

THE MOLECULAR BIOLOGY OF CORONAVIRUSES

Michael M. C. Lai* and David Cavanagh†

*Department of Molecular Microbiology and Immunology
Howard Hughes Medical Institute
University of Southern California School of Medicine
Los Angeles, California 90033-1054

†Institute for Animal Health
Compton Laboratory
Compton, Newbury, Berkshire, RG20 7NN, United Kingdom

- I. Introduction
- II. Taxonomy and the Essential Characteristics of *Coronaviridae*
- III. Structure of Virions
 - A. Virion Morphology
 - B. Structural Proteins
 - C. RNA Genome
- IV. Nonstructural Proteins
 - A. The Polymerase
 - B. Other Nonstructural (ns) Proteins
- V. Replication Cycle
 - A. Viral Host Ranges and Metabolic Requirements of Viral Replication
 - B. Virus Attachment
 - C. Penetration and Uncoating
 - D. Primary Translation
 - E. Transcription of Viral mRNAs
 - F. Replication of Viral Genomic RNA
 - G. Translation of Viral Proteins
 - H. Virus Assembly and Release
- VI. Genetics of Coronaviruses
 - A. Natural Viral Variants and Mutants
 - B. Complementation
 - C. Phenotypic Mixing and Pseudotype Virus Formation
 - D. RNA Recombination
 - E. Defective-Interfering (DI) RNAs
- VII. Perspectives
- References

I. INTRODUCTION

The word “big” is, perhaps, one of the first to come to mind when considering coronaviruses. The nature of the coronavirus genome—nonsegmented, single-stranded, positive-sense RNA—is not remarkable, but its size, 27 to 32 kb, surely is when compared with other RNA

viruses. The coronavirus polymerase gene alone (20–22 kb) is about the same size as the whole of the picornavirus (~8 kb) and vesicular stomatitis virus (~11 kb) genomes added together. The gene encoding the large surface glycoprotein is up to 4.4 kb, encoding an imposing trimeric, highly glycosylated protein. This soars some 20 nm above the virion envelope, giving the virus the appearance—with a little imagination—of a crown or coronet (Latin *corona*, hence the name of the genus).

Coronaviruses are responsible for a number of economically important diseases. Avian infectious bronchitis virus (IBV) was the first coronavirus to be isolated, from the domestic fowl, and propagated in the 1930s. In addition to respiratory disease, which can predispose chickens to possibly lethal secondary bacterial infections, some strains also cause nephritis (King and Cavanagh, 1991; Cook and Mockett, 1995). Porcine transmissible gastroenteritis virus (TGEV) causes devastating disease in newborn pigs, with mortality often approaching 100% (Enjuanes and van der Zeijst, 1995). Intriguingly, there are also naturally occurring mutants [i.e., porcine respiratory coronavirus (PRCV)] of TGEV which cause only mild respiratory disease and no enteritis. Several other coronaviruses also cause enteritis: bovine coronavirus (BCV), turkey coronavirus (TCV; bluecomb virus), feline coronavirus (FCV), canine coronavirus (CCV) and porcine epidemic diarrhea virus (PEDV). FCV may also cause feline infectious peritonitis. An FCV has been isolated from a cheetah and BCVs from wild sambar deer and waterbuck (Tsunemitsu *et al.*, 1995). These BCVs caused enteritis when inoculated into domestic calves. Humans are known to suffer from two very different coronaviruses, human coronavirus (HCV) OC43 and HCV 229E, both of which are a cause of the common cold. There is evidence for the presence of coronaviruses in tissues taken from multiple sclerosis (MS) patients (reviewed by Cavanagh and Macnaughton, 1995). This inflammatory, demyelinating neurological disease is associated with autoreactive T lymphocytes sensitized to myelin components of the central nervous system. Recently, Talbot and colleagues (1996) have demonstrated that many CD4⁺ T-cell lines derived from MS patients showed a human leukocyte antigen-(HLA)-DR-restricted, cross-reactive pattern of antigen activation after *in vitro* selection of either myelin basic protein or HCV-229E proteins, suggesting that molecular mimicry between HCV and myelin may be an immunopathological mechanism in MS. Other coronaviruses [some strains of murine hepatitis virus (MHV) and porcine hemagglutinating encephalomyelitis virus (HEV)] are well-known causes of neurological diseases, and MHV has been studied for many years in this context (Dales and Anderson, 1995), although many MHV strains cause primarily hepatitis.

The 1970s and early 1980s was the period in which coronavirus virion proteins and nested-set arrangements of mRNAs were identified and the discontinuous nature of coronavirus transcription was initially demonstrated. The first published sequence of a coronavirus gene appeared in 1983, starting an era in which the whole of the genomes of four coronaviruses were cloned — in pieces — and sequenced. This decade has seen the manipulation of these clones, and of complementary DNAs (cDNAs) of defective-interfering (DI) RNAs, to study coronavirus RNA replication, transcription, recombination, processing and transport of proteins, virion assembly, identification of cell receptors for coronaviruses, and processing of the polymerase.

This review is largely concerned with these areas. Some topics are notable by their absence, space not permitting their inclusion. For example, the elucidation of the molecular basis of the antigenic properties of the large surface (spike) glycoprotein and its role in tissue tropism has been omitted. For these topics and all others both within and without the compass of our review for which a concurrently comprehensive and in-depth treatise is desired, the reader is referred to the book edited by Siddell (1995a). Individual chapters in that book will be referenced at the appropriate places in this review.

II. TAXONOMY AND THE ESSENTIAL CHARACTERISTICS OF *Coronaviridae*

All coronaviruses belong to one genus, *Coronavirus*, within the family *Coronaviridae* (Cavanagh *et al.*, 1994, 1995). Initially, serological analysis was used to differentiate coronavirus species and showed that they could be divided into four antigenic groups (Holmes, 1990). The species and group divisions were subsequently refined by monoclonal antibody analysis and nucleotide sequencing, which revealed the close relatedness between TCV and BCV, resulting in the current classification of three antigenic groups (Table I). The same groupings emerge regardless of which structural protein sequences are compared (Siddell, 1995b). Within group 1, TGEV, FCV, and CCV are particularly closely related, all the members of group 2 being tightly clustered. The sole member of group 3, IBV, not only differs extensively from all other coronaviruses but also exhibits extensive variation within the species.

The *Coronaviridae* had remained a monogeneric family for a quarter of a century, until an accumulation of observations which showed that many of the features thought to be characteristic of the *Coronaviridae* applied equally well to the genus *Torovirus*, which had not been officially assigned to a family (Figs. 1 and 2, Table II). Therefore, in

TABLE I
SPECIES WITHIN THE GENERA *CORONAVIRUS* AND *TOROVIRUS*

	Species	Antigenic group	Contain HE gene
<i>Coronavirus</i>			
IBV	Avian infectious bronchitis virus	3	-
MHV	Murine hepatitis virus	2	+
BCV	Bovine coronavirus	2	+
HCV (OC43)	Human coronavirus OC43	2	+
HEV	Porcine hemagglutinating encephalomyelitis virus	2	+
TCV	Turkey coronavirus	2	+
TGEV	Porcine transmissible gastroenteritis virus	1	-
FCV	Feline coronavirus and feline infectious peritonitis virus (FIPV)	1	-
CCV	Canine coronavirus	1	-
HCV (229E)	Human coronavirus 229E	1	-
PEDV	Porcine epidemic diarrhea virus	1	-
<i>Torovirus</i>			
BEV	Berne virus (equine)		+
BRV	Breda virus (bovine)		NK ^a

^a NK, not known.

1993, the International Committee for the Taxonomy of Viruses (ICTV) formally expanded the *Coronaviridae* to include *Torovirus* (Cavanagh *et al.*, 1994, 1995).

The bringing together of *Coronavirus* and *Torovirus* was not the end of the taxonomic story; another family, *Arteriviridae*, shared important characteristics in relation to the genome, structure, and strategies of transcription and translation (Table II) (Plagemann and Moennig, 1992; Snijder and Spaan, 1995). However, the distinct morphology of the arteriviruses (Fig. 1), and their underlying differences from the coronaviruses in the size of the genome (Fig. 2) and structural proteins (Table II), precluded their inclusion in the *Coronaviridae*. The common features uniting the two families (Table II) are at the heart of a proposal that an order be created to contain *Coronaviridae* and *Arteriviridae* to reflect their common features and, probably, their evolutionary relationships. The name Nidovirales, from the Latin *nidus*, meaning nest, has been designated for the order, as all members produce mRNAs in an extensive nested-set arrangement.

The remainder of this review is restricted largely to the coronaviruses.

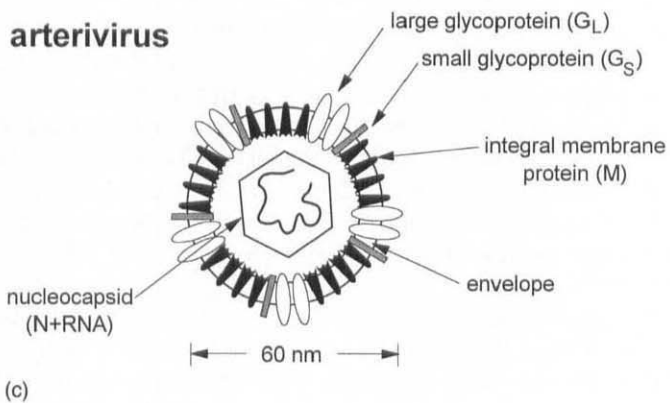
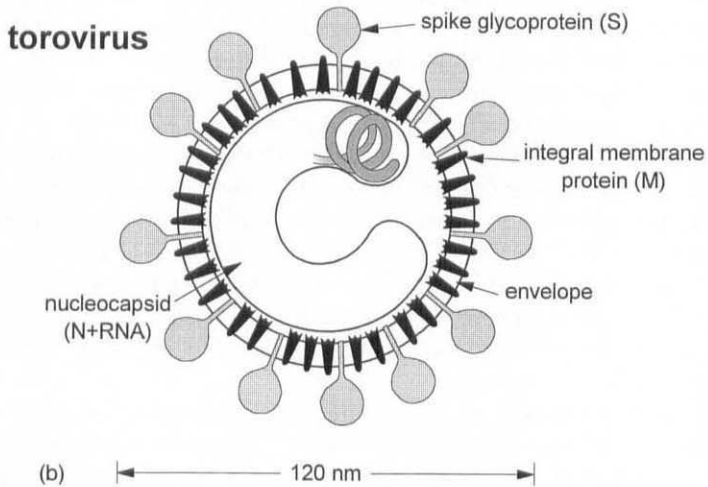
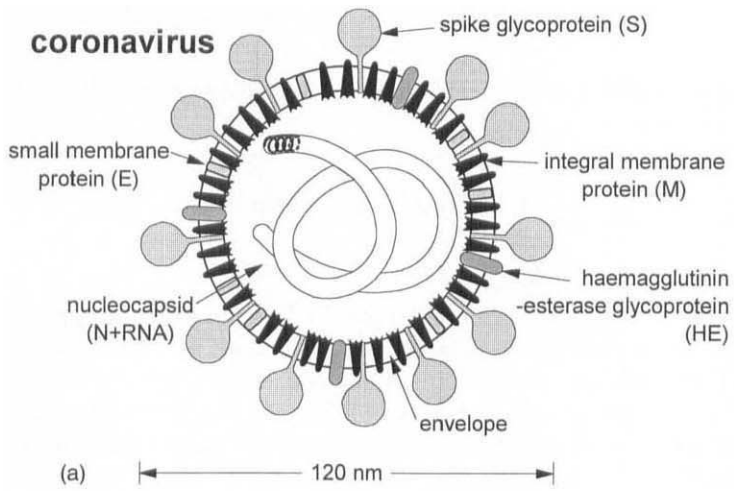
III. STRUCTURE OF VIRIONS

A. *Virion Morphology*

Coronaviruses are enveloped, more or less spherical, approximately 120 nm in diameter, with a prominent fringe of 20-nm-long, petal-shaped surface projections (spikes) composed of a heavily glycosylated type I glycoprotein, spike protein (S) (Fig. 1). A subset of the coronaviruses (Table I) has an additional layer of short spikes (Caul and Egglestone, 1977; Dea and Tijssen, 1988), which consist of hemagglutinin-esterase (HE) protein, also a type I glycoprotein. These small spikes are not essential for viral infectivity. Both the large and small spikes are anchored in the envelope, which is a lipid bilayer formed by virus budding from intracellular membranes. The envelope is associated with, in addition to the S and HE proteins, a smaller type III integral membrane protein (M), which spans the envelope three times. An even smaller protein [envelope (E) or small membrane (sM) protein] has recently been shown to be an integral membrane protein of the viral envelope. Inside the envelope is a ribonucleoprotein (RNP) core, which comprises the RNA genome and a single species of nucleocapsid protein N. Electron microscopic observation of viral RNP showed a long helix of 14 to 16 nm (Macnaughton *et al.*, 1978; Sturman and Holmes, 1983).

A very recent study of intact and detergent-treated TGEV virions (Risco *et al.*, 1996) by *negative-staining*, *ultrathin sectioning*, *freeze-fracture*, *immunogold mapping* and *cryoelectron microscopy* showed a surprising new feature of coronavirus particles, namely, a spherical, probably icosahedral, core inside the virion (Fig. 3). These internal cores comprise not only the N protein and RNA but also the M protein, M being the major core shell component. Disruption of the cores released helical nucleocapsids. The presence of an icosahedral core in the coronavirus virion had heretofore been unsuspected. This core structure was also detected with MHV virion (Risco *et al.*, 1996). This surprising new finding gives us cause to reconsider our view of coronavirus architecture. Thus, the precise structure of the core and RNP inside the virion is not certain.

Toroviruses and coronaviruses have a similar morphology and virion composition (Fig. 1, Table II) but are distinguishable in a number of ways (Table II) (Weiss and Horzinek, 1987; Snijder and Horzinek, 1993,



1995; Koopmans and Horzinek, 1994), necessitating their inclusion in separate genera. The morphology of the arteriviruses is substantially different from that of coronaviruses and toroviruses, particularly in having an icosahedral RNP core (Fig. 1) (Snijder and Spaan, 1995); hence, a separate family is maintained for arteriviruses. However, the recent discovery of the icosahedral core for coronavirus (Risco *et al.*, 1996) may have blurred this distinction.

B. Structural Proteins

1. Spike Glycoprotein (S)

The S glycoprotein is the outermost component of the virion, and is responsible for the attachment of the virus to cells (Collins *et al.*, 1982; Godet *et al.*, 1994; Kubo *et al.*, 1994) and for instigating the fusion of the virus envelope with cell membranes. It is the primary target for the host's immune responses; neutralizing antibodies are induced mainly by S (Collins *et al.*, 1982), and immunization in animals with S alone can induce protection from some coronaviruses (Ignjatovic and Galli, 1994; Torres *et al.*, 1995). Within a coronavirus species, sequence variation is usually exhibited more by S than by any other structural proteins; the variation of the S protein sequence probably confers a selective advantage in immune animals. These and other aspects have recently been reviewed in detail (Cavanagh *et al.*, 1995).

The S protein is large, ranging from some 1160 (IBV) to 1452 amino acids (FCV). There are many potential N-linked glycosylation sites (21 to 35), most of which have glycans attached. The S preproprotein has a N-terminal signal sequence and a membrane-anchoring sequence near the C terminus (Fig. 4). The S protein may be cleaved into S1 and S2 subunits; the extent of its cleavage varies greatly among the species (Cavanagh, 1995). A high proportion, up to 100%, of the S protein is cleaved in some coronaviruses (IBV, MHV, BCV, TCV, PEDV) (Cavanagh, 1983a); none is cleaved in others (TGEV, FCV, CCV) (Garwes and Pocock, 1975); and very little of the S protein of HCV-229E and HCV-OC43 is cleaved, although the S of OC43 is completely

FIG 1. Models of the virions of a coronavirus, a torovirus, and an arterivirus. The HE protein is present only in antigenic group 2 coronaviruses (see Table I). Reproduced with permission from Cavanagh *et al.* (1994).

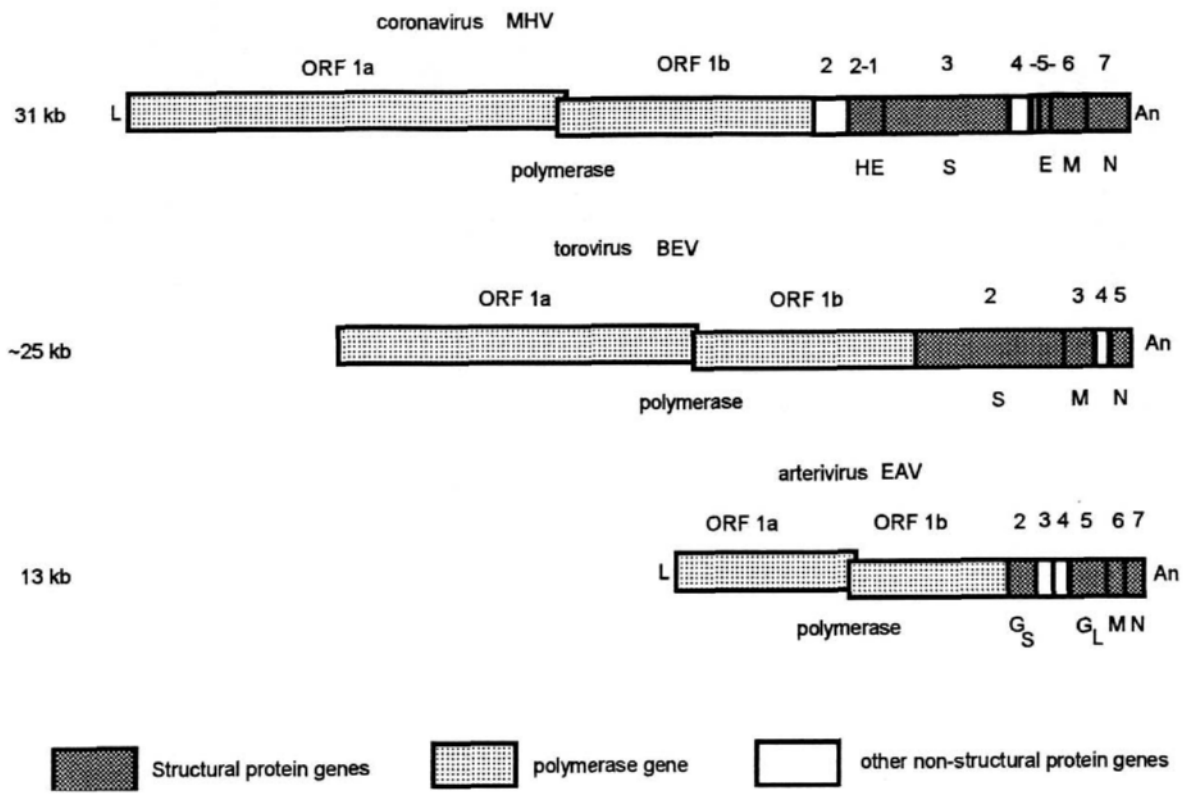


TABLE II
 FEATURES OF CORONAVIRUSES, TOROVIRUSES, AND ARTERIVIRUSES

Feature	Coronavirus	Torovirus	Arterivirus
Enveloped	+	+	+
Linear positive-sense ssRNA with poly(A) tail ^a	+	+	+
Genome organization ^a			
5' polymerase gene-structural protein genes 3'	+	+	+
3' co-terminal nested set of ≥ 4 subgenomic mRNAs ^a	+	+	+
Leader sequence in mRNAs	+	-	+
Only 5' unique region of mRNAs is translationally active ^a	+	+	+
Ribosomal frameshifting in the polymerase gene	+	+	+
M protein with triple membrane-spanning sequences	+	+	+
Intracellular budding	+	+	+
Genome size (kb)	27-31	~25	13-15
Nucleocapsid	Helical ^c	Tubular	Isometric
Prominent spikes	+	+	-
Coiled-coil structure in spikes	+	+	-
Size of virion proteins (kDa)			
Large surface glycoprotein (S or G)	180-220	200	G _L 30-42 G _S 25
Hemagglutinin-esterase protein (HE)	60-65 ^b	^d	^d
Integral membrane protein (M)	25-35	26	18
Small membrane protein (E)	10-12	^e	^e
Nucleocapsid protein	43-50	18	12

^a Primary common characteristics for inclusion of these viruses in the proposed order Nidovirales.

^b Present in only a subset of coronaviruses (Table I).

^c May have an isometric core in addition (Risco *et al.*, 1996).

^d HE pseudogene known for BEV.

^e No such protein described.

FIG 2. Comparison of the genome organization of a coronavirus (MHV), a torovirus (Berne virus, BEV), and an arterivirus (equine arteritis virus, EAV). The genes (numbered) are drawn approximately to scale. The various coronaviruses differ with respect to the possession of an HE gene (see Table I) and with respect to the number and position of nonstructural protein genes. The polymerase genes encode two ORFs, 1a and 1b, which overlap. L, leader sequence; HE, hemagglutinin-esterase; S, spike; E, small membrane protein; M, integral membrane protein; N, nucleocapsid protein; An, poly(A) tail; G_S and G_L, small and large glycoproteins, respectively.

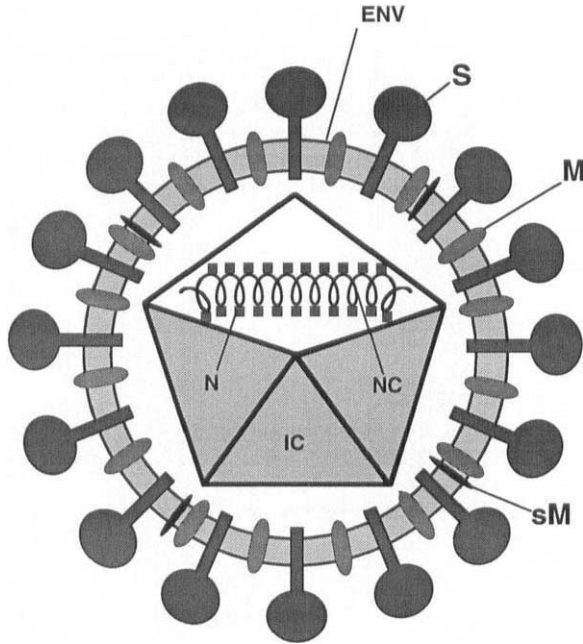


FIG 3. Model of the coronavirus virion based on the data of Risco *et al.* (1996) for TGEV. This model illustrates the observation that internal cores (IC), possibly icosahedral, were observed inside virions of TGEV. The cores comprise the helical ribonucleoprotein (NC) (genome RNA + N protein) and the M protein. Reproduced with permission from Risco *et al.* (1996).

cleaved if trypsin is present (Hogue and Brian, 1986). The extent of S cleavage depends on the cell type (Frana *et al.*, 1985). Cleavage generates two glycopolypeptides, N-terminal S1 and C-terminal S2, the latter being acylated (Sturman *et al.*, 1985). S1 is probably linked to the S2 subunits by noncovalent linkage: trypsin treatment of MHV virions caused cleavage of all S proteins without disrupting the spikes (Sturman *et al.*, 1985); however, S1 can be released from virion by either urea or mild alkali treatment (Cavanagh and Davis, 1986; Sturman *et al.*, 1990; Weismiller *et al.*, 1990).

Among the coronavirus genus as a whole, the S2 polypeptide is much more conserved than S1. Regions of up to 30% amino acid identity (particularly in the transmembrane domain) exist between the S2 polypeptides of coronaviruses in the different antigenic groups, whereas

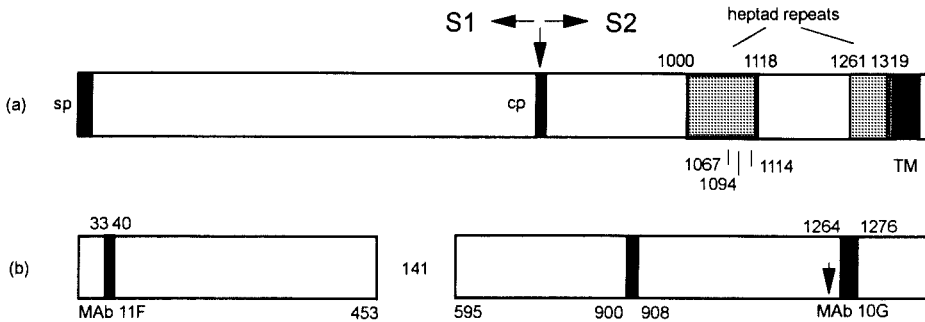


FIG 4. Features of the S protein, based on two MHV-JHM strains, (a) (S. E. Parker *et al.*, 1989) and (b) (Schmidt *et al.*, 1987). The amino acid numbering has been normalized with respect to that of the longest known MHV S protein, that of MHV4 (JHM) (S. E. Parker *et al.*, 1989). (a) The protein has an amino-terminal signal peptide (sp) and a transmembrane (TM) sequence near the C terminus. The glycosylated propolypeptide is cleaved at a basic connecting peptide (cp) to yield glycopolypeptides S1 and S2. The locations shown are those of three mutations present in mutants of MHV4 recovered from a persistently infected neural cell line, the mutants requiring a pH of 5.5–6.0 for membrane fusion (Gallagher *et al.*, 1991). (b) S of another MHV-JHM (Schmidt *et al.*, 1987), which has a 141-amino acid deletion with respect to (a). Bacterial expression products containing residues 33–40 and 1264–1276 bound MAb 11F and 10G, respectively, both of which neutralize virus infectivity and inhibit membrane fusion. The arrow indicates the positions of amino acid substitutions in JHM MAb 11F-resistant mutants (Grosse and Siddell, 1994). A peptide comprising residues 900–908 bound another MAb that neutralized virus and inhibited fusion (Luytjes *et al.*, 1989).

there is almost no conservation of the S1 sequence. Furthermore, comparison of S1 sequences among strains of a given species, or between species of a given group, reveals hypervariable regions, which include frequent deletions, mutations, or recombination (Cavanagh *et al.*, 1988; S. E. Parker *et al.*, 1989; Banner *et al.*, 1990; Gallagher *et al.*, 1990), suggesting that this region is externally exposed and not essential for the structure.

The S2 polypeptide has two regions with a seven-residue periodicity, forming heptad repeats (Fig. 4) indicative of a coiled-coil structure (de Groot *et al.*, 1987). Indeed, current evidence suggests that the mature S protein forms an oligomer; for TGEV, it is probably a trimer (Delmas and Laude, 1990). However, a dimer structure has been proposed for IBV S protein (Cavanagh, 1983c). Therefore, the oligomeric S protein is envisaged as being anchored in the membrane by an α -helical region near to the C terminus of S2. Just beyond the outer membrane surface

is the shorter (minor) repeat structure predicted to be an α helix of 5–7 nm. The major repeat indicates a helix of 10–13 nm, which may form the narrow stalk of the spikes (de Groot *et al.*, 1987). All coronavirus S2 proteins have a highly conserved eight-residue sequence KWPWW/YVWL, the last five residues of which probably form the beginning of the membrane-spanning domain. Terminating 10 residues upstream of KWP is a leucine-zipper motif, the length varying from three to five heptad repeats (Britton, 1991). The highly conserved sequences of S2 may play a role in forming the stalk, which has a more rigid structure. In contrast, the S1 domain is predicted to form the globular portion of the spikes, consistent with its highly variable nature.

The S protein has two important biological activities for the virus:

a. Induction of Membrane Fusion. This activity may be required for viral entry into cells or for cytopathic effects. Expression of the recombinant S gene has provided unequivocal evidence that the S protein alone is sufficient to cause membrane fusion, as shown by syncytium formation (de Groot *et al.*, 1989; Pfeleiderer *et al.*, 1990; Yoo *et al.*, 1991; Taguchi, 1993). Several regions of the S protein, widely separated in a linear sense, have been implicated in the membrane fusion process by the following observations: (1) S2 of BCV expressed in insect cells caused fusion (Yoo *et al.*, 1991). (2) A monoclonal antibody that inhibited cell fusion was shown to bind to the S2 domain of MHV (Fig. 4) (Luytjes *et al.*, 1989). (3) Changes at three S2 residues (1067, 1094, and 1114 in the MHV4 S protein; Fig. 4) were associated with a change from a requirement for a neutral pH to an acidic pH for fusion (Gallagher *et al.*, 1991). (4) Two bacterial expression products containing residues 33–40 (S1) and 1264–1276 (S2) of the JHM strain of MHV induced monoclonal antibodies 11F and 10G, respectively, both of which inhibited fusion (Fig. 4) (Routledge *et al.*, 1991). (5) Chemical modifications of the cysteine residues, specifically residue 1163 in the ectodomain of S2, reduced the fusion activity of the JHM strain of MHV (Gallagher, 1996). This result also suggests strain-specific differences in the conformation of the S protein, since the fusion activity of the A59 strain of MHV was not affected by this modification. (6) Some mutations to cysteine residues within the transmembrane domain of S adversely affected fusion, suggesting that the transmembrane domain is involved in conformational changes that are associated with fusion activity (Bos *et al.*, 1995). These results combined suggest that the S2 ectodomain contains the major determinants for membrane fusion; however, S2 does not contain hydrophobic domains typical of fusion proteins. Thus, several disparate regions, including some in the S1, may contribute to the fusion activity,

probably because some of these regions are juxtaposed in the three-dimensional structure or can affect the overall conformation of the spikes. Interestingly, monoclonal antibody-resistant mutants of the JHM strain of MHV selected with antibody 11F had mutations not at the antibody-binding site (residues 33–40 of S1), but at a distant site, i.e., residues 1109–1116 in the S2 domain (Grosse and Siddell, 1994) (Fig. 4), suggesting that S is folded such that regions which are widely separated in the linear sense are juxtaposed to form functional domains.

Early studies of coronavirus-induced cell–cell fusion suggested that only cleaved S was able to promote cell fusion (Sturman *et al.*, 1985). More recent studies in which MHV S proteins with mutated S1–S2 connecting peptides were expressed have shown that cleavage is not essential for fusogenic activity, although cell–cell fusion is more efficient when the S protein is cleaved (Stauber *et al.*, 1993; Taguchi, 1993; Bos *et al.*, 1995). Furthermore, naturally occurring mutants of MHV, derived from persistently infected mouse cells, which are defective in S cleavage, have delayed fusion activity (Gombold *et al.*, 1993). Expression of the feline infectious peritonitis virus (FIPV) S protein, which is not cleaved at all, also resulted in syncytia formation (de Groot *et al.*, 1989). These results indicate that S protein cleavage is not required for but can enhance membrane fusion. Whether membrane fusion activity, as manifested by syncytia formation, is required for viral infectivity has not been established. There are MHV strains (e.g., MHV-2) that do not cause syncytia formation in cultured cells; however, these viruses may be able to cause virus–cell membrane fusion within the infected cells.

b. Receptor Binding. Monoclonal antibodies (MAb) against the S protein of most coronaviruses can neutralize viral infectivity; thus, it is assumed that the S protein mediates virus binding to the receptors on target cells. Indeed, the S protein or a portion of it can bind to the viral receptor molecules *in vitro*. This has been demonstrated for MHV and TGEV S proteins (Godet *et al.*, 1994; Kubo *et al.*, 1994). The binding domain has been mapped to the N-terminal 330 amino acids of MHV S1 protein. Site-directed mutagenesis of this region showed that mutations of the residues at position 62 and positions 212, 214, and 216 abolished the binding of the protein to the receptor (Suzuki and Taguchi, 1996), suggesting that the receptor-binding site might comprise discontinuous regions in the linear sense. The S2 subunit is not involved in receptor binding (Taguchi, 1995).

The receptor-binding sites of TGEV S protein have been mapped to a 223-residue region (aa 506–729) of the S1 (Godet *et al.*, 1994), which overlaps with an epitope for a neutralizing MAb. This neutralizing MAb was able to block the binding of the 223-residue polypeptide to

the receptor; conversely, the receptor did not block the binding of the MAb to this polypeptide, suggesting that the receptor-binding determinants and the neutralizing epitopes are distinct and are part of a domain of S whose configuration is independent of the remainder of the S protein.

S proteins of BCV and HCV-OC43 bind to 9-*O*-acetylneuraminic acid (Schultze *et al.*, 1991a); this binding is required for viral infection. The significance of this binding will be discussed in Section V,B on virus attachment. Intriguingly, several coronavirus S proteins share some sequence identity with the receptor for the Fc fragment of mammalian immunoglobulins (Fc γ receptor). Thus, MAb to the Fc γ receptor could immunoprecipitate S protein from the MHV-infected cells, and S could bind to the Fc fragment of immunoglobulin. This molecular mimicry was first demonstrated for MHV and, more recently, for BCV and TGEV as well (Oleszak and Leibowitz, 1990; Oleszak *et al.*, 1992, 1995). It may play a role in modulating viral pathogenicity. This potential function is significant because expression of the S protein in the infected cells induces not only humoral antibodies but cellular immunity as well (Welsh *et al.*, 1986); the potential binding of S to the Fc fragment of immunoglobulin may modulate these immune responses.

2. Integral Membrane Glycoprotein (M)

The M protein is one of only two of the structural proteins [the other being the E protein (see below)] that are essential for the production of coronavirus-like particles. The sequence of the M protein reveals that the M polypeptides comprise 225–230 amino acids, except for some members of the TGEV group, which have an additional 30 or so residues at the amino terminus, forming a cleavable membrane insertion signal. The amino-terminal 20 or so residues of the mature M protein of all the coronaviruses are hydrophilic, exposed at the virion surface, and have a small number of glycosylation sites. Glycans are of the N-linked type for IBV and the TGEV group and O-linked for the MHV group (Rottier, 1995). The remainder of the N-terminal half of the molecule forms three helical membrane-spanning domains, although a mutant M protein which lacked all three of the membrane-spanning domains did associate with membranes *in vitro* (Mayer *et al.*, 1988). The structure of the C-terminal half is uncertain, but it is believed to be largely situated on the inside of the viral envelope, based on protease susceptibility (Rottier *et al.*, 1984; Cavanagh *et al.*, 1986b) and sequence-based predictions (Armstrong *et al.*, 1984; Rottier *et al.*, 1986). However, some M molecules of TGEV virions have the C terminus exposed at the virion surface (Laviada *et al.*, 1990; Risco *et al.*, 1995). Moreover, MAb specific

for the C-terminal 46 amino acids of M neutralized TGEV virions in the presence of complement and caused antibody-mediated, complement-dependent cytolysis of TGEV-infected cells (Risco *et al.*, 1995). Studies with mutant MHV M proteins expressed from vaccinia virus recombinants had shown that some had the N terminus and others the C terminus at the luminal side of the endoplasmic reticulum, equivalent to the outer surface of virions (Locker *et al.*, 1992b). Some molecules of one mutant M protein had both termini at the luminal surface, and other molecules had both termini at the cytoplasmic surface (Locker *et al.*, 1992b, 1994). Thus, the precise topology and the structural role of the M protein are still not certain. Recent studies have shown that some M proteins are also associated with the RNP core of TGEV and constitute the outer shell of the internal core (Risco *et al.*, 1996). This core-associated M can be clearly separated from the viral envelope. Therefore, M may play a dual structural role in forming both the envelope and the internal core of the virion.

Several properties of the M protein suggest that it is involved in virus particle assembly: (1) The M protein binds to the purified nucleocapsid *in vitro* (Sturman *et al.*, 1980). (2) When the M protein was expressed alone, it was localized in the Golgi complex, near the location where virus particles bud (Tooze *et al.*, 1984; Tooze and Tooze, 1985). However, recent studies showed that the site of M protein retention in the Golgi was slightly different from that for viral particle budding (Klumperman *et al.*, 1994), suggesting that additional factors are involved in virus particle assembly. This will be discussed in Section V,H on virus assembly.

The M protein of TGEV has an additional biological activity: induction of α -interferon (Charley and Laude, 1988; Laude *et al.*, 1992). Thus, it may play a role in viral pathogenesis. Monoclonal antibodies against the M protein do not neutralize viral infectivity, suggesting that M is not involved in receptor binding. However, some of these antibodies can neutralize viral infectivity in the presence of complement (Collins *et al.*, 1982; Laviada *et al.*, 1990), indicating that part of the M protein is exposed on the virion surface.

3. Hemagglutinin-Esterase Glycoprotein (HE)

The HE glycoprotein — or perhaps one should say the HE gene — of coronaviruses is something of an enigma. Only coronaviruses belonging to the MHV group possess the HE gene (Table I). Even there, not all virus strains within a species express the HE protein (Luytjes *et al.*, 1988; Yokomori *et al.*, 1991). As with many of the so-called nonstructural protein genes of coronaviruses, the product of the HE gene is not

essential for viral replication, certainly not in the cell types used in the laboratory. The HE protein was first detected in BCV (King *et al.*, 1985) and some MHV strains; however, acceptance of it as a legitimate virus-encoded protein was delayed because in one of the MHV strains studied most thoroughly, A59, virions lacked HE. The HE gene of A59 was later shown to lack the initiation codon of the HE open reading frame (ORF); thus, the HE gene is a pseudogene in this (Luytjes *et al.*, 1988) and several other MHV strains (Yokomori *et al.*, 1991). A complete, functional HE gene was subsequently identified in the JHM strain (Shieh *et al.*, 1989) and several others (Yokomori *et al.*, 1991).

The HE glycoprotein, of approximately 65 kDa (424 amino acids in BCV), has been detected in virions of HEV, MHV, HCV-OC43, BCV, and TCV. When analyzed under nonreducing conditions, the HE protein migrates as a dimer of approximately 140 kDa (King *et al.*, 1985). The mature protein is believed to exist in the virion as a dimer, anchored by the C terminus, forming a fringe of short spikes visualized by electron microscopy (Caul and Egglestone, 1977; Dea and Tijssen, 1988). It is not known whether each spike consists of more than one HE dimer.

Those coronaviruses which contain HE in their virions cause hemagglutination much more efficiently than those that do not. Similar to the S protein, HE alone can mediate hemagglutination and hemadsorption (King *et al.*, 1985; Hogue and Brian, 1986; Vlasak *et al.*, 1988b; Deregt *et al.*, 1989; Pfeleiderer *et al.*, 1991; Schultze *et al.*, 1991a); however, HE seems to have weaker activity than S (Schultze *et al.*, 1991a). HE binds to 9-*O*-acetylated neuraminic acid (Vlasak *et al.*, 1988b; Schultze *et al.*, 1991a), which is also a target for S binding. Some HE-specific MAb can neutralize BCV infectivity (Deregt and Babiuk, 1987; Deregt *et al.*, 1989). Thus, HE protein of BCV may participate in virus binding to the receptor. The relative importance of HE and S in hemagglutination and tissue tropism of BCV is not known.

As its name implies, the HE protein also has esterase activity; specifically, it is a neuraminidase-*O*-acetyl esterase. It hydrolyzes the 9-*O*-acetylated sialic acid on erythrocytes, thereby reversing hemagglutination induced by the HE or S protein; thus, HE is considered a receptor-destroying enzyme (Vlasak *et al.*, 1988a,b; Yokomori *et al.*, 1989; Parker *et al.*, 1990). The putative esterase active site is FGDS, encoded by amino acids 19–22 of the mature HE polypeptide of BCV (M. D. Parker *et al.*, 1989; Kienzle *et al.*, 1990). In these respects, it resembles the hemagglutinin-esterase-fusion (HEF) glycoprotein of influenza C viruses, which also has hemagglutinating and 9-*O*-acetylated sialic acid-hydrolyzing esterase activities. Moreover, the HE protein of coronavi-

ruses shares some 29% amino acid identity with the HEF of influenza C virus, including conservation of the position of the putative esterase-active site FGDS and many cysteine residues (Luytjes *et al.*, 1988; S. E. Parker *et al.*, 1989; Kienzle *et al.*, 1990; Zhang *et al.*, 1991). Unlike the HEF protein of influenza C virus, which is cleaved into two subunits (Nakada *et al.*, 1984), the coronavirus HE protein is not cleaved and lacks most of the C-terminal subunit of the HEF of influenza C virus.

Because of the close relatedness between the coronavirus HE protein and the influenza C virus HEF protein, and because the HE gene is present in only one coronavirus group, it was proposed that the HE gene was acquired by a coronavirus as a result of recombination between an ancestral coronavirus and influenza C virus (Luytjes *et al.*, 1988). Interestingly, the torovirus Berne virus also has an HE pseudogene (gene 4; Fig. 2) (Snijder and Horzinek, 1995), the amino acid sequence of which has approximately 30% identity with the C-terminal part of the coronavirus HE.

The functional significance of HE for coronaviruses is not known. Among coronaviruses, only BCV requires HE for infectivity; however, the presence of HE may affect the pathogenicity of some coronaviruses, as evidenced by the findings that passive administration of HE-specific MAb in mice altered MHV pathogenicity and that MHVs with an HE have different neuropathogenicity from those without HE (Yokomori *et al.*, 1992a, 1995). Conceivably, the presence of HE in an MHV may allow the virus to utilize an alternative receptor independently of the S protein. However, this is not the case, as evidenced by the finding that an MAb specific for the murine biliary glycoprotein molecule, which is the major MHV receptor recognized by S, inhibited the infectivity of an HE-containing MHV (Gagneten *et al.*, 1995). Thus, the HE protein does not enable a virus to bypass the primary MHV receptor and may provide only an auxiliary function for virus binding to target cells.

4. Small Membrane Protein (E)

Until recently it was thought that coronaviruses possessed three (S, N, M) or four (including HE) structural proteins. It is now clear that coronaviruses, but not toroviruses, possess an additional virion protein, the E protein. It plays an essential role in virion assembly. It has been shown that the E and M proteins are the only two viral proteins absolutely required for virion assembly (Bos *et al.*, 1996; Vennema *et al.*, 1996). This protein has been demonstrated for IBV (Smith *et al.*, 1990; Liu *et al.*, 1991), TGEV (Godet *et al.*, 1992) and MHV (Yu *et al.*, 1994). When the deduced E proteins of the other coronaviruses are taken into account, it transpires that the E proteins vary from 84 to

109 amino acids, corresponding to molecular weights of 9100 to 12,400 (Siddell, 1995c). Siddell has highlighted a number of features common to all the E proteins, namely, a hydrophobic region of some two dozen residues, starting near the N terminus; a cysteine-rich region immediately downstream from this; a conserved proline residue in the middle of the molecule, and otherwise very low amino acid identity in the genus as a whole; and an abundance of charged residues in the C-terminal half of the protein (Siddell, 1995c).

It is now well established that this protein is associated with highly purified virion preparations (Liu *et al.*, 1991; Godet *et al.*, 1992; Yu *et al.*, 1994). Liu and Inglis calculated the ratio of S:N:M:E proteins in virions of IBV-Beaudette strain to be 1:11:10:2, indicating an amount of E protein similar to that of S protein (Liu *et al.*, 1991). In contrast, Godet *et al.* estimated that the S:M:E protein ratio in virions of TGEV was 20:300:1 (Godet *et al.*, 1992), and Vennema *et al.* (1996) have suggested an M:E ratio of approximately 100:1 for virions of MHV. It is not clear why there is such a wide range of variations.

The E protein in the cells is localized in the perinuclear region, with some migrating to the cell surface (Godet *et al.*, 1992; Yu *et al.*, 1994). Experimental evidence suggests that the E protein is anchored in the membrane by sequence in the N-terminal half of the molecule. Thus antibodies specific for epitopes in the C-terminal half of the TGEV E protein produced cell-surface fluorescence in paraformaldehyde-fixed, TGEV-infected cells (Godet *et al.*, 1992), but the precise topology of the protein has not been elucidated. The role of the E protein in virion assembly will be discussed in Section V,H on virus assembly and release.

The E proteins of IBV and MHV are translated from the third and second ORFs, respectively, of mRNAs 3 and 5 of the respective viruses. Both of these are polycistronic mRNAs (see Figs. 5 and 7 and Section V,G,2). In contrast, in all other viruses, the E protein is derived from a monocistronic mRNA. The mechanism of translation of the IBV and MHV E proteins is discussed in Section V,G.

5. Nucleocapsid Protein (N)

The N protein is a 50- to 60-kDa phosphoprotein which, together with the genomic RNA, forms a helical nucleocapsid (RNP). The RNP of coronaviruses have been reported variously as being from 9–11 to 14–16 nm in diameter (see Laude and Masters, 1995, for references). The N protein in RNP provides only limited protection to the RNA genome against ribonucleases. The N proteins vary from 377 to 455 amino acids in length, are highly basic, and have a high (7–11%) serine

content, potential targets for phosphorylation. Sequence conservation within the genus is low. Thus, the N proteins of IBV and TGEV have only 29% identity with that of BCV. Even within the MHV group, the N proteins of MHV and BCV share only 70% identity, whereas the M proteins of these two viruses have 86% identity (Lapps *et al.*, 1987). Based on sequence comparison, three structural domains in the N protein have been identified (Parker and Masters, 1990). The middle domain is an RNA-binding domain (Masters, 1992; Nelson and Stohlman, 1993) which binds to both coronaviral and nonviral RNA sequences *in vitro* (Robbins *et al.*, 1986; Stohlman *et al.*, 1988; Masters, 1992); however, it does not contain any motifs characteristic of other RNA-binding proteins. Under specific binding conditions, the MHV N protein binds to the leader RNA sequence, particularly nucleotides 56–67 (Stohlman *et al.*, 1988). Furthermore, an anti-N MAb immunoprecipitated all of the MHV RNA molecules which had the leader sequence (Baric *et al.*, 1988). The N protein of IBV also bound to the 3' untranslated region of the IBV RNA *in vitro* (Zhou *et al.*, 1996). These RNA-binding properties are consistent with the fact that the N protein interacts with the viral genomic RNA to form nucleocapsid. This interaction is necessary for the formation of virus particles, as N alone cannot be incorporated into virus particles, whereas the N–RNA complex can (Bos *et al.*, 1996; Vennema *et al.*, 1996). However, the specificity of the RNA–N protein interaction required for nucleocapsid formation has not been elucidated. The N protein also binds to membranes and phospholipid (Anderson and Wong, 1993). This may be another property which facilitates the formation of virus particles.

The finding that the N protein binds to the 5' and 3' ends of viral RNA suggests that the N protein may also modulate viral RNA synthesis because the ends of the RNA are likely involved in the regulation of RNA synthesis. In an *in vitro* RNA replication system, the addition of MHV N-specific antibodies inhibited viral RNA synthesis (Compton *et al.*, 1987), suggesting that the N protein is a component of the RNA-synthesizing machinery. The ability of N to bind to the membrane (Anderson and Wong, 1993) may enable the formation of the RNA replication or transcription complex, in view of the fact that viral RNA synthesis occurs in the membrane fraction of infected cells (Brayton *et al.*, 1982; Dennis and Brian, 1982).

The three structural domains of the N protein are separated by spacer regions, which are not conserved (Masters, 1992). The functions of the N- and C-terminal conserved domains are not yet clear. Using a targeted recombination approach (Koetzner *et al.*, 1992; Masters *et al.*, 1994) to generate recombinant viruses that have a chimeric N gene

containing parts of BCV and MHV sequences, Peng *et al.* (1995a) have shown that there is strict sequence specificity within the conserved structural domains for viable recombinants. Since the N protein constitutes the nucleocapsid, mutations within the N protein will likely affect the stability or viability of the virus. Indeed, several temperature-sensitive and thermolabile mutants of MHV have deletions or mutations within the N protein (Koetzner *et al.*, 1992; Peng *et al.*, 1995b). Viruses with site-specific mutations of the N gene have been generated by targeted recombination techniques; interestingly, revertants of these mutants often have second-site mutations located at different domains, suggesting that there are interactions between different domains of the N protein (Peng *et al.*, 1995b).

The role of phosphorylation in the N protein has not been elucidated.

C. RNA Genome

The coronavirus contains a positive-sense, single-stranded RNA genome, which is the largest viral RNA genome known, ranging from 27.6 to 31 kb. The large size of the viral RNA requires the virus to develop special mechanisms of RNA synthesis to counter the deleterious effects of the possible errors during RNA synthesis. The virion RNA functions as an mRNA and is infectious. It contains approximately 7–10 functional genes, 4 or 5 of which encode structural proteins. The genes are arranged in the order 5'-polymerase-(HE)-S-E-M-N-3', with a variable number of other, mostly nonstructural and largely nonessential, genes interspersed among them (Fig. 5). This gene arrangement also applies to toroviruses and arteriviruses (Fig. 2). The 5' terminus of the coronavirus genome is capped, and the RNA starts with a leader sequence of 65–98 nucleotides, which is also present at the 5' end of mRNAs, followed by a 200- to 400-nucleotide untranslated region (UTR). At the other end of the genome is a 3' UTR of 200–500 nucleotides followed by a poly(A) tail. Almost two-thirds of the entire RNA is occupied by the polymerase gene, which comprises two overlapping ORFs, 1a and 1b. At the overlap region is a specific seven-nucleotide "slippery" sequence and a pseudoknot structure, characteristic of the ribosomal frameshifting signal (Brierley *et al.*, 1987, 1989; Lee *et al.*, 1991; Herold and Siddell, 1993), which is required for the translation of ORF 1b. The architecture of the nonstructural protein genes interspersed between the known structural protein genes varies significantly among different coronavirus species (Fig. 5). For example, in HCV-229E, gene 3 contains two ORFs, whereas in the related virus

PEDV these two ORFs are fused (Duarte *et al.*, 1994). In HCV-OC43, gene 4 is missing altogether (Mounir and Talbot, 1993). Finally, in IBV, two ORFs are inserted between M and N genes. The variability of gene structure indicates the plasticity of coronavirus RNA and the frequent occurrence of recombination and also suggests that there is no strong conservation pressure on these nonstructural proteins. There is a stretch of consensus sequence, UCUAAAC (for MHV), or a related

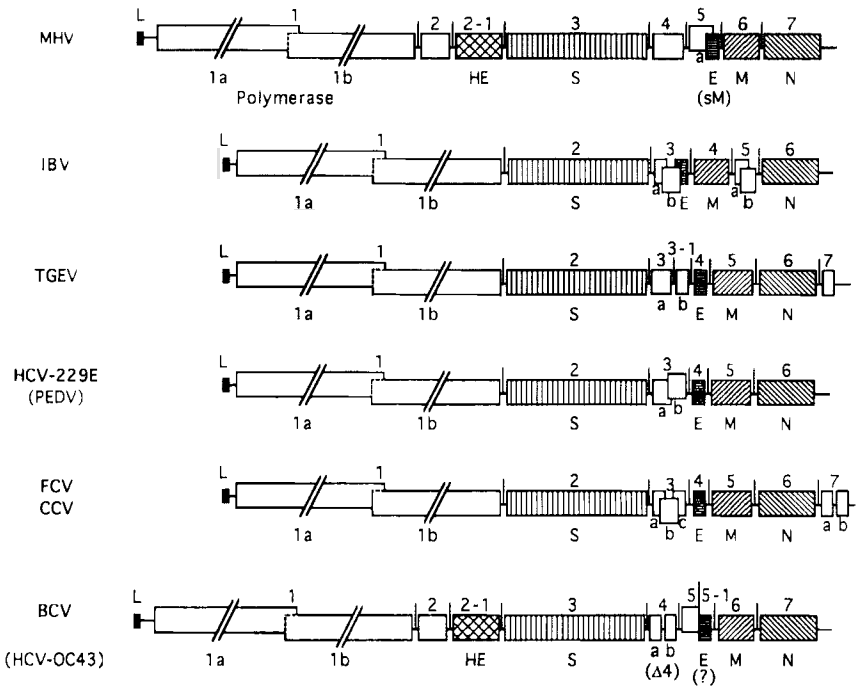


FIG 5. Comparative genome structure of the different coronaviruses. The complete sequences are available for MHV, IBV, TGEV, and HCV-229E. The gene 1 sequences of the remaining viruses have not been completed. Gene 1 sequences are interrupted and shortened to highlight the remaining genes. The vertical lines represent mRNA start sites; thus, each region between two vertical lines represents a separate gene ("transcription unit"). The structural protein genes are marked by various symbols, and nonstructural protein genes are represented by unfilled boxes. The gene arrangements of ns protein genes and E protein gene are very heterogeneous in terms of transcription unit and the relative size and position among different strains of the same virus species; only the representative one is presented. The numbering system for the genes of HCV-229E deviates from the published one (Herold *et al.*, 1993) to be consistent with the other viruses. HCV-OC43 does not have a gene 4.

sequence, at sites immediately upstream of most of the genes. These sequences represent signals for transcription of subgenomic mRNAs (see Section V,E).

Finally, a pseudoknot structure has been shown to be present at the 3' end of the coronaviral RNA (Williams *et al.*, 1995). A characteristic feature of the *Coronaviridae*, and of the *Arteriviridae* as well, is that all known member species generate a 3'-coterminal nested set of five or more mRNAs (see Fig. 7). Each coronavirus and arterivirus subgenomic mRNA has the leader sequence at its 5' end. Curiously, no leader RNA sequence is present in the torovirus RNAs (Fig. 2).

IV. NONSTRUCTURAL PROTEINS

In 1990, the Coronavirus Study Group published its recommendations for the nomenclature of coronavirus genes, mRNAs, and proteins (Cavanagh *et al.*, 1990). At that time it was reluctant to apply the term "nonstructural" to the potential products of genes which were suspected of not being structural proteins. This caution was a consequence of our lack of knowledge of those gene products, a situation which has improved greatly in the last 5 years or so. This has resulted in the term "nonstructural (ns)" being applied more widely to several gene products. Every gene that encodes the ns proteins has been deleted in at least some naturally occurring virus isolates; thus, most of the ns genes are not essential for viral replication. However, some of the ns proteins may play a role in viral tissue tropism or pathogenicity.

A. *The Polymerase*

The polymerase is encoded by gene 1, which accounts for approximately two-thirds of the genome (Fig. 2). The complete polymerase gene of four coronaviruses (IBV, MHV, HCV-229E, and TGEV) covering each of the three coronavirus groups has been sequenced (Bournell *et al.*, 1987; Lee *et al.*, 1991; Herold *et al.*, 1993; Bonilla *et al.*, 1994; Eleouet *et al.*, 1995). Although the polymerase genes vary in size from approximately 18 to 22 kb, the encoded proteins have many structural features in common. The degree of amino acid identity for this gene product is greater than is observed for any other coronavirus gene product.

The polymerase gene is predicted to encode a protein of approximately 740–800 kDa. Proteins of this size have not been detected in coronavirus-infected cells, in part because of co-translational polypro-

tein processing. The *pol* gene encodes two ORFs, 1a and 1b, which overlap by a few dozen nucleotides (Figs. 2 and 6). The second, ORF 1b, is in the -1 reading frame with respect to the upstream ORF 1a and is translated following ribosomal frameshifting in the overlap region. This will be examined in more detail in Section V.G.

There is greater amino acid identity among the 1b than the 1a ORFs. For example, 1a and 1b of IBV, the least typical coronavirus in terms of protein sequences, have amino acid identity/similarity of approximately 30/50% and 55/79%, respectively, compared with those of MHV, HCV, and TGEV. It is the 1a ORF which accounts for the MHV polymerase gene being approximately 1–2 kb longer than those of IBV, HCV, and TGEV.

A number of functional domains within *pol* have been predicted following computer-based motif analyses (Bournsnel *et al.*, 1987; Hodgman, 1988; Gorbalenya *et al.*, 1989a,b; Lee *et al.*, 1991); some of these functional domains have been confirmed by experimental analysis. The location of these motifs is illustrated in Fig. 6. Three motifs have been identified in ORF 1a, indicating the presence of one or two papain-like cysteine proteases (PLP): a chymotrypsin/picornaviral 3C-like protease

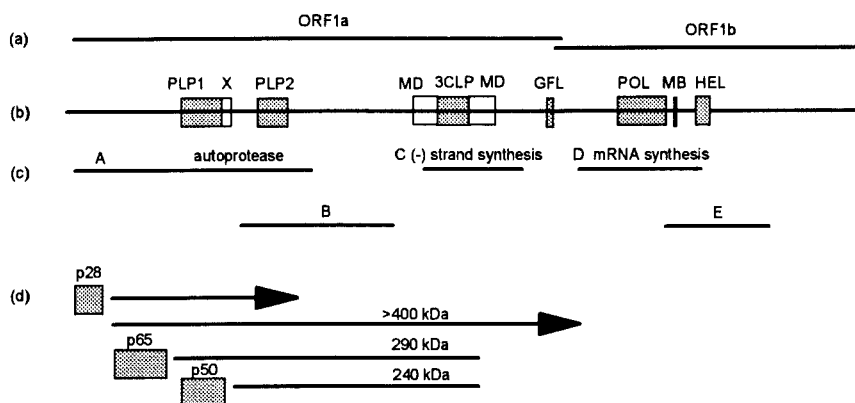


FIG 6. Features of the coronavirus polymerase gene, based on that of MHV (Lee *et al.*, 1991). (a) The polymerase gene comprises two ORFs, 1a and 1b, which overlap, the 1b ORF being translated after ribosomal frameshifting. (b) The positions of motifs: PLP 1 and 2, papain-like protease; X domain, highly conserved between IBV and MHV; 3CLP, picornavirus-3C-like protease; MD, membrane-associated domain; GFL, growth factor-like; POL, RNA-dependent RNA polymerase; MB, metal-binding motif; HEL, helicase. (c) Genetic complementation groups (Schaad *et al.*, 1990; Fu and Baric, 1994). (d) Processing scheme for part of the 1a ORF (Denison *et al.*, 1992, 1995).

(3CLP) and a cysteine-rich growth factor-related protein (GFL). MHV, HCV-229E, and TGEV have two PLP domains (1 and 2), with PLP2 corresponding to the single PLP domain of IBV. Sequence corresponding to a cysteine protease of *Streptococcus pneumoniae* has been identified in 1a of IBV. Upstream of PLP2 is a region termed the X domain, a region of particularly high conservation between IBV and MHV and similar to one near the thiol protease of alpha- and rubiviruses (Gorbalenya *et al.*, 1991). There is no functional evidence so far to link the GFL with known growth factors, but the predictions of most of the protease domains have been confirmed by experimental analysis. The first PLP domain of MHV is responsible for the cleavage of p28/p30 and p65 from the N terminus of the MHV ORF 1a polyprotein (Fig. 6) (Baker *et al.*, 1989, 1993; Bonilla *et al.*, 1995, 1997). This PLP was inhibited by zinc chloride but not by leupeptin (Baker *et al.*, 1989; Denison *et al.*, 1992). Deletion analysis defined this proteinase domain to be within the sequence encoded by the 3.6–4.4-kb region from the 5' end of the genome. Site-directed mutagenesis showed that residues Cys-1137 and His-1288 were essential for protease activity (Baker *et al.*, 1993). Some amino acid sequences between the p28 cleavage site and the PLP domain were also essential for the cis cleavage that generates p28 (Baker *et al.*, 1993; Bonilla *et al.*, 1995). The function of PLP2 has not been demonstrated.

The 3CLP domain extends for approximately 300 amino acids and is homologous to proteases encoded by picornaviruses and several other virus genera. The putative 3CLP domain of HCV-229E has been expressed as a β -galactosidase fusion protein in *Escherichia coli* and shown to have autocatalytic proteolytic activity, releasing an active 3CLP protein (Ziebuhr *et al.*, 1995). An antiserum against this fusion protein immunoprecipitated a 34-kDa protein from HCV-229E-infected cells. Similar activity has been demonstrated for the 3CLPs of MHV (Lu *et al.*, 1995) and IBV (Tibbles *et al.*, 1996). This protease cleaves not only its own boundaries but also several downstream sites within ORF 1a and ORF 1b, probably both in cis and in trans. Computer analysis predicted that the catalytic center of the IBV 3CLP would include Cys-2922, His-2820, and Glu-2843 (Gorbalenya *et al.*, 1989a,b). Site-directed mutagenesis confirmed the role of the Cys and His residues but showed that the Glu residue was not essential (Liu and Brown, 1995). The same approach confirmed that the predicted QS(G) dipeptide bonds in the 1b ORF are the targets for the protease activity of the 3CLP of IBV (discussed further in Section V,G). Similar conclusions were reached for 3CLP activity of MHV and HCV-229E (Lu *et al.*, 1995; Grötzinger *et al.*, 1996). The importance of Cys-3495 in the 3CLP of

MHV has been demonstrated (Seybert *et al.*, 1997). *In vitro* transcription and translation of a cDNA containing the putative 3CLP of MHV produced polypeptides of 38 and 33 kDa, which were subsequently processed to products of 32 and 27 kDa (Lu *et al.*, 1995). The 27-kDa protein possesses the 3C-like protease activity (Lu *et al.*, 1996). The 3CLP domain is flanked by predicted membrane-spanning domains, which may be important for the proteolytic activity (Tibbles *et al.*, 1996) (Fig. 6).

Poor expression of the IBV 3CLP protein *in vitro* led to the discovery that this protease was ubiquitinated and subsequently degraded by an adenosine triphosphate (ATP)-dependent protease present in reticulo-cyte lysate (Tibbles *et al.*, 1995). This is the third example of a viral protein subject to turnover in this manner and involves a different virus class from the previously reported examples, in a picornavirus (Oberst *et al.*, 1993) and an alphavirus (de Groot *et al.*, 1991). The ubiquitin-mediated, ATP-dependent proteolytic pathway is a major cellular, nonlysosomal, protein degradation system, which may cause rapid turnover of the coronaviral polymerase.

The functional domains associated with RNA synthesis are located within the more conserved 1b ORF. These include domains for an RNA-dependent RNA polymerase, a nucleoside triphosphate (NTP)-binding/helicase domain, and a zinc-finger nucleic acid-binding domain (metal binding domain) (Fig. 6). Computer analyses identified the polymerase domain (Bournsnell *et al.*, 1987; Hodgman, 1988; Gorbalenya *et al.*, 1989a,b). Unlike the GDD motif present in many viruses, the corresponding sequence in coronaviruses is SDD. Whether the polymerase gene products contain activities other than proteases and polymerases is not known.

B. Other Nonstructural (ns) Proteins

The coronaviruses exhibit great heterogeneity with respect to the number and genome location of ns protein genes and in regard to the number of ORFs within a gene (Fig. 5). The functions of these ns proteins are still unknown.

1. Genes between the Polymerase and S Gene (Gene 2 of MHV, BCV, and HCV-OC43)

There are two genes located between the polymerase and S genes of these viruses (Fig. 5). Gene 2-1 encodes the HE protein, while gene 2 encodes an ns protein of unknown function. The gene 2 protein com-

prises approximately 260 amino acids (30 kDa) (Luytjes *et al.*, 1988; Shieh *et al.*, 1989; Labonte *et al.*, 1995). The BCV and MHV homologs share 45% amino acid identity, while the homolog of HCV-OC43 has 92% identity with that of BCV. This gene product has been detected in the cytoplasm of MHV-, BCV-, and HCV-OC43-infected cells and may be phosphorylated (Bredenbeek *et al.*, 1990; Zoltick *et al.*, 1990; Cox *et al.*, 1991; Labonte *et al.*, 1995). Computer analysis of its sequence suggested the presence of a nucleotide-binding site (Luytjes *et al.*, 1988). However, no function has been assigned to this protein, and it is not required for virus replication in culture (Schwarz *et al.*, 1990). Interestingly, the C terminus of the torovirus ORF 1a product (polymerase) has 31–36% sequence identity with the gene 2 product of MHV (Snijder *et al.*, 1991). This evolutionary relationship between coronavirus and torovirus suggests that the gene 2 product is probably involved in viral RNA synthesis, since it is expressed as part of the torovirus polymerases.

2. *Genes between S and E (Genes 3 and 3-1 of IBV, TGEV, HCV-229E, and FCV and Gene 4 of MHV and BCV)*

There are two to three ORFs in this region, and their structure and the mechanism of expression of gene products vary markedly among different coronavirus species. They can be expressed as two different genes, i.e., expressed from two separate mRNAs (e.g., mRNAs 4 and 5 of MHV and BCV and mRNAs 3 and 3-1 of TGEV) or localized in one gene, thus requiring internal initiation of translation from a single polycistronic mRNA (e.g., mRNA 3 of the IBV and FCV groups). In IBV, it contains three ORFs (3a, 3b, and 3c); ORF 3c encodes the E protein, which is a viral structural protein, while 3a and 3b encode ns proteins. The gene products of both ORFs 3a and 3b (approximately 7 kDa) have been detected in small quantities in virus-infected cells (Liu *et al.*, 1991). In TGEV, this region contains two ORFs, being separated from the E protein gene. These two ORFs are encoded by mRNAs 3 and 3-1, respectively, the predicted protein products being approximately 8 and 27 kDa, respectively. In a related nonenterogenic strain, PRCV, however, there are multiple deletions in this region, essentially inactivating one or both of the ORFs (Rasschaert *et al.*, 1990; Wesley *et al.*, 1991). It has been suggested that the absence of the 3a product, in addition to a shorter S protein, might be associated with their lack of enteropathogenicity. However, Vaughn *et al.* (1995) have recently described two PRCV strains which have an intact 3a gene (Vaughn *et al.*, 1995).

Canine coronavirus has gene 3 ORFs equivalent to those of TGEV, exhibiting high amino acid identity (>80%), although the second ORF is truncated by a stop codon (Horsburgh *et al.*, 1992). Two other members of the TGEV group exhibit a variation on the same theme. PEDV and HCV-229E lack a homolog of ORF 3a of TGEV and CCV. PEDV has an ORF corresponding to 3b of TGEV, while HCV-229E has two ORFs corresponding to the single ORF of PEDV (Duarte *et al.*, 1994).

Members of the group I coronaviruses also exhibit great heterogeneity in this region. MHV-JHM produces mRNA 4, which encodes a 15-kDa protein. This protein has been detected in virus-infected cells (Ebner *et al.*, 1988). In contrast, HCV-OC43 contains only 11 amino acids in this region (Mounir and Talbot, 1993). Gene 5 of MHV has two ORFs, 5a and 5b. The latter encodes the structural E protein and is the predominant product made from mRNA 5 (Leibowitz *et al.*, 1988). It is not clear whether ORF 5a is translated at all. At least one strain of MHV lacks the 5a ORF (Yokomori and Lai, 1991); also, HCV-OC43 has the 5a ORF but is unable to produce a corresponding mRNA (Mounir and Talbot, 1993).

In summary, there is great heterogeneity with respect to the number, size, and mechanism of expression of ORFs between the S and E genes. These ns proteins probably are not required for viral replication. The lack of necessary function may account for the heterogeneity which arose during evolution.

3. Gene 5 (between M and N Genes) of IBV

IBV is unique in that it has two ORFs (5a and 5b), which encode proteins of 7.4 and 9.5 kDa, respectively. These proteins have been detected in very small amounts in virus-infected cells (Liu and Inglis, 1992a). The function of these ORFs is not clear.

4. ORFs in the Very 3' End

TGEV has an additional gene 7, which encodes a 9.1-kDa protein (Garwes *et al.*, 1989; Tung *et al.*, 1992), in the region corresponding to the 3' end untranslated region of other viruses (Fig. 5). This protein is hydrophobic and is associated with the endoplasmic reticulum and cell surface membranes (Tung *et al.*, 1992), but its nuclear localization has also been reported (Garwes *et al.*, 1989). FCVs and CCV have two ORFs in the same region, the first being analogous to the single ORF of TGEV. The second (7b) ORF encodes a 14-kDa soluble protein containing the sequence KTEL (Vennema *et al.*, 1992), which is similar to the endoplasmic reticulum retention signal, KDEL. The protein is

partially retained in the endoplasmic reticulum but is also slowly secreted out of the cells. The functions of these proteins are not known.

V. REPLICATION CYCLE

A. *Viral Host Ranges and Metabolic Requirements of Viral Replication*

Coronaviruses have relatively restricted host ranges, infecting only their natural hosts and closely related animal species. Occasionally, cross-species infection of coronaviruses occurs, such as the experimental infection of monkey by MHV, which causes central nervous system demyelination (Murray *et al.*, 1992; Cabirac *et al.*, 1994), and the occasional infection of humans by BCV, which causes diarrhea. BCV also infects turkeys and TGEV infects dogs, suggesting some flexibility in their host range. The expansion of viral host range can be achieved by passing the coronavirus in a heterologous cell line, as demonstrated by the emergence of an MHV variant with the ability to infect originally nonpermissive cell lines, such as human cells, after serial passages (Baric *et al.*, 1997). In animals, coronaviruses have restricted tissue tropism; for example, most HCV strains cause only respiratory infections. Different strains of a coronavirus may have distinct tissue specificity; for example, TGEV infects both the gastrointestinal tract, causing fatal diarrhea, and respiratory tract tissues without causing primary respiratory symptoms, whereas PRCV, which is closely related to TGEV, infects the respiratory tract of pigs but replicates poorly in the intestinal tract (Cox *et al.*, 1990). The species and tissue specificity of a coronavirus infection is at least partially dictated by the nature and distribution of cellular receptors and other related molecules that regulate virus entry, as evidenced by the viral replication that results when viral RNA is directly introduced into cell types of other animal species. Thus, coronaviruses have the potential to replicate in many cell types.

The complete coronavirus replication cycle takes place in the cytoplasm. It has been shown that MHV can replicate in enucleated cells and in the presence of actinomycin D and α -amanitin, suggesting that nuclear functions are not required for viral replication (Brayton *et al.*, 1981; Wilhelmsen *et al.*, 1981). There are, however, reports of the inhibition of replication by actinomycin D of some coronaviruses, including feline enteric coronavirus (Lewis *et al.*, 1993), IBV (Evans and Simpson, 1980), HCV-229E (Kennedy and Johnson-Lussenberg, 1978), and MHV in some cell lines (Dupuy and Lamontagne, 1987). Thus,

nuclear functions may be required for viral replication under certain conditions. This issue has not been resolved.

B. Virus Attachment

1. Virus Binding to Erythrocytes

The first step in viral infection is the binding of the virus to target cells. Hemagglutination and hemadsorption have been used as assays for studying virus–cell interaction, although the erythrocyte itself is not a target cell for coronavirus infection. Several coronaviruses, including HEV, IBV, BCoV, and some strains of MHV and HCoV, can cause hemagglutination (Sugiyama and Amano, 1980; Schultze *et al.*, 1990; Zhang *et al.*, 1994a). The binding residue on the cell surface is a 9-*O*-acetylated neuraminic acid of glycoproteins or glycolipids (Schultze *et al.*, 1990), although different coronaviruses may prefer different structural isoforms of 9-*O*-acetylated neuraminic acid. For BCoV, the virus binding to erythrocytes is mediated through either the S or HE protein, both of which have hemagglutinating activities, the S protein having the stronger activity (King *et al.*, 1985; Schultze *et al.*, 1991a,b). The HE protein of BCoV and HEV also recognizes 9-*O*-acetylated neuraminic acid, and its esterase activity is also specific for this molecule; thus, HE protein has both receptor-binding and receptor-destroying activities (Vlasak *et al.*, 1988a,b; Schultze *et al.*, 1991b). Expression of the HE protein of MHV on the cell surface conferred a hemadsorption activity (Pfleiderer *et al.*, 1991); however, even viruses that lack HE protein (e.g., IBV) can cause hemagglutination, suggesting the role of S protein in hemagglutination. Thus, the HE and S proteins of various coronaviruses may have comparable functions, enabling the virus to bind the sialic acid residues; however, only the HE protein confers the receptor-destroying activity. The residue necessary for hemagglutination by IBV is A2,3-linked *N*-acetylneuraminic acid (Schultze *et al.*, 1992). Curiously, the hemagglutinating activity of IBV is not evident until the virus particle is treated with neuraminidase, suggesting that the S protein itself is covered by sialic acid. Although virus binding to erythrocytes provides a good model system for studying virus–cell interactions, it may not necessarily reflect the actual mechanism of virus attachment to target cells.

2. Virus Binding to Target Cells

The classical study of virus attachment to target cells involved the *in vitro* binding of MHV to macrophages from genetically susceptible

and resistant mouse strains (Shif and Bang, 1970). This study showed that MHV bound equally well to cells from resistant and susceptible mice, even though macrophages from resistant mice were resistant to virus infection. Similar observations have been made on splenic lymphocytes, thymocytes (Krzystyniak and Dupuy, 1981), and glial cells (Wilson and Dales, 1988); thus, it appears that genetic resistance is not exerted at the level of virus binding in these cases. Similarly, established tissue culture cell lines, including murine and primate cells, irrespective of their degree of susceptibility or resistance to MHV, bound MHV to the same extent (van Dinter and Flintoff, 1987; Kooi *et al.*, 1988, 1991). Thus, virus may bind to a ubiquitous molecule on the cell surface, which, however, may not lead to virus infection. Whether this ubiquitous molecule is a sialic acid-containing glycoprotein has not been established.

The binding of BCV to its target cells, such as MDCK cells, is mediated by 9-*O*-acetylneuraminic acid residues similar to those on erythrocytes. Removal of the sialic acid by neuraminidase abolished virus attachment, while resialization restored it (Schultze and Herrler, 1992). HCV-OC43 binds to a similar sialic acid residue but prefers a form slightly different from that for BCV (Kunkel and Herrler, 1993). The HE protein of BCV can also mediate virus binding to target cells, and this binding may be required for viral infection, as suggested by the finding that MAb against HE inhibited BCV infectivity (Deregt and Babiuk, 1987; Deregt *et al.*, 1989). One inhibitor of the esterase activity of HE protein, diisopropylfluorophosphate, also inhibited BCV infection (Vlasak *et al.*, 1988a). The S protein of BCV probably also participates in virus binding to target cells, as suggested by the finding that the MAb against S protein can neutralize BCV infectivity (Deregt *et al.*, 1989). The relative importance of S and HE proteins is not clear. In contrast, none of the MAb against the HE protein of MHV inhibited MHV infection (Yokomori *et al.*, 1992a). Despite the finding that the binding of HE and S proteins to target cells is necessary for BCV infection, the binding of BCV or HCV-OC43 to *N*-acetylneuraminic acid in itself is not likely the basis of viral cell tropism because sialic acid is a common cell surface carbohydrate residue; thus, an additional, more cell type-specific molecule is probably required for viral infection.

3. Specific Virus Receptors

The finding that MHV and other coronaviruses bound to resistant as well as susceptible cells indicates that this binding may represent an initial step in the virus attachment process, which is not sufficient for viral infection. It is likely that a more specific binding between

virus and cells is required for the establishment of viral infection. This binding involves a specific virus receptor molecule on the cell surface.

a. MHV Receptor. The MHV receptor was the first coronavirus receptor to be identified. It is the murine homolog of a member of the carcinoembryonic antigen (CEA) family (Dveksler *et al.*, 1991; Williams *et al.*, 1991) and belongs to the biliary glycoprotein (bgp) subfamily. The terminology of MHV receptors in the literature is somewhat controversial, the following terms being used interchangeably: mmCGM1, MHVR-1, and BgpA. It has an immunoglobulin-like structure, consisting of four immunoglobulin-like loops, the N-terminal loop being the virus-binding domain (Dveksler *et al.*, 1993b). The sequence of the C terminus (cytoplasmic domain) of the receptor is not essential. Glycosylation of the protein also is not necessary for its receptor function *in vivo* (Dveksler *et al.*, 1995). The functional significance of the receptor in viral infection *in vivo* was demonstrated by the finding that an MAB against the MHV receptor inhibited viral infection in mice (Smith *et al.*, 1991).

Subsequently, several additional members of CEA family were found to serve as MHV receptors, including an mmCGM2-like protein (also termed MHVR-2 and BgpB), which is the product of an alternatively spliced form of mmCGM1 RNA and is expressed in both the liver and brain, in contrast to the liver-specific expression of mmCGM1 (Yokomori and Lai, 1992a; Dveksler *et al.*, 1993a); an allelic gene product of the bgp gene in SJL mice, a mouse strain resistant to MHV infection (Yokomori and Lai, 1992b; Dveksler *et al.*, 1993a); Bgp-2, which is the product of a new member of the murine Bgp gene (Nedellec *et al.*, 1994); and a novel pregnancy-specific glycoprotein (psg)-like protein, which is expressed in the mouse brain, in contrast to placenta-specific expression of other psg molecules (Chen *et al.*, 1995). All these molecules contain a consensus motif in the virus-binding domain (N-terminal loop). Thus, several different CEA family members, which are differentially expressed in different cells and tissues, can potentially serve as an MHV receptor. Different strains of MHV may use different CEA-related molecules as receptors at different efficiencies (Compton, 1994; Chen *et al.*, 1995). The prototype MHV receptors (MHVR-1) are expressed in the liver, gastrointestinal tract, B cells, macrophages, and endothelial cells but not in T cells (Coutelier *et al.*, 1994; Godfraind *et al.*, 1995), consistent with the target cell specificity of MHV. However, the MHV receptor is also expressed in other tissues, e.g., kidney, which are not targets for MHV infection. Also, SJL mice express a functional MHV receptor (Yokomori and Lai, 1992b; Dveksler *et al.*, 1993a) but are resistant to MHV infection (Knobler *et al.*, 1984). Thus, receptor

expression is not sufficient for viral infection. It is not yet clear which molecules are used by MHV as receptors in cross-species infection (e.g., rats and monkeys) (Murray *et al.*, 1992; Cabirac *et al.*, 1994). Recently it was shown that bgp and CEA molecules of human origin could serve as receptors for some MHV strains (Chen *et al.*, 1997).

The expression of the receptor molecules on the cell surface is necessary for virus infection, and the expression level of the receptor may determine the relative susceptibility or resistance to viral infection in some cells. During persistent viral infection of cultured murine cells, the expression level of the receptor is often reduced, resulting in the relative resistance of the cells to viral superinfection, which could be overcome by the expression of an exogenous receptor (Sawicki *et al.*, 1995; Chen and Baric, 1996). Thus, there is a rough correlation between receptor expression and the susceptibility of a cell type to virus infection. Under certain circumstances, virus may infect cells by a receptor-independent mechanism; for example, MHV-infected murine cells may fuse with human cells, which do not have MHV receptors, and cause the latter cells to become infected (Gallagher *et al.*, 1992). It has been shown that MHV infects polarized epithelial cells through the apical, but not the basolateral, surface (Rossen *et al.*, 1995a, 1997). It is not clear whether the virus receptor is differentially expressed on the two different surfaces.

b. Receptors for TGEV and HCV-229E. The receptors for TGEV and HCV-229E have been identified as aminopeptidase N (APN) of the porcine and human species, respectively (Delmas *et al.*, 1992; Yeager *et al.*, 1992). PRCV also uses porcine APN as a receptor; thus, virus binding to the receptor is not sufficient to explain the differences in tissue tropism between TGEV and PRCV. APN is a member of the membrane-bound metallopeptidase family and is widely distributed on diverse cell types; it is highly expressed on the brush border membrane of enterocytes. Some of the antibodies against human APN can block HCV-229E binding (Yeager *et al.*, 1992); however, the catalytic site of the protease activity of APN is not required for receptor function, and the inhibitors of APN do not block viral infection (Delmas *et al.*, 1994). Similar to MHV, TGEV infects polarized cells through the apical, but not the basolateral, surface (Rossen *et al.*, 1994). Again, it is not clear whether this is restricted by the differential expression of APN on the different sides of the cells.

TGEV has also been shown to bind to a 200-kDa protein on the surface of the enterocytes on the villi of the small intestine (Weingartl and Derbyshire, 1994). PRCV does not bind to this molecule. Both the temporal expression (mainly in the newborn) and spatial distribution

patterns (on the villi of the gastrointestinal tract) of the 200-kDa protein correspond to the pattern of susceptibility of piglets to TGEV infection. Thus, the expression pattern of this molecule appears to have a better correlation than the porcine APN with the tissue tropism of TGEV. This 200-kDa molecule may be an alternative receptor used by TGEV. The relative functional significance of this molecule and APN as TGEV receptors is not yet clear.

The FIPV strains of FCV and canine coronavirus apparently utilize the APN of feline and canine species, respectively, as receptors (Benbacer *et al.*, 1997). Cross-species utilization of feline APN by coronaviruses of different species (canine, feline, and human) has also been reported (Tresnan *et al.*, 1996). FIPV, however, is unique among coronaviruses in that it causes an antibody-dependent enhancement (ADE) phenomenon (Weiss and Scott, 1981), which is the result of the binding of the virus-antibody complex to Fc receptors on the surface of macrophages, leading to enhanced virus uptake and spread. This ADE phenomenon has been attributed to the S protein-antibody complex (Vennema *et al.*, 1990b; Corapi *et al.*, 1992; Olsen *et al.*, 1992). The Fc receptor may be a co-factor or an alternative receptor for FIPV entry into macrophages. In this regard, the S protein of MHV has been shown to have limited sequence homology with the murine Fc receptor and to have the ability to bind to the Fc fragment of immunoglobulin (Oleszak and Leibowitz, 1990; Oleszak *et al.*, 1992). Whether the Fc receptor plays a role in MHV infection is not clear. However, MHV does not exhibit ADE.

c. Receptors for Other Coronaviruses. Sialic acid (*N*-acetyl-9-*O*-acetylneuraminic acid)-containing glycoproteins are probably a component of the cell surface molecules required for BCV and HCV-OC43 infection because the removal of sialic acids inhibits BCV infection and resialylation restores virus infectivity (Schultze and Herrler, 1992); however, it is unlikely that it is the primary receptor molecule used by these viruses since the distribution of these molecules is more widespread than the susceptible target cells. The identity of the specific receptor for these viruses has not been determined. For HCV-OC43, it has been shown that the virus binds to a major histocompatibility complex class I molecule (Collins, 1994). However, the receptor function of this molecule has not been established.

C. Penetration and Uncoating

The mechanism of coronavirus entry into target cells has been controversial. Early electron microscopic studies visualized virus (MHV and

IBV) particles inside lysosome-like vesicles near plasma membranes, suggesting that virus enters cells by endocytosis ("viropexis") (David-Ferreira and Manaker, 1965); however, other studies suggested that virus enters cells by direct fusion between virions and the plasma membrane (Doughri *et al.*, 1976). Lysosomotropic drugs, such as ammonium chloride and chloroquine, inhibited MHV-3 virus entry (Krzystyniak and Dupuy, 1984). Also, MHV-specific antibodies did not lyse virus-infected cells during the virus-entry process, as would be the case if the virus fused with the cell membrane (Krzystyniak and Dupuy, 1984). These results suggested that MHV-3 enters cells by an endocytotic pathway. Similar studies using the A59 strain of MHV, however, showed that ammonium chloride delayed, but did not inhibit, the viral infection of L-2 cells (Mizzen *et al.*, 1985). The effects of ammonium chloride were interpreted to be inhibiting virus uncoating in this case. Recent studies by the same group have further shown that only a small proportion of adsorbed virus enters cells by the endocytotic pathway since ammonium chloride, chloroquine, and dansylcadaverine, all of which inhibit receptor-mediated endocytosis, did not have significant effects on MHV entry (Kooi *et al.*, 1991). The majority of MHV particles enter cells by virus-cell fusion at the plasma membrane. This interpretation is consistent with the finding that the optimum pH for MHV-induced cell fusion is 7.4 (Weismiller *et al.*, 1990; Kooi *et al.*, 1991), rather than the acidic pH expected for a virus that enters cells by an endocytotic pathway (e.g., vesicular stomatitis virus). The optimum pH for BCV- and IBV-induced cell fusion is also neutral (Payne and Storz, 1988; Li and Cavanagh, 1992). These findings suggest that coronavirus enters cells by virus-cell fusion at the plasma membrane. On the other hand, virus internalization by endocytosis may be a nonproductive mechanism which does not depend on virus-receptor interaction, since some MHV-resistant cell lines can internalize MHV particles as efficiently as susceptible cell lines (Kooi *et al.*, 1991). Most surprisingly, even Vero cells, an African monkey kidney cell line which presumably does not have an MHV receptor, can internalize virus (Kooi *et al.*, 1991). Therefore, it is likely that MHV enters cells by both acidic-pH-dependent (endocytosis) and -nondependent pathways (Kooi *et al.*, 1991). The exact mechanism of virus entry may depend on cell types and virus strains. Interestingly, an MHV variant which has mutations in the S protein has an acidic optimum pH of 5.5–6.0, in contrast to the pH of 7.5 for the parental virus (Gallagher *et al.*, 1991). This virus variant probably enters cells by an endocytic pathway, a fact supported by the finding that infection of this variant virus is inhibited by ammonium chloride or chloroquine.

What triggers virus internalization after virus-receptor binding is not clear. It has been shown that a conformational change in the S protein could be induced at pH 8.0 and incubation at 37°C (Sturman *et al.*, 1990). Whether this represents the expected conformational change following virus-receptor binding is not clear. Irrespective of the mechanism of virus internalization, fusion between the viral envelope and cell membrane must occur, either at the cell surface or in the endosome, for viral infection to take place. Virus-induced cell-cell fusion has been used to investigate the ability of a virus to induce fusion. Early studies with MHV indicated that virus-induced fusion from without (caused by virions at the cell surface) or fusion from within (caused by *de novo* synthesized S protein on the cell surface) required cleavage of the S protein (Frana *et al.*, 1985; Sturman *et al.*, 1985). Work on BCoV supported this view (Payne and Storz, 1988; Storz *et al.*, 1991). However, more recent experiments involving the expression of S protein (de Groot *et al.*, 1989; Stauber *et al.*, 1993; Taguchi, 1993) and studies of MHV fusion mutants (Gombold *et al.*, 1993) have indicated that uncleaved S can cause syncytium formation, though less efficiently than the cleaved S. Of course, coronaviruses such as TGEV, which have no cleaved S protein, are infectious, in fact, highly so. Since fusion of the virion envelope with a cell membrane is an essential part of the infection process, these results suggest that TGEV must be able to cause virus-cell fusion. Thus, virus-cell fusion and cell-cell fusion may have different requirements, and, for at least some coronaviruses, S cleavage is not required for the fusion of a virion with a cell membrane. Nevertheless, cleaved S may be more efficient at inducing fusion for some coronaviruses. The concentration of S at the surface of a virion may be higher than at the cell surface, such that even the uncleaved S can induce virion-cell fusion, even though it cannot cause cell-cell fusion. Virus-receptor interaction may also trigger a signal transduction pathway to facilitate the internalization of the virus-receptor complex. One study showed that tyrosine kinase is activated in macrophages immediately following MHV-3 infection (Dackiw *et al.*, 1995). It is not yet known whether this is required for virus entry.

The mechanism of virus uncoating, i.e., the release of virion RNA from the nucleocapsid, after the virus has been internalized remains unclear. One study suggested that virus uncoating may involve an endosomal neutral phosphatase, which preferentially dephosphorylates the nucleocapsid protein (Mohandas and Dales, 1991). Furthermore, while immature oligodendrocytes were sensitive to JHM virus infection, differentiated oligodendrocytes were resistant, probably due to a block in virion uncoating (Beushausen *et al.*, 1987). The factors

responsible for the differences in these two types of cells may involve protein kinases (Wilson *et al.*, 1990). Additional cellular factors may be required for viral penetration and uncoating. Various murine cell lines, all of which express virus receptors, show different degrees of susceptibility to infection by different MHV strains (Kooi *et al.*, 1988; Asanaka and Lai, 1993; Yokomori *et al.*, 1993). Cell-cell and virus-cell fusion studies indicated that virus infection is blocked at different stages of virus entry, including penetration and uncoating, in different cell lines (van Dinter and Flintoff, 1987; Asanaka and Lai, 1993). These cell lines could be grouped into at least three complementation groups with respect to the virus entry process (Flintoff, 1984; Asanaka and Lai, 1993). Thus, virus penetration and uncoating appear to require separate cellular factors. It has been suggested from the studies using recombinant viruses between the A59 and JHM strains of MHV that viral S protein may interact with these cellular factors (Yokomori *et al.*, 1993). The nature of these factors is not yet clear.

D. Primary Translation

Following virus uncoating, the first macromolecular synthetic event is predicted to be the synthesis of an RNA-dependent RNA polymerase(s) from the incoming viral genomic RNA, as is the case for all positive-strand RNA viruses. The polymerase is translated from gene 1 at the 5' end of the genomic RNA, most likely directly from the incoming genomic RNA. The process of primary translation has not been observed experimentally. However, inhibitors of protein synthesis applied early in the infection blocked RNA transcription (Mahy *et al.*, 1983; Perlman *et al.*, 1986; Sawicki and Sawicki, 1986), indicating that protein synthesis, most likely the translation of a viral polymerase, is necessary for viral RNA synthesis. This virus-specific polymerase is responsible for the synthesis of negative-strand RNA from the incoming genomic RNA and subsequent transcription of mRNAs from the negative-strand template. The nature of polymerase is discussed in Section IV,A.

Since the genomic-sized RNA is used for both packaging into virus particles to become virion RNA and as an mRNA for protein translation, the distinction between RNA transcription and RNA replication is often blurred. In this review, we will use the term "transcription" to describe the synthesis of subgenomic mRNAs as well as genomic RNA used for translation; the term "replication" will be used to describe the synthesis of the genomic RNA destined to be packaged into virions.

E. Transcription of Viral mRNAs

Coronavirus RNA synthesis occurs via an RNA-dependent RNA transcription process; thus, RNA synthesis can occur in the presence of actinomycin D (with the exception of some coronaviruses, as discussed in Section V,A). The majority of the virus-specific RNAs in the cells are mRNAs, which are transcribed from a negative-strand RNA template. For clarity of discussion, the structure of the mRNAs will be discussed first.

1. The Structure of mRNAs

Coronavirus mRNAs consist of six to eight species of different sizes, depending on the coronavirus species and strains (Lai, 1990). The largest mRNA is equivalent to the genomic RNA, and the remainder are subgenomic in size. These RNAs are designated mRNAs 1 through 7, in order of decreasing size, according to the recommendations of the Coronavirus Study Group in 1989 (Cavanagh *et al.*, 1990). Some mRNAs have been given a hyphenated name, e.g., mRNA 2-1, because they were discovered after the original set of mRNAs was named. They have a nested-set structure, and all of them contain sequences starting at the 3' terminus and extending to various distances toward the 5' end (Stern and Kennedy, 1980b; Lai *et al.*, 1981; Leibowitz *et al.*, 1981). The smallest mRNA contains only the 3' terminal ORF, while each next larger mRNA contains one additional ORF. The structure of the mRNAs in relation to the genome structure is shown in Fig. 7. Thus, except for the smallest mRNA, all of the mRNAs are structurally poly-

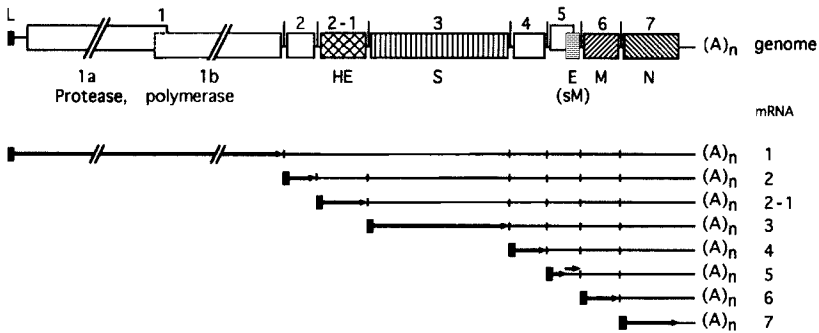


FIG 7. The strategy of transcription and translation of coronavirus (MHV) RNA. The structural relationship between mRNAs and genomic RNA is shown. The arrows indicate the translated portion of each mRNA. Each arrow represents one protein product.

cistronic. In general, each ORF in the genome is represented by an mRNA, whose sequence starts from a consensus signal upstream of the ORF, and only the 5' most ORF of each mRNA can be translated; thus, each mRNA is functionally monocistronic. However, there are exceptions: some mRNAs, e.g., mRNA 5 of MHV and mRNA 3 of IBV, are translated into two or three proteins by different mechanisms (see Section V,G).

Several additional minor mRNA species have been detected, some of which could only be detected by reverse transcription-polymerase chain reaction (La Monica *et al.*, 1992; Schaad and Baric, 1993). These minor RNAs probably represent RNA transcripts from weak or atypical mRNA start signals (see below). Most do not contain a complete ORF at the 5' end; thus, they are probably not functional. Furthermore, in MHV, several mRNAs, e.g., mRNAs 2-1, 2-2, and 3-1, are transcribed only in some virus strains (Shieh *et al.*, 1989; La Monica *et al.*, 1992). The syntheses of these mRNAs appear to be differentially regulated by the sequence at the 5' end of the viral genome (Shieh *et al.*, 1989; La Monica *et al.*, 1992).

Coronavirus mRNAs have another unique structural feature: their 5' ends have a leader sequence of approximately 60–90 nucleotides, which is derived from the 5' end of the genomic RNA (Lai *et al.*, 1982, 1983, 1984; Spaan *et al.*, 1983). The leader sequences of all the mRNAs are identical for a given strain of virus, except for slight variations at some of the leader–mRNA fusion sites, and are identical to the sequence present at the 5' end of the genomic RNA. At the mRNA start sites on the viral genomic RNA, there is a short stretch of sequence that is nearly homologous to the 3' end of the leader RNA (Budzilowicz *et al.*, 1985). This sequence constitutes part of the signal for subgenomic mRNA transcription (Makino *et al.*, 1991). Sequence comparison of viral genomic and mRNAs suggests that subgenomic mRNAs are derived by fusion of the 5' end genomic RNA sequence (leader) to the mRNA start sites on the viral genomic RNA. The mRNA start sites are usually located between the genes; hence, they are termed intergenic (IG) sequences. However, some of the IGs may overlap the coding region of the upstream gene. The core sequence of the IG for MHV is UCUSAAC or a slightly variant form of this sequence at various IG sites (Joo and Makino, 1992). Other virus species also have similar IG sequences.

The leader sequence of MHV ranges in length from 72 to 82 nucleotides, the variation resulting from the heterogeneity of the 3' end sequence, which contains two to four copies of a pentanucleotide (UCUAA) repeat. The homologous nucleotides (UCUAA) at the 3' end of the leader and IG sites serve as fusion sites for the leader and mRNAs. Some of

the MHV mRNAs are heterogeneous, consisting of several subspecies, each containing different copy numbers of the UCUAA repeat (Makino *et al.*, 1988c). This fact suggests that the fusion between the leader RNA and the mRNAs is not very precise.

The length and sequence of the leader RNA in other coronaviruses vary. However, the 3' end of the leader sequence always contains a pentanucleotide UCUAA or a closely related sequence. mRNAs of coronaviruses other than MHV are usually homogeneous in their structure, probably a reflection of the fact that leader RNA at the 5' end of the genome and IG sites in these viruses contain only a single copy of the the UCUAA-like sequence (Hofmann *et al.*, 1993a). The copy number of this pentanucleotide repeat apparently plays an important role in the regulation of mRNA transcription.

2. *The Structure of Negative-Strand RNA*

Coronavirus RNA synthesis is mediated by RNA-dependent RNA synthesis via a negative-strand RNA intermediate (complementary to the genomic RNA). Coronavirus negative-strand RNA represents no more than 1–2% of the total intracellular virus-specific RNA (Perlman *et al.*, 1986; Sawicki and Sawicki, 1986). Both genome-sized and subgenomic negative-strand RNAs, which correspond in number of species and size to those of the virus-specific mRNAs, have been detected (Sethna *et al.*, 1989; Hofmann *et al.*, 1990). The relative molar ratios of the various subgenomic negative-strand RNA species are comparable to those of the positive-strand subgenomic mRNAs. The 5' end of the negative-strand RNA contains poly(U) sequences, which are shorter than the poly(A) sequences present on the positive-strand RNAs (Hofmann and Brian, 1991). At the 3' end of the negative-strand RNA is the complementary sequence of the leader RNA (anti-leader) (Sethna *et al.*, 1991). Structurally speaking, the subgenomic negative-strand RNAs appear to be mirror images of the positive-strand subgenomic mRNAs. All of the negative-strand RNAs in the infected cells are present in the form of double-stranded RNA; no free negative-strand RNA is detected (Perlman *et al.*, 1986).

3. *Kinetics of Viral RNA Synthesis*

In virus-infected cells, virus-specific mRNA synthesis can usually be detected a few hours after infection and throughout most of the viral replication cycle (Stern and Kennedy, 1980a; Leibowitz *et al.*, 1981; Keck *et al.*, 1988a). The molar amounts of the different mRNA species vary; smaller mRNAs are generally more abundant than larger ones, but this rule does not always hold true. Nevertheless, the relative ratio

of different subgenomic mRNA species remains constant throughout, suggesting that the synthesis of the various subgenomic mRNA species is regulated coordinately. Some viruses may show slight variations in the amounts of individual mRNA species present during infection (Hiscox *et al.*, 1995a). Later in infection, there appears to be an enhanced synthesis of the genomic-sized RNA (Keck *et al.*, 1988a).

The kinetics of negative-strand RNA synthesis follows a pattern similar to that of positive-strand mRNA synthesis; however, the peak of negative-strand RNA synthesis appears to occur earlier than for positive-strand RNA (Perlman *et al.*, 1986; Sawicki and Sawicki, 1986). Thereafter, negative-strand RNA synthesis drops significantly, in contrast to that of positive-strand RNA synthesis, and negative-strand RNA appears to be stable (Perlman *et al.*, 1986; Sawicki and Sawicki, 1986). A similar pattern of kinetics of negative-strand RNA synthesis is also seen in the accumulation of the negative-strand RNA of a DI RNA, which very rapidly reaches a steady-state level after transfection (Lin *et al.*, 1994). Therefore, the negative-strand RNA probably functions as a template for multiple rounds of positive-strand RNA synthesis. This conclusion is supported by the study of a *ts* mutant defective in negative-strand RNA synthesis (Schaad and Baric, 1994). However, the ability to synthesize negative-strand RNA seems to be maintained throughout the viral life cycle, as evidenced by the finding that a transfected DI RNA can replicate even when transfected late in the infection (Jeong and Makino, 1992).

4. Mechanism of mRNA Synthesis

Since all subgenomic RNAs consist of a leader RNA derived from the 5' end of the genome and a body sequence derived from various downstream sequences, they must be synthesized by fusion of two discontinuous sequences either during or after transcription. An early study showed that the leader sequence of each mRNA can be exchanged freely between two coinfecting viruses, suggesting that the leader RNA and mRNAs are transcribed independently and can conjoin in a random fashion (Makino *et al.*, 1986b). More recent studies using DI RNA constructs that contain an inserted mRNA start signal (see below) established that the leader RNA and mRNAs are usually derived from two separate RNA molecules (Jeong and Makino, 1994; Zhang *et al.*, 1994b). These studies unequivocally showed that coronaviral mRNA synthesis is carried out by either a discontinuous transcription or trans-splicing process, which fuses sequences from two different RNA molecules. Several transcription models have been proposed, each of which is consistent with some of the experimental data. These models are

not mutually exclusive, as components of each model may operate at different stages of the viral replication cycle. Before presenting these models, we will discuss several findings pertinent to coronaviral RNA transcription.

1. Coronavirus replication takes place entirely in the cytoplasm. Nuclear functions are believed not to be required for RNA synthesis (Brayton *et al.*, 1981; Wilhelmsen *et al.*, 1981); thus, viral RNA transcription does not involve the conventional RNA splicing machinery present in the nucleus.
2. Early ultraviolet (UV) transcriptional mapping studies indicated that in the late stage of viral replication, the UV target size of each subgenomic and genomic mRNA is approximately equivalent to the physical size of the respective mRNA (Jacobs *et al.*, 1981; Stern and Sefton, 1982a); thus, each mRNA is transcribed independently rather than derived by the processing of a large precursor RNA. However, early in infection, the UV target sizes of the subgenomic mRNAs were found to be equivalent to that of the genomic RNA (Yokomori *et al.*, 1992b); thus, at least early in infection, the synthesis of a genomic-length RNA is required for subgenomic mRNA synthesis, although it is not clear whether this requirement is for a positive- or a negative-stranded, full-length RNA. A more recent analysis of the UV target sizes of subgenomic mRNAs of MHV suggested that, even late in the infection, the UV target sizes of some subgenomic mRNAs are slightly larger than their physical lengths but smaller than genomic size (den Boon *et al.*, 1995). Similar observations were made for equine arteritis virus (an arterivirus). This recent result is consistent with either of two interpretations: (a) the subgenomic mRNAs are derived from a slightly longer RNA template or (b) they are derived from a mixture of templates of different sizes (genomic as well as subgenomic). The difference in UV target size between the early and late stages of viral RNA replication suggests that different mechanisms of RNA synthesis may operate at the different stages of the viral replication cycle.
3. The molar ratios of different subgenomic mRNA species and those of subgenomic negative-strand RNAs are similar (Sethna *et al.*, 1989; Hofmann *et al.*, 1990), suggesting that subgenomic mRNAs and subgenomic negative-strand RNAs are derived from each other or under the same transcriptional regulation.
4. The leader RNA at the 5' end of each mRNA is identical in each mRNA and to the leader RNA at the 5' end of genomic RNA. Furthermore, there is sequence homology between the 3' end of the leader

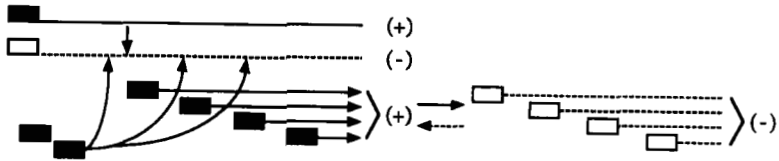
RNA and the mRNA start sites on the genomic RNA (Budzilowicz *et al.*, 1985), where the leader sequence is fused to the mRNAs. There is some sequence divergence between the leader RNA and some of the mRNA start sites; in these cases, the leader RNA of the resulting mRNAs usually mimics the sequence of the mRNA start site rather than the leader at the 5' end of the genome. This finding was used to suggest the possible presence of RNA proofreading activity during coronavirus transcription (Lai, 1986, 1990; van der Most *et al.*, 1994).

The following transcriptional models (Fig. 8) address the possible mechanism of fusion between the leader sequence and mRNAs. Most of the experimental evidence came from MHV studies. The exceptions will be noted.

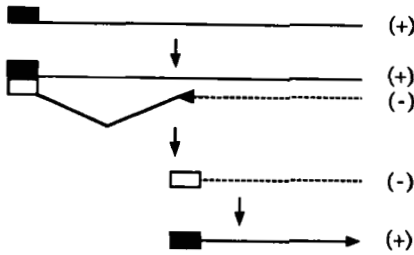
a. Leader-Primed Transcription. This model proposes that the virion genomic RNA is first transcribed into a genomic-length, negative-strand RNA, which, in turn, becomes the template for subsequent subgenomic mRNA synthesis. The leader RNA is transcribed from the 3' end of the negative-strand RNA and dissociated from the template. The free RNA subsequently associates with the template RNA at various mRNA start sites and serves as a primer for transcription of mRNAs. It is proposed that the discontinuous transcription step takes place during positive-strand RNA synthesis. Several pieces of evidence are compatible with this model:

1. Several leader RNAs of approximately 50–90 nucleotides have been detected in the cytoplasm of MHV-infected cells (Baric *et al.*, 1985). Some of these are dissociated from the template RNA and, thus, may serve as a potential source of primers in this transcription model. These RNAs have distinct sizes which are reproducible from cell to cell (Baric *et al.*, 1987); however, they are not exactly the same size as the leader sequence present in the subgenomic mRNAs. Thus, these free leader RNAs must be processed before they are incorporated into mRNAs.
2. A temperature-sensitive mutant of MHV, which synthesizes leader RNA but not mRNAs at the nonpermissive temperature, has been isolated (Baric *et al.*, 1985). The isolation of this mutant suggests that MHV mRNA synthesis is discontinuous, requiring different viral proteins for the synthesis of leader RNA and mRNAs. Thus, a distinction can be made between leader RNA synthesis and mRNA synthesis.
3. During mixed infections with two different MHV strains, as much as 40–50% of the leader sequence on the subgenomic mRNAs of one

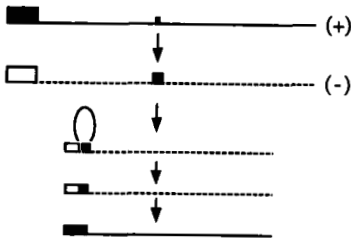
I. Leader-primed transcription:



II. Polymerase jumping during (-) RNA synthesis



III. Trans-splicing of (-) RNA



IV. Amplification of virion-associated mRNAs

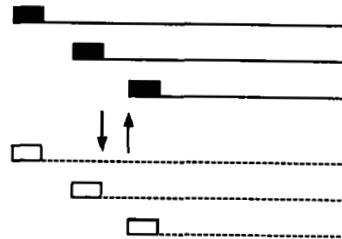


FIG 8. Proposed models of coronavirus mRNA transcription. The solid lines represent positive-strand RNA, the broken lines negative-strand RNA. Boxes represent the leader RNA.

of the viruses is derived from the other coinfecting virus (Makino *et al.*, 1986b). This result suggests that the leader sequence and body sequence of each mRNA are derived from two separate pools. This phenomenon is reminiscent of the RNA reassortment that occurs in RNA viruses with segmented RNA genomes. This result is

best explained by the possibility that free leader RNAs participate in viral RNA synthesis.

4. In an *in vitro* transcription system utilizing cytoplasmic extracts from MHV-infected cells, exogenous leader RNAs can be utilized for mRNA synthesis (Baker and Lai, 1990). The exogenous leader RNA was incorporated into the subgenomic mRNAs at a site that matched precisely that of the endogenous leader RNA present in the viral subgenomic mRNAs, regardless of the length of the exogenous leader RNA used, suggesting that the exogenous leader RNA sequence was processed before being incorporated into mRNAs. Furthermore, the truncated leader RNA which lacked the 3' end UCUAA sequence could not be incorporated into mRNAs, suggesting the importance of this sequence in transcription (Baker and Lai, 1990).
5. The leader RNA sequence, specifically the copy number of the UCUAA repeats at the 3' end of the leader RNA, can affect the transcription of some viral subgenomic mRNAs. For example, whereas an MHV strain containing two UCUAA repeats transcribes mRNA 2-1, a strain with three UCUAA repeats does not, despite identical sequences in the mRNA start sites of these two viruses (Shieh *et al.*, 1989; Yokomori *et al.*, 1991; La Monica *et al.*, 1992). This finding suggests that the leader RNA plays an essential role in the regulation of mRNA transcription.

According to this model, the free leader RNA binds to the mRNA start site (IG) of the full-length negative-strand template via the complementary sequences between the 3' end of the leader (positive-strand) and the IG site of the template RNA (negative-strand) and serves as the primer for RNA transcription. The free leader RNA (primer) may be longer than the leader sequence in the subgenomic mRNAs. There are certain mismatched nucleotides between the leader and template at some mRNA start sites; in the latter case, sequences in the mature mRNAs usually match those of the template instead of the leader. Therefore, the free leader RNA probably undergoes 3' end cleavage before transcription starts to remove the leader nucleotides that are not complementary to the template RNA (Lai, 1986, 1990; van der Most *et al.*, 1994). Transcription is then initiated from the 3' end of the processed leader RNA.

This model is consistent with most of the sequence data of mRNAs. It also explains the curious finding that some mRNAs of MHV are heterogeneous in the copy number (from two to four) of the pentanucleotide (UCUAA) repeats at the leader-mRNA fusion site (Makino *et al.*, 1988c). This heterogeneity is best explained by the imprecise binding

between the leader RNA and template RNA due to the presence of multiple copies of UCUAA (Lai, 1990). Indeed, BCV, which contains only one copy of UCUAA in both the 5' leader and IG sites, does not show this type of heterogeneity in its mRNAs (Hofmann *et al.*, 1993a).

Some recent data, however, cannot be explained by this RNA sequence-homology-driven transcription model. A particular MHV strain (MHV-2C), which has four copies of the UCUAA in the leader RNA, synthesizes some subgenomic mRNAs that are very heterogeneous in length and in leader-mRNA fusion sites (Zhang and Lai, 1994). The sequence data of its mRNAs showed that the leader RNA of this virus is randomly fused to sites where no sequence homology exists between the leader and fusion sites (Zhang *et al.*, 1994b). A similar though less conspicuous heterogeneity in the leader-mRNA fusion sites has also been observed in another MHV strain in a DI RNA-based transcription system (see Section V,E,5) (van der Most *et al.*, 1994). These findings suggest that the sequence complementarity between the leader RNA and IG sites may not be the driving force for mRNA transcription. Thus, a modified version of the leader-primed transcription model proposes that the UCUAAAC sequence provides a recognition signal for viral polymerases and viral or cellular transcription factors. These proteins bind to the leader and IG sites of the template RNA, and the subsequent RNA-protein and protein-protein interactions result in the formation of a transcription complex to initiate mRNA transcription and effect leader-mRNA fusion (Lai *et al.*, 1994; Zhang and Lai, 1995).

The salient feature of this model is that the discontinuous transcription step occurs during positive-strand RNA synthesis; thus, transcriptional regulation is exerted mainly during positive-strand RNA synthesis. This is consistent with current knowledge of the regulation of MHV RNA synthesis. It has been shown that MHV mRNA transcription requires multiple cis-acting RNA sequences (see Section V,E,5). In contrast, the initiation of negative-strand RNA synthesis requires only the 3' end 55-nt plus poly(A) (Lin *et al.*, 1994). Thus, most of the regulatory elements appear to regulate positive-strand RNA synthesis. Since the free leader RNA is the centerpiece of this transcription model, it readily explains why the leader RNA from a different virus can be utilized freely in trans during mixed infections (Makino *et al.*, 1986b). However, this model does not explain the finding that subgenomic replicative-intermediates (RI) and replicative-form (RF) RNAs were detected and were functional during viral RNA synthesis (Sawicki and Sawicki, 1990; Schaad and Baric, 1994) (see Section V,E,4,b). It is possible that the subgenomic mRNAs synthesized can be transcribed

into subgenomic negative-strand RNAs, which, in turn, become the templates for mRNA transcription at a later stage in the viral replication cycle. This would explain why the UV target sizes for mRNAs are nearly equivalent to the physical sizes of mRNAs late in the infection and yet are equivalent to the genomic-sized RNA early in the infection (Yokomori *et al.*, 1992b).

b. Discontinuous Transcription During Negative-Strand RNA Synthesis. In contrast to the leader-primed transcription model, this model proposes that the discontinuous transcription step occurs during negative-strand RNA synthesis, generating subgenomic negative-strand RNAs, which then serve as templates for subgenomic mRNAs in uninterrupted transcription. This model was proposed to account for the detection of subgenomic negative-strand RNAs (Sethna *et al.*, 1989; Hofmann *et al.*, 1990) and subgenomic RIs (Sawicki and Sawicki, 1990) in virus-infected cells. In this model, IG (mRNA start site) sequences on the genomic RNA serve as termination or pausing signals for negative-strand synthesis (Konings *et al.*, 1988), and the nascent subgenomic negative-strand RNA then jumps to the leader RNA sequence at the 5' end of the genomic RNA by an unknown mechanism to continue RNA synthesis. As a result, the nascent negative-strand subgenomic RNA fuses with the negative-strand leader sequence, generating a subgenomic negative-strand RNA that contains an anti-leader sequence at its 3' end and a poly(U) sequence at its 5' end (Hofmann and Brian, 1991; Sethna *et al.*, 1991). Structurally, these negative-strand RNAs are mirror images of the subgenomic mRNAs and, thus, can potentially serve as a template for uninterrupted transcription of subgenomic mRNAs.

In this model, the regulation of subgenomic mRNA transcription would be exerted on negative-strand instead of positive-strand RNA synthesis. This model is consistent with the following observations:

1. Subgenomic negative-strand RNAs have been detected in virus-infected cells (Sethna *et al.*, 1989; Hofmann *et al.*, 1990). These RNAs have structures that are mirror images of those of the completed subgenomic mRNAs. The relative molar ratios of the different subgenomic negative-strand RNAs are similar to those of the corresponding viral mRNAs (Sethna *et al.*, 1989; Hofmann *et al.*, 1990).
2. Subgenomic RI RNAs have been detected in virus-infected cells later in the infection (Sawicki and Sawicki, 1990). The smaller RIs were precursors of the smaller mRNAs and the larger RIs generated the larger mRNAs, suggesting that each subgenomic mRNA was transcribed from the corresponding subgenomic-sized negative-strand

template (Sawicki and Sawicki, 1990). Another study, which analyzed the subgenomic RFs of a temperature-sensitive mutant of MHV, also suggested that subgenomic negative-strand RNAs are functional (Schaad and Baric, 1994); although, in this study, RIs were not directly examined.

3. The UV targets for subgenomic mRNA synthesis at the later stage of viral replication are subgenomic in length (Jacobs *et al.*, 1981; Stern and Sefton, 1982a; Yokomori *et al.*, 1992b), roughly corresponding to the physical lengths of each subgenomic mRNA, suggesting that the templates for these mRNAs are subgenomic.
4. In DI RNA systems (see Section V,E,5), when multiple IG sequences were present, the sequences in the 3' end often had a higher transcription efficiency than those at the 5' end, consistent with the proposal that IGs serve as transcriptional termination sites, which impede the elongation of the negative-strand RNAs (Van Marle *et al.*, 1995; Krishnan *et al.*, 1996). However, in some cases, the higher transcription efficiency of the 3' proximal IG was observed only when the neighboring IGs were very close together, suggesting a spatial constraint rather than sequential interference (Joo and Makino, 1995).

This model, however, cannot explain why the UV targets for subgenomic mRNA synthesis early in infection are of genomic size (Yokomori *et al.*, 1992b) and why, later in the infection, the targets for these same mRNAs are still larger than the respective subgenomic mRNAs but not longer than genomic size (den Boon *et al.*, 1995). It also cannot explain why the nature of the leader sequence can regulate differential transcription of various mRNA species, such as mRNA 2-1 of MHV, inasmuch as the leader sequence on the template RNA is localized downstream of the transcription termination site for negative-strand RNA synthesis. Finally, it is difficult to explain why the leader RNAs are derived in trans.

c. Trans-Splicing of Nascent RNA Transcripts. This model proposes that the full-length positive- or negative-strand RNAs are spliced post-transcriptionally to generate subgenomic RNAs. It was initially considered unlikely because of the findings that coronavirus replicates in the cytoplasm rather than in the nucleus (Brayton *et al.*, 1981; Wilhelmsen *et al.*, 1981), where the splicing machinery is present, and that UV target sizes of subgenomic mRNAs are equivalent to the physical sizes of subgenomic mRNAs (Jacobs *et al.*, 1981). Furthermore, there are no consensus splicing donor and acceptor sequences in the coronavirus genomic RNAs. However, the trans-splicing model is compatible with

recent findings that early in infection, the UV targets for subgenomic mRNA synthesis are of genomic length (Yokomori *et al.*, 1992b), and that both the leader RNA and IG sequence of MHV negative-strand RNA bind to a cellular factor, heterogeneous nuclear RNP (hnRNP) A1, which is involved in alternative RNA splicing (Zhang and Lai, 1995; H.-P. Li and M. M. C. Lai, unpublished observation). A modified splicing model thus can be proposed as follows: a full-length negative-strand RNA is first synthesized. Components of the splicing machinery derived from the nucleus or cytoplasm then bind to the leader sequence and IG sites on the negative-strand RNA and form a splicing complex. The leader and IG can be derived from different RNA molecules. Splicing between the leader and IG generates a subgenomic negative-strand RNA. Once the spliced subgenomic negative-strand RNAs are generated, they are used as templates for subsequent mRNA synthesis. Later in infection, even the subgenomic negative-strand RNAs may be able to participate in RNA splicing to generate smaller subgenomic negative-strand RNAs because they themselves also contain the leader and IG sequences. This model may thus explain why the UV target for mRNA transcription is of genomic length early in infection (Yokomori *et al.*, 1992b) and may shed light on the recent puzzling finding that later in infection, the UV target sizes are still larger than the actual sizes of the subgenomic mRNAs (den Boon *et al.*, 1995). It also explains the functional roles of subgenomic RIs (Sawicki and Sawicki, 1990). This potential splicing, however, must be different from conventional RNA splicing because it occurs in the cytoplasm, and the splicing donor and acceptor sequences must also be different from the conventional ones. Since some of the splicing factors are probably derived from the nucleus, this model predicts that nuclear functions are involved in MHV RNA transcription.

d. Amplification of Virion-Associated Subgenomic RNAs. Based on the findings that some coronaviruses, including BCV, TGEV, and IBV (Sethna *et al.*, 1989; Hofmann *et al.*, 1990; Zhao *et al.*, 1993), contain subgenomic mRNAs in the virion, probably as a result of nonspecific RNA packaging, it was proposed that these virion-associated subgenomic mRNAs can be used directly as templates for the synthesis of subgenomic negative-strand RNAs, which, in turn, serve as templates for the synthesis of additional subgenomic mRNAs (Sethna *et al.*, 1989). This model may explain the presence of subgenomic negative-strand RNAs and RIs in the infected cells, but it cannot explain the genomic-length nature of the UV target sizes for mRNA synthesis early in infection (Yokomori *et al.*, 1992b), nor can it explain how leader RNAs from different virus strains can be randomly incorporated into mRNAs

of a different virus. Furthermore, the virion-associated subgenomic mRNAs have not been detected in all coronavirus species.

The available data cannot unequivocally rule out any of the proposed transcription models. The primary difficulty in experimental analysis is that once the subgenomic mRNAs are synthesized, by whatever mechanism, they are transcribed into negative-strand RNAs because the cis-acting signal for negative-strand RNA synthesis in MHV resides in the 55 nucleotides at the 3' end plus poly(A) (Lin *et al.*, 1994), which is present in every subgenomic RNA. Thus, it is difficult to separate the primary and secondary events of transcription. It is possible that these transcription models are not mutually exclusive. For example, early in infection, a leader-primed transcription or trans-splicing mechanism may operate, generating subgenomic mRNAs, which are then amplified into subgenomic negative-strand RNAs; the latter serve as templates for further amplification of subgenomic mRNAs thereafter. The subgenomic negative-strand RNA can be used for either uninterrupted transcription or leader-primed transcription to generate positive-strand subgenomic RNAs. A combination of these models would be consistent with most of the experimental data. This two-step model of primary and secondary transcription (Jeong and Makino, 1992) may explain the apparent differences in the possible mechanism of transcription between early and late stages of viral infection.

5. Cis- and Trans-Acting Signals for Transcription as Revealed by DI RNA Vectors

Because of the large size of coronavirus RNA, no infectious cDNA or RNA clones are now available for reverse genetics studies. This difficulty has hampered progress in the study of the molecular biology of coronaviruses. DI RNAs of several coronaviruses (see Section VI,E) have been molecularly cloned and used as a substitute for the genomic RNA to study the cis- and trans-acting signals involved in viral RNA synthesis. Although natural DI RNAs do not contain an mRNA start signal and, consequently, cannot transcribe an mRNA, the insertion of such a signal into the DI RNA allows an mRNA to be transcribed from the transfected DI RNA in the virus-infected cells, thus enabling studies of the regulatory sequences for transcription.

Following is a summary of information that has been obtained using this approach. It should be cautioned, however, that regulation of RNA transcription probably depends on overall RNA conformation and that the cis-acting sequence required for RNA synthesis very often varies with the DI RNA vector used; therefore, the results obtained from DI RNA studies may not be directly applicable to the viral genome. A full

understanding of the regulation of viral RNA synthesis still awaits the development of an infectious cDNA clone.

The following cis-acting signals for coronavirus RNA transcription have been determined primarily from MHV DI RNA studies (with some from BCV DI) (Fig. 9).

a. IG Sequence. The IG sequence can be considered to be the promoter element for transcription. It also serves as the mRNA start site and the site of fusion between the leader RNA and body sequence of mRNAs. A seven-nucleotide core sequence, UCUAAAC, is sufficient to initiate mRNA synthesis (Makino *et al.*, 1991). Extensive site-specific mutagenesis studies have shown that most of the single-nucleotide mutations within this core sequence could be tolerated, although the transcription efficiency of some of these mutants was lower (Joo and Makino, 1992; van der Most *et al.*, 1994). These seven nucleotides represent the minimum promoter; deletion of a nucleotide results in complete ablation of mRNA transcription. The effects of the sequences near the promoter on transcription are contradictory: in certain situa-

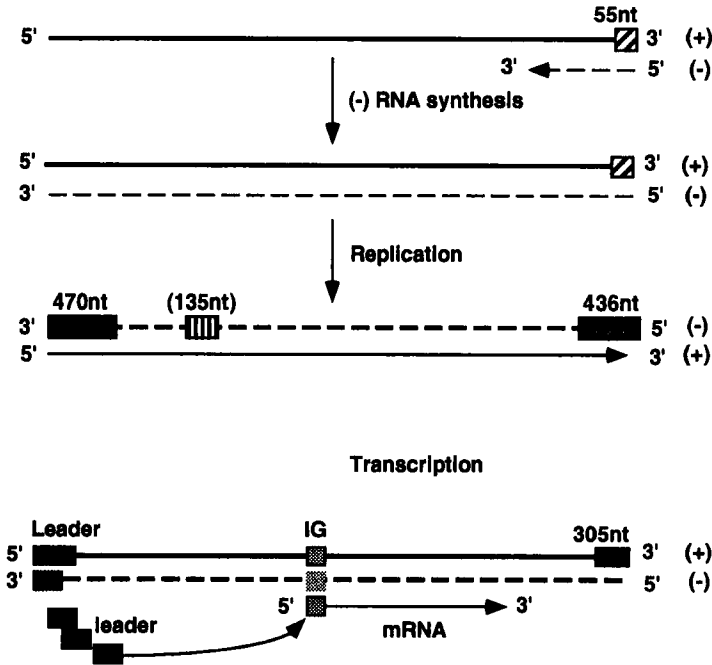


FIG 9. Cis-acting signals for various steps of MHV DI RNA synthesis. The boxed regions represent the cis-acting signals for the indicated steps of RNA synthesis.

tions, the nature of the neighboring sequences did not affect transcription (Makino and Joo, 1993), but under other circumstances, it did (Jeong *et al.*, 1996). Thus, the strength of the promoter appears to depend on the context of the overall RNA sequence and structure.

The relative flexibility of sequence requirement of the promoter sequence in the DI RNA system appears to differ significantly from that seen in the viral genomic RNA. In the MHV genome, there are more than 20 stretches of sequence resembling the UCUAAAC sequence, in addition to the six promoters for the known subgenomic mRNAs (Joo and Makino, 1992). Yet, most of these did not promote mRNA synthesis from the viral RNA genome to any appreciable extent, in contrast to their ability to promote transcription in the DI RNA vector system (Joo and Makino, 1992). In the viral genome, the single-nucleotide substitution of a G residue in the core promoter sequence completely abolished mRNA synthesis (Shieh *et al.*, 1989), whereas this is tolerated in the DI RNA (Joo and Makino, 1992). Thus, there appear to be significant differences between the sequence requirement for mRNA synthesis in the DI RNA and in the natural viral genomic RNA. When there are multiple IG sequences in the DI RNA, the order of the IG sequences may influence transcriptional efficiency. An IG located at the 3' end generally has an advantage in initiating mRNA synthesis (Van Marle *et al.*, 1995; Krishnan *et al.*, 1996). The sequences near the IGs may suppress transcription (Jeong *et al.*, 1996).

b. The Leader Sequence at the 5' End of the DI RNA. The leader sequence at the 5' end of the viral genomic RNA becomes the leader sequence of subgenomic mRNAs; thus, it fills a structural role for mRNA synthesis. However, the leader RNA of the subgenomic mRNAs is not derived exclusively from the leader RNA of the same (DI) RNA; in fact, most are derived in trans from a separate RNA molecule, such as helper virus RNA (Jeong and Makino, 1994; Liao and Lai, 1994; Zhang *et al.*, 1994b). Nevertheless, mRNA transcription from an IG site in the DI RNA still requires the presence of a leader RNA sequence at the 5' end of the DI RNA as a cis-acting sequence (Liao and Lai, 1994). Deletion of this cis-acting leader abolished transcription. Furthermore, the sequence of this leader RNA, particularly its 3' end sequence, can affect the efficiency of transcription from certain IG sequences on the DI RNA (Zhang *et al.*, 1994b). For example, the leader RNA containing two pentanucleotide (UCUAA) repeats transcribes an mRNA from the IG 2-1 site more efficiently than the leader RNA with three UCUAA repeats. Thus, the cis-acting leader RNA plays a role similar to that of an enhancer. These findings suggest that the leader RNA serves two functions (Liao and Lai, 1994): (1) it supplies the leader RNA to the

subgenomic mRNAs, and (2) it serves as an enhancer-like sequence to regulate transcription. This finding also suggests that there is either a direct or an indirect interaction between the leader and IG sequences.

Some additional sequences downstream of the leader may also enhance transcription from an IG site in the DI RNA (Liao and Lai, 1994); however, the precise sequence requirement is not known. This sequence requirement shows some virus sequence specificity, since it cannot be replaced with other viral RNA sequences (Liao and Lai, 1994). It may be needed to maintain overall RNA conformation for the recognition of the IG sequence.

c. The 3' UTR. In an MHV DI RNA construct, partial deletion of the 3' UTR completely abolished transcription from an upstream IG site in the DI RNA (Lin *et al.*, 1996). This stretch of 3' UTR is probably involved in positive-strand RNA synthesis, since the length of this required sequence (305 nt) is significantly longer than that required for negative-strand RNA synthesis (55 nt). The 3' UTR requirement for mRNA transcription is surprising, since positive-strand RNA synthesis starts from the 5' end; thus, the 3' end sequence is the last to be transcribed. This 3' UTR sequence requirement is similar to that for RNA replication (Kim *et al.*, 1993b; Lin and Lai, 1993) (see Section V,F). This finding suggests that the 3' end may interact with the 5' end and possibly with IG sequences during transcription.

d. A Nine-Nucleotide Sequence, UUUUAUAAAC. This sequence, located immediately downstream from the UCUA repeats at the 3' end of the leader RNA in the viral genome, plays a significant role in RNA transcription. It is deleted from the genome of one of the MHV strains and is often deleted in naturally occurring DI RNAs (Lai *et al.*, 1987). In this particular MHV strain (MHV-2C), the leader-mRNA fusion sites are very heterogeneous and do not always occur at the usual UCUAAC sites (Zhang *et al.*, 1994b). This nine-nucleotide sequence can serve as an mRNA start signal, allowing transcription of an almost genomic-length mRNA (Zhang and Lai, 1996). In the DI RNAs, the presence or absence of this nine-nucleotide sequence influences transcription efficiency from the downstream IG site and, most importantly, affects the source of the leader RNA incorporated into subgenomic mRNAs (Zhang *et al.*, 1994b). When this nine-nucleotide sequence is present, the leader sequence in the subgenomic mRNAs is contributed both from the DI RNA in *cis* and from helper virus RNA in *trans*. When this sequence is missing, the leader RNA is derived exclusively from the helper virus RNA (Zhang *et al.*, 1994b). Thus, this nine-nucleotide sequence appears to regulate the mechanism by which the leader RNA is fused to the subgenomic mRNAs.

These results combined suggest that multiple RNA regions are involved in the regulation of mRNA transcription. However, a recent study appears to contradict the need of cis-acting sequences other than the IGs for mRNA transcription. When a negative-strand RNA containing only an IG sequence of TGEV and a reporter gene was transcribed *in situ* from a transfected cDNA by using a recombinant vaccinia virus-T7 RNA polymerase expression system, this RNA was transcribed in the presence of TGEV, generating an mRNA with a correctly fused TGEV leader sequence (Hiscox *et al.*, 1995b). The leader-containing mRNA could have been generated by either of the transcription mechanisms described (Section V,E,4,a or Section V,E,4,b) above. This study suggests that this negative-strand IG site is sufficient for transcription. However, it is possible that this activity represents a basal level of transcription and that other cis-acting sequences may enhance the efficiency of transcription.

6. Proteins Involved in RNA Synthesis

The application of inhibitors of protein synthesis at any time during the viral life cycle inhibits viral RNA synthesis, suggesting that continuous protein synthesis is required for RNA synthesis (Perlman *et al.*, 1986; Sawicki and Sawicki, 1986). A similar observation has been made using an inhibitor of cysteine proteases, which inhibits a specific step of the processing of gene 1a products of MHV (Kim *et al.*, 1995) (see Section V,G,2), suggesting that continuous production of polymerase is required for viral RNA synthesis. The precise nature of the viral proteins involved has yet to be determined. Temperature-sensitive mutants of MHV that are defective in RNA synthesis at the nonpermissive temperature have been divided into at least five complementation groups, indicating that at least five proteins are involved in viral RNA synthesis (Leibowitz *et al.*, 1982a; Baric *et al.*, 1990) (see Fig. 6). All of these complementation groups are mapped within the gene 1 region (including both 1a and 1b). Sequence analysis showed that gene 1b contains an RNA polymerase motif (Gorbalenya *et al.*, 1989b; Lee *et al.*, 1991). Polymerase activities have been demonstrated in membrane fractions of BCV- and MHV-infected cells (Brayton *et al.*, 1982, 1984; Dennis and Brian, 1982), and several *in vitro* RNA synthesis systems have been reported (Compton *et al.*, 1987; Leibowitz and DeVries, 1988; Baker and Lai, 1990); however, the nature of polymerases in these systems has not been identified. In one study, it was demonstrated that the antibodies against the N protein could inhibit RNA synthesis, suggesting that N protein may be involved in RNA synthesis (Compton *et al.*, 1987).

In addition to viral proteins, cellular factors may also be involved in RNA synthesis. Several cellular proteins have been shown to bind to the regulatory elements of MHV RNA, including the 5' and 3' ends of the genomic RNA and the 3' end of the negative-strand RNA and IG sites (Furuya and Lai, 1993; Yu and Leibowitz, 1995a,b; Zhang and Lai, 1995). The binding sites for the cellular proteins at the 5' end of genomic RNA and the 3' end of negative-strand RNA are complementary (Furuya and Lai, 1993). The protein p35, which binds to the negative-strand leader sequence and the IG site, is particularly interesting. Site-specific mutations of the IG site affected the binding of this protein and the efficiency of RNA transcription to the same extent, suggesting that the binding of this protein is required for RNA transcription (Zhang and Lai, 1995). This protein recently has been identified as hnRNP A1 (H.-P. Li and M. M. C. Lai, unpublished observation). The mutations at the 3' end of the viral genomic RNA that abolished the binding of cellular proteins also inhibited both negative-strand and positive-strand RNA synthesis, although the correlation between protein binding and RNA replication was not absolute (Yu and Leibowitz, 1995a). Thus, cellular proteins probably play a significant role in viral RNA replication and transcription. Curiously, viral proteins in the infected cell extract could not be cross-linked to the viral RNA *in vitro*, suggesting that viral proteins may interact with viral RNA only indirectly through cellular proteins. This is in contrast to the finding that the purified N protein can bind to the leader RNA sequence *in vitro* (Baric *et al.*, 1988; Stohlman *et al.*, 1988). The reason for this discrepancy is not clear.

F. Replication of Viral Genomic RNA

The genomic-sized RNA in coronavirus-infected cells theoretically consists of two populations: the messenger RNA (mRNA 1), which is translated to yield gene 1a and 1b products, and the genomic RNA, which is destined to be packaged into virion. Early studies demonstrated that, late in the infection, most (95%) of the genomic-sized RNA in the cells was associated with the viral nucleocapsid, while the remainder (5%) was present in polysomes (Spaan *et al.*, 1981; Perlman *et al.*, 1986). Presumably, early in infection, most of the genomic-sized RNA would be associated with polysomes to serve as mRNAs for the synthesis of polymerase; however, this has not been demonstrated. It is not clear whether there is any difference in structure and mechanism of synthesis between these two RNA populations.

Since genomic RNA requires uninterrupted synthesis from the full-length template, whereas mRNAs involve discontinuous transcription, the two types of genomic-sized RNA (mRNA 1 and virion genome RNA) may be synthesized by two different mechanisms. A recent study suggests that at least some of the MHV genomic-sized RNAs are indeed synthesized by a discontinuous transcription, using the UCUAA repeat in the leader RNA and the nine-nucleotide UUUUAUAAAC immediately downstream of the leader RNA as the transcription start site (Zhang and Lai, 1996). This raised the possibility that mRNA 1 and virion genomic RNA are distinguishable. However, it cannot be inferred from this study that the fate of the genomic-sized RNA products derived from discontinuous transcription is different from the fate of those derived from uninterrupted RNA synthesis.

The possible involvement of discontinuous transcription in generating genomic-sized RNA may explain several interesting findings regarding MHV genomic RNA:

1. The copy number of the UCUAA repeat in the leader sequence of the genomic RNA, which ranges from two to four copies in different MHV strains, rapidly evolves during virus passage (Makino and Lai, 1989a; La Monica *et al.*, 1992). Starting with a pure virus population, the copy number in the viral genomic RNA rapidly becomes heterogeneous during serial passages in tissue culture, and a new virus population with a different copy number of UCUAA repeats emerges (Makino and Lai, 1989a). Since this sequence variation is seen in the leader region but not in the IG regions, where uninterrupted RNA synthesis probably occurs, this finding is best explained by the discontinuous transcription mechanism involving the 5' leader region. The imprecise fusion of the leader RNA to the mRNA start sites would result in heterogeneity of the copy number of the UCUAA repeats (Makino *et al.*, 1988c; Lai, 1990). Such heterogeneity is not observed when the virus, e.g., BCV, contains only one UCUAA copy in the leader RNA (Hofmann *et al.*, 1993a).
2. The UCUAA region at the 5' end of the genomic RNA is a hot spot of RNA recombination during mixed infection of MHVs, resulting in recombinant MHVs with a crossover site at the 3' end of the leader RNA sequence (Keck *et al.*, 1987). This result is best explained by the discontinuous RNA synthesis at the 5' end of the genomic RNA.
3. If the generation of DI RNAs is viewed as an anomaly of RNA replication, the structure of naturally occurring DI RNAs reveals an insight into the mechanism of RNA replication. Most of the naturally

occurring MHV DI RNAs have a copy number of the UCUAA repeat different from that of the parental virus, and most lack a nine-nucleotide sequence downstream of the UCUAA repeats (Lai *et al.*, 1987). As discussed above, this is a reflection of discontinuous transcription in the region.

The understanding of the mechanism of RNA replication has been aided by the use of *in vitro*-transcribed DI RNA generated from cloned cDNA. When DI RNA was transfected into virus-infected cells, the leader RNA was rapidly replaced by that of the helper virus (Makino and Lai, 1989b; Chang *et al.*, 1996). This leader exchange is dependent on the presence of the nine-nucleotide sequence (UUUAUAAAC) in the DI RNA (Makino and Lai, 1989b), consistent with the finding that this sequence serves as an mRNA start signal for discontinuous transcription (Zhang and Lai, 1996). The use of the cloned DI RNA also allowed the determination of the cis-acting signals for RNA replication (Kim *et al.*, 1993b; Lin and Lai, 1993). It was shown that more than 400 nucleotides at both the 5' and 3' ends of the DI RNA are required for RNA replication, and that some MHV DI RNAs also required a stretch (130 nt) of internal sequence in the gene 1 region for RNA replication; however, the requirement for the internal sequence was not observed in other MHV or BCV DI RNA constructs (Chang *et al.*, 1994; Luytjes *et al.*, 1996). Thus, this internal sequence probably plays a role in maintaining the overall RNA conformation for some DI RNAs (Y. N. Kim and Makino, 1995). Again, the requirement of a 3' end sequence (436 nt) that is longer than that required for negative-strand RNA synthesis (55 nt) is a surprise. These 3' end sequences are probably required for positive-strand RNA synthesis during RNA replication. This finding is reminiscent of the sequence requirement for RNA transcription discussed above and suggests that there is a direct or indirect RNA-RNA interaction between the 5' and 3' ends during RNA replication. These DI RNA studies also showed that replication of DI RNA is inhibited when an mRNA is transcribed from an IG site within the same DI RNA, and that the mechanism of inhibition is due not to competition for the same transcription machinery (Jeong and Makino, 1992), but most likely to the overlap of the cis-acting signals for these two different processes. However, the sequence requirements for replication and transcription are different, indicating that these two processes are distinguishable.

The mRNA transcription and genomic RNA replication may be regulated by the same mechanism throughout most of the viral replication cycle. However, the ratio between the genomic RNA and subgenomic

RNAs, as detected by radioactive uridine incorporation, increases during the late stages of the BCV replication cycle (Keck *et al.*, 1988a), suggesting a possible switching mechanism from transcription to replication. It has been shown that genomic RNA replication is coupled to encapsidation, since no free genomic RNA is found (Perlman *et al.*, 1986). Since the encapsidation of RNA requires the N protein, this protein may participate in the regulation of switching between transcription and replication.

G. Translation of Viral Proteins

1. Mechanisms of Translation

The sequences of coronavirus mRNAs usually start from a site immediately upstream of a gene. These mRNAs, except for the smallest mRNA, are structurally polycistronic, containing multiple ORFs. Only the 5' most ORF in the mRNAs is translatable; the remaining ORFs are usually functionally silent. Thus, most of these mRNAs are functionally monocistronic (see Fig. 7). The S, HE, M, and N proteins, and in most coronaviruses the E protein, are translated from separate mRNAs by this mechanism; initiation of their translation is unremarkable, utilizing a cap-dependent translation mechanism. Many ns proteins, however, are translated from truly polycistronic mRNAs, i.e., two or three proteins are translated from the same mRNA. For these mRNAs, the first ORF, e.g., 3a of IBV or 5a of MHV, is probably also translated by the same mechanism as the structural protein genes. For internal ORFs, e.g., E protein of IBV and MHV, an alternative mechanism must be employed to initiate translation internally.

One characteristic of coronavirus mRNAs is the presence of the leader RNA sequence at the 5' end, which not only participates in RNA transcription, but also regulates the efficiency of translation. It has been shown that the presence of the MHV leader sequence on a heterologous mRNA in a chimeric RNA construct can enhance its translation in virus-infected cell lysates but not in uninfected cell lysates (Tahara *et al.*, 1994). This effect conceivably will enable the efficient translation of viral mRNAs in the face of shutoff of translation of cellular mRNAs in the infected cells (Siddell *et al.*, 1980; Hilton *et al.*, 1986). The mechanism of translational enhancement by the leader RNA has not been determined. It has been shown that during persistent infection of BCV, the leader RNA sequence underwent frequent mutations (Hofmann *et al.*, 1993b). One of these mutants had an intraleader short ORF and a lower translation efficiency, indicating that the leader sequence in-

deed can modulate translation. Another region which can potentially regulate the translation of coronavirus mRNAs is the 5' UTR (other than the leader sequence) of mRNAs. The genomic RNA (mRNA 1) has a particularly long 5' UTR (200–400 nt). An MHV with a specific point mutation within the 5' UTR was selected during persistent infection *in vitro* (Chen and Baric, 1995). This mutant had a significantly higher translation efficiency than the wild-type virus. Different subgenomic mRNAs had 5' UTR of various lengths, which may also affect their translation.

For the translation of internal ORFs, several different mechanisms are used by coronaviruses:

a. Ribosomal Frameshifting Within the Polymerase Gene. All of the coronavirus genes 1 (polymerase) sequenced so far contain two overlapping ORFs. Several features of the IBV polymerase gene sequence (Bournsnel *et al.*, 1987), coupled with the absence of a distinct mRNA for ORF 1b, suggested that translation of ORF 1b involved ribosomal frameshifting from ORF 1a, thus synthesizing a large polyprotein containing both 1a and 1b sequences. Subsequently, a highly efficient (30% frequency) –1 frameshift was demonstrated experimentally *in vitro* (Brierley *et al.*, 1987; Somogyi *et al.*, 1993) and *in vivo* (Brierly *et al.*, 1990). This mechanism has been shown to operate in gene 1 of MHV, HCV-229E, and TGEV as well (Bredenbeek *et al.*, 1990; Lee *et al.*, 1991; Herold and Siddell, 1993; Eleouet *et al.*, 1995). In all cases, the mechanism involves two essential elements: a slippery site followed by an RNA pseudoknot (Brierley *et al.*, 1989). The site at which the ribosome slips backward has the sequence UUUAAAC. The pseudoknots of IBV and MHV are similar, comprising two base-paired regions stacked coaxially in a quasi-continuous manner and connected by two single-strand loop regions. The HCV-229E pseudoknot is more complex (Herold and Siddell, 1993). It is the overall shape and stability of the pseudoknot that are important, not the nucleotide sequence per se.

Two reasons have been put forward to explain why coronaviruses should employ ribosomal frameshifting to translate ORF 1b (Brown and Brierly, 1995). One reason is that this is done primarily to control the relative amounts of the 1a and 1b products. That could be achieved in other ways, of course, e.g., by translating ORF 1b from a separate mRNA; this will require that the transcription of 1a and 1b mRNAs is tightly regulated. The other reason may be to avoid making a 1b mRNA. Such an mRNA might be packaged into virions in competition with genomic RNA, as the RNA region corresponding to the 1b ORF of MHV contains a sequence that is essential for packaging into virions (Fosmire *et al.*, 1992) (see Section VI,E,1).

b. Internal Initiation of Translation of the IBV and MHV E Protein mRNA. The E proteins of IBV and MHV are encoded by the third and second ORF, respectively, of the corresponding genes 3 and 5 (Fig. 5). Cells infected with IBV contain the products of all three of the gene 3 ORFs (Liu *et al.*, 1991). Both of the MHV gene 5 ORFs are translated *in vitro* (Budzilowicz and Weiss, 1987), but only the 5b ORF product has been detected in virus-infected cells (Leibowitz *et al.*, 1988). Experiments have shown that the E protein ORF of both IBV and MHV mRNAs is translated by a cap-independent, internal ribosomal entry mechanism (Liu and Inglis, 1992b; Thiel and Siddell, 1994). Furthermore, if the 3a and 3b ORFs were eliminated from the IBV mRNA, translation of the 3c (E) ORF did not occur (Liu and Inglis, 1992b). This suggested that the 3a/3b region contains an internal ribosome entry site (IRES) for the E protein ORF. Le and colleagues have predicted the existence of secondary structures in the 3a/3b region of IBV which resemble the IRES elements of picornaviruses (Le *et al.*, 1994). They predicted a 265-nucleotide sequence in 3a/3b which would fold into five stem-loops, forming a compact structure by the interaction of two pseudoknots.

c. Translation of Nonstructural Proteins. In addition to the ns proteins encoded from the 5'-most ORFs of mRNAs, several other ns proteins are encoded from an internal ORF of some viral mRNAs, e.g., 3b of IBV and HCV-229E, 4b of BCV, and 7b of FCV (Fig. 5). Most of these products have been detected in virus-infected cells; however, the mechanism of the internal initiation of translation has not been elucidated.

BCV and MHV RNA contains an additional internal ORF within the N protein gene. This ORF (termed I) is in a different reading frame from that of N protein and encodes a hydrophobic protein (Senanayake *et al.*, 1992; Fischer *et al.*, 1997). This protein is translated in virus-infected cells by a leaky ribosomal scanning mechanism from the bicistronic mRNA of N gene (Senanayake *et al.*, 1992). It is a nonessential gene. The mechanism of its regulation is not yet clear.

2. Posttranslational Processing and Modifications

a. Processing of Pol Proteins 1a and 1b. The gene 1 product is predicted to be nearly 700–800 kDa. It is probably processed into multiple proteins posttranslationally by its own proteases. The processing pathway has just begun to be explored. Remarkably, the protease domains and potential cleavage sites predicted by computer analysis (Gorbalenya *et al.*, 1989b; Lee *et al.*, 1991) have largely been confirmed by experimental data. Initially, *in vitro* translation of virion RNA of MHV

revealed several polypeptides of more than 200 kDa (Leibowitz *et al.*, 1982b; Denison and Perlman, 1986). In addition, a 28-kDa product was detected and shown to have originated by cleavage from the N terminus of a precursor (Denison and Perlman 1986; Soe *et al.*, 1987), now known to be the beginning of the ORF 1a polyprotein (Fig. 6). The cleavage which generates p28 is carried out by PLP 1 (Fig. 6). It cleaves between residues Gly-247 and Val-248, mutation of either residue resulting in almost no cleavage (Dong and Baker, 1994; Hughes *et al.*, 1995). In addition to p28, the MHV ORF 1a encodes a protein of more than 400 kDa, which is cleaved to a 290-kDa product, which, in turn, is cleaved to produce a 50-kDa and a 240-kDa product (Denison *et al.*, 1992) (Fig. 6). Another protein of 65 kDa is derived from sequence immediately downstream of the p28-encoding region, thus representing the N-terminal part of the large polyprotein initially found in *in vitro* translation (probably more than 400 kDa) (Denison *et al.*, 1995) (Fig. 6). The cleavage of p65 from the polyprotein was also carried out by PLP1 (Bonilla *et al.*, 1995, 1997). Inhibition of the C-terminal cleavage of p65 by E64d, an irreversible inhibitor of cysteine (thio) proteinases, inhibited MHV replication (Kim *et al.*, 1995). In addition, the 3CLP domain is cleaved from the polyprotein by the autocatalytic cleavage activity of 3CLP itself to generate a 27–34 kDa protein, which contains both the trans- and cis-acting proteolytic activities (Lu *et al.*, 1995, 1996; Liu and Brown, 1994; Ziebuhr *et al.*, 1995). E64d also inhibited the 3CLP protease activity.

The processing pathway of the 1b protein sequence is less clear. There is experimental evidence with IBV and HCV-229E that the 1b polyprotein is cleaved in trans by the 3CLP encoded by ORF 1a (Liu *et al.*, 1994; Ziebuhr *et al.*, 1995; Grötzinger *et al.*, 1996). A polypeptide of approximately 100 kDa, representing the extreme C terminus of ORF 1a and the N terminus of the frame shifted ORF 1b, was immunoprecipitated from IBV-infected cells. The cleavage sites of the 100-kDa protein appear to be at the Q/S sites, as predicted from the computer analysis and consistent with the known substrate specificity of the picornavirus 3C protease. A similar observation was recently made with HCV-229E (Grötzinger *et al.*, 1996). This 100-kDa protein contains the putative RNA polymerase motif and thus may represent the coronavirus polymerase. The coding region for this protein belongs to complementation group D, which has been shown to effect mRNA transcription (Fig. 6) (Schaad *et al.*, 1990).

b. Processing of the Structural Proteins

1. *S protein.* The S protein is co-translationally glycosylated with N-linked glycans. Conversion of the high mannose (simple) glycans of

the S protein to complex ones is a slow process, the half-life being one to several hours (Vennema *et al.*, 1990a). The S protein undergoes multiple disulfide linkages to fold into a more complex structure (Opstelten *et al.*, 1993) and oligomerize into a trimer in the Golgi complex (Delmas and Laude, 1990).

The S prepolypeptide is converted to a polypeptide by removal of the N-terminal signal peptide. Whether the polypeptide is cleaved to generate S1 and S2 depends on the virus species and strain and, to some extent, on the cell type in which the virus is grown (Frana *et al.*, 1985). The S1-S2 cleavage site in IBV and MHV is adjacent to several basic residues (Cavanagh *et al.*, 1986a; Luytjes *et al.*, 1987). Those coronaviruses whose S protein is not cleaved, e.g., FCV, TGEV, and CCV, have no such pairs of basic residues. Cleavage of the MHV S protein occurred after conversion of the glycans from simple to complex forms (Vennema *et al.*, 1990a). After cleavage, the S1 and S2 subunits are held together by noncovalent linkages (Cavanagh *et al.*, 1986b; Sturman *et al.*, 1990). The S2 protein of MHV is acylated, possibly involving some of the many cysteine residues in the C-terminal, hydrophilic tail of S (Schmidt, 1982; Sturman *et al.*, 1985; van Berlo *et al.*, 1987). The processing of S proteins is reviewed in greater detail by Cavanagh (Cavanagh *et al.*, 1995).

2. *M protein.* Modification of the M protein depends greatly on the virus species. The major modification is glycosylation. The oligosaccharides of IBV and the TGEV group are of the co-translationally added N-linked glycans (Stern and Sefton, 1982b). The conversion of the high mannose to complex glycans is not very efficient. In contrast, viruses of the MHV group have O-linked glycans which are added posttranslationally (Holmes *et al.*, 1981; Niemann *et al.*, 1982; 1984; Tooze *et al.*, 1988; Locker *et al.*, 1992a; Krijnse-Locker *et al.*, 1994). The M protein of TGEV is also sulfated (Garwes *et al.*, 1976), but whether this is linked directly to the polypeptide or to glycans is unknown. Unlike the M proteins of IBV and the MHV group, which have an internal membrane insertion sequence, those of the TGEV group have an N-terminal membrane insertion sequence that is absent from the mature M protein (Laude *et al.*, 1987). This signal sequence, however, is not an essential requirement for the membrane insertion of the M protein (Kapke *et al.*, 1988; Vennema *et al.*, 1991).
3. *HE protein.* The HE glycoprotein has N-linked glycans which are converted to complex ones in the Golgi complex. The N-terminal signal sequence is cleaved from the mature protein, which then

forms dimers by disulfide bonds (King *et al.*, 1985; Hogue *et al.*, 1989; Kienzle *et al.*, 1990; Yoo *et al.*, 1992).

4. *E protein*. The only known modification for the E protein is acylation of MHV E protein (Yu *et al.*, 1994). However, this was not observed for the E protein of TGEV when expressed in insect cells (Godet *et al.*, 1992).
5. *N protein*. The N protein is phosphorylated, the phosphate linkage being exclusively to serine residues (Stohlman and Lai, 1979). The role of phosphorylation is unknown.

H. Virus Assembly and Release

In virus-infected cells, the assembly of virus particles presumably starts with the formation of RNP, which interacts with the components of viral envelope proteins to form enveloped virus particles and bud into the endoplasmic reticulum (ER) and Golgi complex. Several recent advances shed light on this process:

1. Early studies have shown that the S proteins are not necessary for virus particle formation; thus, denuded virus particles without spikes can be formed in the virus-infected cells treated with tunicamycin, which inhibits *N*-glycosylation and transport of the S and HE proteins (Holmes *et al.*, 1981). Further, recent studies have shown that the minimum requirement for the formation of virus-like particles (VLP), i.e., empty virus particles, is the M and E proteins (Bos *et al.*, 1996; Vennema *et al.*, 1996);
2. The sites of virus budding are in the ER and Golgi, near the sites of accumulation of the M protein (Dubois-Dalcq *et al.*, 1982; Tooze *et al.*, 1984; Tooze and Tooze, 1985; Klumperman *et al.*, 1994); thus, the interaction between the M and E proteins appears to be the key event for virus particle assembly. The incorporation of the nucleocapsids and S and HE proteins into virus particles may involve subsequent interactions of these components with the M-E complex.

The virus assembly and release process has been studied in most detail for MHV (J. Tooze *et al.*, 1984, 1987; Tooze and Tooze, 1985; S. A. Tooze *et al.*, 1988; Krijnse-Locker *et al.*, 1994), and the gross features have recently been confirmed for IBV, TGEV, and FIPV (Klumperman *et al.*, 1994). Recently, an ultrastructural study of the replication of IBV in renal ductotubular epithelial cells of infected chicks has also been very informative (Chen and Itakura, 1996). The first virions form in the perinuclear region, in small, smooth vesicles/

tubules between the rough ER and the cis side of the Golgi stack. Later, the rough ER becomes the major site of virion assembly, extending beyond the perinuclear region. Virions then proceed through the Golgi complex, at the trans side of which they are collected into vesicles of the constitutive exocytic pathway and subsequently released from the cell.

The major determining factor for the site of virus assembly appears to be the site of localization of the M protein, which is in the Golgi complex. There are some points of difference among the coronaviruses. When the M protein of MHV was expressed, it accumulated in the trans-Golgi membranes, consistent with its *O*-linked glycosylation, which occurs efficiently (Locker *et al.*, 1992a; Klumperman *et al.*, 1994). In contrast, expression of the IBV M protein from cDNA resulted in its accumulation in cis-Golgi membranes; consequently the high-mannose *N*-linked glycans of the M protein were not efficiently converted to complex ones (Machamer *et al.*, 1990; Klumperman *et al.*, 1994), in agreement with the properties of the M protein in the IBV virions. Glycosylation of the coronavirus M protein is not essential for its translocation or for virus particle formation. The M protein exists as monomers in the ER, but it oligomerizes to form variously sized complexes during transport through the Golgi and trans-Golgi network (Locker *et al.*, 1995). It is likely that the M molecules in the virus particles are in complexed form.

The sequence requirements for insertion of the nascent M polypeptide into the rough ER have not been precisely defined. With the exception of the TGEV group, the coronavirus M proteins do not have an amino-terminal signal peptide. Even in the case of the TGEV group, the signal peptide is not essential for membrane insertion of the M protein (Kapke *et al.*, 1988; Vennema *et al.*, 1991). Rather, one of the three transmembrane sequences of the M protein is responsible for the insertion of M into the ER and its final localization in the Golgi complex (Machamer and Rose, 1987; Mayer *et al.*, 1988; Armstrong *et al.*, 1990; Locker *et al.*, 1992b). Different domains of the M protein of IBV and MHV have been identified as the sequences responsible for the final localization of the protein. The first membrane-spanning domain of the IBV M protein performs this function, the M protein being concentrated in the cis-Golgi membranes (Machamer and Rose, 1987; Machamer *et al.*, 1990, 1993; Swift and Machamer, 1991). In contrast, the carboxy-terminal domain of the MHV M protein, probably in combination with a middle domain, directs the protein to the trans-Golgi (Armstrong and Patel, 1991; Weisz *et al.*, 1993; Krijnse-Locker *et al.*, 1994). It should be borne in mind, however, that the major site of virus particle formation is

proximal to either of the Golgi compartments, namely, in an intermediate compartment between the ER and the Golgi complex (Klumperman *et al.*, 1994). Thus, it is proximal to the major site of M accumulation. What is responsible for that? The answer would appear to be that the S glycoprotein and the nucleocapsid interact with the M protein molecules before the M proteins have migrated to the Golgi, precipitating virus particle formation.

It has been shown that the coronavirus M protein can interact with nucleocapsids (Sturman *et al.*, 1980). This interaction requires the presence of viral RNA, since the N protein alone cannot be incorporated into the VLPs (Bos *et al.*, 1996; Vennema *et al.*, 1996), suggesting either that M interacts with viral RNA, or that RNA-N protein binding induces a conformational change in the N protein, enabling it to interact with M. Interaction between the M and S proteins has also been demonstrated. The M and S proteins co-sediment under certain ionic conditions after dissolution of virions with mild detergents (Cavanagh, 1983b), and cell-associated complexes containing M and S have been detected (Opstelten *et al.*, 1995). The S protein undergoes certain conformational changes induced by disulfide linkage before it is able to interact with M (Opstelten *et al.*, 1993, 1995). Inhibition of correct oligomerization of S by dithiothreitol prevented interaction of S with M and, as a result, the rate of transport of the M protein to the trans-Golgi increased (Opstelten *et al.*, 1993). This result suggests that S-M interaction can retard the transport of the M protein. The ability of the S or HE protein to interact with the M protein appears to be a prerequisite for their incorporation into virus particles. In this regard, it is interesting to note that MHV ts mutants with a deletion in the ectodomain of the S protein or those with defects in oligomerization of the S protein do not incorporate the S protein (Ricard *et al.*, 1995; Luytjes *et al.*, 1997). Also, partial deletions in the ectodomain of the HE protein prevent its incorporation into virus particles (Liao *et al.*, 1995). These results suggest that the interaction of S or HE with M occurs through the ectodomain or requires the correct protein conformation in the ectodomain. The formation of the S-M complex occurs in the pre-Golgi complex, whereas the S-M complex progresses until the Golgi complex, indicating that this interaction is not sufficient to localize it in the pre-Golgi complex, the ultimate site of virion budding (Opstelten *et al.*, 1995). Thus, M-nucleocapsid interaction may also contribute to the determination of the site of virus assembly. In this regard, it is important to note that the recent discovery that M is present in the viral RNP core, as well as in the envelope (Risco *et al.*, 1996) may further indicate the crucial role of the M protein in the virus assembly process.

Only the M and E proteins are required for the production of VLPs (Bos *et al.*, 1996; Vennema *et al.*, 1996). These particles were formed when the M and E proteins were expressed from transfected plasmids. S protein was incorporated into the VLPs if expressed. In the absence of viral RNA, the N protein also was not incorporated. When all the structural proteins were expressed from plasmids in the presence of an MHV DI RNA, which contains a packaging signal, and in the absence of helper virus, the VLPs contained the DI RNA (Bos *et al.*, 1996). Moreover, these VLP were "infectious," i.e., on transfer of the released VLPs to a new cell culture, they were able to infect the cells, as revealed by the rescue of the DI RNA by helper virus. These results show that N is dispensable for the formation of VLPs but the packaging of RNA into virion requires an interaction between M and the N-containing ribonucleoprotein, as previously demonstrated (Sturman *et al.*, 1980). The expression of the M protein alone in the cells did not lead to VLP formation or induction of curvature in the M-containing intracellular membranes. The presence of the E protein together with the M protein triggered both events, but the ratio of M:E in virions was as high as 100:1 (Vennema *et al.*, 1996). This has led to the suggestion that E does not have frequent, regular positions in the lattice formed by M but rather occupies strategic positions within the lattice to cause membrane curvature. Alternatively, its role may be to close the neck of the virus particle as it pinches off from the membrane in the final stage of budding.

What determines the site of virion budding? It is possible that the E protein dictates the site of budding, since this protein is also localized in the perinuclear region and associated with membrane (Godet *et al.*, 1992; Yu *et al.*, 1994). Alternatively, it may be the interaction of the RNP-nucleocapsid with the S-M complexes which halts the migration of the latter and promotes budding. Relevant to this notion is the observation that the nucleocapsids and free N protein have affinity for membranes (Anderson and Wong, 1993). It should be remembered, however, that in the absence of S, HE, and nucleocapsids, the E and M proteins alone can induce budding to form VLPs (Bos *et al.*, 1996; Vennema *et al.*, 1996). It is not yet clear whether the budding site of VLP containing only M and E is the same as that for the complete virion. Empty virus particles have previously been isolated from IBV, which were grown in embryonated fowl eggs (Macnaughton and Davies, 1980). This supports the view that even during natural infection, virus budding can be induced without involvement of the viral nucleocapsid. Parallels have been drawn between the E protein of coronaviruses, the

M₂ protein of orthomyxoviruses, and the 6K protein of alphaviruses. All are minor envelope proteins that play a role in virus assembly.

Once the virus particles bud into the pre-Golgi compartment, they are transported through the Golgi complex. Whether the Golgi-associated posttranslational modifications occur before or after incorporation of the proteins into virus particles is not known. Retrograde transport of the proteins may be required for some steps of the virus assembly process. Finally, the release of virus particles from the cells appears to be restricted to certain areas of cells. TGEV grown in polarized LLC-PK1 cells both enter and exit by the apical surface (Rossen *et al.*, 1994), whereas MHV-A59 enters polarized murine kidney cells (mTAL) by the apical surface but is released via the basolateral surface (Rossen *et al.*, 1995a). However, the site of virus release varies with different cell lines (Rossen *et al.*, 1997). The factors governing this process are not known (Rossen *et al.*, 1995b).

VI. GENETICS OF CORONAVIRUSES

Probably because of the large size of their RNA genomes, coronaviruses have developed a variety of genetic mechanisms, among which are RNA recombination and generation of DI RNA, to maintain their genetic stability and, as a side product, generate diversity. Coronaviruses also readily undergo genetic mutation, a characteristic common to all RNA viruses. Thus, they evolve rapidly and are heterogeneous. These genetic phenomena provide virologists with useful tools for understanding coronavirus biology, particularly because reverse genetics studies for coronaviruses are not yet feasible.

A. Natural Virus Variants and Mutants

1. Temperature-Sensitive Mutants

Using a variety of chemical mutagens, several laboratories have isolated MHV temperature-sensitive (ts) mutants which cannot produce infectious virus particles or cause different plaque morphology at the nonpermissive temperature (Haspel *et al.*, 1978; Robb *et al.*, 1979; Wege *et al.*, 1981; Koolen *et al.*, 1983; Schaad *et al.*, 1990). Some of these mutants have been characterized with respect to their ability to synthesize RNA and have been grouped into at least seven complementation groups (Leibowitz *et al.*, 1982a), five of which have the RNA (-) phenotype (i.e., cannot synthesize RNA at the nonpermissive temperature) (Leibowitz *et al.*, 1982a; Schaad *et al.*, 1990) (see Fig. 6). With the

use of recombination analysis (see below), the possible genetic defects of the mutants were mapped on the RNA genome (Baric *et al.*, 1990; Fu and Baric, 1994). It appears that all of the RNA (-) mutants have genetic defects within gene 1, suggesting that gene 1 encodes RNA polymerase and other proteins involved directly or indirectly in viral RNA synthesis. The genetic defects of some of these mutants have been confirmed by RNA sequence analysis of the mutants and their revertants (Fu and Baric, 1994). These five different complementation groups have been demonstrated to affect different steps of RNA synthesis, including the synthesis of leader RNA, negative-strand RNA, and positive-strand RNA (Fig. 6), suggesting that different steps of RNA synthesis require different viral proteins (Baric *et al.*, 1990; Schaad *et al.*, 1990). It is still not possible, however, to correlate the genetic defects definitively with the known processed products of the gene 1 polyprotein.

Among the RNA (+) mutants, two complementation groups have been assigned to the gene encoding the S protein (Baric *et al.*, 1990; Fu and Baric, 1994), but the phenotype of these mutants has not been well characterized. Another RNA (+) mutant, *Alb 18*, has a single amino acid substitution in the N-terminal domain of S protein and cannot incorporate S protein into the virus particles (Ricard *et al.*, 1995). Still another group of RNA (+) mutants have a defective N protein (Koetzner *et al.*, 1992; Masters *et al.*, 1994; Peng *et al.*, 1995a) and produce smaller plaques at the nonpermissive temperature; several of these mutants have a deletion in the N gene (Masters, 1992) and are defective in RNA-binding activity (Peng *et al.*, 1995a). Most wild-type revertants have a second-site mutation in the N protein and restored RNA-binding activity (Peng *et al.*, 1995a).

2. Neutralization-Escape Mutants

Another class of viral mutants was obtained by a specific selection scheme, e.g., by treating viruses with neutralizing MAb and selecting mutant viruses resistant to neutralization. Since neutralizing antibodies are usually directed against the S protein, all of the neutralization-escape mutants were presumed to have defects in the S gene. This was indeed the case (reviewed by Cavanagh *et al.*, 1995). Depending on the neutralizing MAb used for selection, the mutants obtained had either deletions or point mutations in the neutralization epitopes of the S protein (Gallagher *et al.*, 1990; Wang *et al.*, 1992). These mutants generally retain growth properties very similar to those of the parental virus but often have significantly different pathogenic properties with

altered tissue tropism (Dalziel *et al.*, 1986; Fleming *et al.*, 1986, 1987; Wege *et al.*, 1988).

3. Other Nonconditional Deletions or Substitutions

During serial virus passages in tissue culture or in animals, coronaviruses often undergo various deletions or substitutions even in the absence of experimentally applied selection pressure. These genetic changes probably provide the emerging virus variants with evolutionary advantages under experimental conditions or in natural infection. The deletions occur most frequently within the S gene, particularly within a hypervariable region encoding the S1 subunit (S. E. Parker *et al.*, 1989; Wang *et al.*, 1992). In fact, some natural isolates of MHV have a deletion of 150–460 nucleotides in this region (Fig. 4). Similar deletions have been detected in virus variants during central nervous system (CNS) infections of rats (La Monica *et al.*, 1991). In persistent infections of cultured cells of CNS origin, viruses with point mutations or deletions in the gene encoding S protein are frequently selected (Gallagher *et al.*, 1991; Gombold *et al.*, 1993; Rowe *et al.*, 1997). These viruses often have altered cell fusion and pathogenic properties.

The most striking effect of deletions during natural virus infection is illustrated by the emergence of PRCV from TGEV. TGEV causes epizootic enteric infection in pigs, resulting in a very high mortality rate in newborn pigs. An attenuated virus strain that is related to TGEV but infects only respiratory tissues was isolated in Western Europe in the early 1980s (Pensaert *et al.*, 1986). An independent isolate of PRCV was subsequently obtained in the United States (Wesley *et al.*, 1990). Both of these PRCV isolates have similar extents of deletion in the N terminus of the S1 protein, in addition to smaller deletions in gene 3, which eliminates its expression (Rasschaert *et al.*, 1990; Wesley *et al.*, 1991; Laude, 1993). Although it is not yet possible to link the changes in viral pathogenicity to the deletions in the S gene or gene 3, the TGEV-PRCV evolution illustrates the power of deletions in coronavirus evolution.

B. Complementation

Different ts mutants with defects in different coronaviral genes have been demonstrated to complement each other. The available ts mutants of MHV have been divided into at least seven complementation groups, five of which have an RNA (–) phenotype (Leibowitz *et al.*, 1982a) (Fig. 6). It is worth noting that these five RNA (–) complementation groups

have been mapped in gene 1 (Baric *et al.*, 1990), which is translated into a polyprotein. The existence of five complementation groups within this gene indicates that this polyprotein is processed into at least five different proteins that function independently. It is not possible, however, to complement the genetic defects of a virus by expressing a wild-type viral protein from an exogenous vector.

C. Phenotypic Mixing and Pseudotype Virus Formation

Mixed infection with MHV and murine leukemia virus in tissue culture cells yielded a pseudotype MHV which contained a murine leukemia virus envelope protein and was neutralized by antibodies against both murine leukemia virus and MHV (Yoshikura and Taguchi, 1978). This phenotypic mixing of viral proteins suggests the lack of a stringent requirement for a virus-specific spike protein for the formation of coronavirus particles. Pseudotype formation of virus particles has also been achieved by expressing a viral protein, e.g., HE protein, from a DI RNA vector (see Section VI, E), which was incorporated into virus particles (Liao *et al.*, 1995).

D. RNA Recombination

One unique genetic feature of coronaviruses is their ability to undergo RNA recombination at a very high frequency; this is particularly true of MHV, in which recombinant viruses containing parts of the genomic sequences of both parental viruses could be isolated at high frequency when two strains of MHV with defined genetic markers were co-infected into culture cells or animals. This genetic phenomenon was first discovered using two *ts* mutants of MHV (Lai *et al.*, 1985). Subsequently, many different recombinant MHVs were isolated (Keck *et al.*, 1987, 1988b,c; Makino *et al.*, 1987) using a combination of selection markers, such as *ts* markers, resistance to neutralizing antibodies, and cytopathic effects (the ability of the virus to cause fusion). Based on the distribution of the crossover sites on the viral RNA genome, it appears that recombination can occur practically anywhere on the viral genome, although some combinations of virus strains favor selection of viruses with certain recombination sites (Lai, 1992). For example, between the MHV A59 and JHM strains, recombination occurs mostly at the 5' end of the genome and rarely at the 3' end. In contrast, recombination between the MHV-2 and JHM strains occurs readily at the 3' end (Keck *et al.*, 1988c). The most surprising finding with regard

to MHV recombination is the extremely high frequency of recombination, which has been calculated to be nearly 25% for the entire MHV genome (Baric *et al.*, 1990). This high frequency of recombination is reminiscent of the reassortment of segmented RNA genomes in viruses such as influenza virus and reovirus. The recombination map for MHV is nearly linear, suggesting the random occurrence of recombination (Baric *et al.*, 1990); however, more careful analysis of the recombination frequency showed that there is an increasing gradient of recombination frequency (in the direction of 5'→3') across the genome, suggesting that subgenomic mRNAs, which represent preferentially the 3' end sequences, may participate in RNA recombination (Fu and Baric, 1992, 1994). Recombination has now been demonstrated experimentally for IBV (Kotier *et al.*, 1995) and TGEV (Ballesteros *et al.*, 1997) in embryonated eggs or tissue culture; however, the recombination frequency for these viruses has not been determined.

Recombination can provide a powerful tool for virus evolution. For example, in a study in which ts mutants of the A59 strain of MHV were co-infected with a wild-type JHM strain, the majority of the progeny viruses after a single passage were recombinants which contained the 5' end of the A59 genome (Makino *et al.*, 1986a), suggesting that this recombinant virus has evolutionary advantages. Recombination has also been demonstrated during virus infection in animals (Keck *et al.*, 1988b).

Similar to the situation in other RNA viruses, coronavirus recombination probably occurs by a copy-choice mechanism (Lai, 1992). It has been shown that MHV RNA synthesis normally pauses at certain sites on the RNA genome (Baric *et al.*, 1987). The nascent, incomplete RNA transcripts may dissociate from the template RNA and then rebind to the template to resume RNA synthesis. When the nascent RNA binds to a different template, the resumed RNA synthesis will result in a recombinant RNA. Whether coronavirus recombination occurs more frequently at certain RNA sites with more complex secondary structure is not yet known. When RNA recombination was examined under nonselective conditions (by reverse transcription-polymerase chain reaction detection of the intracellular RNA from virus-infected cells), recombination sites appeared to be random; only after serial passages did "hot spots" of RNA recombination become apparent (Banner and Lai, 1991). This finding indicates that the recombination hot spots may be the result of selection.

Recombination has been detected during natural infections of coronaviruses, most notably IBV. Sequence analysis of natural IBV strains has provided convincing evidence that some IBV strains are recombinants

between different IBV strains; recombination sites have been detected so far in the 5' half of the S gene and at the 3' end of viral RNA (Kusters *et al.*, 1989; Cavanagh and Davis, 1992; Wang *et al.*, 1993, 1994; Jia *et al.*, 1995). Thus, recombination is a natural evolutionary strategy for coronaviruses.

RNA recombination may also explain the difference in genome structure among different coronaviruses. For example, IBV contains an additional gene, gene 5 (a nonstructural protein gene) inserted between gene M and gene N (Fig. 5). This insertion could be the result of a recombination mechanism involving the consensus IG sequence, which provides a favored recombination site. Since all of the coronavirus genes are flanked by consensus IG sequences, each gene can be considered a gene "cassette," which can be rearranged by homologous recombination involving the consensus IG sequence. A nonhomologous recombination event between coronavirus RNAs and other virus or cellular RNAs may also explain the gene insertions in some coronaviruses. For example, MHV and BCV contain an additional gene, HE, which is similar in sequence to the HE gene of influenza C virus (Luytjes *et al.*, 1988). This gene may have been derived by recombination between a coronavirus and influenza C virus. Comparison between genome structures of coronavirus and torovirus also suggests that several recombination events may have been involved in rearranging the order of several genes during the evolution of these viruses (Snijder *et al.*, 1991).

Recombination has been demonstrated to occur between viral RNA and a transfected RNA fragment derived from the viral genome (Koetzner *et al.*, 1992; Liao and Lai, 1992). Since transfection of both the positive- and negative-strand RNA fragments led to recombination, these results suggested that recombination can occur during both positive- and negative-strand RNA synthesis (Liao and Lai, 1992). Recombination can also take place between DI RNAs and viral RNA reciprocally, i.e., the viral RNA sequence can be incorporated into DI RNA, and vice versa, during viral RNA replication. The incorporation of a helper viral RNA sequence into DI RNA accounts at least partially for the continuous evolution of MHV DI RNA species during serial passages in cultured cells (Furuya *et al.*, 1993) (see the next section). This phenomenon also explains why some genetic markers in the DI RNA were rapidly replaced by the helper viral RNA sequences during DI RNA replication (de Groot *et al.*, 1992; Kim *et al.*, 1993a). On the other hand, the incorporation of DI RNA sequences into viral RNA by recombination provides an important tool to introduce desired sequences into the viral genome. For example, when an mRNA 7 or DI RNA containing the N gene of MHV was transfected into cells infected

with an MHV ts mutant containing a defective N protein, recombination occurred between the DI RNA and the wild-type viral RNA, resulting in recombinant viruses which had a wild-type RNA sequence derived from the transfected RNA in place of the defective N gene (Koetzner *et al.*, 1992; van der Most *et al.*, 1992; Masters *et al.*, 1994; Peng *et al.*, 1995a). An MHV recombinant containing a chimeric N protein of BCV and MHV has also been derived by this RNA recombination strategy (Peng *et al.*, 1995b). This targeted RNA recombination promises to be a powerful tool.

Recombination is thus one of the most unique aspects of coronavirus biology. It can potentially provide a genetic mechanism by which coronaviruses maintain their sequence integrity. In view of the large size of the coronavirus RNA, it is predictable that most of the viral RNA molecules would contain mutations due to the high error frequencies of RNA polymerases; recombination may provide a repair mechanism for the virus (Lai, 1992).

E. Defective-Interfering (DI) RNAs

Similar to most RNA viruses, coronaviruses can readily generate DI particles when viruses are passaged in tissue culture at a high multiplicity of infection. This has been demonstrated for MHV, IBV, and TGEV. When MHV was serially passaged, different types of DI RNA appeared at different passage levels, suggesting that DI RNAs continue to evolve and that new DIs have a selective advantage under the evolving cellular conditions (Makino *et al.*, 1985). However, the IBV and TGEV DIs appear to be more stable (Penzes *et al.*, 1994; Mendez *et al.*, 1996). The generation of DI RNAs is probably caused by polymerase jumping during RNA replication or nonhomologous RNA recombination. Although no sequence homology exists at the fusion sites of different RNA regions within the DI RNA, a high degree of potential secondary structure does exist at some of its RNA fusion sites (Makino *et al.*, 1988b), which may have facilitated the pausing and template switching of RNA polymerase during synthesis. If nonhomologous recombination is involved in generating DI RNA, it probably occurs between two different RNA molecules because DI RNAs are generated only at high multiplicity of infection. Recombination between an existing DI RNA and helper virus RNA has been shown to contribute to the evolution of MHV DI RNAs during virus passages (Furuya *et al.*, 1993).

The coronavirus DI RNAs can be grouped into three types. The first type is of nearly genomic size and is typified by DIssA RNA of MHV

(Makino *et al.*, 1985). This DI RNA is efficiently packaged into virus particles and contains several deletions in the viral genome, but it contains a functional gene 1, which encodes RNA polymerase, and a functional gene 7, which encodes N protein. These two functional gene products are sufficient to support DI RNA replication (K. H. Kim and Makino, 1995); thus, this type of DI RNA can replicate without a helper virus (Makino *et al.*, 1988a; K. H. Kim and Makino, 1995). By definition, it is not a DI RNA, inasmuch as it is not defective in replication; however, because it is smaller than the genomic RNA and is produced at a high multiplicity of infection, it is classified as a DI RNA. This type of DI is unique to coronavirus. A 22-kb DI RNA has been described for TGEV (Mendez *et al.*, 1996), but whether it can replicate in the absence of a helper virus has not been examined.

The second type DI RNA is typified by DIssE of MHV (Makino *et al.*, 1988b). This DI RNA is truly defective and can replicate only in the presence of helper viruses. It replicates very efficiently, but is poorly packaged into virus particles because it lacks a specific RNA-packaging signal. This type of DI RNA typically contains both the 5' and 3' ends of the wild-type viral RNA and one or several discontinuous regions of the wild-type RNAs. Because of the high efficiency of replication, this type of DI can still be serially passaged in tissue culture for at least several passages, probably because a small amount of DI RNA can be nonspecifically packaged into the virion.

The third type of DI RNA is represented by DIssF of MHV-JHM (Makino *et al.*, 1990) and DI-a of MHV A59 (van der Most *et al.*, 1991). It is similar to the second type but contains an RNA-packaging signal and is thus packaged efficiently into virus particles. This type of DI RNA has been detected in IBV (Penzes *et al.*, 1994) and TGEV (Mendez *et al.*, 1996). A small DI RNA (2.2 kb) of BCV may also belong to this type (Chang and Brian, 1996), but whether this DI RNA can be specifically packaged into virion is not certain.

All three types of DI RNAs contain an ORF, which encodes a protein fused from two different viral proteins. This ORF is not required for the replication of MHV DI RNA (Liao and Lai, 1995); nevertheless, MHV DI RNAs with a functional ORF usually have an evolutionary advantage over those without one or with a smaller ORF (de Groot *et al.*, 1992; Kim *et al.*, 1993a). Therefore, a DI RNA containing a short ORF was often rapidly replaced by DI RNAs containing a longer ORF that had been generated by recombination or mutation (de Groot *et al.*, 1992; Kim *et al.*, 1993a). The translatability of the ORF may be more important than the nature of the actual protein translated from this ORF (van der Most *et al.*, 1995), suggesting that translation of RNA

may facilitate RNA replication. Reduction of the ORF of an IBV DI RNA to just 20 amino acids did not diminish its capacity to be replicated or packaged (Penzes *et al.*, 1996). However, it has been shown for a BCV DI RNA that a BCV-specific N protein translated from the DI ORF (a cis-acting protein) is required for efficient DI RNA replication (Chang and Brian, 1996). The variation in the sequence requirement for RNA replication of these DI RNAs may be related to their overall RNA conformation. The significance of DI RNA in the biology and natural evolution of coronaviruses is not known.

DI RNAs provide useful tools for studying the sequence and structural requirements for various functions of viral genomic RNA. As they contain cis-acting signals for RNA replication, they are mini-versions of the viral genomic RNA. However, it should be cautioned again that because of the small size of the DI RNA compared to the genomic RNA, the structural requirements for various RNA functions, as determined from the use of DI RNA constructs, may be different from those of the whole viral genome.

The following cis-acting signals for various RNA functions have been determined using various DI RNAs:

1. *RNA-packaging signal.* In a comparison of MHV DI RNAs that are efficiently and inefficiently packaged, it was determined that the packaging signal for MHV DI RNA is localized near the 3' end of gene 1 (in the 1b region, approximately 20 kb from the 5' end) (Makino *et al.*, 1990; van der Most *et al.*, 1991; Fosmire *et al.*, 1992). This packaging signal forms a stem-loop structure which may be required for the RNA-packaging activity (Fosmire *et al.*, 1992). It is necessary and sufficient for the packaging of DI RNA or a heterologous RNA into the virions (Woo *et al.*, 1997). The fact that this packaging signal is localized in gene 1, which is present in genomic but not subgenomic RNAs, is consistent with the packaging of genomic but not subgenomic RNAs in virus particles. The packaging signal for DI RNAs of other coronaviruses has not been determined. However, some coronaviruses have been shown to package subgenomic mRNAs at low efficiency (Sethna *et al.*, 1989; Hofmann *et al.*, 1990; Zhao *et al.*, 1993). These are probably packaged nonspecifically; however, the possibility that these viruses may have a different RNA packaging signal cannot yet be ruled out. Similarly, DI RNAs that do not contain this packaging signal, such as DIssE RNA of MHV (Makino *et al.*, 1988b) and DI RNA of BCV (Chang and Brian, 1996; Chang *et al.*, 1996), can be packaged at low efficiency, thereby maintaining themselves for at least several passages in tissue culture.

2. *Negative-strand RNA synthesis.* For MHV DI RNA, it has been shown that only 55 nucleotides at the 3' end plus a stretch of poly(A) sequence are required for negative-strand RNA synthesis (Lin *et al.*, 1994); no specific upstream RNA sequences are required. However, when an mRNA is transcribed from an IG site in the same DI RNA, the negative-strand RNA synthesis from this DI RNA is inhibited, suggesting a common element involved in mRNA transcription and negative-strand RNA synthesis (Lin *et al.*, 1994). One unanswered question is whether or not the sequence requirements for the synthesis of genomic and subgenomic negative-strand RNA are identical.
3. *Replication signal.* Sequential deletion analysis has shown that the replication (i.e., complete cycles of negative- and positive-strand RNA synthesis) of MHV DIssE or DIssF RNAs requires approximately 400–800 nucleotides from both the 5' and 3' ends. The minimum sequence requirement for RNA replication may vary with different DI RNAs. These issues have been discussed in Section V,F.
4. *Transcriptional signal.* DI RNAs normally do not transcribe subgenomic mRNAs because they do not have IG sequences. Thus, natural DI RNAs can synthesize only the full-sized DI RNA. However, by introduction of the consensus IG sequences into DI RNA (Makino *et al.*, 1991), it has been possible to use DI RNA as a vector for determining the sequence requirement for subgenomic RNA transcription. The cis- and trans-acting signals for transcription have been described in Section V,E.
5. *Recombination.* DI RNAs of MHV have been demonstrated to undergo a high frequency of recombination with helper virus RNA. As discussed above, this accounts for the evolution of MHV DI RNA species during serial passages of viruses (Furuya *et al.*, 1993). Furthermore, MHV DI RNAs with a smaller ORF are frequently replaced by a DI RNA with a larger ORF by recombination with the helper virus RNA (de Groot *et al.*, 1992; Kim *et al.*, 1993a), suggesting that recombination between DI RNAs and helper virus RNAs occurs readily. The reciprocal recombination between DI RNA and helper virus RNA, i.e., the transfer of DI RNA sequences to the helper virus RNA, also has been observed. As a result, the genetic markers on the DI RNA can be incorporated into the helper virus RNA (Koetzner *et al.*, 1992). Recombination between two DI RNAs, however, has not been described. Sequence requirements for RNA recombination also have not been studied. BCV DI RNAs also undergo frequent recombination (Chang *et al.*, 1996). However, DI RNAs of IBV and TGEV appear to be more stable.

VII. PERSPECTIVES

Coronavirus research has made tremendous progress in the last decade. The virus family has grown in size, and many of the features thought to be unique to coronaviruses have now been found to be shared by some other viruses. Since the last time this serial publication published the first comprehensive review of the molecular biology of coronaviruses (Sturman and Holmes, 1983), the literature on this virus has grown to exceed anyone's ability to do a comprehensive review of every topic relating to coronaviruses. In this review, we have concentrated on areas which have shown the most progress and which present the most challenges. Our choice of literature was meant to be representative but is by no means comprehensive. Notably missing from this review are the molecular studies related to viral pathogenesis and the interactions between the virus and cells.

Coronavirus research has contributed to the understanding of many aspects of molecular biology in general, such as the mechanism of RNA synthesis, translational control, and protein transport and processing. It remains a treasure capable of generating unexpected insights. Despite two decades of studies on the molecular biology of this virus, there are still many problems to be solved:

1. With regard to the mechanism of RNA transcription, many conflicting data remain. Coronavirus undoubtedly utilizes a unique, discontinuous transcription mechanism, but how it acts is a subject of debate. An *in vitro* RNA transcription system, so necessary for an understanding of RNA synthesis, is still in its infancy. Related to this question is the nature of RNA polymerase. The sheer size of the polymerase gene presents a daunting task. The availability of the cDNA clones and expression vectors for this gene has just begun to allow this black box to be cracked open. This will undoubtedly be a fruitful area of future research.
2. The last two years have seen the unraveling of the mechanism of coronavirus assembly, which, as it turns out, involves a little-characterized E protein. How the various viral structural proteins interact with each other in the various subcellular compartments to form a complete virus particle is an exciting frontier.
3. After more than 30 years since the first coronavirus was seen under electron microscope, an unexpected new feature of the virus, namely, an icosahedral core with a helical nucleoprotein, was recently uncovered. This structure places coronavirus in a unique position among RNA viruses because it takes on the characteristics of positive-,

- negative- and double-strand RNA viruses in morphology. This recent finding challenges us to reevaluate the structure of coronaviruses.
4. The ability to perform reverse genetic studies of coronavirus is still very limited. Expression of individual viral genes and targeted recombination of very limited RNA regions are the only available genetic means for examining the structure and function of the coronavirus genome. Perhaps it is an unrealistic dream, but progress in polymerase chain reaction technology may one day allow an infectious cDNA for coronavirus RNA to be made.
 5. The early events of viral replication have so far been largely ignored. Identification of the cellular receptors for the viruses may finally provide penetrating molecular tools to allow these issues to be examined. It will not be a surprise to discover that virus penetration and uncoating play defining roles in the cellular tropism of viruses.
 6. Are nonstructural protein genes really unnecessary? Even if they are auxiliary genes, they may prove to play significant roles in the biology of the virus.
 7. Finally, what of the potential interaction between the virus and host, which has been one of the major themes of virology in recent years? It may be a little premature to conclude that cellular factors play major roles in coronavirus replication, but there is little doubt that cells are playing more active roles than was previously suspected. Is the nucleus contributing to the coronavirus replication? This may require reexamination.

These are but some of the exciting challenges for the coronavirologists to tackle. The next decade should bring us an even better understanding of the various aspects of the molecular biology of coronaviruses.

ACKNOWLEDGMENTS

We would like to thank the following individuals for making data available prior to publication: Luis Enjuanes, Hubert Laude, Peter Rottier, and Pierre Talbot. We would also like to thank Daphne Shimoda for her tremendous help with editorial revisions, preparation of the figures, and the typing of the manuscript. M.M.C.L. is Investigator of the Howard Hughes Medical Institute.

REFERENCES

- Anderson, R., and Wong, F. (1993). Membrane and phospholipid binding by murine coronavirus nucleocapsid N protein. *Virology* **194**, 224–232.
- Armstrong, J., and Patel, S. (1991). The Golgi sorting domain of coronavirus E1 protein. *J. Cell Sci.* **98**, 567–575.

- Armstrong, J., Niemann, H., Smeekens, S., Rottier, P., and Warren, G. (1984). Sequence and topology of a model intracellular membrane protein, E1 glycoprotein, from a coronavirus. *Nature (London)* **308**, 751–752.
- Armstrong, J., Patel, S., and Riddle, P. (1990). Lysosomal sorting mutants of coronavirus E1 protein, a Golgi membrane protein. *J. Cell Sci.* **95**, 191–197.
- Asanaka, M., and Lai, M. M. C. (1993). Cell fusion studies identified multiple cellular factors involved in mouse hepatitis virus entry. *Virology* **197**, 732–741.
- Baker, S. C., Shieh, C.-K., Soe, L. H., Chang, M.-F., Vannier, D. M., and Lai, M. M. C. (1989). Identification of a domain required for the autoproteolytic cleavage of murine coronavirus gene A polyprotein. *J. Virol.* **63**, 3693–3699.
- Baker, S. C., Yokomori, K., Dong, S., Carlisle, R., Gorbalenya, A. E., Koonin, E. V., and Lai, M. M. C. (1993). Identification of the catalytic sites of a papain-like cysteine proteinase of murine coronavirus. *J. Virol.* **67**, 6056–6063.
- Ballesteros, M. L., Sanchez, C. M., and Enjuanes, L. (1997). Two amino acid changes at the N-terminus of transmissible gastroenteritis coronavirus spike protein result in the loss of enteric tropism. *Virology* **227**(2), 378–388.
- Banner, L. R., and Lai, M. M. C. (1991). Random nature of coronavirus RNA recombination in the absence of selection pressure. *Virology* **185**, 441–445.
- Banner, L. R., Keck, J. G., and Lai, M. M. C. (1990). A clustering of RNA recombination sites adjacent to a hypervariable region of the peplomer gene of murine coronavirus. *Virology* **175**, 548–555.
- Baric, R. S., Stohlman, S. A., Razavi, M. K., and Lai, M. M. C. (1985). Characterization of leader-related small RNAs in coronavirus-infected cells: Further evidence for leader-primed mechanism of transcription. *Virus Res.* **3**, 19–33.
- Baric, R. S., Shieh, C.-K., Stohlman, S. A., and Lai, M. M. C. (1987). Analysis of intracellular small RNAs of mouse hepatitis virus: Evidence for discontinuous transcription. *Virology* **156**, 342–354.
- Baric, R. S., Nelson, G. W., Fleming, J. O., Deans, R. J., Keck, J. G., Casteel, N., and Stohlman, S. A. (1988). Interactions between coronavirus nucleocapsid protein and viral RNAs: Implications for viral transcription. *J. Virol.* **62**, 4280–4287.
- Baric, R. S., Fu, K., Schaad, M. C., and Stohlman, S. A. (1990). Establishing a genetic recombination map for murine coronavirus strain A59 complementation groups. *Virology* **177**, 646–656.
- Baric, R. S., Yount, B., Hensley, L., Peel, S. A., and Chen, W. (1997). Episodic evolution mediates interspecies transfer of a murine coronavirus. *J. Virol.* **71**, 1946–1955.
- Benbacer, L., Kut, E., Besnardeau, L., Laude, H., and Delmas, B. (1997). Interspecies aminopeptidase-N chimeras reveal species-specific receptor recognition by canine coronavirus, feline infectious peritonitis virus, and transmissible gastroenteritis virus. *J. Virol.* **71**, 734–737.
- Beushausen, S., Narindrasorasak, S., Sanwal, B. D., and Dales, S. (1987). In vivo and in vitro models of demyelinating disease: Activation of the adenylate cyclase system influences JHM virus expression in explanted rat oligodendrocytes. *J. Virol.* **61**, 3795–3803.
- Bonilla, P. J., Gorbalenya, A. E., and Weiss, S. R. (1994). Mouse hepatitis virus strain A59 RNA polymerase gene ORF 1a: Heterogeneity among MHV strains. *Virology* **198**, 736–740.
- Bonilla, P. J., Hughes, S. A., Pinon, J. D., and Weiss, S. R. (1995). Characterization of the leader papain-like proteinase of MHV-A59: Identification of a new in vitro cleavage site. *Virology* **209**, 489–497.
- Bonilla, P. J., Hughes, S. A., and Weiss, S. R. (1997). Characterization of a second cleavage site and demonstration of activity in trans by the papain-like proteinase of the murine coronavirus mouse hepatitis virus strain A59. *J. Virol.* **71**, 900–909.

- Bos, E. C. W., Heunen, L., Luytjes, W., and Spaan, W. J. M. (1995). Mutational analysis of the murine coronavirus spike protein: Effect on cell-to-cell fusion. *Virology* **214**, 453-463.
- Bos, E. C. W., Luytjes, W., van der Meulen, H., Koerten, H. K., and Spaan, W. J. M. (1996). The production of recombinant infectious DI-particles of a murine coronavirus in the absence of helper virus. *Virology* **218**, 52-60.
- Bournsnel, M. E. G., Brown, T. D. K., Foulds, I. J., Green, P. F., Tomley, F. M., and Binns, M. M. (1987). Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *J. Gen. Virol.* **68**, 57-77.
- Brayton, P. R., Ganges, R. G., and Stohman, S. A. (1981). Host cell nuclear function and murine hepatitis virus replication. *J. Gen. Virol.* **56**, 457-460.
- Brayton, P. R., Lai, M. M. C., Patton, C. D., and Stohman, S. A. (1982). Characterization of two RNA polymerase activities induced by mouse hepatitis virus. *J. Virol.* **42**, 847-853.
- Brayton, P. R., Stohman, S. A., and Lai, M. M. C. (1984). Further characterization of mouse hepatitis virus RNA dependent RNA polymerases. *Virology* **133**, 197-201.
- Bredenbeek, P. J., Pachuk, C. J., Noten, A. F. H., Charite, J., Luytjes, W., Weiss, S. R., and Spaan, W. J. M. (1990). The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59: A highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. *Nucleic Acids Res.* **18**, 1825-1832.
- Brierley, I., Bournsnel, M. E. G., Binns, M. M., Bilimoria, B., Blok, V. C., Brown, T. D. K., and Inglis, S. C. (1987). An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J.* **6**, 3779-3785.
- Brierley, I., Digard, P., and Inglis, S. C. (1989). Characterization of an efficient coronavirus ribosomal frame-shifting signal: Requirement for an RNA pseudoknot. *Cell (Cambridge, Mass.)* **57**, 537-547.
- Brierly, I., Bournsnel, M. E., Binns, M. M., Bilimoria, B., Rolley, N. J., Brown, T. D., and Inglis, S. C. (1990). Products of the polymerase-encoding region of the coronavirus IBV. *Adv. Exp. Med. Biol.* **276**, 275-281.
- Britton, P. (1991). Coronavirus motif [letter]. *Nature (London)* **353**, 394.
- Brown, T. D. K., and Brierly, I. (1995). The coronavirus nonstructural proteins. In "The Coronaviridae" (S. G. Siddell, ed.), pp. 191-218. Plenum, New York.
- Budzilowicz, C. J., and 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., and 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. *J. Virol.* **53**, 834-840.
- Cabirac, G. F., Soike, K. F., Zhang, J. Y., Hoel, K., Butunoi, C., Cai, G. Y., Johnson, S., and Murray, R. S. (1994). Entry of coronavirus into primate CNS following peripheral infection. *Microb. Pathog.* **16**, 349-357.
- Caul, E. O., and Egglestone, S. I. (1977). Further studies on human enteric coronavirus. *Arch. Virol.* **54**, 107-117.
- Cavanagh, D. (1983a). Coronavirus IBV glycopolypeptides: Size of their polypeptide moieties and nature of their oligosaccharides. *J. Gen. Virol.* **64**, 1187-1191.
- Cavanagh, D. (1983b). Coronavirus IBV: Further evidence that the surface projections are associated with two glycopolypeptides. *J. Gen. Virol.* **64**, 1787-1791.
- Cavanagh, D. (1983c). Coronavirus IBV: Structural characterization of the spike protein. *J. Gen. Virol.* **64**, 2577-2583.

- Cavanagh, D. (1995). The coronavirus surface glycoprotein. In "The Coronaviridae" (S. G. Siddell, ed.), pp. 73–113. Plenum, New York.
- Cavanagh, D., and Davis, P. J. (1986). Coronavirus IBV: Removal of spike glycopolypeptide S1 by urea abolishes infectivity and haemagglutination but not attachment to cells. *J. Gen. Virol.* **67**, 1443–1448.
- Cavanagh, D., and Davis, P. J. (1992). Sequence analysis of strains of avian infectious bronchitis coronavirus isolated during the 1960s in the UK. *Arch. Virol.* **130**, 471–472.
- Cavanagh, D., and Macnaughton, M. (1995). Coronaviruses. In "Principles and Practice of Clinical Virology" (A. J. Zuckerman, J. E., Bantvala, and J. R. Pattison, eds.), pp. 325–336. Wiley, Chichester.
- Cavanagh, D., Davis, P. J., Pappin, D. J., Binns, M. M., Bournsnel, M. E., and Brown, T. D. (1986a). Coronavirus IBV: Partial amino terminal sequencing of the 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 Res.* **4**, 133–143.
- Cavanagh, D., Davis, P. J., and Pappin, D. J. C. (1986b). Coronavirus IBV glycopolypeptides: Locational studies using proteases and saponin, a membrane permeabilizer. *Virus Res.* **4**, 145–156.
- Cavanagh, D., Davis, P. J., and Mockett, A. P. (1988). Amino acids within hypervariable region 1 of avian coronavirus IBV (Massachusetts serotype) spike glycoprotein are associated with neutralization epitopes. *Virus Res.* **11**, 141–150.
- Cavanagh, D., Brian, D. A., Enjuanes, L., Holmes, K. V., Lai, M. M. C., Laude, H., Siddell, S. G., Spaan, W., Taguchi, F., and Talbot, P. J. (1990). Recommendations of the Coronavirus Study Group for the nomenclature of the structural proteins, mRNAs and genes of coronaviruses. *Virology* **176**, 306–307.
- Cavanagh, D., Brian, D. A., Brinton, M. A., Enjuanes, L., Holmes, K. V., Horzinek, M. C., Lai, M. M. C., Laude, H., Plegemann, P. G. W., Siddell, S. G., Spaan, W., Taguchi, F., and Talbot, P. J. (1994). Revision of the taxonomy of the Coronavirus, Torovirus and Arterivirus genera. *Arch. Virol.* **135**, 227–237.
- Cavanagh, D., Brian, D. A., Brinton, M. A., Enjuanes, L., Holmes, K. V., Horzinek, M. C., Lai, M. M. C., Laude, H., Plegemann, P. G. W., Siddell, S. G., Spaan, W., Taguchi, F., and Talbot, P. J. (1995). Coronaviridae. In "Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses" (F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers, eds.), pp. 407–411. Springer-Verlag, Vienna and New York.
- Chang, R. Y., and Brian, D. A. (1996). Cis requirement for N-specific protein sequence in bovine coronavirus defective interfering RNA replication. *J. Virol.* **70**, 2201–2207.
- Chang, R. Y., Hofman, M. A., Sethna, P. B., and Brian, D. A. (1994). A cis-acting function for the coronavirus leader in defective-interfering RNA replication. *J. Virol.* **68**, 8223–8231.
- Chang, R. Y., Krishnan, R., and Brian, D. A. (1996). The UCUAAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA. *J. Virol.* **70**, 2720–2729.
- Charley, B., and Laude, H. (1988). Induction of alpha interferon by transmissible gastroenteritis coronavirus: Role of transmembrane glycoprotein E1. *J. Virol.* **62**, 8–11.
- Chen, B. Y., and Itakura, C. (1996). Cytopathology of chick renal epithelial cells experimentally infected with avian infectious bronchitis virus. *Avian Pathol.* **25**, 675–690.
- Chen, D. S., Asanaka, M., Yokomori, K., Wang, F.-I., Hwang, S. B., Li, H.-P., and Lai, M. M. C. (1995). A pregnancy-specific glycoprotein is expressed in the brain and serves as a receptor for mouse hepatitis virus. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 12095–12099.

- Chen, D. S., Asanaka, M., Chen, F. S., Shively, J. E., and Lai, M. M. (1997). Human carcinoembryonic antigen and biliary glycoprotein can serve as mouse hepatitis virus receptors. *J. Virol.* **71**, 1688–1691.
- Chen, W., and Baric, R. S. (1995). Function of a 5'-end genomic RNA mutation that evolves during persistent mouse hepatitis virus infection in vitro. *J. Virol.* **69**, 7529–7540.
- Chen, W., and Baric, R. S. (1996). Molecular anatomy of mouse hepatitis virus persistence: Coevolution of increased host cell resistance and virus virulence. *J. Virol.* **70**, 3947–3960.
- Collins, A. R. (1994). Human coronavirus OC43 interacts with major histocompatibility complex class I molecules at the cell surface to establish infection. *Immunol. Invest.* **23**, 313–321.
- Collins, A. R., Knobler, R. L., Powell, H., and Buchmeier, M. J. (1982). Monoclonal antibodies to murine hepatitis virus-4 (strain JHM) define the viral glycoprotein responsible for attachment and cell-cell fusion. *Virology* **119**, 358–371.
- Compton, S. R. (1994). Enterotropic strains of mouse coronavirus differ in their use of murine carcinoembryonic antigen-related glycoprotein receptors. *Virology* **203**, 197–201.
- Compton, S. R., Rogers, D. B., Holmes, K. V., Fertsch, D., Remenick, J., and McGowan, J. J. (1987). In vitro replication of mouse hepatitis virus strain A59. *J. Virol.* **61**, 1814–1820.
- Cook, J. K. A., and Mockett, A. P. A. (1995). Epidemiology of infectious bronchitis virus. In "The Coronaviridae" (S. G. Siddell, ed.), pp. 317–336. Plenum, New York.
- Corapi, W. V., Olsen, C. W., and Scott, F. W. (1992). Monoclonal antibody analysis of neutralization and antibody-dependent enhancement of feline infectious peritonitis virus. *J. Virol.* **66**, 6695–6705.
- Coutelier, J. P., Godfraind, C., Dveksler, G. S., Wysocka, M., Cardellichio, C. B., Noel, H., and Holmes, K. V. (1994). B lymphocyte and macrophage expression of carcinoembryonic antigen-related adhesion molecules that serve as receptors for murine coronavirus. *Eur. J. Immunol.* **24**, 1383–1390.
- Cox, E., Pensaert, M. B., Callebaut, P., and van Deun, K. (1990). Intestinal replication of a respiratory coronavirus closely related antigenically to the enteric transmissible gastroenteritis virus. *Vet. Microbiol.* **23**, 237–243.
- Cox, G. J., Parker, M. D., and Babiuk, L. A. (1991). Bovine coronavirus nonstructural protein ns2 is a phosphoprotein. *Virology* **185**, 509–512.
- Dackiw, A. P. B., Zakrezewski, K., Nathens, A. B., Cheung, P. Y. C., Fingerote, R., Levy, G. A., and Rotstein, O. D. (1995). Induction of macrophage procoagulant activity by murine hepatitis virus strain 3: Role of tyrosine phosphorylation. *J. Virol.* **69**, 5824–5828.
- Dales, S., and Anderson, R. (1995). Pathogenesis and diseases of the central nervous system caused by murine coronaviruses. In "The Coronaviridae" (S. G. Siddell, ed.), pp. 257–292. Plenum, New York.
- Dalziel, R. G., Lampert, P. W., Talbot, P. J., and Buchmeier, M. J. (1986). Site-specific alteration of murine hepatitis virus type-4 (MHV-4) peplomer glycoprotein E2 results in reduced neurovirulence. *J. Virol.* **59**, 463–471.
- David-Ferreira, J. F., and Manaker, R. A. (1965). An electron microscope study of the development of a mouse hepatitis virus in tissue culture cells. *J. Cell Biol.* **24**, 57–78.
- Dea, S., and Tijssen, P. (1988). Identification of the structural proteins of turkey enteric coronavirus. *Arch. Virol.* **99**, 173–186.
- de Groot, R. J., Luytjes, W., Horzinek, M. C., van der Zeijst, B. A. M., Spaan, W. J. M., and Lenstra, J. A. (1987). Evidence for a coiled-coil structure in the spike proteins of coronavirus. *J. Mol. Biol.* **196**, 963–966.

- de Groot, R. J., Van Leer, R. W., Dalderup, M. J., Vennema, H., Horzinek, M. C., and Spaan, W. J. (1989). Stably expressed FIPV peplomer protein induces cell fusion and elicits neutralizing antibodies in mice. *Virology* **171**, 493–502.
- de Groot, R. J., Rumenapf, T., Kuhn, R. J., Strauss, E. G., and Strauss, J. H. (1991). Sindbis virus RNA polymerase is degraded by the N-end rule pathway. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8967–8971.
- de Groot, R. M., van der Most, R. G., and Spaan, W. J. (1992). The fitness of defective interfering murine coronavirus DI-a and its derivatives is decreased by nonsense and frameshift mutations. *J. Virol.* **66**, 5898–5905.
- Delmas, B., and Laude, H. (1990). Assembly of coronavirus spike protein into trimers and its role in epitope expression. *J. Virol.* **64**, 5367–5375.
- Delmas, B., Gelfi, J., Haridon, R., Vogel, L. K., Sjöström, H., Noren, O., and Laude, H. (1992). Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. *Nature (London)* **357**, 417–420.
- Delmas, B., Gelfi, J., Kut, E., Sjöström, H., Noren, O., and Laude, H. (1994). Determinants essential for the transmissible gastroenteritis virus-receptor interaction reside within a domain of aminopeptidase-N that is distinct from the enzymatic site. *J. Virol.* **68**, 5216–5224.
- den Boon, J. A., Spaan, W. J., and Snijder, E. J. (1995). Equine arteritis virus subgenomic RNA transcription: UV inactivation and translation inhibition studies. *Virology* **213**, 364–372.
- Denison, M. R., and Perlman, S. (1986). Translation and processing of mouse hepatitis virus virion RNA in a cell-free system. *J. Virol.* **60**, 12–18.
- Denison, M. R., Zoltick, P. W., Hughes, S. A., Giangreco, B., Olson, A. L., Perlman, S., Leibowitz, J. L., and Weiss, S. R. (1992). Intracellular processing of the N-terminal ORF 1a proteins of the coronavirus MHV-A59 requires multiple proteolytic events. *Virology* **189**, 274–284.
- Denison, M. R., Hughes, S. R., and Weiss, S. R. (1995). Identification and characterization of a 65-kDa protein processed from the gene 1 polyprotein of the murine coronavirus MHV-A59. *Virology* **207**, 316–320.
- Dennis, D. E., and Brian, D. A. (1982). RNA-dependent RNA polymerase activity in coronavirus-infected cells. *J. Virol.* **42**, 153–164.
- Deregt, D., and Babiuk, L. A. (1987). Monoclonal antibodies to bovine coronavirus: Characteristics and topographical mapping of neutralizing epitopes on the E2 and E3 glycoproteins. *Virology* **161**, 410–420.
- Deregt, D., Gifford, G. A., Khalid Ijaz, M., Watts, T. C., Gilchrist, J. E., Haines, D. M., and Babiuk, L. A. (1989). Monoclonal antibodies to bovine coronavirus glycoproteins E2 and E3: Demonstration of in vivo neutralizing activity. *J. Gen. Virol.* **70**, 993–998.
- Dong, S. H., and Baker, S. C. (1994). Determinants of the p28 cleavage site recognized by the first papain-like cysteine proteinase of murine coronavirus. *Virology* **204**, 541–549.
- Doughri, A. M., Storz, J., Hajer, I., and Fernando, H. S. (1976). Morphology and morphogenesis of a coronavirus infecting intestinal epithelial cells of newborn calves. *Exp. Mol. Pathol.* **25**, 355–370.
- Duarte, M., Tobler, K., Bridgen, A., Rasschaert, D., Ackermann, M., and Laude, H. (1994). Sequence analysis of the porcine epidemic diarrhea virus genome between the nucleocapsid and spike protein genes reveals a polymorphic ORF. *Virology* **198**, 466–476.
- Dubois-Dalcq, M. E., Doller, E. W., Haspel, M. V., and Holmes, K. V. (1982). Cell tropism and expression of mouse hepatitis viruses (MHV) in mouse spinal cord cultures. *Virology* **119**, 317–331.

- Dupuy, J. M., and Lamontagne, L. (1987). Genetically-determined sensitivity to MHV3 infections is expressed in vitro in lymphoid cells and macrophages. *Adv. Exp. Med. Biol.* **218**, 455-463.
- Dveksler, G. S., Pensiero, M. N., Cardellichio, C. B., Williams, R. K., Jiang, G., Holmes, K. V., and Dieffenbach, C. W. (1991). Cloning of the mouse hepatitis virus (MHV) receptor: Expression in human and hamster cell lines confers susceptibility to MHV. *J. Virol.* **65**, 6881-6891.
- Dveksler, G. S., Dieffenbach, C. W., Cardellichio, C. B., McCuaig, K., Pensiero, M. N., Jiang, G., Beauchemin, N., and Holmes, K. V. (1993a). Several members of the mouse carcinoembryonic antigen-related glycoprotein family are functional receptors for the coronavirus mouse hepatitis virus-A59. *J. Virol.* **67**, 1-8.
- Dveksler, G. S., Pensiero, M. N., Dieffenbach, C. W., Cardellichio, C. B., Basile, A. A., Elia, P. E., and Holmes, K. V. (1993b). Mouse hepatitis virus strain A59 and blocking antireceptor monoclonal antibody bind to the N-terminal domain of cellular receptor. *Proc. Natl. Acad. Sci. USA* **90**, 1716-1720.
- Dveksler, G. S., Basile, A. A., Cardellichio, C. B., and Holmes, K. V. (1995). Mouse hepatitis virus receptor activities of an MHVR/mph chimera and MHVR mutants lacking N-linked glycosylation of the N-terminal domain. *J. Virol.* **69**, 543-546.
- Ebner, D., Raabe, T., and Siddell, S. G. (1988). Identification of the coronavirus MHV-JHM mRNA 4 product. *J. Gen. Virol.* **69**, 1041-1050.
- Eleouet, J. F., Rasschaert, D., Lambert, P., Levy, L., Vende, P., and Laude, H. (1995). Complete sequence (20 kilobases) of the polyprotein-encoding gene 1 of transmissible gastroenteritis virus. *Virology* **206**, 817-822.
- Enjuanes, L., and van der Zeijst, B. A. M. (1995). Molecular basis of transmissible gastroenteritis virus epidemiology. In "The Coronaviridae" (S. G. Siddell, ed.), pp. 337-376. Plenum, New York.
- Evans, M. R., and Simpson, R. W. (1980). The coronavirus avian infectious bronchitis virus requires the cell nucleus and host transcriptional factors. *Virology* **105**, 582-591.
- Fischer, F., Peng, D., Hingley, S. T., Weiss, S. R., and Masters, P. S. (1997). The internal open reading frame within the nucleocapsid gene of mouse hepatitis virus encodes a structural protein that is not essential for viral replication. *J. Virol.* **71**, 996-1003.
- Fleming, J. O., Trousdale, M. D., El-Zaatari, F. A. K., Stohlman, S. A., and Weiner, L. P. (1986). Pathogenicity of antigenic variants of murine coronavirus JHM selected with monoclonal antibodies. *J. Virol.* **58**, 869-875.
- Fleming, J. O., Trousdale, M. D., Bradbury, J., Stohlman, S. A., and Weiner, L. P. (1987). Experimental demyelination induced by coronavirus JHM (MHV-4): Molecular identification of a viral determinant of paralytic disease. *Microb. Pathog.* **3**, 9-20.
- Flintoff, W. F. (1984). Replication of murine coronaviruses in somatic cell hybrids between murine fibroblasts and rat schwannoma cells. *Virology* **134**, 450-459.
- Fosmire, J. A., Hwang, K., and Makino, S. (1992). Identification and characterization of a coronavirus packaging signal. *J. Virol.* **66**, 3522-3530.
- Frana, M. F., Behnke, J. N., Sturman, L. S., and Holmes, K. V. (1985). Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: Host-dependent differences in proteolytic cleavage and cell fusion. *J. Virol.* **56**, 912-920.
- Fu, K., and Baric, R. S. (1992). Evidence for variable rates of recombination in the MHV genome. *Virology* **189**, 88-102.
- Fu, K., and Baric, R. S. (1994). Map locations of mouse hepatitis virus temperature-sensitive mutants: Confirmation of variable rates of recombination. *J. Virol.* **68**, 7458-7466.
- Furuya, T., and Lai, M. M. C. (1993). Three different cellular proteins bind to the complementary sites on the 5'-end positive- and 3'-end negative-strands of mouse hepatitis virus RNA. *J. Virol.* **67**, 7215-7222.

- Furuya, T., Macnaughton, T. B., La Monica, N., and Lai, M. M. C. (1993). Natural evolution of coronavirus defective-interfering RNA involves RNA recombination. *Virology* **194**, 408–413.
- Gagneten, S., Gout, O., Dubois-Dalcq, M., Rottier, P., and Rossen, J. (1995). Interaction of mouse hepatitis virus (MHV) spike glycoprotein with receptor glycoprotein MHVR is required for infection with an MHV strain that expresses the hemagglutinin-esterase glycoprotein. *J. Virol.* **69**, 889–895.
- Gallagher, T. M. (1996). Murine coronavirus membrane fusion is blocked by modification of thiols buried within the spike protein. *J. Virol.* **70**, 4683–4690.
- Gallagher, T. M., Parker, S. E., and Buchmeier, M. J. (1990). Neutralization-resistant variants of a neurotropic coronavirus are generated by deletions within the amino-terminal half of the spike glycoprotein. *J. Virol.* **64**, 731–741.
- Gallagher, T. M., Escarmis, C., and Buchmeier, M. J. (1991). Alteration of the pH dependence of coronavirus-induced cell fusion: Effect of mutations in the spike glycoprotein. *J. Virol.* **65**, 1916–1928.
- Gallagher, T. M., Buchmeier, M. J., and Perlman, S. (1992). Cell receptor-independent infection by a neurotropic murine coronavirus. *Virology* **191**, 517–522.
- Garwes, D. J., and Pocock, D. H. (1975). The polypeptide structure of transmissible gastroenteritis virus. *J. Gen. Virol.* **29**, 25–34.
- Garwes, D. J., Pocock, D. H., and Pike, B. V. (1976). Isolation of subviral components from transmissible gastroenteritis virus. *J. Gen. Virol.* **32**, 283–294.
- Garwes, D. J., Stewart, F., and Britton, P. (1989). The polypeptide of M, 14,000 of porcine transmissible gastroenteritis virus: Gene assignment and intracellular location. *J. Gen. Virol.* **70**, 2495–2499.
- Godet, M., Haridon, R., Vautherot, J. F., and Laude, H. (1992). TGEV coronavirus ORF4 encodes a membrane protein that is incorporated into virions. *Virology* **188**, 666–675.
- Godet, M., Grosclaude, J., Delmas, B., and Laude, H. (1994). Major receptor-binding and neutralization determinants are located within the same domain of the transmissible gastroenteritis virus (coronavirus) spike protein. *J. Virol.* **68**, 8008–8016.
- Godfraind, C., Langreth, S. G., Cardellichio, C. B., Knobler, R., Coutelier, J. P., Dubois-Dalcq, M., and Holmes, K. V. (1995). Tissue and cellular distribution of an adhesion molecule in the carcinoembryonic antigen family that serves as a receptor for mouse hepatitis virus. *Lab. Invest.* **73**, 615–627.
- Gombold, J. L., Hingley, S. T., and Weiss, S. R. (1993). Fusion-defective mutants of mouse hepatitis virus A59 contain a mutation in the spike protein cleavage signal. *J. Virol.* **67**, 4504–4512.
- Gorbalenya, A. E., Blinov, V. M., Donchenko, A. P., and Koonin, E. V. (1989a). An NTP-binding motif is the most conserved sequence in a highly diverged monophyletic group of proteins involved in positive strand RNA viral replication. *J. Mol. Evol.* **28**, 256–258.
- Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1989b). Coronavirus genome: Prediction of putative functional domains in the nonstructural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Res.* **17**, 4847–4861.
- Gorbalenya, A. E., Koonin, E. V., and Lai, M. M. C. (1991). Putative papain-related thiol proteases of positive-strand RNA viruses. *FEBS Lett.* **288**, 201–205.
- Grosse, B., and Siddell, S. G. (1994). Single amino acid changes in the S2 subunit of the MHV surface glycoprotein confer resistance to neutralization by S1 subunit-specific monoclonal antibody. *Virology* **202**, 814–824.
- Gröttinger, C., Heusipp, G., Ziebuhr, J., Harms, U., Suss, J., and Siddell, S. G. (1996). Characterization of a 105-kDa polypeptide encoded in gene 1 of the human coronavirus HCV 229E. *Virology* **222**, 227–235.

- Haspel, M. V., Lampert, P. W., and Oldstone, M. B. A. (1978). Temperature-sensitive mutants of mouse hepatitis virus produce a high incidence of demyelination. *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4033-4036.
- Herold, J., and Siddell, S. G. (1993). An "elaborated" pseudoknot is required for high frequency frameshifting during translation of HCV 229E polymerase mRNA. *Nucleic Acids Res.* **21**, 5838-5842.
- Herold, J., Raabe, T., Schelke-Prinz, B., and Siddell, S. G. (1993). Nucleotide sequence of the human coronavirus 229E RNA polymerase locus. *Virology* **195**, 680-691.
- Hilton, A., Mizzen, L., MacIntyre, G., Cheley, S., and Anderson, R. (1986). Translational control in murine hepatitis virus infection. *J. Gen. Virol.* **67**, 923-932.
- Hiscox, J. A., Cavanagh, D., and Britton, P. (1995a). Quantification of individual subgenomic mRNA species during replication of the coronavirus transmissible gastroenteritis virus. *Virus Res.* **36**, 119-130.
- Hiscox, J. A., Mawditt, K. L., Cavanagh, D., and Britton, P. (1995b). Investigation of the control of coronavirus subgenomic mRNA transcription by using T7-generated negative-sense RNA transcripts. *J. Virol.* **69**, 6219-6227.
- Hodgman, T. C. (1988). A new superfamily of replicative proteins [letter]. *Nature (London)* **333**, 22-23; (erratum published in *Nature (London)* **333**, 1578 (1988).
- Hofmann, M. A., and Brian, D. A. (1991). The 5'-end of coronavirus minus-strand RNAs contains a short poly(U) tract. *J. Virol.* **65**, 6331-6333.
- Hofmann, M. A., Sethna, P. B., and Brian, D. A. (1990). Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture. *J. Virol.* **64**, 4108-4114.
- Hofmann, M. A., Chang, R. Y., Ku, S., and Brian, D. A. (1993a). Leader-mRNA junction sequences are unique for each subgenomic mRNA species in the bovine coronavirus and remain so throughout persistent infection. *Virology* **196**, 163-171.
- Hofmann, M. A., Senanayake, S. D., and Brian, D. A. (1993b). A translation-attenuating intraleader open reading frame is selected on coronavirus mRNAs during persistent infection. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 11733-11737.
- Hogue, B. G., and Brian, D. A. (1986). Structural proteins of human respiratory coronavirus OC43. *Virus Res.* **5**, 131-144.
- Hogue, B. G., Kienzle, T. E., and Brian, D. A. (1989). Synthesis and processing of the bovine enteric coronavirus hemagglutinin protein. *J. Gen. Virol.* **70**, 345-352.
- Holmes, K. V. (1990). Coronaviridae and their replication. In "Fields Virology" (B. N. Fields and D. M. Knipe, eds.), pp. 841-856. Raven Press, New York.
- Holmes, K. V., Doller, E. W., and Sturman, L. S. (1981). Tunicamycin-resistant glycosylation of coronavirus glycoprotein: Demonstration of a novel type of viral glycoprotein. *Virology* **115**, 334-344.
- Horsburgh, B. C., Brierley, I., and Brown, T. D. (1992). Analysis of a 9.6 kb sequence from the 3' end of canine coronavirus genomic RNA. *J. Gen. Virol.* **73**, 2849-2862.
- Hughes, S. A., Bonilla, P. J., and Weiss, S. R. (1995). Identification of the murine coronavirus p28 cleavage site. *J. Virol.* **69**, 809-813.
- Ignjatovic, J., and Galli, L. (1994). The S1 glycoprotein but not the N or M proteins of an avian infectious bronchitis virus induces protection in vaccinated chickens. *Arch. Virol.* **138**, 117-134.
- Jacobs, L., Spaan, W. J., Horzinek, M. C., and van der Zeijst, B. A. (1981). Synthesis of subgenomic mRNA's of mouse hepatitis virus is initiated independently: Evidence from UV transcription mapping. *J. Virol.* **39**, 401-406.
- Jeong, Y. S., and Makino, S. (1992). Mechanism of coronavirus transcription: Duration of primary transcription initiation activity and effects of subgenomic RNA transcription on RNA replication. *J. Virol.* **66**, 3339-3346.

- Jeong, Y. S., and Makino, S. (1994). Evidence for coronavirus discontinuous transcription. *J. Virol.* **68**, 2615–2623.
- Jeong, Y. S., Repass, J. F., Kim, Y. N., Hwang, S. M., and Makino, S. (1996). Coronavirus transcription mediated by sequences flanking the transcription consensus sequence. *Virology* **217**, 311–322.
- Jia, W., Karaca, K., Parrish, C. R., and Naqi, S. A. (1995). A novel variant of avian infectious bronchitis virus resulting from recombination among three different strains. *Arch. Virol.* **140**, 259–271.
- Joo, M., and Makino, S. (1992). Mutagenic analysis of the coronavirus intergenic consensus sequence. *J. Virol.* **66**, 6330–6337.
- Joo, M., and Makino, S. (1995). The effect of two closely inserted transcription consensus sequences on coronavirus transcription. *J. Virol.* **69**, 272–280.
- Kapke, P. A., Tung, F. Y. T., Hogue, B. G., Brian, D. A., Woods, R. D., and Wesley, R. (1988). The amino-terminal signal peptide on the porcine transmissible gastroenteritis coronavirus matrix protein is not an absolute requirement for membrane translocation and glycosylation. *Virology* **165**, 367–376.
- Keck, J. G., Stohlman, S. A., Soe, L. H., Makino, S., and 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., and Lai, M. M. C. (1988a). Temporal regulation of bovine coronavirus RNA synthesis. *Virus Res.* **9**, 343–356.
- Keck, J. G., Matsushima, G. K., Makino, S., Fleming, J. O., Vannier, D. M., Stohlman, S. A., and Lai, M. M. C. (1988b). In vivo RNA-RNA recombination of coronavirus in mouse brain. *J. Virol.* **62**, 1810–1813.
- Keck, J. G., Soe, L. H., Makino, S., Stohlman, S. A., and Lai, M. M. C. (1988c). RNA recombination of murine coronaviruses: Recombination between fusion-positive MHV-A59 and fusion-negative MHV-2. *J. Virol.* **62**, 1989–1998.
- Kennedy, D. A., and Johnson-Lussenberg, C. M. (1978). Inhibition of coronavirus 229E replication by actinomycin D. *J. Virol.* **29**, 401–404.
- Kienzle, T. E., Abraham, S., Hogue, B. G., and Brian, D. A. (1990). Structure and orientation of expressed bovine coronavirus hemagglutinin-esterase protein. *J. Virol.* **64**, 1834–1838.
- Kim, J. C., Spence, R. A., Currier, P. F., Lu, X. T., and Denison, M. R. (1995). Coronavirus protein processing and RNA synthesis is inhibited by the cysteine proteinase inhibitor E64d. *Virology* **208**, 1–8.
- Kim, K. H., and Makino, S. (1995). Two murine coronavirus genes suffice for viral RNA synthesis. *J. Virol.* **69**, 2313–2321.
- Kim, Y. N., and Makino, S. (1995). Characterization of a murine coronavirus defective interfering RNA internal cis-acting replication signal. *J. Virol.* **69**, 4963–4971.
- Kim, Y.-N., Lai, M. M. C., and Makino, S. (1993a). Generation and selection of coronavirus defective interfering RNA with large open reading frame by RNA recombination and possible editing. *Virology* **194**, 244–253.
- Kim, Y.-N., Jeong, Y. S., and Makino, S. (1993b). Analysis of cis-acting sequences essential for coronavirus defective interfering RNA replication. *Virology* **197**, 53–63.
- King, B., Potts, B. J., and Brian, D. A. (1985). Bovine coronavirus hemagglutinin protein. *Virus Res.* **2**, 53–59.
- King, D. J., and Cavanagh, D. (1991). Infectious bronchitis. In "Diseases of Poultry" (B. W. Calnek, H. J. Barnes, C. W. Beard, W. M. Reid, and H. W. Yoder, eds.), pp. 471–484. Iowa State University Press, Ames.
- Klumperman, J., Locker, J. K., Meijer, A., Horzinek, M. C., Geuze, H. J., and Rottier, P. J. (1994). Coronavirus M proteins accumulate in the Golgi complex beyond the site of virion budding. *J. Virol.* **68**, 6523–6534.

- Knobler, R. L., Taylor, B. A., Wooddell, M. K., Beamer, W. G., and Oldstone, M. B. A. (1984). Host genetic control of mouse hepatitis virus type-4 (JHM strain) replication. II. The gene locus for susceptibility is linked to the Svp-2 locus on mouse chromosome 7. *Exp. Clin. Immunogenet.* **1**, 217-222.
- Koetzner, C. A., Parker, M. M., Ricard, C. S., Sturman, L. S., and Masters, P. S. (1992). Repair and mutagenesis of the genome of a deletion mutant of the coronavirus mouse hepatitis virus by targeted RNA recombination. *J. Virol.* **66**, 1841-1848.
- Konings, D. A. M., Bredenbeek, P. J., Noten, J. F. H., Hogeweg, P., and Spaan, W. J. M. (1988). Differential premature termination of transcription as a proposed mechanism for the regulation of coronavirus gene expression. *Nucleic Acids Res.* **16**, 10849-10860.
- Kooi, C., Mizzen, L., Alderson, C., Daya, M., and Anderson, R. (1988). Early events of importance in determining host cell permissiveness to mouse hepatitis virus infection. *J. Gen. Virol.* **69**, 1125-1135.
- Kooi, C., Cervin, M., and Anderson, R. (1991). Differentiation of acid-pH-dependent and -nondependent entry pathways for mouse hepatitis virus. *Virology* **180**, 108-119.
- Koolen, M. J., Osterhaus, A. D. M. E., van Steenis, G., Horzinek, M. C., and van der Zeijst, B. A. M. (1983). Temperature-sensitive mutants of mouse hepatitis virus strain A59: Isolation, characterization and neuropathogenic properties. *Virology* **125**, 393-402.
- Koopmans, M., and Horzinek, M. C. (1994). Toroviruses of animals and humans: A review. *Adv. Virus Res.* **43**, 233-273.
- Kotier, S. A., Cavanagh, D., and Britton, P. (1995). Experimental evidence of recombination in coronavirus infectious bronchitis virus. *Virology* **213**, 569-580.
- Krijnse-Locker, J., Ericsson, M., Rottier, P. J. M., and Griffiths, G. (1994). Characterization of the budding compartment of mouse hepatitis virus: Evidence that transport from the RER to the Golgi complex requires only one vesicular transport step. *J. Cell Biol.* **124**, 55-70.
- Krishnan, R., Chang, R. Y., and Brian, D. A. (1996). Tandem placement of a coronavirus promoter results in enhanced mRNA synthesis from the downstream-most initiation site. *Virology* **218**, 400-405.
- Krzystyniak, K., and Dupuy, J. M. (1981). Early interaction between mouse hepatitis virus 3 and cells. *J. Gen. Virol.* **57**, 53-61.
- Krzystyniak, K., and Dupuy, J. M. (1984). Entry of mouse hepatitis virus 3 into cells. *J. Gen. Virol.* **65**, 227-231.
- Kubo, H., Yamada, Y. K., and Taguchi, F. (1994). Localization of neutralizing epitopes and the receptor-binding site within the amino-terminal 330 amino acids of the murine coronavirus spike protein. *J. Virol.* **68**, 5403-5410.
- Kunkel, F., and Herrler, G. (1993). Structural and functional analysis of the surface protein of human coronavirus OC43. *Virology* **195**, 195-202.
- Kusters, J. G., Niesters, H. G. M., Lenstra, J. A., Horzinek, M. C., and van der Zeijst, B. A. M. (1989). Phylogeny of antigenic variants of avian coronavirus IBV. *Virology* **169**, 217-221.
- Labonte, P., Mounir, S., and Talbot, P. J. (1995). Sequence and expression of the ns2 protein gene of human coronavirus OC43. *J. Gen. Virol.* **76**, 431-435.
- Lai, M. M. C. (1986). Coronavirus leader RNA-primed transcription: An alternative mechanism to RNA splicing. *BioEssays* **5**, 257-260.
- Lai, M. M. C. (1990). Coronavirus: Organization, replication and expression of genome. *Annu. Rev. Microbiol.* **44**, 303-333.
- Lai, M. M. C. (1992). RNA recombination in animal and plant viruses. *Microbiol. Rev.* **56**, 61-79.

- Lai, M. M. C., Brayton, P. R., Armen, R. C., Patton, C. D., Pugh, C., and Stohlman, S. A. (1981). Mouse hepatitis virus A59: Messenger RNA structure and genetic localization of the sequence divergence from the hepatotropic strain MHV 3. *J. Virol.* **39**, 823-834.
- Lai, M. M. C., Patton, C. D., and Stohlman, S. A. (1982). Further characterization of mRNAs of mouse hepatitis virus: Presence of common 5'-end nucleotides. *J. Virol.* **41**, 557-656.
- Lai, M. M. C., Patton, C. D., Baric, R. S., and Stohlman, S. A. (1983). The presence of leader sequences in the mRNA of mouse hepatitis virus. *J. Virol.* **46**, 1027-1033.
- Lai, M. M. C., Baric, R. S., Brayton, P. R., and Stohlman, S. A. (1984). Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus-a cytoplasmic RNA virus. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3626-3630.
- Lai, M. M. C., Baric, R. S., Makino, S., Keck, J. G., Egbert, J., Leibowitz, J. L., and Stohlman, S. A. (1985). Recombination between nonsegmented RNA genomes of murine coronaviruses. *J. Virol.* **56**, 449-456.
- Lai, M. M. C., Makino, S., Soe, L. H., Shieh, C.-K., Keck, J. G., and Fleming, J. O. (1987). Coronavirus: A jumping RNA transcription. *Cold Spring Harbor Lab. Symp.* **52**, 359-365.
- Lai, M. M. C., Liao, C.-L., Lin, Y.-J., and Zhang, X. (1994). Coronavirus: How a large RNA viral genome is replicated and transcribed. *Infect. Agents Dis.* **3**, 98-105.
- La Monica, N., Banner, L. R., Morris, V. L., and Lai, M. M. C. (1991). Localization of extensive deletions in the structural genes of two neurotropic variants of murine coronavirus JHM. *Virology* **182**, 883-888.
- La Monica, N., Yokomori, K., and Lai, M. M. C. (1992). Coronavirus mRNA synthesis: Identification of novel transcription initiation signals which are differentially regulated by different leader sequences. *Virology* **188**, 402-407.
- Lapps, W., Hogue, B. G., and Brian, D. A. (1987). Sequence analysis of the bovine coronavirus nucleocapsid and matrix protein genes. *Virology* **157**, 47-57.
- Laude, H. (1993). Porcine respiratory coronavirus: Molecular features and virus-host interactions. *Vet. Res.* **24**, 125-150.
- Laude, H., and Masters, P. S. (1995). The coronavirus nucleocapsid protein. *In* "The Coronaviridae" (S. G. Siddell, ed.), pp. 141-158. Plenum, New York.
- Laude, H., Rasschaert, D., and Huet, J. C. (1987). Sequence and N-terminal processing of the transmembrane protein E1 of the coronavirus transmissible gastroenteritis virus. *J. Gen. Virol.* **68**, 1687-1693.
- Laude, H., Gelfi, J., Lavenant, L., and Charley, B. (1992). Single amino acid changes in the viral glycoprotein M affect induction of alpha interferon by the coronavirus transmissible gastroenteritis virus. *J. Virol.* **66**, 743-749.
- Laviada, M. D., Videgain, S. P., Moreno, L., Alonso, F., Enjuanes, L., and Escribano, J. M. (1990). Expression of swine transmissible gastroenteritis virus envelope antigens on the surface of infected cells: Epitopes externally exposed. *Virus Res.* **16**, 247-254.
- Le, S. Y., Sonenberg, N., and Maizel, J. V. J. (1994). Distinct structural elements and internal entry of ribosomes in mRNA3 encoded by infectious bronchitis virus. *Virology* **198**, 405-411.
- Lee, H.-J., Shieh, C.-K., Gorbalenya, A. E., Koonin, E. V., La Monica, N., Tuler, J., Bagdzyahdzhyan, A., and Lai, M. M. C. (1991). The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* **180**, 567-582.
- Leibowitz, J. L., and DeVries, J. R. (1988). Synthesis of virus-specific RNA in permeabilized murine coronavirus-infected cells. *Virology* **166**, 66-75.

- Leibowitz, J. L., Wilhelmsen, K. C., and Bond, C. W. (1981). The virus specific intracellular RNA species of two murine coronaviruses: MHV A59 and MHV-JHM. *Virology* **114**, 39–51.
- Leibowitz, J. L., DeVries, J. R., and Haspel, M. V. (1982a). Genetic analysis of murine hepatitis virus strain JHM. *J. Virol.* **42**, 1080–1087.
- Leibowitz, J. L., Weiss, S. R., Paavola, E., and Bond, C. W. (1982b). Cell-free translation of murine coronavirus RNA. *J. Virol.* **43**, 905–913.
- Leibowitz, J. L., Perlman, S., Weinstock, G., DeVries, J. R., Budzillowicz, C., Weissemann, J. M., and Weiss, S. R. (1988). Detection of a murine coronavirus nonstructural protein encoded in a downstream open reading frame. *Virology* **164**, 156–164.
- Lewis, E. L., Harbour, D. A., Beringer, J. E., and Grinstead, J. (1993). Differential in vitro inhibition of feline enteric coronavirus and feline infectious peritonitis virus by actinomycin D. *J. Gen. Virol.* **73**, 3285–3288.
- Li, D., and Cavanagh, D. (1992). Coronavirus IBV-induced membrane fusion occurs at near-neutral pH. *Arch. Virol.* **122**, 307–316.
- Liao, C.-L., and Lai, M. M. C. (1992). RNA recombination in a coronavirus: Recombination between viral genomic RNA and transfected RNA fragments. *J. Virol.* **66**, 6117–6124.
- Liao, C.-L., and Lai, M. M. C. (1994). Requirement of the 5'-end genomic sequence as an upstream cis-acting element for coronavirus subgenomic mRNA transcription. *J. Virol.* **68**, 4727–4737.
- Liao, C. L., and Lai, M. M. C. (1995). A cis-acting viral protein is not required for the replication of a coronavirus defective-interfering RNA. *Virology* **209**, 428–436.
- Liao, C. L., Zhang, X., and Lai, M. M. C. (1995). Coronavirus defective-interfering RNA as an expression vector: The generation of a pseudorecombinant mouse hepatitis virus expressing hemagglutinin-esterase. *Virology* **208**, 319–327.
- Lin, Y.-J., and Lai, M. M. C. (1993). Deletion mapping of a mouse hepatitis virus defective-interfering RNA reveals the requirement of an internal and discontinuous sequence for replication. *J. Virol.* **67**, 6110–6118.
- Lin, Y.-J., Liao, C. L., and Lai, M. M. C. (1994). Identification of the cis-acting signal for minus-strand RNA synthesis of a murine coronavirus: Implications for the role of minus-strand RNA in RNA replication and transcription. *J. Virol.* **68**, 8131–8140.
- Lin, Y.-J., Zhang, X., Wu, R.-C., and Lai, M. M. C. (1996). The 3' untranslated region of the coronaviral RNA is required for subgenomic mRNA transcription from a defective-interfering RNA. *J. Virol.* **70**, 7236–7240.
- Liu, D. X., and Brown, T. D. K. (1995). Characterisation and mutational analysis of an ORF 1a-encoding proteinase domain responsible for proteolytic processing of the infectious bronchitis virus 1a/1b polyprotein. *Virology* **209**, 420–427.
- Liu, D. X., and Inglis, S. C. (1992a). Identification of two new polypeptides encoded by mRNA5 of the coronavirus infectious bronchitis virus. *Virology* **186**, 342–347.
- Liu, D. X., and Inglis, S. C. (1992b). Internal entry of ribosomes on a tricistronic mRNA encoded by infectious bronchitis virus. *J. Virol.* **66**, 6143–6154; erratum published in *J. Virol.* **66**, 6840 (1992).
- Liu, D. X., Cavanagh, D., Green, P., and Inglis, S. C. (1991). A polycistronic mRNA specified by the coronavirus infectious bronchitis virus. *Virology* **184**, 531–544.
- Liu, D. X., Brierley, I., Tibbles, K. W., and Brown, T. D. K. (1994). A 100-kilodalton polypeptide encoded by open reading frame (ORF) 1b of the coronavirus infectious bronchitis virus is processed by ORF 1a products. *J. Virol.* **68**, 5772–5780.
- Locker, J. K., Griffiths, G., Horzinek, M. C., and Rottier, P. J. M. (1992a). O-glycosylation of the coronavirus M protein. Differential localization of sialyltransferases in N- and O-linked glycosylation. *J. Biol. Chem.* **267**, 14094–14101.

- Locker, J. K., Rose, J. K., Horzinek, M. C., and Rottier, P. J. M. (1992b). Membrane assembly of the triple-spanning coronavirus M protein. Individual transmembrane domains show preferred orientation. *J. Biol. Chem.* **267**, 21911–21918.
- Locker, J. K., Klumperman, J., Oorschot, V., Horzinek, M. C., Geuze, H. J., and Rottier, P. J. (1994). The cytoplasmic tail of mouse hepatitis virus M protein is essential but not sufficient for its retention in the Golgi complex. *J. Biol. Chem.* **269**, 28263–28269.
- Locker, J. K., Opstelten, D. J. E., Ericsson, M., Horzinek, M. C., and Rottier, P. J. M. (1995). Oligomerization of a trans-Golgi/trans-Golgi network retained protein occurs in the Golgi complex and may be part of its retention. *J. Biol. Chem.* **270**, 8815–8821.
- Lu, X. T., Lu, Y. Q., and Denison, M. R. (1996). Intracellular and in vitro-translated 27-kDa proteins contain the 3C-like proteinase activity of the coronavirus MHV-A59. *Virology* **222**, 375–382.
- Lu, Y. Q., Lu, X. T., and Denison, M. R. (1995). Identification and characterization of a serine-like proteinase of the murine coronavirus MHV-A59. *J. Virol.* **69**, 3554–3559.
- Luytjes, W., Sturman, L. S., Bredenbeek, P. J., Charite, J., van der Zeijst, B. A. M., Horzinek, M. C., and Spaan, W. J. M. (1987). Primary structure of the glycoprotein E2 of coronavirus MHV-A59 and identification of the trypsin cleavage site. *Virology* **161**, 479–487.
- Luytjes, W., Bredenbeek, P. J., Noten, A. F. H., Horzinek, M. C., and Spaan, W. J. M. (1988). Sequence of mouse hepatitis virus A59 mRNA 2: Indications for RNA-recombination between coronavirus and influenza C virus. *Virology* **166**, 415–422.
- Luytjes, W., Geerts, D., Posthumus, W., Melen, R., and Spaan, W. (1989). Amino acid sequence of a conserved neutralizing epitope of murine coronaviruses. *J. Virol.* **63**, 1408–1412.
- Luytjes, W., Gerritsma, H., and Spaan, W. J. M. (1996). Replication of synthetic defective interfering RNAs derived from coronavirus mouse hepatitis virus-A59. *Virology* **216**, 174–183.
- Luytjes, W., Gerritsma, H., Bos, E., and Spaan, W. (1997). Characterization of two temperature-sensitive mutants of coronavirus mouse hepatitis virus strain A59 with maturation defects in the spike protein. *J. Virol.* **71**, 949–955.
- Machamer, C. E., and Rose, J. K. (1987). A specific transmembrane domain of a coronavirus E1 glycoprotein is required for its retention in the Golgi region. *J. Cell Biol.* **105**, 1205–1214.
- Machamer, C. E., Mentone, S. A., Rose, J. K., and Farquhar, M. G. (1990). The E1 glycoprotein of an avian coronavirus is targeted to the cis Golgi complex. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6944–6948.
- Machamer, C. E., Grim, M. G., Esquela, A., Chung, S. W., Rolls, M., Ryan, K., and Swift, A. M. (1993). Retention of a cis Golgi protein requires polar residues on one face of a predicted alpha-helix in the transmembrane domain. *Mol. Biol. Cell* **4**, 695–704.
- Macnaughton, M. R., and Davies, H. A. (1980). Two particle types of avian infectious bronchitis virus. *J. Gen. Virol.* **47**, 365–372.
- Macnaughton, M. R., Davies, H. A., and Nermut, M. V. (1978). Ribonucleoprotein-like structures from coronavirus particles. *J. Gen. Virol.* **39**, 545–549.
- Mahy, B. W., Siddell, S., Wege, H., and ter Meulen, V. (1983). RNA-dependent RNA polymerase activity in murine coronavirus-infected cells. *J. Gen. Virol.* **64**, 103–111.
- Makino, S., and Joo, M. (1993). Effect of intergenic consensus sequence flanking sequences on coronavirus transcription. *J. Virol.* **67**, 3304–3311.
- Makino, S., and Lai, M. M. C. (1989a). Evolution of the 5'-end of genomic RNA of murine coronaviruses during passages in vitro. *Virology* **169**, 227–232.
- Makino, S., and Lai, M. M. C. (1989b). High-frequency leader sequence switching during coronavirus defective interfering RNA replication. *J. Virol.* **63**, 5285–5292.

- Makino, S., Fujioka, N., and Fujiwara, K. (1985). Structure of the intracellular defective viral RNAs of defective interfering particles of mouse hepatitis virus. *J. Virol.* **54**, 329–336.
- Makino, S., Keck, J. G., Stohman, S. A., and Lai, M. M. C. (1986a). High-frequency RNA recombination of murine coronaviruses. *J. Virol.* **57**, 729–737.
- Makino, S., Stohman, S. A., and Lai, M. M. C. (1986b). Leader sequences of murine coronavirus mRNAs can be freely reassorted: Evidence for the role of free leader RNA in transcription. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 4204–4208.
- Makino, S., Fleming, J. O., Keck, J. G., Stohman, S. A., and Lai, M. M. C. (1987). RNA recombination of coronaviruses: Localization of neutralizing epitopes and neuropathogenic determinants on the carboxyl terminus of peplomers. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6567–6571.
- Makino, S., Shieh, C.-K., Keck, J. G., and Lai, M. M. C. (1988a). Defective interfering particles of murine coronavirus: Mechanism of synthesis of defective viral RNAs. *Virology* **163**, 104–111.
- Makino, S., Shieh, C.-K., Soe, L. H., Baker, S. C., and Lai, M. M. C. (1988b). Primary structure and translation of a defective-interfering RNA of murine coronavirus. *Virology* **166**, 550–560.
- Makino, S., Soe, L. H., Shieh, C.-K., and Lai, M. M. C. (1988c). Discontinuous transcription generates heterogeneity at the leader fusion sites of coronavirus mRNAs. *J. Virol.* **62**, 3870–3873.
- Makino, S., Yokomori, K., and Lai, M. M. C. (1990). Analysis of efficiently packaged defective-interfering RNAs of murine coronavirus: Localization of a possible RNA-packaging signal. *J. Virol.* **64**, 6045–6053.
- Makino, S., Joo, M., and Makino, J. K. (1991). A system for study of coronavirus mRNA synthesis: A regulated, expressed subgenomic defective-interfering RNA results from intergenic site insertion. *J. Virol.* **65**, 6031–6041.
- Masters, P. S. (1992). Localization of an RNA-binding domain in the nucleocapsid protein of the coronavirus mouse hepatitis virus. *Arch. Virol.* **125**, 141–160.
- Masters, P. S., Koetzner, C. A., Kerr, C. A., and Heo, Y. (1994). Optimization of targeted RNA recombination and mapping of a novel nucleocapsid gene mutation in the coronavirus mouse hepatitis virus. *J. Virol.* **68**, 328–337.
- Mayer, T., Tamura, T., Falk, M., and Niemann, H. (1988). Membrane integration and intracellular transport of the coronavirus glycoprotein E1, a class III membrane glycoprotein. *J. Biol. Chem.* **263**, 14956–14963.
- Mendez, A., Smerdou, C., Izeta, A., Gebauer, F., and Enjuanes, L. (1996). Molecular characterization of transmissible gastroenteritis coronavirus defective interfering genomes: Packaging and heterogeneity. *Virology* **217**, 495–507.
- Mizzen, L., Hilton, A., Cheley, S., and Anderson, R. (1985). Attenuation of murine coronavirus infection by ammonium chloride. *Virology* **142**, 378–388.
- Mohandas, D. V., and Dales, S. (1991). Endosomal association of a protein phosphatase with high dephosphorylating activity against a coronavirus nucleocapsid protein. *FEBS Lett.* **282**, 419–424.
- Mounir, S., and Talbot, P. J. (1993). Human coronavirus OC43 RNA 4 lacks two open reading frames located downstream of the S gene of bovine coronavirus. *Virology* **192**, 355–360.
- Murray, R. S., Cai, G., Hoel, K., Zhang, J., Soike, K. F., and Cabirac, G. F. (1992). Coronavirus infects and causes demyelination in primate central nervous system. *Virology* **188**, 274–284.
- Nakada, S., Creager, R. S., Krystal, M., Aaronson, R. P., and Palese, P. (1984). Influenza C virus hemagglutinin: Comparison with influenza A and B virus hemagglutinins. *J. Virol.* **50**, 118.

- Nedellec, P., Dveksler, G. S., Daniels, E., Turbide, C., Chow, B., Basile, A. A., Holmes, K. V., and Beauchemin, N. (1994). Bgp2, a new member of the carcinoembryonic antigen-related gene family, encodes an alternative receptor for mouse hepatitis virus. *J. Virol.* **68**, 4525–4537.
- Nelson, G. W., and Stohlman, S. A. (1993). Localization of the RNA-binding domain of MHV nucleocapsid protein. *J. Gen. Virol.* **74**, 1975–1979.
- Niemann, H., Boschek, B., Evans, D., Rosing, M., Tamura, T., and Klenk, H. D. (1982). Post-translational glycosylation of coronavirus glycoprotein E1: Inhibition by monensin. *EMBO J.* **1**, 1499–1504.
- Niemann, H., Geyer, R., Klenk, H. D., Linder, D., Stirm, S., and Wirth, M. (1984). The carbohydrates of mouse hepatitis virus (MHV) A59: Structures of the O-glycosidically linked oligosaccharides of glycoprotein E1. *EMBO J.* **3**, 665–670.
- Oberst, M. D., Gollan, T. J., Gupta, M., Peura, S. R., Zydlewski, J. D., Sudarsanan, P., and Lawson, T. G. (1993). The encephalomyocarditis virus 3C protease is rapidly degraded by an ATP-dependent proteolytic system in reticulocyte lysate. *Virology* **193**, 28–40.
- Oleszak, E. L., and Leibowitz, J. L. (1990). Immunoglobulin Fc binding activity is associated with the mouse hepatitis virus E2 peplomer protein. *Virology* **176**, 70–80.
- Oleszak, E. L., Perlman, S., and Leibowitz, J. L. (1992). MHV S peplomer protein expressed by a recombinant vaccinia virus vector exhibits IgG Fc-receptor activity. *Virology* **186**, 122–132.
- Oleszak, E. L., Kuzmak, J., Hogue, B., Parr, R., Collisson, E. W., Rodkey, L. S., and Leibowitz, J. L. (1995). Molecular mimicry between Fc receptor and S peplomer protein of mouse hepatitis virus, bovine coronavirus, and transmissible gastroenteritis virus. *Hybridoma* **14**, 1–8.
- Olsen, C. W., Corapi, W. V., Ngichabe, C. K., Baines, J. D., and Scott, F. W. (1992). Monoclonal antibodies to the spike protein of feline infectious peritonitis virus mediate antibody-dependent enhancement of infection of feline macrophages. *J. Virol.* **66**, 956–965.
- Opstelten, D. J. E., de Groote, P., Horzinek, M. C., Vennema, H., and Rottier, P. J. M. (1993). Disulfide bonds in folding and transport of mouse hepatitis coronavirus glycoproteins. *J. Virol.* **67**, 7394–7401.
- Opstelten, D. J. E., Raamsman, M. J., Wolfs, K., Horzinek, M. C., and Rottier, P. J. M. (1995). Envelope glycoprotein interactions in coronavirus assembly. *J. Cell Biol.* **131**, 339–349.
- Parker, M. D., Cox, G. J., Deregt, D., Fitzpatrick, D. R., and Babiuk, L. A. (1989). Cloning and in vitro expression of the gene for the E3 haemagglutinin glycoprotein of bovine coronavirus. *J. Gen. Virol.* **70**, 155–164.
- Parker, M. D., Yoo, D., and Babiuk, L. A. (1990). Expression and secretion of the bovine coronavirus hemagglutinin-esterase glycoprotein by insect cells infected with recombinant baculoviruses. *J. Virol.* **64**, 1625–1629.
- Parker, M. M., and Masters, P. S. (1990). Sequence comparison of the N genes of five strains of the coronavirus mouse hepatitis virus suggests a three domain structure for the nucleocapsid protein. *Virology* **179**, 463–468.
- Parker, S. E., Gallagher, T. M., and Buchmeier, M. J. (1989). Sequence analysis reveals extensive polymorphism and evidence of deletions within the E2 glycoprotein gene of several strains of murine hepatitis virus. *Virology* **173**, 664–673.
- Payne, H. R., and Storz, J. (1988). Analysis of cell fusion induced by bovine coronavirus infection. *Arch. Virol.* **103**, 27–33.
- Peng, D., Koetzner, C. A., and Masters, P. S. (1995a). Analysis of second-site revertants of a murine coronavirus nucleocapsid protein deletion mutant and construction of nucleocapsid protein mutants by targeted RNA recombination. *J. Virol.* **69**, 3449–3457.

- Peng, D., Koetzner, C. A., McMahon, T., Zhu, Y., and Masters, P. S. (1995b). Construction of murine coronavirus mutants containing interspecies chimeric nucleocapsid proteins. *J. Virol.* **69**, 5475–5484.
- Pensaert, M., Callebaut, P., and Vengote, J. (1986). Isolation of a porcine respiratory, non-enteric coronavirus related to transmissible gastroenteritis. *Vet. Q.* **8**, 257–261.
- Penzes, Z., Tibbles, K., Shaw, K., Britton, P., Brown, T. D. K., and Cavanagh, D. (1994). Characterization of a replicating and packaged defective RNA of avian coronavirus infectious bronchitis virus. *Virology* **203**, 286–293.
- Penzes, Z., Wroe, C., Brown, T. D. K., Britton, P., and Cavanagh, D. (1996). Replication and packaging of coronavirus infectious bronchitis virus defective RNAs lacking a long open reading frame. *J. Virol.* **70**, 8660–8668.
- Perlman, S., Ries, D., Bolger, E., Chang, L. J., and Stoltzfus, C. M. (1986). MHV nucleocapsid synthesis in the presence of cyclohexamide and accumulation of negative-strand MHV RNA. *Virus Res.* **6**, 261–272.
- Pfleiderer, M., Routledge, E., and Siddell, S. G. (1990). Functional analysis of the coronavirus MHV-JHM surface glycoproteins in vaccinia virus recombinants. *Adv. Exp. Med. Biol.* **276**, 21–31.
- Pfleiderer, M., Routledge, E., Herrler, G., and Siddell, S. G. (1991). High level transient expression of the murine coronavirus haemagglutinin-esterase. *J. Gen. Virol.* **72**, 1309–1315.
- Plagemann, P. G. W., and Moennig, V. (1992). Lactate dehydrogenase-elevating virus, equine arteritis virus and simian haemorrhagic fever virus, a new group of positive strand RNA viruses. *Adv. Virus Res.* **41**, 99–192.
- Rasschaert, D., Duarte, M., and Laude, H. (1990). Porcine respiratory coronavirus differs from transmissible gastroenteritis virus by a few genomic deletions. *J. Gen. Virol.* **71**, 2599–2607.
- Ricard, C. S., Koetzner, C. A., Sturman, L. S., and Masters, P. S. (1995). A conditional-lethal murine coronavirus mutant that fails to incorporate the spike glycoprotein into assembled virions. *Virus Res.* **39**, 261–276.
- Risco, C., Anton, I. M., Sune, C., Pedregosa, A. M., Martin-Alonso, J. M., Parra, F., Carrascosa, J. L., and Enjuanes, L. (1995). Membrane protein molecules of transmissible gastroenteritis coronavirus also expose the carboxy-terminal region on the external surface of the virion. *J. Virol.* **69**, 5269–5277.
- Risco, C., Anton, I. M., Enjuanes, L., and Carrascosa, J. L. (1996). The transmissible gastroenteritis coronavirus contains a spherical core shell consisting of M and N proteins. *J. Virol.* **70**, 4773–4777.
- Robb, J. A., Bond, C. W., and Leibowitz, J. L. (1979). Pathogenic murine coronaviruses. III. Biological and biochemical characterization of temperature-sensitive mutants of JHMV. *Virology* **94**, 385–399.
- Robbins, S. G., Frana, M. F., McGowan, J. J., Boyle, J. F., and Holmes, K. V. (1986). RNA-binding proteins of coronavirus MHV: Detection of monomeric and multimeric N protein with an RNA overlay-protein blot assay. *Virology* **150**, 402–410.
- Rossen, J. W. A., Bekker, C. P. J., Voorhout, W. F., Strous, G. J. A. M., van der Ende, A., and Rottier, P. J. M. (1994). Entry and release of transmissible gastroenteritis coronavirus are restricted to apical surfaces of polarized epithelial cells. *J. Virol.* **68**, 7966–7973.
- Rossen, J. W. A., Voorhout, W. F., Horzinek, M. C., van der Ende, A., Strous, G. J. A. M., and Rottier, P. J. M. (1995a). MHV-A59 enters polarized murine epithelial cells through the apical surface but is released basolaterally. *Virology* **210**, 54–66.
- Rossen, J. W. A., Horzinek, M. C., and Rottier, P. J. M. (1995b). Coronavirus infection of polarised epithelial cells. *Trends Microbiol.* **3**, 486–490.

- Rossen, J. W., Strous, G. J. A. M., Horzinek, M. C., and Rottier, P. J. M. (1997). Mouse hepatitis virus strain A59 is released from opposite sides of different epithelial cell types. *J. Gen. Virol.* **78**, 61–69.
- Rottier, P., Branderberg, D., Armstrong, J., van der Zeijst, B., and Warren, G. (1984). Assembly in vitro of a spanning membrane protein of the endoplasmic reticulum: The E1 glycoprotein of coronavirus mouse hepatitis virus A59. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1421–1425.
- Rottier, P. J. M., Welling, G. W., Welling-Wester, S., Niesters, H. G. M., Lenstra, J. A., and van der Zeijst, B. A. M. (1986). Predicted membrane topology of the coronavirus protein E1. *Biochemistry* **25**, 1335–1339.
- Rottier, P. J. M. (1995). The coronavirus membrane glycoprotein. In “The Coronaviridae” (S. G. Siddell, ed.) pp. 115–139. Plenum, New York.
- Routledge, E., Stauber, R., Pfeleiderer, M., and Siddell, S. G. (1991). Analysis of murine coronavirus surface glycoprotein functions by using monoclonal antibodies. *J. Virol.* **65**, 254–262.
- Rowe, C. L., Baker, S. C., Nathan, M. J., and Fleming, J. O. (1997). Evolution of mouse hepatitis virus: Detection and characterization of spike deletion variants during persistent infection. *J. Virol.* **71**, 2959–2969.
- Sawicki, S. G., and Sawicki, D. L. (1986). Coronavirus minus-strand RNA synthesis and effect of cycloheximide on coronavirus RNA synthesis. *J. Virol.* **57**, 328–334.
- Sawicki, S. G., and Sawicki, D. L. (1990). Coronavirus transcription: Subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J. Virol.* **64**, 1050–1056.
- Sawicki, S. G., Lu, J. H., and Holmes, K. V. (1995). Persistent infection of cultured cells with mouse hepatitis virus (MHV) results from epigenetic expression of the MHV receptor. *J. Virol.* **69**, 5535–5543.
- Schaad, M. C., and Baric, R. S. (1993). Evidence for new transcriptional units encoded at the 3' end of the mouse hepatitis virus genome. *Virology* **196**, 190–198.
- Schaad, M. C., and Baric, R. S. (1994). Genetics of mouse hepatitis virus transcription: Evidence that subgenomic negative strands are functional templates. *J. Virol.* **68**, 8169–8179.
- Schaad, M. C., Stohman, S. A., Egbert, J., Lum, K., Fu, K., Wei, T., and Baric, R. S. (1990). Genetics of mouse hepatitis virus transcription: Identification of cistrons which may function in positive and negative strand RNA synthesis. *Virology* **177**, 634–645.
- Schmidt, I., Skinner, M., and Siddell, S. (1987). Nucleotide sequence of the gene encoding the surface projection glycoprotein of coronavirus MHV-JHM. *J. Gen. Virol.* **68**, 47–56.
- Schmidt, M. F. G. (1982). Acylation of viral spike glycoproteins: A feature of enveloped RNA viruses. *Virology* **116**, 327–338.
- Schultze, B., and Herrler, G. (1992). Bovine coronavirus uses *N*-acetyl-9-*O*-acetylneuraminic acid as a receptor determinant to initiate the infection of cultured cells. *J. Gen. Virol.* **73**, 901–906.
- Schultze, B., Gross, H. J., Brossmer, R., Klenk, H. D., and Herrler, G. (1990). Hemagglutinating encephalomyelitis virus attaches to *N*-acetyl-9-*O*-acetylneuraminic acid-containing receptors on erythrocytes: Comparison with bovine coronavirus and influenza C virus. *Virus Res.* **16**, 185–194.
- Schultze, B., Gross, H. J., Brossmer, R., and Herrler, G. (1991a). The S protein of bovine coronavirus is a hemagglutinin recognizing 9-*O*-acetylated sialic acid as a receptor determinant. *J. Virol.* **65**, 6232–6237.
- Schultze, B., Wahn, K., Klenk, H.-D., and Herrler, G. (1991b). Isolated HE-protein from hemagglutinating encephalomyelitis virus and bovine coronavirus has receptor-destroying and receptor-binding activity. *Virology* **180**, 221–228.

- Schultze, B., Cavanagh, D., and Herrler, G. (1992). Neuraminidase treatment of avian infectious bronchitis coronavirus reveals a hemagglutinating activity that is dependent on sialic acid-containing receptors on erythrocytes. *Virology* **189**, 792-794.
- Schwarz, B., Routledge, E., and Siddell, S. G. (1990). Murine coronavirus nonstructural protein ns2 is not essential for virus replication in transformed cells. *J. Virol.* **64**, 4784-4791.
- Senanayake, S. D., Hofmann, M. A., Maki, J. L., and Brian, D. A. (1992). The nucleocapsid protein gene of bovine coronavirus is bicistronic. *J. Virol.* **66**, 5277-5283.
- Sethna, P. B., Hung, S. L., and Brian, D. A. (1989). Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5626-5630.
- Sethna, P. B., Hofmann, M. A., and Brian, D. A. (1991). Minus-strand copies of replicating coronavirus mRNAs contain antileaders. *J. Virol.* **65**, 320-325.
- Seybert, A., Ziebuhr, J., and Siddell, S. G. (1997). Expression and characterization of a recombinant murine coronavirus 3C-like proteinase. *J. Gen. Virol.* **78**, 71-75.
- Shieh, C.-K., Lee, H.-J., Yokomori, K., La Monica, N., Makino, S., and Lai, M. M. C. (1989). Identification of a new transcriptional initiation site and the corresponding functional gene 2b in the murine coronavirus RNA genome. *J. Virol.* **63**, 3729-3736.
- Shif, I., and Bang, F. B. (1970). In vitro interaction of mouse hepatitis virus and macrophages from genetically resistant mice. I. Adsorption of virus and growth curves. *J. Exp. Med.* **131**, 843-862.
- Siddell, S. G., ed. (1995a). "The Coronaviridae." Plenum, New York.
- Siddell, S. G. (1995b). The Coronaviridae: An introduction. In "The Coronaviridae" (S. G. Siddell, ed.), pp. 1-10. Plenum, New York.
- Siddell, S. G. (1995c). The small membrane protein. In "The Coronaviridae" (S. G. Siddell, ed.), pp. 181-190. Plenum, New York.
- Siddell, S., Wege, H., Barthel, A., and ter Meulen, V. (1980). Coronavirus JHM: Intracellular protein synthesis. *J. Gen. Virol.* **53**, 145-155.
- Smith, A. L., Cardellicchio, C. B., Winograd, D. F., de Souza, M. S., Barthold, S. W., and Holmes, K. V. (1991). Monoclonal antibody to the receptor for murine coronavirus MHV-A59 inhibits viral replication in vivo. *J. Infect. Dis.* **163**, 879-882.
- Smith, A. R., Bournsnel, M. E., Binns, M. M., Brown, T. D., and Inglis, S. C. (1990). Identification of a new membrane-associated polypeptide specified by the coronavirus infectious bronchitis virus. *J. Gen. Virol.* **71**, 3-11.
- Snijder, E. J., and Horzinek, M. C. (1993). Toroviruses: Replication, evolution and comparison with other members of the coronavirus-like superfamily. *J. Gen. Virol.* **74**, 2305-2316.
- Snijder, E. J., and Horzinek, M. C. (1995). The molecular biology of toroviruses. In "The Coronaviridae" (S. G. Siddell, ed.), pp. 219-238. Plenum, New York.
- Snijder, E. J., and Spaan, W. J. M. (1995). The coronavirus superfamily. In "The Coronaviridae" (S. G. Siddell, ed.). Plenum, New York.
- Snijder, E. J., den Boon, J. A., Horzinek, M. C., and Spaan, W. J. (1991). Comparison of the genome organization of toro- and coronaviruses: Evidence for two nonhomologous RNA recombination events during Berne virus evolution. *Virology* **180**, 448-452.
- Soe, L. H., Shieh, C.-K., Baker, S. C., Chang, M.-F., and 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. *J. Virol.* **61**, 3968-3976.
- Somogyi, P., Jenner, A. J., Brierley, I., and Inglis, S. C. (1993). Ribosomal pausing during translation of an RNA pseudoknot. *Mol. Cell. Biol.* **13**, 6931-6940.
- Spaan, W. J. M., Rottier, P. J. M., Horzinek, M. C., and van der Zeijst, B. A. M. (1981). Isolation and identification of virus-specific mRNA in cells infected with mouth hepatitis virus (MHV-A59). *Virology* **108**, 424-434.

- Spaan, W. J. M., Delius, H., Skinner, M., Armstrong, J., Rottier, P., Smeekens, S., van der Zeijst, B. A. M., and Siddell, S. G. (1983). Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO J.* **2**, 1839-1844.
- Stauber, R., Pfeleiderer, M., and Siddell, S. (1993). Proteolytic cleavage of the murine coronavirus surface glycoprotein is not required for fusion activity. *J. Gen. Virol.* **74**, 183-191.
- Stern, D. F., and Kennedy, S. I. T. (1980a). Coronavirus multiplication strategy. I. Identification and characterization of virus-specified RNA. *J. Virol.* **34**, 665-674.
- Stern, D. F., and Kennedy, S. I. T. (1980b). Coronavirus multiplication strategy. II. Mapping the avian infectious bronchitis virus intracellular RNA species to the genome. *J. Virol.* **36**, 440-449.
- Stern, D. F., and Sefton, B. M. (1982a). Synthesis of coronavirus mRNAs: Kinetics of inactivation of infectious bronchitis virus RNA synthesis by UV light. *J. Virol.* **42**, 755-759.
- Stern, D. F., and Sefton, B. M. (1982b). Coronavirus proteins: Structure and function of the oligosaccharides of the avian infectious bronchitis virus glycoproteins. *J. Virol.* **44**, 804-812.
- Stohlman, S. A., and Lai, M. M. C. (1979). Phosphoproteins of murine hepatitis viruses. *J. Virol.* **32**, 672-675.
- Stohlman, S. A., Baric, R. S., Nelson, G. N., Soe, L. H., Welter, L. M., and Deans, R. J. (1988). Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J. Virol.* **62**, 4288-4295.
- Storz, J., Herrler, G., Snodgrass, D. R., Hussain, K. A., and Zhang, X. M. (1991). Monoclonal antibodies differentiate between the haemagglutinating and the receptor-destroying activities of bovine coronavirus. *J. Gen. Virol.* **72**, 2817-2820.
- Sturman, L. S., and Holmes, K. V. (1983). The molecular biology of coronaviruses. *Adv. Virus Res.* **28**, 35-112.
- Sturman, L. S., Holmes, K. V., and Behnke, J. (1980). Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. *J. Virol.* **33**, 449-462.
- Sturman, L. S., Ricard, C. S., and Holmes, K. V. (1985). Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: Activation of cell fusing activity of virions by trypsin and separation of two different 90K cleavage fragments. *J. Virol.* **56**, 904-911.
- Sturman, L. S., Ricard, C. S., and Holmes, K. V. (1990). Conformational change of the coronavirus peplomer glycoprotein at pH 8.0 and 37 degrees C correlates with virus aggregation and virus-induced cell fusion. *J. Virol.* **64**, 3042-3050.
- Sugiyama, K., and Amano, Y. (1980). Hemagglutination and structural polypeptides of a new coronavirus associated with diarrhea in infant mice. *Arch. Virol.* **66**, 95-105.
- Suzuki, H., and Taguchi, F. (1996). Analysis of receptor-binding site of murine coronavirus spike protein. *J. Virol.* **70**, 2632-2636.
- Swift, A. M., and Machamer, C. E. (1991). A Golgi retention signal in a membrane-spanning domain of coronavirus E1 protein. *J. Cell Biol.* **115**, 19-30.
- Taguchi, F. (1993). Fusion formation by the uncleaved spike protein of murine coronavirus JHMV variant c1-2. *J. Virol.* **67**, 1195-1202.
- Taguchi, F. (1995). The S2 subunit of the murine coronavirus spike protein is not involved in receptor binding. *J. Virol.* **70**, 7260-7263.
- Tahara, S. M., Dietlin, T. A., Bergmann, C. C., Nelson, G. W., Kyuwa, S., Anthony, R. P., and Stohlman, S. A. (1994). Coronavirus translational regulation: Leader affects mRNA efficiency. *Virology* **202**, 621-630.
- Talbot, P. J., Paquette, J. S., Ciurli, C., Antel, J. P., and Ouellet, F. (1996). Myelin basic protein and human coronavirus 229E cross-reactive T cells in multiple sclerosis. *Ann. Neurol.* **39**, 233-240.

- Thiel, V., and Siddell, S. G. (1994). Internal ribosome entry in the coding region of murine hepatitis virus mRNA 5. *J. Gen. Virol.* **75**, 3041–3046.
- Tibbles, K. W., Brierley, I., Cavanagh, D., and Brown, T. D. K. (1995). A region of the coronavirus infectious bronchitis virus 1a polyprotein encoding the 3C-like protease domain is subject to rapid turnover when expressed in rabbit reticulocyte lysate. *J. Gen. Virol.* **76**, 3059–3070.
- Tibbles, K. W., Brierly, I., Cavanagh, D., and Brown, T. D. K. (1996). Characterization in vitro of an autocatalytic processing activity associated with the predicted 3C-like proteinase domain of the coronavirus avian infectious bronchitis virus. *J. Virol.* **70**, 1923–1930.
- Tooze, J., and Tooze, S. A. (1985). Infection of AtT20 murine pituitary tumour cells by mouse hepatitis virus strain A59: Virus budding is restricted to the Golgi region. *Eur. J. Cell Biol.* **37**, 203–212.
- Tooze, J., Tooze, S. A., and Warren, G. (1984). Replication of coronavirus MHV-A59 in sac-cells: Determination of the first site of budding of progeny virions. *Eur. J. Cell Biol.* **33**, 281–293.
- Tooze, J., Tooze, S. A., and Fuller, S. D. (1987). Sorting of progeny coronavirus from condensed secretory proteins at the exit from the trans-Golgi network of AtT20 cells. *J. Cell Biol.* **105**, 1215–1226.
- Tooze, S. A., Tooze, J., and Warren, G. (1988). Site of addition of *N*-acetyl-galactosamine to the E1 glycoprotein of mouse hepatitis virus-A59. *J. Cell Biol.* **106**, 1475–1487.
- Torres, J. M., Sanchez, C., Sune, C., Smerdou, C., Prevec, L., Graham, F., and Enjuanes, L. (1995). Induction of antibodies protecting against transmissible gastroenteritis coronavirus (TGEV) by recombinant adenovirus expressing TGEV spike protein. *Virology* **213**, 503–516.
- Tresnan, D. B., Levis, R., and Holmes, K. V. (1996). Feline aminopeptidase N serves as a receptor for feline, canine, porcine, and human coronaviruses in serogroup I. *J. Virol.* **70**, 8669–8674.
- Tsunemitsu, H., Elkanawati, Z. R., Smith, D. R., Reed, H. H., and Saif, L. J. (1995). Isolation of coronaviruses antigenically indistinguishable from bovine coronavirus from wild ruminants with diarrhea. *J. Clin. Microbiol.* **33**, 3264–3269.
- Tung, F. Y. T., Abraham, S., Sethna, M., Hung, S.-L., Sethna, P., Hogue, B. G., and Brian, D. A. (1992). The 9-kDa hydrophobic protein encoded at the 3'-end of the porcine transmissible gastroenteritis coronavirus genome is membrane-associated. *Virology* **186**, 676–683.
- van Berlo, M. F., van den Brink, W. J., Horzinek, M. C., and van der Zeijst, B. A. M. (1987). Fatty acid acylation of viral proteins murine hepatitis virus-infected cells. Brief report. *Arch. Virol.* **95**, 123–128.
- van der Most, R. G., Bredenbeek, P. J., and Spaan, W. J. M. (1991). A domain at the 3'-end of the polymerase gene is essential for encapsidation of coronavirus defective interfering RNAs. *J. Virol.* **65**, 3219–3226.
- van der Most, R. G., Heijnen, L., Spaan, W. J. M., and de Groot, R. J. (1992). Homologous RNA recombination allows efficient introduction of site-specific mutations into the genome of coronavirus MHV-A59 via synthetic co-replicating RNAs. *Nucleic Acids Res.* **20**, 3375–3381.
- van der Most, R. G., de Groot, R. J., and Spaan, W. J. M. (1994). Subgenomic RNA synthesis directed by a synthetic defective interfering RNA of mouse hepatitis virus: A study of coronavirus transcription initiation. *J. Virol.* **68**, 3656–3666.
- van der Most, R. G., Luytjes, W., Rutjes, S., and Spaan, W. J. M. (1995). Translation but not the encoded sequence is essential for the efficient propagation of the defective interfering RNAs of the coronavirus mouse hepatitis virus. *J. Virol.* **69**, 3744–3751.

- van Dinter, S., and Flintoff, W. F. (1987). Rat glial C6 cells are defective in murine coronavirus internalization. *J. Gen. Virol.* **68**, 1677-1685.
- Van Marle, G., Luytjes, W., van der Most, R. G., van der Straaten, T., and Spaan, W. J. M. (1995). Regulation of coronavirus mRNA transcription. *J. Virol.* **69**, 7851-7856.
- Vaughn, E. M., Halbur, P. G., and Paul, P. S. (1995). Sequence comparison of porcine respiratory coronavirus isolates reveals heterogeneity in the S, 3, and 3-1 genes. *J. Virol.* **69**, 3176-3184.
- Vennema, H., Heijnen, L., Zijderveld, A., Horzinek, M. C., and Spaan, W. J. M. (1990a). Intracellular transport of recombinant coronavirus spike proteins: Implications for virus assembly. *J. Virol.* **64**, 339-346.
- Vennema, H., de Groot, R. J., Harbour, D. A., Dalderup, M., Gruffydd-Jones, T., Horzinek, M. C., and Spaan, W. J. M. (1990b). Early death after feline infectious peritonitis virus challenge due to recombinant vaccinia virus immunization. *J. Virol.* **64**, 1407-1409.
- Vennema, H., de Groot, R. J., Harbour, D. A., Horzinek, M. C., and Spaan, W. J. M. (1991). Primary structure of the membrane and nucleocapsid protein genes of feline infectious peritonitis virus and immunogenicity of recombinant vaccinia viruses in kittens. *Virology* **181**, 327-335.
- Vennema, H., Heijnen, L., Rottier, P. J. M., Horzinek, M. C., and Spaan, W. J. M. (1992). A novel glycoprotein of feline infectious peritonitis coronavirus contains a KDEL-like endoplasmic reticulum retention signal. *J. Virol.* **66**, 4951-4956.
- Vennema, H., Godeke, G. J., Rossen, J. W. A., Voorhout, W. F., Horzinek, M. C., Opstelten, D. J. E., and Rottier, P. J. M. (1996). Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. *EMBO J.* **15**, 2020-2028.
- Vlasak, R., Luytjes, W., Lieder, J., Spaan, W. J. M., and Palese, P. (1988a). The E3 protein of bovine coronavirus is a receptor-destroying enzyme with acetyltransferase activity. *J. Virol.* **62**, 4686-4690.
- Vlasak, R., Luytjes, W., Spaan, W. J. M., and Palese, P. (1988b). Human and bovine coronaviruses recognize sialic acid-containing receptors similar to those of influenza C viruses. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4526-4529.
- Wang, F.-I., Fleming, J. O., and Lai, M. M. C. (1992). Sequence analysis of the spike protein gene of murine coronavirus variants: Study of genetic sites affecting neuropathogenicity. *Virology* **186**, 742-749.
- Wang, L., Junker, D., and Collison, E. W. (1993). Evidence of natural recombination within the S1 gene of the infectious bronchitis virus. *Virology* **192**, 710-716.
- Wang, L., Junker, D., Hock, L., Ebiary, E., and Collisson, E. W. (1994). Evolutionary implications of genetic variations in the S1 gene of infectious bronchitis virus. *Virus Res.* **34**, 327-338.
- Wege, H., Stephenson, J. R., Koga, M., Wege, H., and ter Meulen, V. (1981). Genetic variation of neurotropic and non-neurotropic murine coronaviruses. *J. Gen. Virol.* **54**, 67-74.
- Wege, H., Winter, J., and Meyermann, R. (1988). The peplomer protein E2 of coronavirus JHM as a determinant of neurovirulence: Definition of critical epitopes by variant analysis. *J. Gen. Virol.* **69**, 87-98.
- Weingartl, H. M., and Derbyshire, J. B. (1994). Evidence for a putative second receptor for porcine transmissible gastroenteritis virus on the villous enterocytes of newborn pigs. *J. Virol.* **68**, 7253-7259.
- Weismiller, D. G., Sturman, L. S., Buchmeier, M. J., Fleming, J. O., and Holmes, K. V. (1990). Monoclonal antibodies to the peplomer glycoprotein of coronavirus mouse

- hepatitis virus identify two subunits and detect a conformational change in the subunit released under mild alkaline conditions. *J. Virol.* **64**, 3051–3055.
- Weiss, M., and Horzinek, M. C. (1987). The proposed family Toroviridae: agents of enteric infections. Brief review. *Arch. Virol.* **92**, 1–15.
- Weiss, R. C., and Scott, F. W. (1981). Pathogenesis of feline infectious peritonitis: nature and development of viremia. *Am. J. Vet. Res.* **42**, 382–390.
- Weisz, O. A., Swift, A. M., and Machamer, C. E. (1993). Oligomerization of a membrane protein correlates with its retention in the Golgi complex. *J. Cell Biol.* **122**, 1185–1196.
- Welsh, R. M., Haspel, M. V., Parker, D. C., and Holmes, K. V. (1986). Natural cytotoxicity against mouse hepatitis virus-infected cells. II. A cytotoxic effector cell with a B lymphocyte phenotype. *J. Immunol.* **136**, 1454–1460.
- Wesley, R. D., Woods, R. D., Hill, H. T., and Biwer, J. D. (1990). Evidence for a porcine respiratory coronavirus, antigenically similar to transmissible gastroenteritis virus, in the United States. *J. Vet. Diagn. Invest.* **2**, 312–317.
- Wesley, R. D., Woods, R. D., and Cheung, A. K. (1991). Genetic analysis of porcine respiratory coronavirus, an attenuated variant of transmissible gastroenteritis virus. *J. Virol.* **65**, 3369–3373.
- Wilhelmsen, K. C., Leibowitz, J. L., Bond, C. W., and Robb, J. A. (1981). The replication of murine coronaviruses in enucleated cells. *Virology* **110**, 225–230.
- Williams, G. D., Chang, R.-Y., and Brian, D. A. (1995). Evidence for a pseudoknot in the 3'-untranslated region of the bovine coronavirus genome. In "Corona and Related Viruses" (P. J. Talbot and G. A. Levy, eds.), pp. 511–514. Plenum, New York.
- Williams, R. K., Jiang, G., and Holmes, K. V. (1991). Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5533–5536.
- Wilson, G. A. R., and Dales, S. (1988). In vivo and in vitro models of demyelinating disease: Efficiency of virus spread and formation of infectious centers among glial cells is genetically determined by the murine host. *J. Virol.* **62**, 3371–3377.
- Wilson, G. A. R., Mohandas, D. V., and Dales, S. (1990). In vivo and in vitro models of demyelinating disease. Possible relationship between induction of regulatory subunit from cAMP dependent protein kinases and inhibition of JHMV replication in cultured oligodendrocytes. *Adv. Exp. Med. Biol.* **276**, 261–266.
- Woo, K., Joo, M., Narayanan, K., Kim, K. H., and Makino, S. (1997). Murine coronavirus packaging signal confers packaging to nonviral RNA. *J. Virol.* **71**, 824–827.
- Yeager, C. L., Ashum, R. A., Williams, R. K., Cardellicchio, C. B., Shapiro, L. H., Look, A. T., and Holmes, K. V. (1992). Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature (London)* **357**, 420–422.
- Yokomori, K., and Lai, M. M. C. (1991). Mouse hepatitis virus S RNA sequence reveals that nonstructural proteins NS 4 and NS 5a are not essential for murine coronavirus replication. *J. Virol.* **65**, 5605–5608.
- Yokomori, K., and Lai, M. M. C. (1992a). Mouse hepatitis virus utilizes two carcinoembryonic antigens as alternative receptors. *J. Virol.* **66**, 6194–6199.
- Yokomori, K., and Lai, M. M. C. (1992b). The receptor for mouse hepatitis virus in the resistant mouse strain SJL is functional: Implications for the requirement of a second factor for viral infection. *J. Virol.* **66**, 6931–6938.
- Yokomori, K., La Monica, N., Makino, S., Shieh, C.-K., and Lai, M. M. C. (1989). Biosynthesis, structure, and biological activities of envelope protein gp65 of murine coronavirus. *Virology* **173**, 683–691.
- Yokomori, K., Banner, L. R., and Lai, M. M. C. (1991). Heterogeneity of gene expression of hemagglutinin-esterase (HE) protein of murine coronaviruses. *Virology* **183**, 647–657.

- Yokomori, K., Baker, S. C., Stohlman, S. A., and Lai, M. M. C. (1992a). Hemagglutinin-esterase (HE)-specific monoclonal antibodies alter the neuropathogenicity of mouse hepatitis virus. *J. Virol.* **66**, 2865-2874.
- Yokomori, K., Banner, L. R., and Lai, M. M. C. (1992b). Coronavirus mRNA transcription: UV light transcriptional mapping studies suggest an early requirement for a genomic-length template. *J. Virol.* **66**, 4671-4678.
- Yokomori, K., Asanaka, M., Stohlman, S. A., and Lai, M. M. C. (1993). A spike protein-dependent cellular factor other than the viral receptor is required for mouse hepatitis virus entry. *Virology* **196**, 45-56.
- Yokomori, K., Asanaka, M., Stohlman, S. A., Makino, S., Shubin, R. A., Gilmore, W., Weiner, L. P., Wang, F. I., and Lai, M. M. C. (1995). Neuropathogenicity of mouse hepatitis virus JHM isolates differing in hemagglutinin-esterase protein expression. *J. Neurovirol.* **1**, 330-339.
- Yoo, D., Graham, F. L., Prevec, L., Parker, M. D., Benko, M., Zamb, T., and Babiuk, L. A. (1992). Synthesis and processing of the haemagglutinin-esterase glycoprotein of bovine coronavirus encoded in the E3 region of adenovirus. *J. Gen. Virol.* **73**, 2591-2600.
- Yoo, D. W., Parker, M. D., and Babiuk, L. A. (1991). The S2 subunit of the spike glycoprotein of bovine coronavirus mediates membrane fusion in insect cells. *Virology* **180**, 395-399.
- Yoshikura, H., and Taguchi, F. (1978). Mouse hepatitis virus strain MHV-S: Formation of pseudotypes with a murine leukemia virus envelope. *Intervirology* **10**, 132-136.
- Yu, W., and Leibowitz, J. L. (1995a). A conserved motif at the 3' end of mouse hepatitis virus genomic RNA required for host protein binding and viral RNA replication. *Virology* **214**, 128-138.
- Yu, W., and Leibowitz, J. L. (1995b). Specific binding of host cellular proteins to multiple sites within the 3'-end of mouse hepatitis virus genomic RNA. *J. Virol.* **69**, 2016-2023.
- Yu, X., Bi, W., Weiss, S. R., and Leibowitz, J. L. (1994). Mouse hepatitis virus gene 5b protein is a new virion envelope protein. *Virology* **202**, 1018-1023.
- Zhang, X., and Lai, M. M. C. (1994). Unusual heterogeneity of leader-mRNA fusion in a murine coronavirus: Implications for the mechanism of RNA transcription and recombination. *J. Virol.* **68**, 6626-6633.
- Zhang, X., and Lai, M. M. C. (1995). Interactions between the cytoplasmic proteins and the intergenic (promoter) sequence of mouse hepatitis virus RNA: Correlation with the amounts of subgenomic mRNA transcribed. *J. Virol.* **69**, 1637-1644.
- Zhang, X., and Lai, M. M. C. (1996). A 5'-proximal RNA sequence of murine coronavirus as a potential initiation site for genomic-length mRNA transcription. *J. Virol.* **70**, 705-711.
- Zhang, X., Kousoulas, K. G., and Storz, J. (1991). The hemagglutinin/esterase glycoprotein of bovine coronaviruses: Sequence and functional comparisons between virulent and avirulent strains. *Virology* **185**, 847-852.
- Zhang, X., Herbst, W., Kousoulas, K. G., and Storz, J. (1994a). Biological and genetic characterization of a hemagglutinating coronavirus isolated from a diarrhoeic child. *J. Med. Virol.* **44**, 152-161.
- Zhang, X., Liao, C.-L., and Lai, M. M. C. (1994b). Coronavirus leader RNA regulates and initiates subgenomic mRNA transcription, both in trans and in cis. *J. Virol.* **68**, 4738-4746.
- Zhao, X., Shaw, K., and Cavanagh, D. (1993). Presence of subgenomic mRNAs in virions of coronavirus IBV. *Virology* **196**, 172-178.
- Zhou, M., Williams, A. K., Chung, S.-I., Wang, L., and Collisson, E. W. (1996). The infectious bronchitis virus nucleocapsid protein binds RNA sequences in the 3' terminus of the genome. *Virology* **217**, 191-199.
- Ziebuhr, J., Herold, J., and Siddell, S. G. (1995). Characterization of a human coronavirus (strain 229E) 3C-like proteinase activity. *J. Virol.* **69**, 4331-4338.
- Zoltick, P. W., Leibowitz, J. L., Oleszak, E. L., and Weiss, S. R. (1990). Mouse hepatitis virus ORF 2a is expressed in the cytosol of infected mouse fibroblasts. *Virology* **174**, 605-607.