J. gen. Virol. (1988), **69**, 2939–2952. *Printed in Great Britain*

Key words: *coronaviruses/genome expression/protein structure*

REVIEW ARTICLE

Coronaviruses: Structure and Genome Expression

By W. SPAAN,[1]* D. CAVANAGH[2] AND M. C. HORZINEK[1]

[1] *Institute of Virology, Veterinary Faculty, State University of Utrecht, Yalelaan 1, 3508 TD Utrecht, The Netherlands and* [2] *AFRC Institute for Animal Health, Houghton Laboratory, Huntingdon, Cambridgeshire PE17 2DA, U.K.*

INTRODUCTION

Progress in coronavirology is illustrated by the number of workshops convened and reviews written. International meetings have been held in Germany (1980), the Netherlands (1983) and the U.S.A. (1986), and the Fourth Coronavirus Symposium will be organized by one of us (D.C.) in Cambridge, U.K. in July 1989. In addition, reviews have appeared which highlighted particularly interesting characteristics of the family, e.g. the replication strategy (Lai, 1986) and the glycoproteins (Sturman & Holmes, 1985). As the last general accounts were published some 5 years ago (Siddell *et al.*, 1983; Sturman & Holmes, 1983) an update is timely. The present article is based on the large amount of sequence data accumulated in these years and focuses on the viral nucleic acids and proteins and their function.

Coronaviruses cause infections in man, other mammals and birds. Most experimental data have been obtained from studies of mouse hepatitis virus (MHV) and infectious bronchitis virus of chickens (IBV). Additional representatives of the family reviewed in this article are the human (HCV) and bovine (BCV) coronaviruses, transmissible gastroenteritis virus (TGEV), haemagglutinating encephalitis virus (HEV) and feline infectious peritonitis virus (FIPV).

VIRION PROTEINS

Coronaviruses possess three major structural proteins: a nucleocapsid protein (N), a small integral membrane glycoprotein (M, E1) and a large spike glycoprotein (S, E2); for the sake of uniformity we use the letters N, M and S in this review. While all coronaviruses contains these proteins, a subset (HEV, HCV-OC43 and BCV) is now recognized to possess an additional glycopolypeptide (gp65), which is unrelated to S or M.

N protein

The number of amino acids in the N protein has been determined by cloning and sequencing for MHV strains A59 (Armstrong *et al.*, 1983) and JHM (Skinner & Siddell, 1983), IBV strains Beaudette and M41 (Boursnell *et al.*, 1985*a*), for TGEV (Kapke & Brian, 1986) and BCV (Lapps *et al.*, 1987; Table 1). These proteins are basic, the basic residues occurring in clusters; the C terminus is acidic. Serine residues account for 8 to 10% of the total number of amino acids; their clustering may correlate with the fact that N is phosphorylated specifically on serines. The homology of the BCV N protein with that of MHV is 70% (72% base homology), of TGEV 29% (37% base homology) and of IBV 29% (43% base homology; Kapke & Brian, 1986; Lapps *et al.*, 1987). One prominent region of homology is a stretch of about 68 amino acids, which exhibits between 51 and 79% similarity depending on the pairs of viruses compared (Kapke & Brian, 1986; Lapps *et al.*, 1987).

M glycoprotein

As is the case with the N protein, the M glycoprotein of the various coronaviruses also exhibits different M_r values in polyacrylamide gels (see review by Siddell *et al.*, 1983; also Resta *et al.*, 1985; Hogue & Brian, 1986; Sugiyama *et al.*, 1986; Cavanagh & Davis, 1987). These variations are not only due to differences in the number of amino acid residues, but also to the extent of

Table 1. *Properties of virion proteins*

	IBV	MHV	FIPV	TGEV	BCV
Nucleocapsid protein (N)					
M_r ($\times 10^{-3}$)	45	50	–*	43	49
No. of amino acids	409	455	–	382	448
Matrix glycoprotein (M)					
M_r ($\times 10^{-3}$)	25	26	–	30	26
No. of amino acids					
Total, mature protein	224	227	–	245	229
Hydrophilic N terminus	21	24	–	29	23
Membrane-embedded domain	76	81	–	88	81
Luminal C terminus	127	122	–	128	125
No. potential glycosylation sites	1, 2	4	–	1	6
Spike protein (S)					
M_r ($\times 10^{-3}$)	128	137	159	158	–
No. of amino acids†					
Total, with signal	1162	1235‡ 1324§	1452	1447	–
S1, including signal sequence	537	628‡ 717§	NA‖	NA	–
S2	625	607‡ 606§	NA	NA	
Signal sequence	18	–	–	16	–
No. potential glycosylation sites	28	21	35	32	–
gp65/130					
Present in virion	–	JHM+/A59–	–	–	+

* Information not available.
† Approximate, because of strain variation.
‡ MHV-JHM.
§ MHV-A59.
‖ NA, Not applicable; spike not cleaved.

glycosylation, the type of linkage of the glycans (*O*-linked or *N*-linked) and the degree to which *N*-linked high-mannose (simple) glycans have been converted to complex glycans (Table 1). Nucleotide sequencing of the M gene of MHV-A59 and MHV-JHM (Armstrong *et al.*, 1984; Pfleiderer *et al.*, 1986), IBV-Beaudette and IBV-6/82 (Boursnell *et al.*, 1984; Binns *et al.*, 1986*a*), BCV (Lapps *et al.*, 1987) and TGEV (Laude *et al.*, 1987) has revealed many interesting features of the M glycoprotein. Computer predictions of its secondary structure have led to a model in which approximately 10% of the N-terminal part of the molecule is exposed on the outer surface of the virus membrane (see references above and Rottier *et al.*, 1986). This view is supported by experimental evidence (Rottier *et al.*, 1984; Cavanagh *et al.*, 1986*a*). The next 80 or so residues, approximately one-third of the molecule, form three hydrophobic α-helices, which span the membrane three times. The C-terminal half of the protein has neither strong hydrophobic nor hydrophilic properties and is located in the interior of the virus particle.

Although the M protein of MHV, IBV and BCV does not possess an N-terminal signal sequence, M needs the signal recognition particle for membrane insertion, like many secretory and membrane proteins (Rottier *et al.*, 1985). The first (amino-terminal) and/or third membrane-spanning helices can function as signal sequences (Machamer & Rose, 1987; Mayer *et al.*, 1988). However, this feature is not universal among the coronaviruses; a putative N-terminal signal peptide of 17 residues has been identified for the TGEV M protein (Laude *et al.*, 1987). The MHV M protein shares 86% homology with BCV but only 35% and 38% with IBV and TGEV, respectively (Lapps *et al.*, 1987; Laude *et al.*, 1987).

Previous studies (see Siddell *et al.*, 1983) have shown that the glycans of MHV (Niemann *et al.*, 1984), and of BCV are of the *O*-linked type, in contrast to the *N*-linked glycans of IBV and TGEV (Stern & Sefton, 1982*a*; Cavanagh, 1983*a*; Garwes *et al.*, 1984; Jacobs *et al.*, 1986; Cavanagh & Davis, 1987, 1988). Most M protein glycan molecules of IBV are of the simple type but a proportion are converted to complex glycans (Stern & Sefton, 1982*a*; Cavanagh, 1983*a*).

S protein

The S protein gene has been sequenced for several strains of IBV (Binns *et al.*, 1985; Niesters *et al.*, 1986; Binns *et al.*, 1986*b*), MHV-JHM (Schmidt *et al.*, 1987) and MHV-A59 (Luytjes *et al.*, 1987), TGEV (Rasschaert & Laude, 1987; Jacobs *et al.*, 1987) and FIPV (de Groot *et al.*, 1987*b*). These S glycoproteins possess an overall hydrophobic hydropathicity profile, an N-terminal signal sequence, a C-terminal hydrophilic sequence preceded by a membrane-spanning domain, and a large number of potential *N*-linked glycosylation sites, most of which appear to be glycosylated (Table 1). In contrast, there are large differences in the number of amino acids of S between different coronaviruses and different isolates of a given virus (see MHV in Table 1; also Taguchi *et al.*, 1985).

The extent to which S is cleaved into S1 and S2 depends on the virus in question and the cells infected. Most authors have reported the S protein of IBV (Stern & Sefton, 1982*b*; Cavanagh, 1983*b, c*; Cavanagh & Davis, 1987) and BCV (Hogue *et al.*, 1984; Deregt *et al.*, 1987) to be in a cleaved form, but no S protein of TGEV, FIPV or canine coronavirus (Garwes & Reynolds, 1981; Horzinek *et al.*, 1982) and only a little of HCV (Hogue & Brian, 1986) occurs cleaved. Cleavage of MHV S protein varies from 0 to 100%, depending on the virus strain and cell type (Sturman *et al.*, 1985; Sugiyama *et al.*, 1986).

The order of S1 and S2 within the S0 protein of IBV and MHV is: N terminus–S1–S2–C terminus (in MHV 90B and 90A are equivalent to S1 and S2, respectively; Binns *et al.*, 1985; Cavanagh *et al.*, 1986*b*; Luytjes *et al.*, 1987). For IBV and MHV the cleavage site is adjacent to the amino acid sequence RRFRR, RRSRR or RRHRR (IBV, eight strains sequenced; Cavanagh *et al.*, 1986*b*; Binns *et al.*, 1986*b*; J. G. Kusters *et al.*, unpublished data), RRAHR (MHV-A59; Luytjes *et al.*, 1987) and RRARR (MHV-JHM; Schmidt *et al.*, 1987). As would be expected from such basic sequences, S0 of both IBV and MHV can be cleaved *in vitro* by trypsin and, in the case of S0 with the connecting peptide RRFRR, by chymotrypsin (Sturman *et al.*, 1985; Frana *et al.*, 1985; Cavanagh *et al.*, 1986*a*). The S0 protein of TGEV and FIPV lacks such pairs of basic residues.

Sedimentation studies have indicated that the peplomer of IBV is an oligomer comprising two or three molecules of S1 + S2 (Cavanagh, 1983*c*). Interpeptide disulphide bonds are not involved in maintaining quaternary structure, permitting S1 to be removed from virions by urea treatment, with S2 left in place. This led to the proposal that the outer, bulbous part of S might be formed largely by S1, with S2 being anchored by its C terminus in the virus envelope (Cavanagh, 1983*c*). Computer-aided analysis of the S protein primary structure of IBV, MHV and FIPV (de Groot *et al.*, 1987*a*) has identified two heptad repeats in the C-terminal domain, which indicate an intra-chain coiled coil structure. These results have been confirmed for TGEV (Rasschaert & Laude, 1987). The major repeat suggests a helix occupying more than half the length of the peplomer. In the oligomer the major helices are probably involved in an inter-chain coiled coil, reminiscent of structures in the haemagglutinin glycoprotein trimer of influenza virus. The S2 protein of IBV has been found to be susceptible to hydrolysis by several proteases within a region adjacent to the N-terminal side of the membrane-spanning hydrophobic domain (Cavanagh *et al.*, 1986*a*).

Comparison of the peplomer protein amino acid sequences of IBV, MHV and FIPV has shown homologies in the S2 half of the molecule of 35, 30 and 29% for IBV–FIPV, IBV–MHV and MHV-FIPV, respectively (de Groot *et al.*, 1987*a*). In contrast, S1 (or the equivalent region in FIPV) exhibits very little conservation. In S2 several regions of more than 30% homology, including sequences of seven to ten identical amino acids, have been identified (Schmidt *et al.*, 1987; Rasschaert & Laude, 1987). The unusually high number of cysteine residues in the vicinity of the transmembrane domain is also conserved in the S proteins sequenced so far. Some of these residues may be involved in S2 acylation (Sturman *et al.*, 1985) which can occur in the absence of glycosylation (Van Berlo *et al.*, 1987).

gp65 glycopolypeptide

Early studies have shown that HCV-229E, HCV-OC43, HEV, MHV-JHM and BCV possess a glycopolypeptide in addition to S and M (see Siddell *et al.*, 1983; Makino *et al.*, 1983). More

recent studies with BCV (King *et al.*, 1985; Deregt *et al.*, 1987), HCV-OC43 (Hogue & Brian, 1986) and diarrhoea virus of infant mice (Sugiyama *et al.*, 1986) have confirmed that these viruses contain a glycopolypeptide of about 65K (gp65) which, in the absence of mercaptoethanol, runs as a dimer of 130K to 140K (gp130-140) in PAGE. Siddell (1982) has shown by tryptic peptide fingerprinting that gp65 is structurally unrelated to S.

Functions of the virion proteins

In addition to its role in encapsidating genomic RNA and facilitating its incorporation into virions (by the formation of ribonucleoprotein, RNP), the N protein has been implicated in the process of RNA replication. Addition of antiserum raised against N (but not against S or M) to an *in vitro* replication system inhibited the synthesis of genome-sized RNA by 90% (Compton *et al.*, 1987). In the presence of tunicamycin, MHV and IBV formed virions which lacked S but contained M and RNP (Holmes *et al.*, 1981; Rottier *et al.*, 1981; Stern & Sefton, 1982*a*). In infected cells virion budding occurs at the site of M accumulation (Tooze *et al.*, 1984). These data indicate that M is necessary for virus maturation and that it determines the site at which virus particles are assembled. The intracellular accumulation of the M protein is a property of the protein itself (Rottier & Rose, 1987; Machamer & Rose, 1987). A mutant M protein of IBV possessing only the first transmembrane domain accumulated intracellularly, while another one with only the third domain was transported to the plasma membrane (Machamer & Rose, 1987). The domain of M that interacts with the RNP is not known; it has been shown that M itself has an affinity for RNA (Sturman *et al.*, 1980).

It has been assumed that attachment of virions to cells is mediated by the S protein or by both S and gp65 (in HEV, BCV and HCV-OC43, which possess the latter). Recently, a 110K plasma membrane glycoprotein with an affinity for MHV-A59 was described, and its presence was correlated with the virus susceptibility of target cells (Boyle *et al.*, 1987). Neuraminidase and glycosidase treatment did not destroy its S protein-binding properties. The host cell protein probably serves as a virus receptor as cells were protected against infection after incubation with homologous polyclonal and monoclonal antibody (MAb) (K. Holmes, personal communication). Spikeless virions of IBV and MHV are non-infectious (Cavanagh, 1981; Stern & Sefton, 1982*a*; Holmes *et al.*, 1981; Rottier *et al.*, 1981). Removal from IBV of S1 (but not of S2) by urea abolished both infectivity and haemagglutinating activity (HA; Cavanagh & Davis, 1986). Although this result indicated that attachment to erythrocytes had been affected qualitatively, the amount of virus that attached to red blood cells and chicken embryo kidney cells was not reduced. The inference that HA is mediated by S1 is supported by the finding that MAbs to S1 inhibit HA (Mockett *et al.*, 1984). In those coronaviruses which possess gp65 it is this protein, probably as a dimer, which is associated with HA (see Siddell *et al.*, 1983; King *et al.*, 1985). Haemagglutination studies indicated that HCV-OC43 and BCV recognize *O*-acetylated sialic acid or a similar derivative as the red blood cell receptor, as do influenza C viruses (Vlasak *et al.*, 1988). It will be necessary to examine whether viral binding *in vivo* also involves modified sialic acids. BCV also exhibits an acetylesterase receptor-destroying activity similar to the enzymic activity found in influenza C viruses (Vlasak *et al.*, 1988). This activity is associated with the gp65 spike protein (R. Vlasak, W. Luytjes, W. Spaan & P. Palese, unpublished observations).

The failure of the virus lacking S1 to replicate despite attachment to cells suggested that some other function had been lost, most probably the fusion activity. The S protein induces membrane fusion, and this has been demonstrated in three ways: first, antiserum to S but not to M (Sturman *et al.*, 1985) and MAbs specific for S (Collins *et al.*, 1982; Wege *et al.*, 1984) inhibited cell fusion after infection with MHV (fusion from within); second, MHV-A59 grown in the 17 Cl 1 line of spontaneously transformed BALB/c 3T3 cells did not cause fusion of L cells when added at high multiplicity (fusion from without) unless S had previously been cleaved by trypsin (Sturman *et al.*, 1985); third, vaccinia virus recombinants containing either the MHV S gene or the FIPV S gene were able to induce cell fusion (H. Vennema, L. Heijnen & W. Spaan, unpublished observation). Although S is responsible for the induction of membrane fusion, its cleavage is not always a necessary precondition. In FIPV and TGEV, a cleaved S protein has not been observed, yet replication results in syncytium formation (de Groot *et al.*, 1987*c*). Similarly,

HCV-OC43 and a small plaque variant of MHV-A59 (Sawicki, 1987) possess uncleaved S but nevertheless undergo multiple cycles of replication. The requirement of S cleavage for membrane fusion may therefore depend not only on the virus strain and host cell membrane properties (Frana *et al.*, 1985) but also on the type of fusion in question: from within, from without or virus–endosome fusion. Concentrated MHV caused rapid fusion from without, the optimum pH being above 7 (Sturman *et al.*, 1985; Frana *et al.*, 1985). In contrast, agents that increased endosomal pH adversely affected MHV replication (Krzystyniak & Dupuy, 1984; Mizzen *et al.*, 1985). Clearly more work is required to define the pH requirements for S protein activity in fusion.

GENOMIC RNA

Genome organization

The genomic RNA of coronaviruses is the largest among RNA viruses, approximately 27 to 30 kb. The genome is organized into six or seven regions, each containing one or more open reading frames (ORFs) which are separated by junction sequences that contain the signal(s) for the transcription of multiple subgenomic mRNAs.

The organization of the coronaviral genome was elucidated on the basis of sequence relationships between the subgenomic RNAs and of *in vitro* translation studies using the individual mRNAs (reviewed by Siddell *et al.*, 1983; Siddell, 1983; Stern & Sefton, 1984; de Groot *et al.*, 1987*c*). In recent years much sequence data have been obtained; amongst others, the complete sequence of IBV genomic RNA, and the sequences of approximately 12 kb and 8·3 kb (extending from the 3′ end) of the MHV and TGEV genomic RNAs, respectively (Fig. 1; for references see legend).

From the sequence data several ORFs could be deduced. The 5′ two-thirds of the IBV genome (region F) encode non-structural protein(s), probably the replicase/transcriptase. The ORFs encoding the structural viral proteins have been identified; from their location a consensus gene order 5′ S–M–N 3′ can be inferred. The number and location of the other ORFs are different between IBV, MHV and TGEV (Fig. 1).

Translation strategy

Expression of the genome of coronaviruses involves the production of seven (BCV), six (MHV and TGEV) or five (IBV and FIPV) subgenomic mRNAs, which together with the virion RNA form a 3′ coterminal nested set (reviewed by Siddell *et al.*, 1983; Keck *et al.*, 1988*a*; de Groot *et al.*, 1987*c*; Rasschaert *et al.*, 1987). Except for the smallest subgenomic mRNA, the virus-specific intracellular mRNAs are polygenic, but only the unique region of each mRNA is translationally active. The unique regions of the mRNAs encoding the N, M and S proteins of IBV, TGEV and MHV (see references in the section on virion proteins), the non-structural 15K protein of MHV (Skinner & Siddell, 1985) and mRNAs 4 and 7 of TGEV (Rasschaert *et al.*, 1987) each comprise only one ORF. In contrast, two ORFs are present in the unique region of the genomic RNA (Boursnell *et al.*, 1987) and of mRNA B (Boursnell & Brown, 1984) of IBV, mRNAs 2 (Luytjes *et al.*, 1988) and 5 (Skinner *et al.*, 1985; Budzilowicz & Weiss, 1987) of MHV and in mRNA 3 of TGEV (Rasschaert *et al.*, 1987), while three ORFs have been identified at the unique 5′ end of IBV mRNA D (Boursnell *et al.*, 1985*b*). These mRNAs may therefore be functionally polycistronic.

ORFs F1 and F2 which are present in the IBV genome-size mRNA have the capacity to encode 400K and 350K polypeptides, respectively; they overlap by 42 nucleotides, but F2 starts in a different reading frame (Boursnell *et al.*, 1987). A cDNA fragment spanning the F1/F2 overlap was able to direct ribosomal frameshifting *in vitro* (Brierly *et al.*, 1987); it remains to be determined whether the same mechanism operates *in vivo*.

The third ORF of IBV mRNA D has been expressed, and an antiserum was raised against its product, D3, a polypeptide of M_r 12·4K (Smith *et al.*, 1987). Immunoprecipitation has shown that the D3 polypeptide occurs in IBV-infected cells, is unglycosylated, membrane-associated and co-fractionates with virions in sucrose gradients. A molar ratio of 1 : 2 : 2 : 7 : 11 has been estimated for D3 : S1 : S2 : N : M in virion preparations (A. R. Smith, M. E. G. Boursnell, M. M.

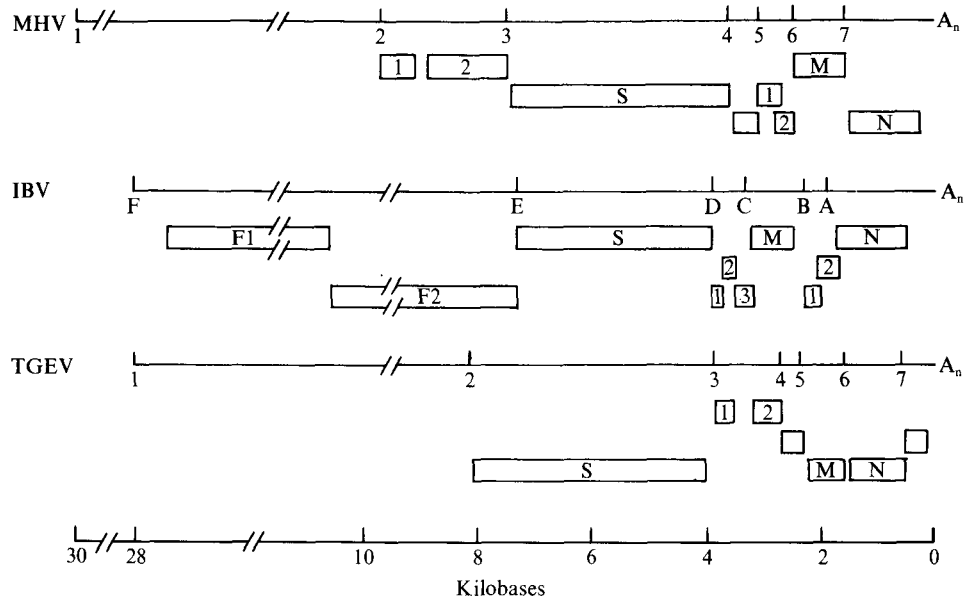

Fig. 1. The organization and expression of the MHV, IBV and TGEV genomes. The genomes are presented from 5′ (left) to 3′ (right) by horizontal lines; short vertical lines show the location of the conserved junction sequences that contains the signal(s) for the transcription of multiple subgenomic mRNAs. The numbers (1 to 7) and letters (A to F) below the junction regions indicate the nomenclature of the corresponding mRNAs for MHV or TGEV and IBV, respectively. The mRNAs (not indicated) form a 3′ coterminal nested set and their length can be deduced from the nucleotide kb map at the bottom of the figure. The symbol A_n indicates the poly(A) tract at the 3′ end of the genome. ORFs are shown as open rectangles, and the position on the genome of the ORFs encoding the replicase/transcriptase, the peplomer, the small integral membrane and nucleocapsid protein are indicated by F1 and F2, S, M and N respectively. The rectangles which include numbers (1 to 3) indicate that the unique regions (enclosed by the junction sequences) of the corresponding mRNAs contain more than one ORF (for details see text). For IBV the data have been compiled from Boursnell *et al.* (1987), for MHV from Armstrong (1983, 1984), Skinner & Siddell (1983, 1985), Skinner *et al.* (1985), Budzilowicz & Weiss (1987), Schmidt *et al.* (1987), Luytjes *et al.* (1987, 1988), P. J. Bredenbeek & W. Spaan (unpublished data), and for TGEV from Rasschaert *et al.* (1987).

Binns, T. D. K. Brown & S. C. Inglis, personal communication). Whether the other two small ORFs of mRNA D remain silent is unknown, but the unusual codon usage and the presence of weak translation initiation codons (Kozak, 1987) suggest that they are translationally inactive.

In vitro translation of the MHV genomic RNA has revealed the synthesis of a 250K polyprotein which is subsequently cleaved into a p28 and a p220 (28K and 220K respectively) protein (Denison & Perlman, 1986). The p28 protein could be labelled with *N*-formyl-[^{35}S]methionyl tRNA indicating its N-terminal location; the protein was also identified in infected cells (Denison & Perlman, 1987). Tryptic peptide maps of p28 synthesized *in vitro* by translation of genomic RNA and of RNA transcribed from a cDNA clone covering 1·1 kb of the 5′ end of the MHV genome were similar, confirming that the p28 protein is the N-terminal cleavage product of the putative MHV RNA polymerase (Soe *et al.*, 1987). Except for this cleavage and that of some spike protein precursors into two similarly sized subunits no processing of polyproteins has so far been observed in coronaviruses.

In vitro translation of MHV mRNA 2 gave rise to a 30K polypeptide (Siddell, 1983). Sequence analysis of the unique region of mRNA 2 of MHV strain A59 has revealed two ORFs (ORF 1 and ORF 2; Luytjes *et al.*, 1988). A 30K protein has been predicted on the basis of the ORF 1 sequence and was detected in MHV-infected cells by antiserum raised against an expression product of ORF 1 (P. Bredenbeek, A. Noten & W. Spaan, unpublished observations). The second reading frame has the potential to encode a 43K protein, but it is unlikely to be expressed

since in the first 109 triplets an AUG codon is lacking. Strikingly, the predicted amino acid sequence of ORF 2 has 30% homology with the haemagglutinin subunit 1 of influenza C virus. In strain JHM an ORF of which the predicted amino acid sequence is almost identical to ORF 2 of MHV-A59 and which includes an AUG translation start codon has been identified (E. Routledge & S. Siddell, personal communication).

An antiserum raised against the carboxy-terminal region of the mRNA 4 ORF product of MHV-JHM reacted specifically with a 15K protein synthesized in JHM-infected cells (Ebner *et al.*, 1988). *In vitro* translation of RNA transcribed from cDNA containing both ORFs of MHV-A59 mRNA 5 indicated that the downstream reading frame is preferentially translated (Budzilowicz & Weiss, 1987). This ORF was shown to be expressed in infected cells (Leibowitz *et al.*, 1988).

In conclusion, coronaviruses have developed more than one strategy to produce their proteins. Subgenomic mRNAs are synthesized to position internal genes at the unique 5′ end of an mRNA. Though most of these mRNAs are functionally monocistronic, several do contain more than one ORF within the unique region. In these mRNAs internal initiation occurs, e.g. at the AUG codon of the second ORF of MHV mRNA 5 and the third ORF of IBV mRNA D. Finally, the expression of the 5′ two-thirds of the IBV and MHV genomes involves ribosomal frameshifting and post-translational cleavage of a polyprotein, respectively.

Transcription and replication

The coronavirus genome serves as a template for the synthesis of a full-length negative-strand RNA (Lai *et al.*, 1982). This RNA was found exclusively in the viral replicative intermediates (RI) and the rate of its synthesis declined 5 to 6 h after infection (Sawicki & Sawicki, 1986). Continuing protein synthesis is a prerequisite for both negative and positive-strand RNA synthesis, although the former is comparatively more sensitive to cycloheximide (Sawicki & Sawicki, 1986). The negative-strand RNA serves as a template for transcription of genomic RNA and subgenomic mRNAs. The viral proteins involved in positive- and negative-strand RNA synthesis have not yet been identified. Six complementation groups are involved in RNA synthesis (Leibowitz *et al.*, 1982; B. A. M. van der Zeijst, personal communication). An RNA-dependent RNA polymerase activity has been detected in TGEV- and MHV-infected cells (Dennis & Brian, 1982; Brayton *et al.*, 1982; Mahy *et al.*, 1983). Brayton *et al.* (1982, 1984) have described two enzymically distinct RNA polymerase activities (acting early and late during infection); the early polymerase was involved in negative-strand RNA synthesis, whereas the late polymerase synthesized positive-stranded RNA. In contrast, only one polymerase activity was detected by Mahy *et al.* (1983) and Compton *et al.* (1987). Inhibition of *in vitro* RNA transcription by antibody to the N protein suggests a role in coronavirus replication (Compton *et al.*, 1987); whether this inhibition reflects a direct function of N in RNA synthesis or is the result of rapid RNA degradation is not known. Since the synthesis of mRNA 7 and of N protein is not inhibited in cells infected with defective interfering particles, in contrast to the other mRNAs and their translation products (Makino *et al.*, 1985), the N protein may well have a role in the replication of viral RNA.

The subgenomic mRNAs which form a 3′ coterminal nested set are synthesized in non-equimolar, but constant, amounts during the replication cycle (reviewed by Siddell *et al.*, 1983). RNase T1 fingerprinting of MHV genomic and subgenomic mRNAs revealed unique oligonucleotides that do not fit into the nested set structure, suggesting that these oligonucleotides are derived from a leader sequence which all mRNAs might share (Spaan *et al.*, 1982; Lai *et al.*, 1983). Direct evidence for a common leader sequence was obtained by nucleotide sequence analysis of the 5′ ends of IBV and MHV mRNAs (Spaan *et al.*, 1983; Lai *et al.*, 1984; Brown *et al.*, 1984). The IBV and MHV leader sequences of about 60 and 72 nucleotides, respectively, are transcribed from the 3′ end of the negative-stranded template (Spaan *et al.*, 1983; Lai *et al.*, 1984; Brown *et al.*, 1986; Bredenbeek *et al.*, 1987; Shieh *et al.*, 1987). However, the coronavirus mRNAs are not generated by splicing; replication occurs exclusively in the cytoplasm, and u.v. inactivation studies have shown that the mRNAs are transcribed independently (reviewed by Siddell *et al.*, 1983).

MHV-A59	AAUCUAAAC (c u)
BCV	AUCUAAAC (c)
FIPV	AACUAAAC
TGEV	AACUAAAC
IBV	CUUAACAA (g)

Fig. 2. Consensus sequences (in the sense orientation) of the reinitiation sites involved in the synthesis of subgenomic mRNAs of MHV, BCV, FIPV, TGEV and IBV. For details see text.

To explain the presence of a common leader sequence, a discontinuous transcription of coronavirus mRNAs was proposed (Baric *et al.*, 1983; Lai *et al.*, 1983; Spaan *et al.*, 1983). Several models for the required fusion of non-contiguous sequences have been suggested: jumping of the polymerase caused by the looping out of the negative-stranded template, post- or cotranscriptional ligation (or trans-splicing) and priming of the mRNA body transcription by the leader. Looping out of the template is unlikely to occur since loop structures have not been encountered in the RI isolated from MHV-infected cells (Baric *et al.*, 1983). Also, the leader RNAs can be exchanged between the mRNAs of co-infecting coronaviruses during a mixed infection (Makino *et al.*, 1986*a*).

Oligonucleotide fingerprints of non-denatured RI contained the oligonucleotides 10 and 19 (markers for the leader sequence), which would exclude the post-transcriptional trans-splicing model (Baric *et al.*, 1983). Both the leader-primed transcription and the cotranscriptional trans-splicing model are compatible with the findings that leader sequences can be freely exchanged (Makino *et al.*, 1986*a*) and that leader-containing transcripts of various sizes are present both in cells infected with wild-type virus and with an RNA$^-$ temperature-sensitive (*ts*) mutant (Baric *et al.*, 1985, 1987). However, the leader RNA transcripts which function in mRNA body transcription have not been identified. In both models the leader RNA is transcribed independently of the mRNA body and, after termination, is translocated to conserved sequences (reinitiation sites) on the negative-stranded template to serve as a primer for the mRNA body synthesis. Alternatively, it may be spliced cotranscriptionally to mRNA body transcripts.

Several reinitiation sites (i.e. junction or homology sequences) have been identified by comparing the sequences of the 5′ ends of the mRNAs and the corresponding regions on the genome (Spaan *et al.*, 1983; Bredenbeek *et al.*, 1987) or by S1 mapping (Brown & Boursnell, 1984; de Groot *et al.*, 1987*b*). The remaining reinitiation sites have been recognized by homology searches and by their position in the genome (Boursnell & Brown, 1984; Binns *et al.*, 1985; Boursnell *et al.*, 1985*b*; Bredenbeek *et al.*, 1986, 1987; Budzilowicz *et al.*, 1985; Rasschaert *et al.*, 1987; Jacobs *et al.*, 1987; Skinner & Siddell, 1985; Skinner *et al.*, 1985; Kapke & Brian, 1986; Lapps *et al.*, 1987; Schmidt *et al.*, 1987; Rasschaert & Laude, 1987; Luytjes *et al.*, 1987, 1988). The consensus sequences of the reinitiation sites of IBV, MHV, BCV, TGEV and FIPV are listed in Fig. 2. It is evident that there is also a limited extent of reinitiation site homology between these viruses. Additionally, in IBV, MHV and TGEV there is a conserved sequence of 10 nucleotides about 80 bases from the 3′ end of the genomic RNA (Kapke & Brian, 1986). This site may be critical for the synthesis of the negative-stranded template.

Sequence data obtained from IBV and MHV genomic RNA demonstrate a complementarity between the transcription reinitiation sites and the 3′ end of the leader transcript (Brown *et al.*, 1986; Bredenbeek *et al.*, 1986, 1987; Shieh *et al.*, 1987). This complementarity would allow base-pairing between the leader transcript and the internal transcription initiation sites. For the leader-primed transcription model it has been suggested that base-pairing is a prerequisite for the initiation of body RNA synthesis (Spaan *et al.*, 1983; Brown *et al.*, 1986; Shieh *et al.*, 1987; Bredenbeek *et al.*, 1986, 1987), and that the degree of complementarity between the free leader

and the different reinitiation sites may regulate the expression of the mRNAs (Budzilowicz *et al.*, 1985; Shieh *et al.*, 1987).

At the time it was formulated, the leader-primed transcription regulation model was based on limited sequence data. Subsequently, computer-assisted analyses of two completely sequenced 'intergenic' regions of IBV strain M42 (Brown *et al.*, 1986) and MHV-A59 (Bredenbeek *et al.*, 1987; Luytjes *et al.*, 1988) have shown that stability differences between intermolecular base-pairings involving the leader and the different reinitiation sites are not sufficient to explain the discrete mRNA moieties in infected cells (D. A. M. Konings, P. J. Bredenbeek, J. F. H. Noten, P. Hogeweg & W. J. M. Spaan, unpublished data).

RNA recombination

The RNA of coronaviruses can undergo recombination (Lai *et al.*, 1985). A high frequency of recombination was detected in cells infected with a *ts* mutant of MHV strain A59 and wild-type JHM virus at the non-permissive temperature (Makino *et al.*, 1986*b*). Multiple recombination sites have been detected at the 5′ end of MHV genomic RNA (Keck *et al.*, 1987). Although sequence data of the crossover sites are not available, the close relationship between the parental genomic RNAs suggests homologous recombination.

Results from studies of poliovirus recombination strongly suggest a copy-choice (or polymerase jumping) mechanism (Kirkegaard & Baltimore, 1986). The ability of the coronavirus polymerase to switch templates during the discontinuous mRNA transcription supports the copy-choice concept. The pool of leader RNA containing incomplete transcripts that have been detected in MHV-infected cells (Baric *et al.*, 1985, 1987) can be the result of a discontinuous and non-processive replication mechanism. These RNA intermediates may play an important role in homologous recombination but also in the generation of defective interfering RNAs (Keck *et al.*, 1987; Makino *et al.*, 1984, 1985).

Recombination is an important feature of coronavirus evolution. Not only has recombination between MHV strains been shown to occur *in vitro* (see above) but also in mouse brain (Keck *et al.*, 1988*b*). Comparison of IBV M protein sequences (Cavanagh & Davis, 1988) with the S1 sequences of the same strains (Niesters, 1987) suggests that some IBV field strains are recombinants. The high sequence divergence at the N terminus between the spike proteins of TGEV and FIPV could be the result of an RNA recombination event (Jacobs *et al.*, 1987). Sequence analysis of mRNA 2 of MHV-A59 has revealed a high similarity between the predicted amino acid sequence of a second ORF located at its 5′ end and the HA1 subunit of the influenza C spike protein. This could be the result of a non-homologous recombination event (Luytjes *et al.*, 1988).

These features and the differences among coronaviruses with respect to the number and order of the genes (de Groot *et al.*, 1987*c*), the length of the intergenic region between the ORFs encoding D3 and M of IBV (Cavanagh & Davis, 1988) and the presence of a translatable gp65 ORF all testify to the inconstant nature of coronavirus genomes. This is likely to be a major area of future research.

CONCLUDING REMARKS

Progress has been rapid in coronavirology during the last few years, and interesting insights have been gained for virology in general. Thus recombination may be a major mechanism responsible for biological variation within the family, including expansion of the host spectrum. 'New' viruses may arise when members of different families occupying the same ecological niche exchange genetic information; the similarity between a predicted amino acid sequence in mRNA 2 of MHV-A59 and the HA1 subunit of the influenza C spike protein could be the result of such a non-homologous recombination event (Luytjes *et al.*, 1988). Apart from evolutionary implications this insight must have consequences for the use of modified 'live' virus vaccines.

There are some areas to which future research will have to be directed. The role of spike protein cleavage and gp65 in the infection process is unclear. Apart from the molecular mechanisms of coronavirus pathogenesis the non-structural proteins, e.g. the polymerase, await

identification, while the details of mRNA transcription, RNA replication and recombination need to be elucidated.

We would like to thank Jacqueline Dumernit for typing the manuscript, and Drs Peter Bredenbeek, Peter Rottier, Willem Luytjes, Raoul de Groot and Harry Vennema for many helpful discussions. We are grateful to Drs K. V. Holmes, P. Palese, S. G. Siddell, B. A. M. van der Zeijst and S. Inglis for making data available to us prior to their publication and to Dr E. Chirnside for critical reading of the manuscript.

REFERENCES

ARMSTRONG, J., SMEEKENS, S. & ROTTIER, P. (1983). Sequence of the nucleocapsid gene from murine coronavirus MHV-A59. *Nucleic Acids Research* **11**, 883–891.

ARMSTRONG, J., NIEMANN, H., SMEEKENS, S., ROTTIER, P. & WARREN, G. (1984). Sequence and topology of a model intracellular membrane protein, E1 glycoprotein, from a coronavirus. *Nature, London* **308**, 751–752.

BARIC, R. S., STOHLMAN, S. A. & LAI, M. M. C. (1983). Characterization of replicative intermediate RNA of mouse hepatitis virus: presence of leader RNA sequences on nascent chains. *Journal of Virology* **48**, 633–640.

BARIC, R. S., STOHLMAN, S. A., RAZAVI, M. K. & LAI, M. M. C. (1985). Characterization of leader-related small RNAs in coronavirus-infected cells: further evidence for leader-primed mechanism of transcription. *Virus Research* **3**, 19–33.

BARIC, R. S., SHIEH, C. K., STOHLMAN, S. A. & LAI, M. M. C. (1987). Analysis of intracellular small RNAs of mouse hepatitis virus: evidence for discontinuous transcription. *Virology* **156**, 342–354.

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

BINNS, M. M., BOURSNELL, M. E. G., TOMLEY, F. M. & BROWN, T. D. K. (1986*a*). Nucleotide sequence encoding the membrane protein of the IBV strain 6/82. *Nucleic Acids Research* **14**, 5558.

BINNS, M. M. BOURSNELL, M. E. G., TOMLEY, F. M. & BROWN, T. D. K. (1986*b*). Comparison of the spike precursor sequences of coronavirus IBV strains M41 and 6/82 with that of IBV Beaudette. *Journal of General Virology* **67**, 2825–2831.

BOURSNELL, M. E. G. & BROWN, T. D. K. (1984). Sequencing of coronavirus IBV genomic RNA: a 195-base open reading frame encoded by mRNA B. *Gene* **29**, 87–92.

BOURSNELL, M. E. G., BROWN, T. D. K. & BINNS, M. M. (1984). Sequence of the membrane protein gene from avian coronavirus IBV. *Virus Research* **1**, 303–313.

BOURSNELL, M. E. G., BINNS, M. M., FOULDS, I. J. & BROWN, T. D. K. (1985*a*). Sequences of the nucleocapsid genes from two strains of avian infectious bronchitis virus. *Journal of General Virology* **66**, 573–580.

BOURSNELL, M. E. G., BINNS, M. M. & BROWN, T. D. K. (1985*b*). Sequencing of coronavirus IBV genomic RNA: three open reading frames in the 5′ 'unique' region of mRNA D. *Journal of General Virology* **66**, 2253–2258.

BOURSNELL, M. E. G., BROWN, T. D. K., FOULDS, I. J., GREEN, P. F., TOMLEY, F. M. & BINNS, M. M. (1987). Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *Journal of General Virology* **68**, 57–77.

BOYLE, J. F., WEISMILLER, D. G. & HOLMES, K. V. (1987). Genetic resistance to mouse hepatitis virus correlates with absence of virus-binding activity on target tissues. *Journal of Virology* **61**, 185–189.

BRAYTON, P. R., LAI, M. M. C., PATTON, D. F. & STOHLMAN, S. A. (1982). Characterization of two RNA polymerase activities induced by mouse hepatitis virus. *Journal of Virology* **42**, 847–853.

BRAYTON, P. R., STOHLMAN, S. A. & LAI, M. M. C. (1984). Further characterization of mouse hepatitis virus RNA-dependent RNA polymerases. *Virology* **133**, 197–201.

BREDENBEEK, P. J., NOTEN, J. F., LENSTRA, J. A., HORZINEK, M. C., VAN DER ZEIJST, B. A. M. & SPAAN, W. J. M. (1986). The nucleotide sequence of the extreme 5′ end of the avian coronavirus genome; implications for the discontinuous mRNA synthesis. *Nucleic Acids Research* **14**, 7806.

BREDENBEEK, P. J., CHARITE, J., NOTEN, J. F., LUYTJES, W., HORZINEK, M. C., VAN DER ZEIJST, B. A. M. & SPAAN, W. J. M. (1987). Sequences involved in the replication of coronaviruses. *Advances in Experimental Medicine and Biology* **218**, 65–72.

BRIERLEY, I., BOURSNELL, M. E. G., BINNS, M. M., BILIMORIA, B., BLOK, V. C., BROWN, T. D. K. & INGLIS, S. C. (1987). An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO Journal* **6**, 3779–3785.

BROWN, T. D. K. & BOURSNELL, M. E. G. (1984). Avian infectious bronchitis virus genomic RNA contains sequence homologies at the intergenic boundaries. *Virus Research* **1**, 15–24.

BROWN, T. D. K., BOURSNELL, M. E. G. & BINNS, M. M. (1984). A leader sequence is present on mRNA A of avian infectious bronchitis virus. *Journal of General Virology* **65**, 1437–1442.

BROWN, T. D. K., BOURSNELL, M. E. G., BINNS, M. M. & TOMLEY, F. M. (1986). Cloning and sequencing of 5′ terminal sequences from avian infectious bronchitis virus genomic RNA. *Journal of General Virology* **67**, 221–228.

BUDZILOWICZ, C. J. & WEISS, S. R. (1987). In vitro synthesis of two polypeptides from a nonstructural gene of coronavirus mouse hepatitis virus strain A59. *Virology* **157**, 509–515.

BUDZILOWICZ, C. J., WILCZYNSKI, S. P. & WEISS, S. R. (1985). Three intergenic regions of coronavirus mouse hepatitis virus strain A59 genome RNA contain a common nucleotide sequence that is homologous to the 3′ end of the viral mRNA leader sequence. *Journal of Virology* **53**, 834–840.

CAVANAGH, D. (1981). Structural polypeptides of coronavirus IBV. *Journal of General Virology* **53**, 93–103.

CAVANAGH, D. (1983*a*). Coronavirus IBV glycopolypeptides: size of their polypeptide moieties and nature of their oligosaccharides. *Journal of General Virology* **64**, 1187–1191.

CAVANAGH, D. (1983*b*). Coronavirus IBV: further evidence that the surface projections are associated with two glycopolypeptides. *Journal of General Virology* **64**, 1787–1791.

CAVANAGH, D. (1983*c*). Coronavirus IBV: structural characterization of the spike protein. *Journal of General Virology* **64**, 2577–2583.

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

CAVANAGH, D. & DAVIS, P. J. (1987). Coronavirus IBV: relationship among recent European isolates studied by limited proteolysis of the virion glycopolypeptides. *Avian Pathology* **16**, 1–13.

CAVANAGH, D. & DAVIS, P. J. (1988). Evolution of avian coronavirus IBV: sequence of the matrix glycoprotein gene and intergenic region of several serotypes. *Journal of General Virology* **69**, 621–629.

CAVANAGH, D., DAVIS, P. J. & PAPPIN, D. J. C. (1986*a*). Coronavirus IBV glycopolypeptides: locational studies using proteases and saponin, a membrane permeabilizer. *Virus Research* **4**, 145–156.

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

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

COMPTON, S. R., ROGERS, D. B., HOLMES, K. V., FERTSCH, D., REMENICK, J. & McGOWAN, J. J. (1987). In vitro replication of mouse hepatitis virus strain A59. *Journal of Virology* **61**, 1814–1820.

DE GROOT, R. J., LUYTJES, W., HORZINEK, M. C., VAN DER ZEIJST, B. A. M., SPAAN, W. J. M. & LENSTRA, J. A. (1987*a*). Evidence for a coiled-coil structure in the spike proteins of coronaviruses. *Journal of Molecular Biology* **196**, 963–966.

DE GROOT, R. J., MADURO, J., LENSTRA, J. A., HORZINEK, M. C., VAN DER ZEIJST, B. A. M. & SPAAN, W. J. M. (1987*b*). cDNA cloning and sequence analysis of the gene encoding the peplomer protein of feline infectious peritonitis virus. *Journal of General Virology* **68**, 2639–2646.

DE GROOT, R. J., TER HAAR, R. J., HORZINEK, M. C. & VAN DER ZEIJST, B .A. M. (1987*c*). Intracellular RNAs of the feline infectious peritonitis coronavirus strain 79-1146. *Journal of General Virology* **68**, 995–1002.

DENISON, M. R. & PERLMAN, S. (1986). Translation and processing of mouse hepatitis virus virion RNA in a cell-free system. *Journal of Virology* **60**, 12–18.

DENISON, M. & PERLMAN, S. (1987). Identification of putative polymerase gene product in cells infected with murine coronavirus A59. *Virology* **157**, 565–568.

DENNIS, D. E. & BRIAN, D. A. (1982). RNA-dependent RNA polymerase activity in coronavirus-infected cells. *Journal of Virology* **42**, 153–164.

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

EBNER, D., RAABE, T. & SIDDELL, S. G. (1988). Identification of the coronavirus MHV-JHM mRNA 4 product. *Journal of General Virology* **69**, 1041–1050.

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

GARWES, D. J. & REYNOLDS, D. J. (1981). The polypeptide structure of canine coronavirus and its relationship to porcine transmissible gastroenteritis virus. *Journal of General Virology* **52**, 153–157.

GARWES, D. J., BOUNTIFF, L., MILLSON, G. C. & ELLEMAN, C. J. (1984). Defective replication of porcine transmissible gastroenteritis virus in a continuous cell line. *Advances in Experimental Medicine and Biology* **173**, 79–93.

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

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

HOLMES, K. V., DOLLER, E. W. & STURMAN, L. S. (1981). Tunicamycin-resistant glycosylation of coronavirus glycoprotein: demonstration of a novel type of viral glycoprotein. *Virology* **115**, 334–344.

HORZINEK, M. C., LUTZ, H. & PEDERSEN, N. C. (1982). Antigenic relationships among homologous structural polypeptides of porcine, feline, and canine coronaviruses. *Infection and Immunity* **37**, 1148–1155.

JACOBS, L., VAN DER ZEIJST, B. A. M. & HORZINEK, M. C. (1986). Characterization and translation of transmissible gastroenteritis virus mRNAs. *Journal of Virology* **57**, 1010–1015.

JACOBS, L., DE GROOT, R., VAN DER ZEIJST, B. A. M., HORZINEK, M. C. & SPAAN, W. (1987). The nucleotide sequence of the peplomer gene of porcine transmissible gastroenteritis virus (TGEV): comparison with the sequence of the peplomer protein of feline infectious peritonitis virus (FIPV). *Virus Research* **8**, 363–371.

KAPKE, P. A. & BRIAN, D. A. (1986). Sequence analysis of the porcine transmissible gastroenteritis coronavirus nucleocapsid protein gene. *Virology* **151**, 41–49.

KECK, J. G., STOHLMAN, S. A., SOE, L. H., MAKINO, S. & LAI, M. M. C. (1987). Multiple recombination sites at the 5′-end of murine coronavirus RNA. *Virology* **156**, 331–341.

KECK, J. G., HOGUE, B. G., BRIAN, D. A. & LAI, M. M. C. (1988*a*). Temporal regulation of bovine coronavirus RNA synthesis. *Virus Research* **9**, 343–356.

KECK, J. G., MATSUSHIMA, G. K., MAKINO, S., FLEMING, J. O., VANNIER, D. M., STOHLMAN, S. A. & LAI, M. M. C. (1988*b*). In vivo RNA-RNA recombination of coronavirus in mouse brain. *Journal of Virology* **62**, 1810–1813.

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

KIRKEGAARD, K. & BALTIMORE, D. (1986). The mechanism of RNA recombination in poliovirus. *Cell* **47**, 433–443.

KOZAK, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research* **15**, 8125–8133.

KRZYSTYNIAK, K. & DUPUY, J. M. (1984). Entry of mouse hepatitis virus 3 into cells. *Journal of General Virology* **65**, 227–231.

LAI, M. M. C. (1986). Coronavirus leader-RNA-primed transcription: an alternative mechanism to RNA splicing. *BioEssays* **5**, 257–260.

LAI, M. M. C., PATTON, C. D. & STOHLMAN, S. A. (1982). Replication of mouse hepatitis virus: negative-stranded RNA and replicative form RNA are of genome length. *Journal of Virology* **44**, 487–492.

LAI, M. M. C., PATTON, C. D., BARIC, R. S. & STOHLMAN, S. A. (1983). Presence of leader sequences in the mRNA of mouse hepatitis virus. *Journal of Virology* **46**, 1027–1033.

LAI, M. M. C., BARIC, R. S., BRAYTON, P. R. & STOHLMAN, S. A. (1984). Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. *Proceedings of the National Academy of Sciences, U.S.A.* **81**, 3626–3630.

LAI, M. M. C., BARIC, R. S., MAKINO, S., KECK, J. G., EGBERT, J., LEIBOWITZ, J. L. & STOHLMAN, S. A. (1985). Recombination between nonsegmented RNA genomes of murine coronaviruses. *Journal of Virology* **56**, 449–456.

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

LAUDE, H., RASSCHAERT, D. & HUET, J.-C. (1987). Sequence and N-terminal processing of the transmembrane protein E1 of the coronavirus transmissible gastroenteritis virus. *Journal of General Virology* **68**, 1687–1693.

LEIBOWITZ, J. L., DE VRIES, J. R. & HASPEL, M. V. (1982). Genetic analysis of murine hepatitis virus strain JHM. *Journal of Virology* **42**, 1080–1087.

LEIBOWITZ, J. L., PERLMAN, S., WEINSTOCK, G., DE VRIES, J. R., BUDZILOWICZ, C., WEISSEMANN, J. M. & WEISS, S. R. (1988). Detection of a murine coronavirus nonstructural protein encoded in a downstream open reading frame. *Virology* **164**, 156–164.

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

LUYTJES, W., BREDENBEEK, P. J., NOTEN, A. F., HORZINEK, M. C. & SPAAN, W. J. (1988). Sequence of mouse hepatitis virus A59 mRNA2: indications for RNA-recombination between coronaviruses and influenza C virus. *Virology* (in press).

MACHAMER, C. E. & ROSE, J. K. (1987). A specific transmembrane domain of a coronavirus E1 glycoprotein is required for its retention in the Golgi region. *Journal of Cell Biology* **105**, 1205–1214.

MAHY, B. W. J., SIDDELL, S., WEGE, H. & TER MEULEN, V. (1983). RNA-dependent RNA polymerase activity in murine coronavirus-infected cells. *Journal of General Virology* **64**, 103–111.

MAKINO, S., TAGUCHI, F., HAYAMI, M. & FUJIWARA, K. (1983). Characterization of small plaque mutants of mouse hepatitis virus, JHM strain. *Microbiology and Immunology* **27**, 445–454.

MAKINO, S., TAGUCHI, F. & FUJIWARA, K. (1984). Defective interfering particles of mouse hepatitis virus. *Virology* **133**, 9–17.

MAKINO, S., FUJIOKA, N. & FUJIWARA, K. (1985). Structure of the intracellular defective viral RNAs of defective interfering particles of mouse hepatitis virus. *Journal of Virology* **54**, 329–336.

MAKINO, S., STOHLMAN, S. A. & LAI, M. M. C. (1986*a*). Leader sequences of murine coronavirus mRNAs can be freely reassorted: evidence for the role of free leader RNA in transcription. *Proceedings of the National Academy of Sciences, U.S.A.* **83**, 4204–4208.

MAKINO, S., KECK, J. G., STOHLMAN, S. A. & LAI, M. M. C. (1986*b*). High-frequency RNA recombination of murine coronaviruses. *Journal of Virology* **57**, 729–737.

MAYER, T., TAMURA, T., FALK, M. & NEIMANN, H. (1988). Membrane integration and intracellular transport of the coronavirus glycoprotein E1, a class III membrane glycoprotein. *Journal of Biological Chemistry* (in press).

MIZZEN, L., HILTON, A., CHELEY, S. & ANDERSON, R. (1985). Attenuation of murine coronavirus infection by ammonium chloride. *Virology* **142**, 378–388.

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

NIEMANN, H., GEYER, R., KLENK, H. D., LINDER, D., STIRM, S. & WIRTH, M. (1984). The carbohydrates of mouse hepatitis virus (MHV) A59: structures of the *O*-glycosidically linked oligosaccharides of glycoprotein E1. *EMBO Journal* **3**, 665–670.

NIESTERS, H. G. (1987). *Molecular epidemiology of infectious bronchitis virus*. Ph.D. thesis, University of Utrecht.

NIESTERS, H. G., LENSTRA, J. A., SPAAN, W. J. M., ZIJDERVELD, A. J., BLEUMINK-PLUYM, N. M., HONG, F., VAN SCHARRENBURG, G. J., HORZINEK, M. C. & VAN DER ZEIJST, B. A. M. (1986). The peplomer protein sequence of the M41 strain of coronavirus IBV and its comparison with Beaudette strains. *Virus Research* **5**, 253–263.

PFLEIDERER, M., SKINNER, M. A. & SIDDELL, S. G. (1986). Coronavirus MHV-JHM: nucleotide sequence of the mRNA that encodes the membrane protein. *Nucleic Acids Research* **14**, 6338.

RASSCHAERT, D. & LAUDE, H. (1987). The predicted primary structure of the peplomer protein E2 of the porcine coronavirus transmissible gastroenteritis virus. *Journal of General Virology* **68**, 1883–1890.

RASSCHAERT, D., GELFI, J. & LAUDE, H. (1987). Enteric coronavirus TGEV: partial sequence of the genomic RNA, its organization and expression. *Biochimie* **69**, 591–600.

RESTA, S., LUBY, J. P., ROSENFELD, C. R. & SIEGEL, J. D. (1985). Isolation and propagation of a human enteric coronavirus. *Science* **229**, 978–981.

ROTTIER, P. J. & ROSE, J. K. (1987). Coronavirus E1 glycoprotein expressed from cloned cDNA localizes in the Golgi region. *Journal of Virology* **61**, 2042–2045.

ROTTIER, P. J., HORZINEK, M. C. & VAN DER ZEIJST, B. A. M. (1981). Viral protein synthesis in mouse hepatitis virus strain A59-infected cells: effect of tunicamycin. *Journal of Virology* **40**, 350–357.

ROTTIER, P., BRANDENBURG, D., ARMSTRONG, J., VAN DER ZEIJST, B. & WARREN, G. (1984). Assembly in vitro of a spanning membrane protein of the endoplasmic reticulum: the E1 glycoprotein of coronavirus mouse hepatitis virus A59. *Proceedings of the National Academy of Sciences, U.S.A.* **81**, 1421–1425.

ROTTIER, P., ARMSTRONG, J. & MEYER, D. I. (1985). Signal recognition particle-dependent insertion of coronavirus E1, an intracellular membrane glycoprotein. *Journal of Biological Chemistry* **260**, 4648–4652.

ROTTIER, P. J., WELLING, G. W., WELLING-WESTER, S., NIESTERS, H. G., LENSTRA, J. A. & VAN DER ZEIJST, B. A. M. (1986). Predicted membrane topology of the coronavirus protein E1. *Biochemistry* **25**, 1335–1339.

SAWICKI, S. G. (1987). Characterization of a small plaque mutant of the A59 strain of mouse hepatitis virus defective in cell fusion. *Advances in Experimental Medicine and Biology* **218**, 169–174.

SAWICKI, S. G. & SAWICKI, D. L. (1986). Coronavirus minus-strand RNA synthesis and effect of cycloheximide on coronavirus RNA synthesis. *Journal of Virology* **57**, 328–334.

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

SHIEH, C. K., SOE, L. H., MAKINO, S., CHANG, M. F., STOHLMAN, S. A. & LAI, M. M. C. (1987). The 5′-end sequence of the murine coronavirus genome: implications for multiple fusion sites in leader-primed transcription. *Virology* **156**, 321–330.

SIDDELL, S. G. (1982). Coronavirus JHM: tryptic peptide fingerprinting of virion proteins and intracellular polypeptides. *Journal of General Virology* **62**, 259–269.

SIDDELL, S. (1983). Coronavirus JHM: coding assignments of subgenomic mRNAs. *Journal of General Virology* **64**, 113–125.

SIDDELL, S., WEGE, H. & TER MEULEN, V. (1983). The biology of coronaviruses. *Journal of General Virology* **64**, 761–776.

SKINNER, M. A. & SIDDELL, S. G. (1983). Coronavirus JHM: nucleotide sequence of the mRNA that encodes nucleocapsid protein. *Nucleic Acids Research* **11**, 5045–5054.

SKINNER, M. A. & SIDDELL, S. G. (1985). Coding sequence of coronavirus MHV-JHM mRNA 4. *Journal of General Virology* **66**, 593–596.

SKINNER, M. A., EBNER, D. & SIDDELL, S. G. (1985). Coronavirus MHV-JHM mRNA 5 has a sequence arrangement which potentially allows translation of a second, downstream open reading frame. *Journal of General Virology* **66**, 581–592.

SMITH, A. R., BOURSNELL, M. E. G., BINNS, M. M., BROWN, T. D. K. & INGLIS, S. C. (1987). Identification of a new gene product encoded by mRNA D of infectious bronchitis virus. *Advances in Experimental Medicine and Biology* **218**, 47–54.

SOE, L. H., SHIEH, C. K., BAKER, S. C., CHANG, M. F. & LAI, M. M. C. (1987). Sequence and translation of the murine coronavirus 5′-end genomic RNA reveals the N-terminal structure of the putative RNA polymerase. *Journal of Virology* **61**, 3968–3976.

SPAAN, W. J., ROTTIER, P. J., HORZINEK, M. C. & VAN DER ZEIJST, B. A. M. (1982). Sequence relationships between the genome and the intracellular RNA species 1, 3, 6, and 7 of mouse hepatitis virus strain A59. *Journal of Virology* **42**, 432–439.

SPAAN, W., DELIUS, H., SKINNER, M., ARMSTRONG, J., ROTTIER, P., SMEEKENS, S., VAN DER ZEIJST, B. A. M. & SIDDELL, S. G. (1983). Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO Journal* **2**, 1839–1844.

STERN, D. F. & SEFTON, B. M. (1982*a*). Coronavirus proteins: structure and function of the oligosaccharides of the avian infectious bronchitis virus glycoproteins. *Journal of Virology* **44**, 804–812.

STERN, D. F. & SEFTON, B. M. (1982*b*). Coronavirus proteins: biogenesis of avian infectious bronchitis virus virion proteins. *Journal of Virology* **44**, 794–803.

STERN, D. F. & SEFTON, B. M. (1984). Coronavirus multiplication: locations of genes for virion proteins on the avian infectious bronchitis virus genome. *Journal of Virology* **50**, 22–29.

STURMAN, L. S. & HOLMES, K. V. (1983). The molecular biology of coronaviruses. *Advances in Virus Research* **28**, 35–112.

STURMAN, L. & HOLMES, K. (1985). The novel glycoproteins of coronaviruses. *Trends in Biochemical Sciences* **10**, 17–20.

STURMAN, L. S., HOLMES, K. V. & BEHNKE, J. (1980). Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. *Journal of Virology* **33**, 449–462.

STURMAN, L. S., RICARD, C. S. & HOLMES, K. V. (1985). Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: activation of cell-fusing activity of virions by trypsin and separation of two different 90K cleavage fragments. *Journal of Virology* **56**, 904–911.

SUGIYAMA, K., ISHIKAWA, R. & FUKUHARA, N. (1986). Structural polypeptides of the murine coronavirus DVIM. *Archives of Virology* **89**, 245–254.

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. *Journal of Virology* **54**, 429–435.

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

VAN BERLO, M. F., VAN DEN BRINK, W. J., HORZINEK, M. C. & VAN DER ZEIJST, B. A. M. (1987). Fatty acid acylation of viral proteins in murine hepatitis virus-infected cells. Brief report. *Archives of Virology* **95**, 123–128.

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

WEGE, H., DÖRRIES, R. & WEGE, H. (1984). Hybridoma antibodies to the murine coronavirus JHM: characterization of epitopes on the peplomer protein (E2). *Journal of General Virology* **65**, 1931–1942.