels of free acetate were detected. Of the three monoclonals which bind three distinct epitopes, the monoclonal antibodies KD9-40 and HC10-5 completely inhibited esterase activity. The third antibody, BD9-8C, which shows minimal neutralizing and haemagglutination inhibiting (HI) activity, partially inhibited the esterase activity. The specificity of the inhibition was indicated by the lack of inhibition of non-immune sera or a pool of monoclonals

Table I
Antibody Inhibition of BCV Acetylesterase

Antibody	Specificity Antigenic Group	Neutralization	HI	Acetate Released (ug)
None	-	-	-	26
Pooled	E3	+++	+++	2
KD9-40	E3 (A2)	+++	+++	2
HC10-5	E3 (A1)	+++	+++	2
BD9-8C	E3 (C)	+	-	15
Anti-BHV gIII	-	-	-	25

specific for the bovine herpesvirus gIII glycoprotein. Similar results were obtained with monoclonal antibodies specific for the BCV N, E1 and E2 polypeptides. Therefore, these data further support the notion that the acetylcysteine activity of BCV is a property of the E3 polypeptide.

In order to determine if the acetylcysteine activity of BCV is an intrinsic property of only the E3 polypeptide, a cDNA clone representing the E3 gene (Parker et al., 1989) was inserted into the genome of the baculovirus *Autographa californica*. Infection of insect cells (SF9) with the recombinant virus, BVLE3, results in the production of a 130K dimer which dissociates into 58K monomers under reducing conditions (Parker et al., submitted). Intact *S. frugiperda* cells infected with the recombinant virus were assayed for acetylcysteine activity. As shown in Table II, 6×10^6 BCV-infected MDBK cells were capable of hydrolyzing approximately 15 μ g of acetate from submaxillary mucin. While uninfected insect cells failed to release acetate from BSM, BVLE3 infection of insect cells resulted in significant esterase activity. The esterase activity on BVLE3-infected insect cells was inhibited by the same antibody which inhibited the esterase in purified virus (data not shown). The ability of the E3 polypeptide produced in BVLE3-infected insect cells to hydrolyze the substrate in the absence of other viral components proves that the acetylcysteine of bovine coronavirus is an intrinsic property of the E3 polypeptide.

Comparison of the BCV E3 Amino Acid Sequence with the HE of Type C Influenza Virus

In agreement with previous reports (Vlasak et al., 1988a, 1988b), the data above indicate that the receptor-destroying enzyme of BCV is similar in substrate specificity to that of the HE of type C influenza. A search of the protein sequence database of the National Biomedical Research Foundation indicated that extensive amino acid sequence homology existed between the E3 and HE of type C influenza.

Figure 2 shows a comparison of the sequence of the BCV E3 polypeptide and the HE of C/Cal/78 influenza virus (Nakada et al., 1984). A similar result was observed in a comparison between C/Cal/78 and an apparent pseudogene located in the MHV-A59 genome at a position analogous to the location of the BCV E3 gene (Luytjes et al., 1988).

The similarity becomes apparent beginning at amino acid 19 of BCV E3 and amino acid 47 of the C/Cal/78 HE. Amino acid 19 is the amino terminus of the mature E3 polypeptide as the preceding 18 amino acids represent the signal sequence and are removed during maturation of the

Table II
Expression of BCV Acetylcysteine in Virus-Infected Cells

Cells	μ g Acetate Released
MDBK	
uninfected	4
BCV-infected	15
<i>Spodoptera frugiperda</i>	
uninfected	3
BVLE3-infected	
-cells	37
-medium ^a	6

^a Immunoprecipitated from culture medium with Mab BD9-3C prior to assay

of the polypeptide (Hogue et al., 1989). The similarity extends to position 384 of the E3 sequence which is six amino acids upstream of a hydrophobic sequence thought to be the membrane anchor of the polypeptide (Parker et al., 1989). The similar portion of the C/Cal/78 HE extends to position 427 which is 20 amino acids upstream from the cleavage site of the HE precursor. Within these boundaries, the similarity is 27% when only identical residues are considered. However, optimization of the alignment through the use of the Dayhoff scoring matrix increases the similarity to 52% as indicated by the boxed positions shown in Figure 2.

Several aspects of the C/Cal/78 sequence emphasize the differences between the polypeptides being compared. The HE is a disulfide-linked heterodimer composed of subunits derived by proteolytic cleavage of a common precursor. As a result, the cysteine residues at positions 20 and 582, thought to be involved in the formation of the intermolecular linkages (Nakada et al., 1984) are located in regions lacking similarity to the E3 and are not shown in Figure 2. In addition, the HE is anchored in the virion envelope by a hydrophobic domain near the carboxy-terminus of the HA2 subunit which is also outside of the regions compared in Figure 2.

Of the six cysteines in the HE thought to be involved in intramolecular disulfide formation, only 3 are conserved in the BCV E3. The cysteines at positions 347 and 371 are conserved in both viruses and may be involved in intramolecular linkages in the HE of C/Cal/78. However, at present the only disulfide known to exist in E3 is in the linkage which is responsible for dimer formation.

With one exception at the asparagine at position 104 of E3, the location of the potential N-glycosylation sites are not conserved between the two polypeptides.

Although both polypeptides exhibit similar receptor specificities and enzymatic activities, only the identity of the active site of the HE has been investigated. After radiolabeling with DFP, Vlasak et al. (1989) determined that the labeled tryptic peptide of the HE contained the sequence TWIGFGDSR with the radiolabel attached to the serine residue. Examination of the E3 sequence shows that the sequence is very highly conserved from positions 33 to 41 which is analogous to the position of the sequence in the C/Cal/78. Experiments are underway to determine if this region comprises the active site of the BCV esterase.

Inspection of the sequence comparison in Figure 2 clearly indicates that the two polypeptides are related. The lack of significant nucleotide sequence similarity (data not shown) as well as the dissimilarity in the locations of glycosylation sites, potential positions of disulfide linkages and the different pathways leading to the formation of the dimeric structures of the final gene products strongly suggests that any proposed genetic interaction between these viruses undoubtedly occurred between an ancient influenza-like virus and a coronavirus progenitor. Reports that an E3-like molecule may be found in variants of murine hepatitis virus (Taguchi et al., 1985, 1986) also suggests that E3 may be a non-essential polypeptide in those viruses which contain it. It may also be argued, based upon these same considerations that the two polypeptides represent convergent evolution. However, based upon the similarities of the binding and enzymatic properties of the two polypeptides, it seems appropriate that the term HE indicating haemagglutinin-acetylsterase be adopted as the designation for the BCV haemagglutinin.

BCVE3	F D - N P P T N V V S H L N - G D W F L F G D S R S D C N H V V N T N P R N Y	55
HICHE1	F e L v k P k a g a S v L N q S t W i g F G D S R t D - - q s n s a f P R - -	82
BCVE3	S - Y M D L N P A L C D S G K I S S K A G N S I F R S F H F T D F Y N Y T G E	93
HICHE1	S - l M s a k t A d k f r s l s g g s l m l S m F g p p g k v D - Y l Y q G c	119
BCVE3	G Q Q I I F Y E G L N F T P Y H A F K C T T S G S N D I W M Q - N K G L F Y T	131
HICHE1	G k h k v F Y E G v N w s P h a A i d C y r k n w t D i k l n f q K s i y e l	158
BCVE3	Q V Y K N - M A V Y R S L T F V N V P Y - V Y N G S A Q S - - T A L C K S G S	166
HICHE1	a s q s h c M s l v n a L d - k t i P l q V t k G v A k n c n n s f L K n p a	196
BCVE3	L V L N N P A Y I A R E A N F G D Y - Y Y K V E A D F Y L S G C D E Y I V P L	202
HICHE1	L y t q e v k p l e g i c g e e n l a f f t l p t q F g t y e C k l h l v a s	235
BCVE3	C - - I F N G K F L S N T K Y Y D D - S Q Y Y F N K D T G V I Y G L N S T E T	240
HICHE1	C y f I y d s K e v y N k r g c g n y f Q v i y d s s g k V v g G L d n r v s	274
BCVE3	I T T G - - - - F D F N C H Y L V L P S G N Y L A I S N E L L L T V P T K A	274
HICHE1	p y T G n s q d t p t m q C d m L q L k p G r Y s v r S s p r f L l m p e r s	312
BCVE3	I C L N K R K D F T P V Q V V D S R W N N A R Q S D N M T A V A C - Q P P Y C	312
HICHE1	y C f d - m K e k g P v t a V q S i W g k g r k S D y a v d q A C l s t P g C	351
BCVE3	Y F R N S T T N Y V G V Y D I N H G D A G F T S I L S G L L Y D S P C F S Q Q	351
HICHE1	m l i q k q k p Y i G e a D d h H G D q e m r e l L S G L d Y e a r C i s Q s	390
BCVE3	G - V F R Y D N V S S V W P L Y S - Y G R C P T A A - D I N T P D V P	384
HICHE1	G w V n e t s p f t e e y l L p p k f G R C P l A A k e e s i P k i p	426

Figure 2. Amino acid sequence comparison of bovine coronavirus E3 glycoprotein and C/Cal/78 influenza virus HE. Similar amino acids determined by Dayhoff scoring matrix are boxed. "-" indicates a single residue gap introduced to maximize sequence alignment.

REFERENCES

- Deregt, D., Gifford, G. A., Ijaz, M. K., Watts, T. C., Gilchrist, J. E., Haines, D. M., Babiuk, L. A., 1989, Monoclonal antibodies to bovine coronavirus glycoproteins E2 and E3: Demonstration of in vivo virus neutralizing activity, *J. Gen. Virol.*, 70:993.
- Hogue, B. G., Brian, D. A., 1986, Structural proteins of human respiratory coronavirus OC-43, *Virus Res.*, 5:131.
- Hogue, B. G., Green, T. E., Brian, D. A., 1989, Synthesis and processing of the bovine enteric coronavirus hemagglutinin protein, *J. Gen. Virol.*, 70:345.
- King, B., Potts, B. J., Brian, D. A., 1985, Bovine coronavirus hemagglutinin protein, *Virus Res.*, 2:53.
- Lapps, W., Hogue, B. G., Brian, D. A., 1987, Sequence analysis of the bovine coronavirus nucleocapsid and matrix protein genes, *Virology*, 157:47.
- Leibowitz, J. L., Weiss, S. R., Paavola, E., Bond, C. W., 1982, Cell free translation of murine coronavirus RNA, *J. Virol.*, 43:905.
- Luytjes, W., Breenbeek, P. J., Noten, A. F. H., Horzinek, M. C., Spaan, W. J., 1988, Sequence of mouse hepatitis virus A59 mRNA 2: Indications for RNA recombination between coronaviruses and influenza C virus, *Virology*, 166:415.
- Nakada, S., Creager, R. S., Krystal, M., Aaronson, R. P., Palese, P., 1984, Influenza C virus hemagglutinin: Comparison with influenza A and B virus hemagglutinins, *J. Virol.*, 50:118.
- 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 hemagglutinin of bovine coronavirus, *J. Gen. Virol.*, 70:155.

- Pocock, D. H., Garwes, D. J., 1977, The polypeptides of haemagglutinating encephalomyelitis virus and isolated subviral particles, *J. Gen. Virol.*, 37:487.
- Taguchi, F., Siddell, S. G., Wege, H., Ter Meulen, V., 1985, Characterization of a variant virus selected in rat brains after infection by coronavirus mouse hepatitis virus JHM, *J. Virol.*, 54:429.
- Taguchi, F., Massa, P. T., Ter Meulen, V., 1986, Characterization of a variant virus isolated from neural cell culture after infection of mouse coronavirus JHMV, *Virology*, 155:267.
- Vlasak, R., Luytjes, W., Leider, J., Spaan, W, Palese, P., 1988a, The E3 protein of bovine coronavirus is a receptor-destroying enzyme with acetylcholinesterase activity, *J. Virol.*, 62:468.
- Vlasak, R., Luytjes, W., Spaan, W., Palese, P., 1988b, Human and bovine coronaviruses recognize sialic acid-containing receptors similar to those of influenza C viruses, *Proc. Natl. Aca. Sci.*, 85:4526.