JOURNAL OF VIROLOGY, Apr. 1990, p. 1834–1838
0022-538X/90/041834-05$02.00/0

Vol. 64, No. 4

Structure and Orientation of Expressed Bovine Coronavirus Hemagglutinin-Esterase Protein

THOMAS E. KIENZLE,† SUSHMA ABRAHAM, BRENDA G. HOGUE,‡ AND DAVID A. BRIAN*

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

Received 6 September 1989/Accepted 2 January 1990

The sequence of the hemagglutinin-esterase (*HE*) gene for the Mebus strain of bovine coronavirus was obtained from cDNA clones, and its deduced product is a 47,700-kilodalton apoprotein of 424 amino acids. Expression of the HE protein in vitro in the presence of microsomes revealed N-terminal signal peptide cleavage and C-terminal anchorage but not disulfide-linked dimerization. Dimerization was observed only after expression in vivo, during which HE was also transported to the cell surface.

Hemagglutination is a distinct and readily assayable quality of some but not all coronaviruses (3, 4, 17, 23, 36). The property of hemagglutination has been correlated in the bovine coronavirus (BCV) with a 140-kilodalton (kDa) structural glycoprotein called the hemagglutinin (gp140, H, or E3 in earlier papers [7, 19, 24]). Recently the BCV hemagglutinin has been shown to induce neutralizing antibodies (5, 6) and to possess an esterase receptor-destroying activity that may be important for virus entry (38, 39). This glycoprotein may therefore be called the hemagglutinin-esterase (HE), and it may serve as a second viral attachment protein (in addition to the large spike protein) for initiating infection. The 140-kDa BCV HE protein is a disulfide-linked dimer of separately synthesized, identical 65-kDa glycosylated subunits (7, 18). Glycosylation is only of the asparagine-linked type, and the unglycosylated subunit electrophoretically migrates as a 42.5-kDa protein (18).

Here we report the sequence of the *HE* gene from the Mebus strain of BCV and compare it to the recently published sequence of the Quebec strain (31). We also document its membrane orientation and demonstrate that it has the potential to become an independently anchored glycoprotein on the surface of infected cells.

The sequence of the *HE* gene was obtained from two cDNA clones, I1 [extending between bases 7609 and 6698 from the 3′ end of the genome, exclusive of the poly(A) tail], and LA6 (extending between bases 8955 and 7483). cDNA cloning was done from genomic RNA as previously described (12, 26) except that random 5-mer oligodeoxynucleotides (Pharmacia) (for I1) or a synthetic primer (5′AT TATGACCGCAACACC3′) (for LA6) was used for first-strand synthesis. DNA sequencing of both strands was done by the dideoxynucleotide-induced-chain termination method (25, 33) with a nested set of subclones (13) or synthetic primers. To form a contiguous *HE* gene sequence, LA6 was spliced at the *Hae*III site (Fig. 1, between bases 1104 and 1105) to the 5′ *Hae*III-*Sty*I fragment of clone I1 (the *Sty*I site occurs 343 bases downstream from the termination codon of *HE*). Subclones of this construct were used for expression analyses.

For in vitro expression, the gene sequence beginning with the CTAAAC consensus sequence (15 bases upstream of the putative ATG initiation codon; Fig. 1) and ending 343 bases downstream of the TAG termination codon was placed under the control of the SP6 promoter in pGEM4Z (Promega Biotec). The resulting plasmid, pHESP6, was linearized with *Hin*dIII to yield full-length transcripts or with *Eco*RV (which cuts between bases 993 and 994 in the sequence shown in Fig. 1) to yield transcripts lacking the putative C-terminal anchor. Capped transcripts were translated in the presence or absence of canine pancreatic microsomes (Promega) in wheat germ extract (Promega) that contained 1 mCi of [^{35}S]methionine (>800 Ci/mmol; ICN Pharmaceuticals Inc.) per ml. Immunoprecipitates were prepared with polyclonal rabbit anti-gp65 (anti-HE) (1, 19). The method of Scheele (34) was used for protease digestion experiments, and that of Fujiki et al. (11) was used for carbonate extractions at pH 11. Competitive inhibition of N-linked glycosylation was accomplished with 30 μM octanoyl-Asp-Leu-Thr (a gift from F. Naider, City University of New York) in the microsome-containing translation mixture (27).

For in vivo expression, the sequence shown in Fig. 1 (including 49 bases downstream of the *HE* stop codon) was cloned into pGEM3Z under the control of the T7 promoter to construct plasmid pHET7. For immunofluorescence, HRT cells were infected with vaccinia virus vTF7-3 (with a multiplicity of infection of 30), which expresses T7 polymerase (10), transfected with 1.0 μg of pHET7 DNA per cm^2 by using Lipofectin (Bethesda Research Laboratories, Inc.) (8), and prepared at 36 h postinfection for internal or surface immunofluorescence (21). For immunoprecipitation, CV-1 cells infected with vaccinia virus vTF7-3 (with a multiplicity of infection of 2) were transfected with 0.4 μg of plasmid pHET7 DNA per cm^2, labeled with 100 μCi each of [^{35}S] cysteine (>600 μCi/mmol; Amersham Corp.) and [^{35}S]methionine (800 μCi/mmol; ICN) per 60-mm dish, and incubated for 16 h before harvest. Immunoprecipitation (21) was done on cell lysates.

The *HE* gene sequence of the Mebus strain of BCV (Fig. 1) differs by only two bases from that of the Quebec strain (31). In the Quebec strain, base 322 is C, making amino acid 103 leucine, and base 522 is A, resulting in no amino acid difference. The Mebus HE protein thus has the features described for the Quebec strain HE protein, including a predicted signal peptide, nine potential sites for N-linked

* Corresponding author.

† Present address: Department of Molecular Biology and Microbiology, School of Medicine, Case Western Reserve University, Cleveland, OH 44106.

‡ Present address: Department of Microbiology and Immunology, University of California, Los Angeles, School of Medicine, Los Angeles, CA 90024.

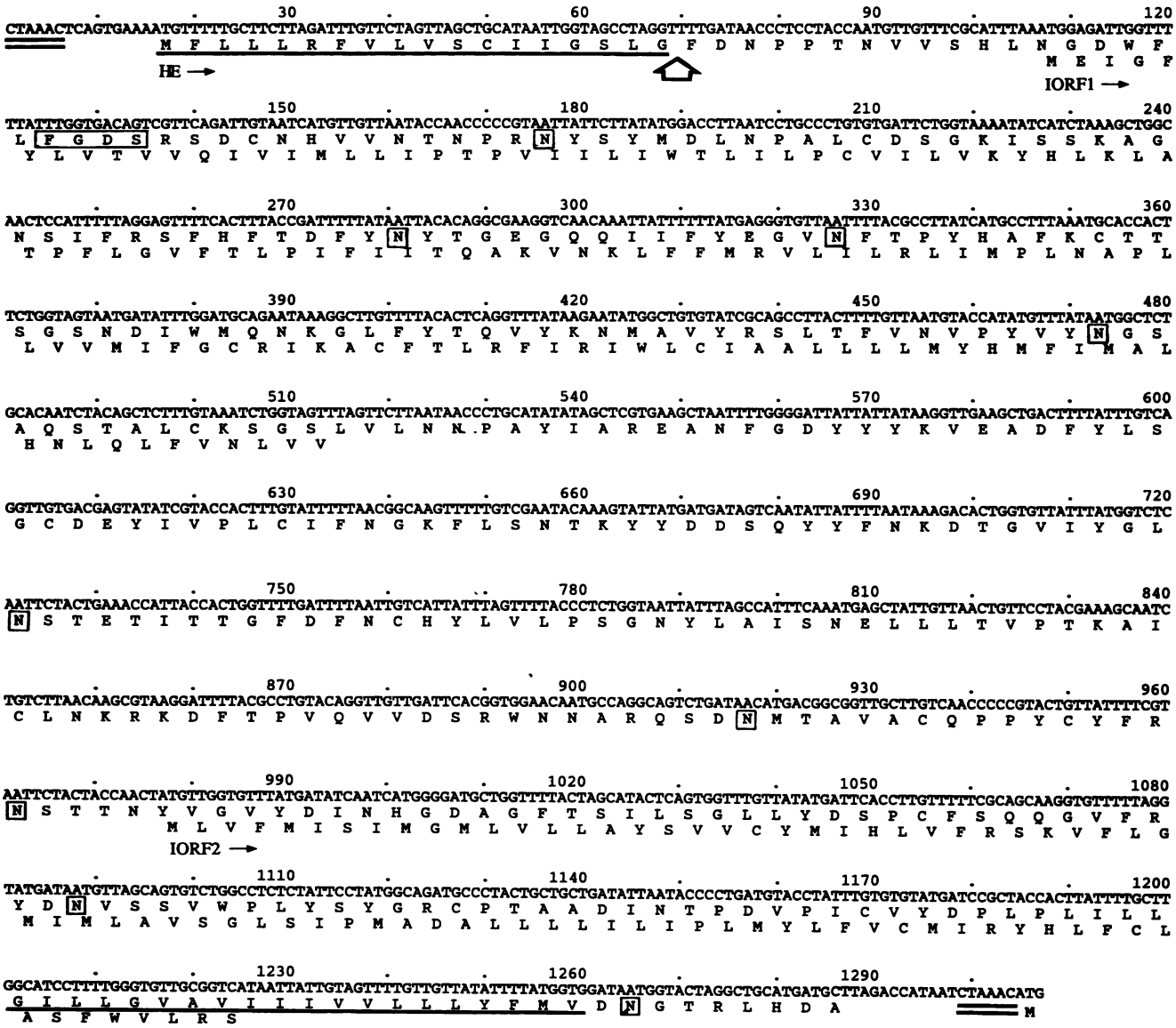

FIG. 1. Sequence of the *HE* gene and deduced amino acid sequences of the HE, IORF1, and IORF2 proteins. The nucleotide sequence begins 15 bases upstream from the presumed start codon of the *HE* gene [8,694 bases from the 3′ end of the genome, excluding the poly(A) tail] CTAAAC consensus sequence (doubly underlined) and ends with the ATG initiation codon of the spike protein gene. The amino-terminal signal peptide and presumed carboxy-terminal anchor sequences are underlined. The amino-terminal signal cleavage site is indicated by an arrow (⌓). The putative esterase active site (F-G-D-S) and potential N-linked glycosylation sites (NXS or NXT, where X ≠ P) are boxed. The CTAAAC consensus sequence preceding the spike protein gene is also doubly underlined. The GenBank accession number of the nucleotide sequence is M31052.

glycosylation, 14 cysteine residues, large nonoverlapping internal open reading frames of 408 (IORF1) and 249 (IORF2) bases, and a potential anchor region of 26 amino acids near the C terminus (Fig. 1). The N-terminal signal peptide cleavage site is predicted to be between amino acids 18 and 19 on the basis of the virion HE protein sequence data (18), and the putative active site for neuraminate *O*-acetylesterase activity, F-G-D-S (39, 40), is found beginning at base 124.

On the basis of the identification of an N-terminal signal peptide (18), each subunit of the virion-anchored HE is predicted to have an N-terminal ectodomain and a C-terminal anchor. To test this prediction, membrane orientation of HE synthesized in vitro in the presence of microsomes was determined. HE produced in vitro appeared authentic, since full-length transcripts translated in the absence of microsomes produced an unglycosylated product (HE^U) with a molecular mass of 43 kDa (Fig. 2A, lane 1) that was similar in size to the 42.5-kDa molecule identified in tunicamycin-treated cells (18) and since those translated in the presence of microsomes produced a glycosylated product (HE^G) with a molecular mass of 65 kDa (Fig. 2A, lane 2) that appeared by electrophoresis to be identical to the virion-derived HE subunit (18, 24). The authenticity of both products was also established by immunoprecipitation with HE-specific polyclonal antiserum (data not shown). Furthermore, HE^G produced in the presence of microsomes was completely or nearly completely protected from protease

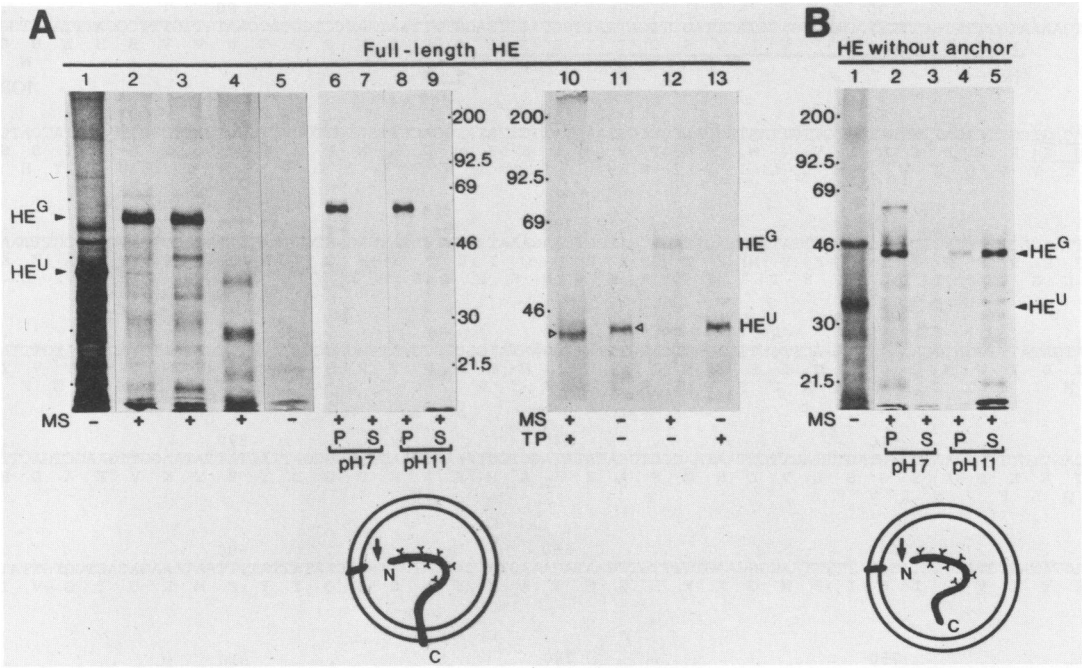

FIG. 2. Translocation, signal cleavage, and membrane anchorage of the HE protein monomer. (A) Full-length transcripts of the *HE* gene were translated without (−) or with (+) microsomes (MS), electrophoresed on sodium dodecyl sulfate-polyacrylamide slab gels, and autoradiographed. Lanes 1 and 2, Products directly from translation mix; lane 3, microsomes incubated without protease; lane 4, microsomes incubated with protease; lane 5, translation of transcripts generated from *Hin*dIII-cut pGEM4Z vector; lanes 6 and 7, membrane pellet (P) and supernatant (S) from protein extractions done at pH 7; lanes 8 and 9, membrane pellet and supernatant from protein extractions done at pH 11; lanes 10 through 13, transcripts translated as indicated. Open arrowheads mark the positions of the major products. TP, Octanoyl-Asp-Leu-Thr tripeptide. (B) Transcripts of the carboxy-terminally truncated *HE* gene were translated without or with microsomes, and the products were treated as indicated (lanes 1 through 5). The properties of products from full-length and truncated transcripts are schematically depicted in the lower part of the figure. Gels were either 9% (A [lanes 1 to 9] and B) or 12% (A [lanes 10 to 13]) polyacrylamide (19). All samples were treated with 5% 2-mercaptoethanol. HE^U and HE^G identify the unglycosylated and glycosylated HE monomers, respectively.

digestion (Fig. 2A, lanes 3 and 4), indicating complete or nearly complete translocation into the microsomal lumen.

HE^G was determined to be anchored in microsomal membranes by extraction experiments in sodium carbonate at pH 11 in which HE^G was found to remain bound to the pelleted membrane fraction (Fig. 2A, lanes 8 and 9). Two experimental approaches demonstrated that the membrane anchorage was C terminal. In the first, N-terminal signal peptide cleavage in microsomes was demonstrated to rule out the possibility that an uncleaved hydrophobic signal peptide was serving as an anchor. For this, a competitive inhibitor of N-linked glycosylation (27) was used so that a change in molecular weight resulting from signal cleavage would not be obscured by glycosylation. HE^U made in microsomes (Fig. 2A, lane 10) was approximately 2 kDa smaller than HE^U made in the absence of microsomes (Fig. 2A, lane 11), indicating that signal peptide cleavage had taken place. The increase in the migration rate observed in lane 10 of the gel shown in Fig. 2A was not due merely to the presence of the tripeptide (Fig. 2A, lane 13). The second approach tested for anchorage of a C-terminally truncated protein (made from a truncated transcript). The truncated protein did not become anchored, even though it was successfully translocated into microsomes (Fig. 2B, lanes 1 through 5).

At no time could disulfide-linked dimeric forms of HE be demonstrated during in vitro expression experiments. Attempts to enhance redox conditions for disulfide isomerase by varying oxidized glutathione concentrations between 0.5 and 7 mM (9, 35) did not help. Possibly, the orientation or proximity of the nascent chains was not proper for dimerization. A curious species of approximately 46 kDa was routinely observed when the truncated *HE* gene was translated (Fig. 2B, lane 1) (or 60 kDa when this gene was translated in the presence of microsomes [Fig. 2B, lane 2]), but it was not a disulfide-linked oligomer, nor was it precipitable with HE-specific antiserum (data not shown). Its identity is not known.

HE in infected cells becomes rapidly disulfide linked and glycosylated, and much of it becomes incorporated into virions during virus assembly at the endoplasmic reticulum and *cis* Golgi (7, 18). Some, however, may escape incorporation into virions and subsequently migrate to the cell surface. Migration to the cell surface is suggested by the fact that some cells in an infected monolayer show surface immunofluorescence (Fig. 3B) and adsorb erythrocytes (22; T. E. Kienzle, S. Abraham, B. G. Hogue, and D. A. Brian, unpublished data). Alternatively, hemadsorption has been proposed to be the result of virions remaining adhered to cells at their sites of release and acting as bridges to bind erythrocytes to the cell (2). Expression experiments in vivo were done to establish whether cloned HE expressed in the absence of other coronavirus proteins can undergo dimerization and migrate to the cell surface as a membrane-anchored protein. The dimeric form of HE was produced in vivo (Fig. 3E, lanes 1 and 2). Furthermore, both intracytoplasmic fluorescence (data not shown) and surface fluorescence (Fig. 3D) were observed, indicating that HE, when expressed alone, does migrate to the cell surface. HE-specific hemadsorption could not be tested, since vaccinia

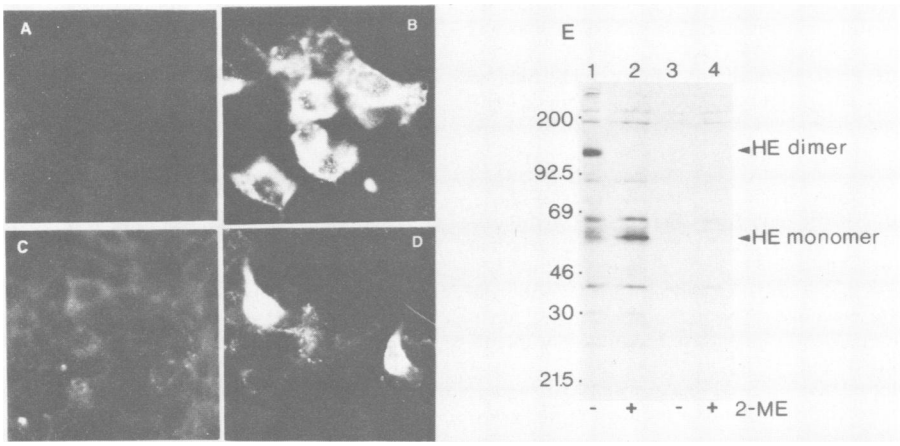

FIG. 3. Surface immunofluorescence and immunoprecipitation of HE protein. (A) Uninfected HRT cells. (B) BCV-infected HRT cells. (C) Vaccinia virus (vTF7-3)-infected HRT cells. (D) HRT cells infected with vaccinia virus vTF7-3 and transfected with pHET7 plasmid DNA. (E) Polyacrylamide gel electrophoresis of immunoprecipitated products from CV-1 cells infected with vaccinia virus vTF7-3 and transfected with pHET7 plasmid DNA (lanes 1 and 2) or pGEM3Z plasmid DNA (lanes 3 and 4). +, Treated with 2-mercaptoethanol (2-ME); −, not treated.

virus also produces a hemagglutinin that causes hemadsorption (20).

One biological consequence of HE anchorage on the surface of infected cells is that HE may become a target for both the humoral and the cellular branches of the immune system. Its usefulness in recombinant vaccines may therefore be twofold, since both humoral immunity and cellular immunity have been shown to be important protection mechanisms against viral infections of the gastrointestinal tract (30).

A presumably incomplete, possibly vestigial, form of the *HE* gene is found in mouse hepatitis coronavirus A59, and it, too, maps on the 5′ side of the peplomer protein gene (28). The mouse hepatitis coronavirus A59 "HE" protein, although apparently not expressed during infection (19), possesses the putative active site (F-G-D-S) for neuraminate *O*-acetylesterase activity (28). An amino acid sequence homology of 30% between the mouse hepatitis coronavirus A59 "HE" protein and the HA1 (HE1) subunit of the hemagglutinin of influenza C virus (28, 29, 32), which also possesses the F-G-D-S site and neuraminate *O*-acetylesterase activity (14, 16, 37, 40), has suggested an evolutionary relationship between these two proteins (28). This view is supported by a comparison of the BCV HE protein (amino acids 30 through 389) with the HA1 (HE1) subunit of influenza C virus HE protein (amino acids 67 through 438 [29]) in which the BCV HE protein shows a sequence identity of 29.7% with the HA1 (HE1) subunit. In this alignment, 10 of the 14 BCV HE cysteine residue positions and the position of the F-G-D-S esterase active site are conserved between the two proteins. Assuming there is a common ancestry, it is interesting to note that the influenza C virus HE glycoprotein derives from a cleaved precursor and apparently does not form homodimeric structures (15), whereas the BCV HE glycoprotein undergoes rapid dimerization and, except for signal peptide removal, is not derived by proteolytic cleavage of a larger precursor.

This work was supported by Public Health Service grant AI14367 from the National Institute of Allergy and Infectious Diseases and by U.S. Department of Agriculture grant 82-CRSR-2-1090. T.E.K. was a predoctoral fellow supported by Public Health Service training grant AI7123 from the National Institute of Allergy and Infectious Diseases.

LITERATURE CITED

1. **Anderson, D. J., and G. Blobel.** 1983. Immunoprecipitation of protein from cell-free translations. Methods Enzymol. **96:**111–120.
2. **Bucknall, R., A. Kalica, and R. Chanock.** 1972. Intracellular development and mechanism of hemadsorption of human coronavirus, OC43. Proc. Soc. Exp. Biol. Med. **139:**811–817.
3. **Callebaut, P. E., and M. B. Pensaert.** 1980. Characterization and isolation of structural polypeptides in hemagglutinating encephalomyelitis virus. J. Gen. Virol. **48:**193–204.
4. **Dea, S., and P. Tijssen.** 1988. Identification of the structural proteins of the turkey enteric coronavirus. Arch. Virol. **99:** 173–186.
5. **Deregt, D., and L. A. Babiuk.** 1987. Monoclonal antibodies to bovine coronavirus: characteristics and topographical mapping of neutralizing epitopes on the E2 and E3 glycoproteins. Virology **161:**410–420.
6. **Deregt, D., G. A. Gifford, M. K. Ijaz, T. C. Watts, J. E. Gilchrist, D. M. Haines, and L. A. Babiuk.** 1989. Monoclonal antibodies to bovine coronavirus glycoproteins E2 and E3: demonstration of *in vivo* virus-neutralizing activity. J. Gen. Virol. **70:**993–998.
7. **Deregt, D., M. Sabara, and L. A. Babiuk.** 1987. Structural proteins of bovine coronavirus and their intracellular processing. J. Gen. Virol. **68:**2863–2877.
8. **Felgner, P., T. Gadek, M. Holm, R. Roman, H. Chou, M. Wenz, J. Northrop, G. Ringold, and M. Danielson.** 1987. Lipofectin: a highly efficient lipid-mediated DNA transfection procedure. Proc. Natl. Acad. Sci. USA **84:**7413–7417.
9. **Freedman, R.** 1984. Native disulphide bond formation in protein biosynthesis: evidence for the role of protein disulphide isomerase. Trends Biochem. Sci. **9:**438–441.
10. **Fuerst, T. R., P. L. Earl, and B. Moss.** 1987. Use of a hybrid vaccinia virus-T7 RNA polymerase system for expression of target genes. Mol. Cell. Biol. **7:**2538–2544.
11. **Fujiki, Y., A. Hubbard, S. Fowler, and P. Lasarow.** 1982. Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. J. Cell Biol. **93:**97–102.
12. **Gubler, U., and B. Hoffman.** 1983. A simple and very efficient method for generating cDNA libraries. Gene **25:**263–269.
13. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:** 351–359.
14. **Herrler, G., I. Durkop, H. Becht, and H. Klenk.** 1988. The glycoprotein of influenza C virus is the hemagglutinin, esterase, and fusion factor. J. Gen. Virol. **69:**839–846.

15. **Herrler, G., A. Nagele, H. Meier-Ewert, A. S. Bhown, and R. W. Compans.** 1981. Isolation and structural analysis of influenza C virion glycoproteins. Virology **113**:439–451.
16. **Herrler, G., R. Rott, H.-D. Klenk, H.-P. Muller, A. K. Shukla, and R. Shauer.** 1985. The receptor-destroying enzyme of influenza C virus is neuraminate-O-acetylesterase. EMBO J. **4**: 1503–1506.
17. **Hogue, B. G., and D. A. Brian.** 1986. Structural proteins of human respiratory coronavirus OC43. Virus Res. **5**:131–144.
18. **Hogue, B. G., T. E. Kienzle, and D. A. Brian.** 1989. Synthesis and processing of the bovine enteric coronavirus hemagglutinin protein. J. Gen. Virol. **70**:345–352.
19. **Hogue, B. G., B. King, and D. A. Brian.** 1984. Antigenic relationships among proteins of bovine coronavirus, human respiratory coronavirus OC43, and mouse hepatitis coronavirus A59. J. Virol. **51**:384–388.
20. **Ichihashi, Y., and S. Dales.** 1971. Biogenesis of poxviruses: interrelationship between hemagglutinin production and polykaryocytosis. Virology **46**:533–543.
21. **Kaariainen, L., I. Virtanen, J. Saraste, and S. Keranen.** 1983. Transport of virus membrane glycoproteins, use of temperature-sensitive mutants and organelle-specific lectins. Methods Enzymol. **96**:453–465.
22. **Kapikian, A., H. James, S. Kelly, L. King, A. Vaughn, and R. Chanock.** 1972. Hemadsorption by coronavirus strain OC43. Proc. Soc. Exp. Biol. Med. **139**:179–186.
23. **King, B., and D. A. Brian.** 1982. Bovine coronavirus structural proteins. J. Virol. **42**:700–707.
24. **King, B., B. J. Potts, and D. A. Brian.** 1985. Bovine coronavirus hemagglutinin protein. Virus Res. **2**:53–59.
25. **Kraft, R., J. Tardiff, K. Krauter, and L. Leinward.** 1988. Using mini-prep plasmid DNA for sequencing double-stranded templates with sequenase. Biotechniques **6**:544–549.
26. **Lapps, W., B. G. Hogue, and D. A. Brian.** 1987. Sequence analysis of the bovine coronavirus nucleocapsid and matrix protein genes. Virology **157**:47–57.
27. **Lau, J. T. Y., J. K. Welply, P. Shenbagamurthi, F. Naider, and W. J. Lennarz.** 1983. Substrate recognition of oligosaccharyl transferase: inhibition of cotranslational glycosylation by acceptor peptides. J. Biol. Chem. **258**:15255–15260.
28. **Luytjes, W., P. Bredenbeek, A. Noten, M. Horzinek, and W. Spaan.** 1988. Sequence of mouse hepatitis virus A59 mRNA2: indications for RNA recombination between coronaviruses and influenza C virus. Virology **166**:415–422.
29. **Nakada, S., R. S. Craeger, M. Krystal, R. P. Aaronson, and P. Palese.** 1984. Influenza C virus hemagglutinin: comparison with influenza A and B virus hemagglutinins. J. Virol. **50**:118–124.
30. **Offit, P. A., and K. I. Dudzik.** 1989. Rotavirus-specific cytotoxic T lymphocytes appear at the intestinal mucosal surface after rotavirus infection. J. Virol. **63**:3507–3512.
31. **Parker, M. D., G. J. Cox, D. Deregt, D. R. Fitzpatrick, and L. A. Babiuk.** 1989. Cloning and *in vitro* expression of the gene for the E3 hemagglutinin glycoprotein of bovine coronavirus. J. Gen. Virol. **70**:155–164.
32. **Pfeifer, J. B., and R. W. Compans.** 1984. Structure of the influenza C glycoprotein gene as determined from cloned DNA. Virus Res. **1**:281–296.
33. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
34. **Scheele, G.** 1983. Methods for the study of protein translocation across the RER membrane using the reticulocyte lysate translation system and canine pancreatic microsomal membranes. Methods Enzymol. **96**:94–111.
35. **Scheele, G., and R. Jacoby.** 1982. Conformational changes associated with proteolytic processing of presecretory proteins allow glutathione-catalyzed formation of native disulfide bonds. J. Biol. Chem. **257**:12277–12282.
36. **Sugiyama, K., R. Ishikawa, and N. Fukuhara.** 1986. Structural polypeptides of the murine coronavirus DVIM. Arch. Virol. **89**:245–254.
37. **Vlasak, R., M. Krystal, M. Nacht, and P. Palese.** 1987. The influenza C virus glycoprotein (HE) exhibits receptor-binding (hemagglutinin) and receptor-destroying (esterase) activities. Virology **160**:419–425.
38. **Vlasak, R., W. Luytjes, J. Leider, W. Spaan, and P. Palese.** 1988. The E3 protein of bovine coronavirus is a receptor-destroying enzyme with acetylesterase activity. J. Virol. **62**: 4686–4690.
39. **Vlasak, R., W. Luytjes, W. Spaan, and P. Palese.** 1988. Human and bovine coronaviruses recognize sialic acid-containing receptors similar to those of influenza C viruses. Proc. Natl. Acad. Sci. USA **85**:4526–4529.
40. **Vlasak, R., T. Muster, A. M. Lauro, J. C. Powers, and P. Palese.** 1989. Influenza C virus esterase: analysis of catalytic site, inhibition, and possible function. J. Virol. **63**:2056–2062.