Annu. Rev. Microbiol. 1990. 44:303–33

CORONAVIRUS: ORGANIZATION, REPLICATION AND EXPRESSION OF GENOME

Michael M. C. Lai

Department of Microbiology, University of Southern California, School of Medicine, Los Angeles, California 90033

KEY WORDS: leader-primed RNA transcription, RNA recombination, mouse hepatitis virus, RNA synthesis, defective interfering RNA

CONTENTS

OVERVIEW

Coronaviruses include a group of RNA viruses of medical and veterinary importance, all of which are characterized by spherical, enveloped virus particles with prominent surface projections, resembling the corona of the

0066-4227/90/1001-0303$02.00

sun. They infect various animal species, causing respiratory, gastrointestinal, cardiovascular, and neurological diseases. In humans, coronaviruses have been associated with common colds, diarrhea, and possibly multiple sclerosis. The prototypical coronaviruses include avian infectious bronchitis virus (IBV), mouse hepatitis virus (MHV), porcine transmissible gastroenteritis virus (TGEV), bovine coronavirus (BCV), human coronavirus (HCV), feline infectious peritonitis virus (FIPV), rat sialodacryoadenitis virus (SDAV), turkey coronavirus (TCV), rabbit coronavirus (RCV), and several viruses of other animal species. The majority of the studies on the molecular biology of coronavirus have been carried out with MHV, IBV, TGEV, BCV, and, to a lesser extent, with HCV and FIPV. These studies have revealed many unique features of coronavirus biology. The scope of this review focuses mainly on the organization, transcription, and replication of coronavirus RNA. Other aspects of coronavirus biology have been dealt with in earlier reviews of Wege et al (126), Sturman & Holmes (116), and Lai (52, 53). A comprehensive summary of the coronavirus gene structure and gene products can also be found in a review by Spaan et al (104).

STRUCTURE OF CORONAVIRUS VIRION

Coronavirus virions are spherical, enveloped virus particles, ranging from 80 to 160 nm in diameter. Prominent surface projections of up to 20 nm in length cover the entire virion surface (117), giving it the corona appearance. In some species of coronavirus, e. g. TCV, smaller spikes are also seen on the virion surface (27). Inside the envelope resides a helical nucleocapsid of 6–8 nm in diameter, which can be released from virus particles by mild treatments with either NP-40 or Triton X-100 (115, 117). The formation of a helical nucleocapsid in coronavirus is unexpected, since the helical nucleocapsid is usually associated with viruses containing a negative-stranded RNA genome; yet, coronavirus contains a positive-stranded RNA genome (see below). Typically, positive-stranded RNA viruses have icosahedral nucleocapsids. This fact might have functional implications in the mechanism of coronavirus RNA synthesis.

All coronavirus virion particles contain three to four structural proteins. The first one is the spike protein (S; old terms: E2 or gp180), which constitutes the spikes, or peplomers, on the virion envelope. It weighs roughly 180 kd and is frequently cleaved into two proteins of 90 kd in most of coronaviruses (23, 33, 118). In MHV, this cleavage has been shown to be mediated by cellular proteases and is required for the activation of viral infectivity (118). In some other coronaviruses, e. g. FIPV and TGEV, the S protein is not cleaved, however. The predicted sequence of the S protein

suggests that it contains an N-terminal signal sequence and a C-terminal hydrophobic membrane-anchoring domain (8, 70, 90, 95) and that it may assume a coil-to-coil structure (28). It contains N-linked carbohydrate chains (108, 114) and may be further modified by palmitic acid (118). This protein is responsible for virus binding to virus receptors on the target cells, induction of cell-to-cell fusion, elicitation of neutralizing antibodies, and cell-mediated immunity (114, 116). The monoclonal antibodies specific for the S protein can neutralize virus infectivity and inhibit fusion-inducing ability of the virus (25, 124).

The second viral structural protein, designated M (old terms: E1, gp23) is also an integral membrane glycoprotein. Sequence and biochemical analysis suggests that most of the protein is hydrophobic and spans the envelope several times (1, 24, 92, 93). The protein is glycosylated by either N-linked carbohydrates, as in IBV (108, 109), or O-linked carbohydrates, as in MHV (41, 84). The function of this protein is unclear: it is unlike the matrix protein present in other enveloped viruses. The inside domain of the M protein may interact with the virion nucleocapsid (117). This interaction could be the focal point for the assembly of virus particles because virus budding appears to occur at the site of M accumulation (120). The monoclonal antibodies specific for the M protein do not neutralize virus infectivity (25, 124).

The third glycoprotein on the virion surface is HE (old terms: E3 or gp65). The presence of this glycoprotein in coronavirus virion appears to be optional: it is present in BCV, TCV, HCV, and some strains of MHV, but is absent in IBV and TGEV (33, 37, 39, 49, 88, 108, 129). This protein most likely constitutes the smaller spikes observed on virus particles in some electron micrographs of coronaviruses (27). It is glycosylated by N-linked carbohydrates (129). An interesting aspect of this glycoprotein is that it shares some sequence homology with the hemagglutinin protein of influenza C virus (69). Correspondingly, the HE protein of BCV also exhibits hemagglutinating and esterase activity (122, 123) and some strains of MHV show esterase activity (129). The latter activity represents the receptor-destroying activity in influenza C virus (121). Furthermore, monoclonal antibodies specific for the HE protein of BCV can inhibit virus-induced hemagglutination and neutralize viral infectivity (32). Thus, HE protein may play a significant role in coronavirus biology. However, the optional nature of this protein suggests that it is dispensable, and may serve an unknown luxury function for the virus.

The fourth structural protein is an internal component of the virus. This protein, N, is a phosphoprotein of 50 kd (113), which constitutes the nucleocapsid protein of the virus. The protein binds to virion RNA, providing the structural basis for the helical nucleocapsid. The RNA-binding activity of the N protein has been demonstrated in vitro (4, 112). The protein appears to bind

specifically to a segment of the leader RNA (see below) (112). It may provide not only a structural role, but possibly also regulatory functions for viral RNA synthesis.

GENOME STRUCTURE AND GENE ASSIGNMENTS

Coronavirus contains a single piece of nonsegmented RNA genome with an estimated molecular weight of 6×10^6 to 8×10^6 daltons (60, 68, 125). The RNA contains a 5' cap structure and 3' poly(A) tail (61, 128) and is infectious upon transfection of the naked RNA into a susceptible cell line. The RNA also serves as a template for in vitro translation of viral proteins (66, 100). Thus, this RNA is a typical positive-stranded RNA. No negative-stranded RNA has been detected in the virion (61). The complete sequence of genomic RNA for IBV RNA (11) revealed a big surprise, i. e. the genomic RNA of IBV is 27.6 kilobases (kb) long, much bigger than the previous size estimates for coronavirus RNAs. Current estimates of MHV genomic RNA by cDNA cloning and sequencing showed that it has an even larger genome, i. e. 32 kb (2, 86). These RNAs are considerably larger than any other known viral RNA (the next smaller viral RNA is paramyxovirus RNA, which is approximately 14 kb long). The fact that these genomic RNAs are larger than the previous estimates was not a surprise because no suitable RNA markers of this size range were available for the determination of RNA sizes by gel electrophoresis. The surprise was that RNA of such a large size can exist at all, considering the high error frequency of RNA-dependent RNA synthesis (40). This large size of the virion RNA entails that the virus must have a unique mechanism of synthesis to counter the deleterious effects of possible errors in RNA synthesis.

Sequencing analysis of the virion genomic RNAs reveals the presence of at least 10 open reading frames (ORFs), some of which correspond to a distinct mRNA species (see below). The ORFs corresponding to the same mRNA are considered to belong to the same gene. Based on the number of mRNA species (see below), coronaviruses are thought to have six to eight separate genes. These ORFs encode various structural and nonstructural proteins of coronaviruses. Figure 1 shows the gene arrangements in different coronaviruses. Several striking features are noticeable: 1. The gene orders in the genome of different coronaviruses are comparable. The order of genes 4 and 5 of IBV is reversed, however, when compared with other coronaviruses. TGEV probably contains an additional gene (gene 7) at the 3' end, which has a small ORF and a corresponding mRNA (31, 130). 2. MHV and BCV contain two additional genes, encoding the HE protein and a nonstructural protein, p30, neither of which is present in IBV or TGEV. 3. Several overlapping ORFs are included within a gene. For instance, genes 1 and 5 of MHV contain two ORFs each and genes 1 and 3 of

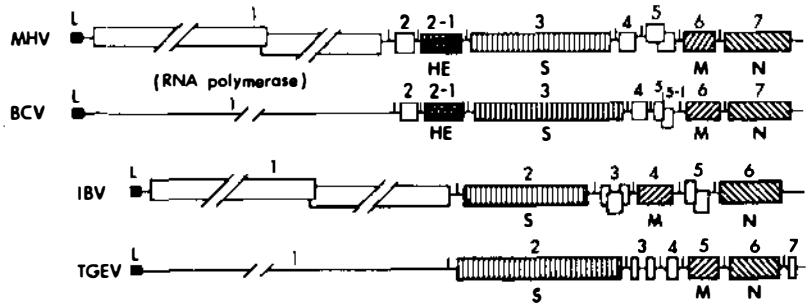


Figure 1 Gene organization of coronavirus genome. Capital letters indicate the structural protein genes. The numbers on top of the genes are designations of mRNAs corresponding to each gene. L stands for the leader RNA. The structure of gene 1 of BCV and TGEV has not been determined. The organization of MHV RNA is deduced from published sequences (98, 101) and unpublished sequence (C.-K. Shieh & M. M. C. Lai); that of IBV from published sequence (10, 11, 102). The structure of BCV and TGEV are from Refs. 62 and 90 and unpublished data of D. Brian.

IBV contain two and three ORFs, respectively. The significance of gene rearrangement and insertion of genes in different coronavirus RNA genomes is discussed in later sections. 4. Gene 1, which probably encodes nonstructural proteins including RNA polymerase, accounts for more than two-thirds (17–23 kb) of the genome.

IBV and MHV RNA genomes contain approximately 200–500 nucleotides of the untranslated sequence at the 5' end (11, 86, 103). The terminal 60–70 nucleotides are termed *leader sequence* and are transposed to the 5' ends of every mRNA species (19, 54, 105). This leader sequence is important in the regulation and transcription of coronavirus genomic RNA and subgenomic mRNAs. At every intergenic region, there is a small stretch of consensus sequence of approximately 8–10 nucleotides (18, 21, 99), which is important for the transcription of the downstream genes (see below).

GENERAL PATHWAY OF CORONAVIRUS REPLICATION

Coronaviruses, in general, have very restricted host ranges, infecting only cells of their own host species. However, some cross-species infections do occur. For instance, BCV can infect bovine, human, and rat cells. In animals, coronaviruses also have restricted tissue specificity. For instance, MHV infects only cells of gastrointestinal tracts, reticuloendothelial systems, and occasionally nervous systems. The species and tissue specificity may be controlled at the cellular receptor level. This control is demonstrated by a classic experiment that showed that MHV does not infect COS cell, a monkey

kidney cell line, but the transfection of purified genomic RNA of MHV into COS cells led to the production of infectious viruses (L. Banner & M. M. C. Lai, unpublished data). A candidate membrane receptor protein for the MHV has been identified in the gastrointestinal tract of a susceptible mouse strain (12). The presence or absence of this protein correlates with the sensitivity or resistance of the mouse strains. A monoclonal antibody against this protein can block viral infections (K. Holmes, personal communication). Thus, the receptor molecules on the surface of the target cells appear to provide the entry points for the viral infection. The S (peplomer) protein of the coronavirus is probably the viral component that interacts with the receptors (116). Studies showing that monoclonal antibodies against the S protein neutralized virus infectivities indirectly suggest the S protein's role as the viral component (25, 124). However, direct biochemical interactions between the S protein and the receptors have not been demonstrated. Monoclonal antibodies against the HE protein of BCV can also block infections of BCV (32). Thus, the HE protein may also interact with the viral receptors. Whether S and HE proteins have distinct receptors is a matter of speculation. The presence of two different receptor-interacting proteins may allow the virus to infect different types of cell. After virus binding, the virus penetrates into cells probably by viropexis and uncoats through endosomal fusion (50, 83). However, virus entry by cell fusion has not been ruled out.

Once the viral RNA genome is released into the cells, the subsequent biochemical events in viral replication cycle take place entirely in the cytoplasm. The nuclear functions are apparently not required for the growth of viruses, inasmuch as coronaviruses can grow in enucleated cells and in the presence of actinomycin D and α-amanitin (13, 127). The replication of IBV may require some nuclear functions (35), although this report has not been confirmed. The biosynthetic events of coronavirus replication in the infected cells start with the synthesis of a virus-encoded, RNA-dependent RNA polymerase. This enzyme is not carried in the virion particle; thus, it has to be synthesized de novo from the incoming viral genomic RNA. Indeed, the replication of coronavirus RNA requires continued protein synthesis almost from the very beginning of viral infections (71, 94). Inhibition of protein synthesis at any point of the viral replication cycle will interrupt viral RNA synthesis. This virus-encoded RNA-dependent RNA polymerase transcribes the incoming viral genomic RNA into full-length (58) and subgenomic (94a, 97) negative-stranded RNA, which then serve, in turn, as the template for the synthesis of subgenomic and genomic mRNAs. The transcription of these mRNAs may be catalyzed by a different RNA-dependent RNA polymerase, which is either modified from the earlier viral RNA polymerase or represents a completely new one (14). The mRNAs are then used for the synthesis of structural and nonstructural proteins. The functions of the nonstructural

proteins are not yet clear; some of them probably participate in viral RNA synthesis and others may be involved in virus assembly, and still others may be responsible for the shut-off of host macromolecular synthesis.

Accumulation of viral structural proteins is a prerequisite for the assembly of virus particles and may also signal the switch of viral RNA synthesis from making mRNAs to making genomic RNA that is destined to be packaged into virion. The assembly of virus particles is probably initiated by the interaction between the virion genomic RNA and structural proteins, such as N protein, thus forming a nucleocapsid. This interaction may trigger a cascade of additional viral protein-protein and protein-RNA interactions such as the binding of N to M (117). The virus nucleocapsid buds into the endoplasmic reticulum in the perinuclear membranes (34, 82, 85) and, in the process, acquires lipid bilayer and S and/or HE proteins. The mature virus particles are transported across the golgi complex and eventually released from the cells. A final step of virus maturation probably occurs extracellularly as the S protein is cleaved by cellular proteases into two disulfide-linked subunits. This cleavage is required only for certain coronaviruses, such as MHV (118), but not for FIPV. The infected cells develop cytopathic effects (CPE), which may result from the viral inhibition of host macromolecular synthesis and from the S protein-induced cell fusion. In most of the cell lines studied, coronaviruses cause cytocidal infection, resulting in cell death. The kinetics of virus growth vary with cell lines and virus strains. In some cells, coronavirus infections result in persistent infections with minimal CPE. In general, persistent infection can be readily established in most of the coronavirus infections. The mechanisms of persistent infection are largely unknown.

TRANSCRIPTION OF CORONAVIRUS mRNAS

The Structure and Gene Products of mRNAs

Coronavirus RNA synthesis occurs independently of DNA-dependent RNA synthesis. Thus, the coronavirus-specific RNA synthesis in the infected cells is usually studied in the presence of actinomycin D, which inhibits host RNA synthesis. Depending on the cell lines and viruses involved, virus-specific RNAs can usually be detected a few hours after virus infections. The vast majority of these RNAs are poly(A)-containing mRNAs, which can usually be separated into 6 to 8 species, ranging in size from 0.6×10^6 to more than 8×10^6 molecular weight (55, 67, 106, 107). These RNAs are named mRNAs 1 to 7 in the order of decreasing size. [IBV RNAs were previously named A to F in the order of increasing size (107). The recommendation that these RNAs be renamed 1 to 6 placed them in conformity with the nomenclature of other coronaviruses (23a). An MHV mRNA was named mRNA 2b or 2-1 (98) because it was

species were detected on polysomes (59, 106), thus indicating that they are functional mRNAs. The largest mRNA (mRNA 1) corresponds in size to the genomic RNA of the virion. The rest of the mRNAs are subgenomic in size. The structure of these mRNAs has been determined initially by oligonucleotide fingerprinting studies (55, 67, 107), which showed that these mRNAs have a nested-set RNA structure, i. e. the sequence of each RNA represents a subset of the next larger mRNA, and the largest mRNA is identical to the virion genomic RNA. Furthermore, all sequences of the mRNAs start from the 3' end of the genome and extend for various distances in the 5' direction. Thus, all the mRNAs overlap in sequence, and, except for the smallest mRNA, all are polycistronic in their structure. However, in vitro translation studies showed that only the 5' portion of each mRNA that does not overlap with the next smaller mRNA is translatable (43, 66, 100, 111). Thus, these mRNAs are functionally monocistronic. Whether the untranslated downstream sequence in each mRNA serves any structural or functional role is not clear.

The number of mRNAs varies among different coronaviruses. For instance, IBV has only six major mRNAs compared with seven for MHV and eight for BCV (44). The number of these mRNAs corresponds to the number of functional genes in each coronavirus species. Thus, IBV has two fewer genes than BCV. Interestingly, some MHV strains contain eight mRNAs, with an additional mRNA species of a size between that of mRNAs 2 and 3 (75, 98). This gene (gene HE) is expressed only in some MHV strains (72, 75). The transcriptional control of this gene is discussed in a later section. Several additional mRNA species, which are either few in number or present in only certain MHV strains, have also been detected (72, 75). Whether they are functional mRNAs and encode any gene products is not clear. Nevertheless, these mRNAs are interesting from the standpoint of transcriptional control.

In vitro translation studies of these mRNAs have helped determine the genetic map of coronavirus RNA genome (Figure 1). These genetic maps have been further defined by cDNA cloning and sequencing of the genomic RNA. The sequence analysis shows that, in general, the coding region of each mRNA contains one ORF. However, the 5' unique regions of several mRNAs encompass more than one ORF. An example is mRNA 5 of MHV, whose coding region covers two ORFs capable of encoding two proteins p13 and p10, respectively (101). These two proteins were synthesized from the same mRNA probably under differential translational regulation (20, 65). No separate mRNA has been detected for the downstream ORF, although the downstream protein (p10) is the major gene product detected in infected cells (65). The second example is mRNA 3 of IBV, which encompasses three ORFs in its coding region (10). At least two of these ORFs have been shown to be translationally active in in vitro translation using a single mRNA species (102). The mechanism of translation in these cases, particularly that of the

downstream ORF, is not clear. The most unusual case is mRNA 1 of IBV and MHV, which has a coding region as large as 17–23 kb (2, 10, 86). The 5' unique region of this mRNA has two slightly overlapping ORFs. These two ORFs are translated into a large protein by a mechanism of ribosomal frameshifting (16). This translation mechanism requires a pseudo-knot structure around the overlapping region (17). The primary protein product of this mRNA is most likely processed into multiple proteins post-translationally. Already, a cleavage product from the N-terminus of this protein has been detected in MHV-infected cells (29, 30). This product is most likely cleaved by an autoprotease inherent in this protein (3, 103). The protease domain appears to reside in a protein domain corresponding to a region 3.9 kb to 5.3 kb from the 5' end of the genome (3). These results suggest that the expression of coronavirus gene products is more complex than the simple monocistronic mRNAs.

Oligonucleotide fingerprinting analysis of the MHV subgenomic mRNAs revealed an additional feature of unusual structure; i. e. a T1-oligonucleotide is present only in subgenomic mRNAs but not in the genomic RNA (59, 67). This finding, along with the 5' end sequence analysis, led us to propose that the mRNAs contain a leader sequence (59). This prediction was subsequently confirmed by additional oligonucleotide mapping (57) and cDNA sequencing (54, 105). Both MHV and IBV mRNAs contain a stretch of approximately 60–70 nucleotides of leader sequence at the 5' end (19, 54). This sequence is identical for every mRNA and is also present at the 5' end of genomic RNA. It is not present in the internal regions of genomic RNA, although a small stretch of 8–10 nucleotides are homologous between the 3' end of the leader sequence and the intergenic region preceding each gene (18, 21, 99). Thus, the leader sequences of subgenomic mRNAs are most likely derived from the 5' end of the genome. This leader RNA sequence plays a major role in mRNA transcription (see below).

The molar amounts of each mRNA are different. In general, the smaller mRNAs are more abundant than the larger ones (67). For instance, mRNA 7 of MHV is the most abundant RNA species. Next in abundance is mRNA 6. However, this relative ratio does not always parallel the mRNA size; mRNAs 4 or 5 are usually synthesized in smaller quantities and tend to vary among different coronavirus strains, even between different MHV strains. The genomic-sized RNA is more abundant than some of the mRNAs, particularly later in the infection. Other than the genomic RNA, the relative ratio of the subgenomic mRNAs remains constant throughout the viral replication cycle.

Mechanism of Subgenomic mRNA Transcription

The coronavirus mRNA synthesis can be detected a few hours after virus infection in most virus-cell systems and continues until the infected cells expire. The nested-set, 3'-coterminal structure of mRNAs suggests that

different mRNAs have separate initiation sites but a common termination point, which is probably the end of the template RNA. Alternatively, these mRNAs could be derived from the cleavage and splicing of a precursor full-length RNA product; the different mRNA species represent cleavage at different points of this hypothetical precursor RNA. A third possibility is that they are synthesized using subgenomic negative-stranded template RNA, all of which are expected to have the same 5' end but terminate at different points. UV transcriptional mapping studies (42, 110) have addressed these alternatives. These studies show that the UV target size of each mRNA is equivalent to its physical size, indicating that each mRNA is transcribed independently, but not derived from the processing of a large precursor RNA. Thus, the discovery of the leader sequence on mRNAs, which are derived from the 5' end of the genomic RNA, posed a puzzling theoretical dilemma. This mRNA structure is similar to most of the eukaryotic mRNAs, which are normally derived by an RNA splicing mechanism. However, the UV transcriptional mapping study ruled out the involvement of the conventional RNA splicing in coronavirus mRNA transcription. Furthermore, coronaviruses replicate exclusively in the cytoplasm of infected cells (13, 127)—again inconsistent with the conventional RNA splicing mechanism, which occurs in the nucleus. These dilemmas led us to propose several possible mechanisms for coronavirus subgenomic mRNA transcription (6). All of these models involve a discontinuous transcriptional process, which fuses the leader sequence at the 5' end of the genomic RNA to the subgenomic mRNAs, which initiate at sites distant from the 5' end.

In the first model, the looping out of the template RNA allows RNA polymerases to jump from the leader region to the internal initiation sites of subgenomic mRNAs during the transcription process. The second model is a post-transcriptional processing mechanism, i. e. a *trans*-splicing mechanism that fuses the leader RNA to the mRNAs by a traditional splicing mechanism. This mechanism would be similar to that used for trypanosome mRNA synthesis (119). This model was considered less likely because the leader sequence is present on the nascent RNA chains in the replicative intermediate (RI) RNA complex (6). Furthermore, no consensus splicing donor-acceptor sequences are presented at the leader and mRNA junction sites (see below).

The third model is the leader-primed transcription, in which a leader RNA is transcribed from one end of the RNA template, dissociates from the template, and then rejoins the template RNA at downstream transcription initiation sites to serve as a primer for transcription (Figure 2). This model is unlike any of the discontinuous transcription mechanisms known to exist.

These three models are based on the finding that the template RNA is of genomic-length (58). However, subgenomic negative-strand template RNAs have recently been detected in some coronavirus-infected cells (94a, 97).

Thus, the discontinuous transcription step could potentially also occur during negative-strand RNA synthesis, instead of positive-strand RNA synthesis. Nevertheless, the prevailing body of evidence supports the direct involvement of a free leader RNA in the transcriptional initiation of mRNAs, irrespective of the actual size of the template RNA. This supporting evidence includes the following:

1. Several small leader sequence-related RNA species have been detected in the cytoplasm of MHV-infected cells (7). These RNAs range from 50 to 90 nucleotides in length and are separated from the template RNA and from the membrane-associated transcription complex (see below). They fit the requirement of a free leader RNA for the leader-primed transcription and may be the RNA used for mRNA synthesis. These leader RNAs are distinct in size and reproducible in different cell types, suggesting that they are synthesized and terminated at distinct termination points, which have been suggested to be sites of strong secondary structure (5). Most of these leader RNAs, however, are either larger or smaller than the leader sequence present at the 5' end of the mRNAs. Thus, additional processing of these leader RNAs likely occurred during subsequent mRNA transcription.

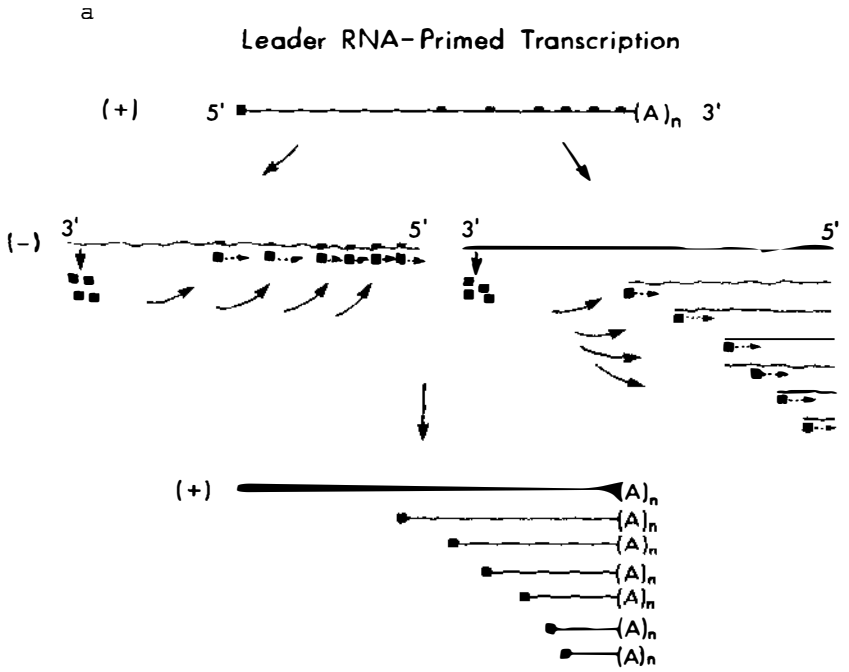
2. A temperature-sensitive (ts) mutant has been isolated that synthesizes only the small leader-related RNAs but not mRNAs at the nonpermissive temperature (7). The isolation of this ts mutant indicates that the synthesis of the leader RNA and mRNAs are two separate and discontinuous events, and that they are mediated by at least two different virus-specific proteins.

3. The evidence in support of the functional role of the free leader RNA in mRNA transcription came from the analysis of the leader sequence on mRNAs during a mixed infection of different MHV strains (80). Some of the leader sequences on the mRNA of a particular MHV strain are derived from the coinfecting virus. This class of mRNA can account for as high as 40–50% of the mRNAs, suggesting that the leader RNA and mRNAs are two separate transcription units that can be joined together almost at random. This phenomenon is reminiscent of the RNA reassortment of segmented RNA viruses, such as influenza virus (87) or reovirus (36). This study established that the free leader RNAs detected in the MHV-infected cells indeed participate in transcription and are not the abortive transcription products.

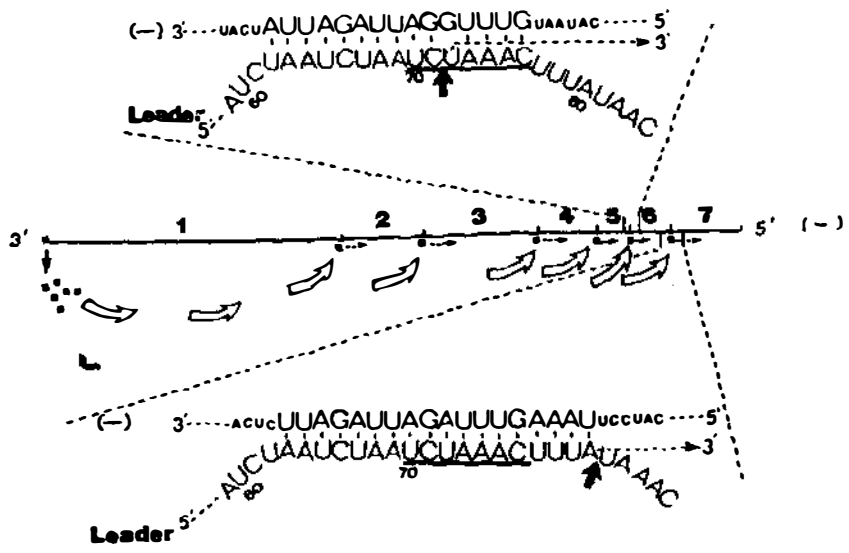
4. A direct biochemical evidence for the involvement of the leader RNA in transcription has recently been obtained from an in vitro transcription system (S. Baker & M. M. C. Lai, unpublished observations). When an exogenous leader RNA was introduced into infected cell lysate, the exogenous leader RNA was incorporated in trans into subgenomic mRNAs at precisely the leader-mRNA junction sites. These data support the leader-primed transcription mechanisms, regardless of the structure of the template RNA.

The understanding of the actual mechanism of the leader-primed transcrip-

a


Leader RNA-Primed Transcription
(+)
5'
(A)n 3'
(-)
3'
5'
3'
5'
(+)
(A)n
(A)n
(A)n
(A)n
(A)n
(A)n
(A)n

b


(—) 3'
UACUAUUAGAUUAGGUUUGUAAUAC
5'
3'
AUCUAAUCUAAUCUAAACUUUAUAAC
70
60
Leader
5'
1
2
3
4
5
6
7
3'
5'
(-)
(—) 3'
ACUCUUAGAUUAGAUUUGAAAUUCCUAC
5'
3'
AUCUAAUCUAAUCUAAACUUUAUAAC
70
60
Leader
5'

tion came from the sequence analysis of the leader RNA and the intergenic regions (i. e. transcription initiation sites). Such sequence analysis has been available for MHV, IBV, and, to a limited extent, for HCV (11, 96, 99). The intergenic regions have also been sequenced for several genes in BCV, TGEV, and FIPV, although their leader sequences have not been determined. The leader of MHV is roughly 72–77 nucleotides long while that of IBV is roughly 60 nucleotides (19, 54). A stretch of 7 to 18 nucleotides is homologous between the 3' end of the leader RNA and the intergenic sequences (19, 99). The 5' ends of mRNAs and the intergenic sequences begin to diverge, in the 5' direction, at the point of this homology. Thus, this homologous region is assumed to be the point of leader RNA binding and priming. Because of this homology, the precise length of the leader RNA cannot be defined precisely; the actual leader can start anywhere within this stretch of homologous sequence. Within this homologous region, seven nucleotides (UCUAAAC) are conserved among most of coronaviruses, except IBV, which has UUAACAA instead. These conserved nucleotides vary slightly among different intergenic sequences of the same coronavirus genome. In other words, the homology between the 3' end of the leader RNA and the intergenic regions is not perfect for every gene. But this mismatch is only minor, and does not involve more than one nucleotide at any intergenic region. This conserved region is probably the consensus sequence recognized by RNA polymerases for mRNA transcription. The nucleotides homologous between the leader and the intergenic sites both upstream and downstream of this consensus sequence probably also contribute to the stability of the leader RNA binding. Note that the consensus sequence and sequences around it are extremely AU-rich. Furthermore, the leader sequence, particularly its 3' half, is relatively conserved among different coronaviruses. Particularly, the secondary structure of the leader RNA is conserved (99). Even in IBV leader RNA, which has a highly diverged leader sequence, this secondary structure is relatively conserved. The functional significance of this secondary structure is not yet clear.

Heterogeneity of Subgenomic mRNAs and Possible RNA Proofreading

The 3' end of the leader sequence in coronavirus subgenomic mRNAs has been defined as the point of homology between the leader RNA and the intergenic regions, but the size of the leader RNA species utilized for RNA

Figure 2 Models of leader-primed transcription. (*a*) Two general models in which the template RNA is either genomic or subgenomic in size, respectively. The *solid squares* represent leader RNA. (*b*) The intergenic sequences preceding genes 6 and 7 of MHV are shown. The arrows indicate the sites of possible nuclease digestion (reproduced from Ref. 52 with permission).

transcription has not been determined. There are several leader RNA species of different sizes in the MHV-infected cells (7), none of which are exactly the same size as the leader sequence present in mRNAs. Presumably only the larger leader RNAs can be used as a primer for mRNA transcription. These larger leader RNAs may be synthesized using a transcriptional termination signal (AU-rich sequence following a stem-and-loop structure) localized downstream of the consensus sequences at the 5' end of genome (99). If these leader RNAs are used for mRNA transcription, a mechanism must cleave the 3' end overhang sequence of the leader RNA before RNA synthesis (priming) can occur. Furthermore, as pointed out earlier, some nucleotides are mismatched between the leader RNA and some of the intergenic regions within the consensus sequence. In these cases, the mature mRNAs usually contain sequences homologous to the intergenic sequences rather than to the leader RNA at the 5' end of the genome. Thus, the coronavirus RNA polymerases must have an ability to recognize these mismatches and correct them or, alternatively, the leader RNA is cleaved at the point of mismatches and RNA priming occurs immediately at the point of cleavage. Figure 2 presents a model of leader-primed transcription. This model proposes that the cleavage of the leader (primer) RNA occurs at the 5' most mismatched nucleotides, and RNA synthesis starts at the 3' OH end of the cleaved primer. Some mismatches localized at the 5' side of the consensus sequence are, in rare situations, not corrected (79). The rule of mismatch repair is not completely understood yet.

Another interesting feature of coronavirus leader RNA is that two to four pentanucleotide (UCUAA in the case of MHV) repeats partially overlap with the consensus sequence (UCUAAAC) at the 3' end of the leader. The number of repeats is variable among different MHV strains and changes rapidly with virus passages in tissue culture (75). For instance, serial passages of JHM strain in tissue-culture cells decreased repeat number from three to two after 10–15 passages (75). IBV also contains an imperfect repeat (UUUAA and CUUAA) within the leader RNA (11). This pentanucleotide sequence is also imperfectly repeated in some intergenic regions. The presence of these repeat nucleotides in both the leader RNA and the intergenic regions within the consensus sequences allowed us to make an interesting prediction that the binding of the leader RNA to the intergenic regions may be imprecise, allowing some slippage between the different pentanucleotide repeats. This results in mRNAs with different numbers of repeats (Figure 3). Primer extension analysis of the purified mRNAs of MHV has indeed shown that every mRNA is heterogeneous, consisting of several subspecies, each of which differs by the repeat number of UCUAA sequence (79). Because the leader sequence at the 5' end of the genomic RNA is homogeneous, this result suggests that the pentanucleotide repeat within the consensus sequence is

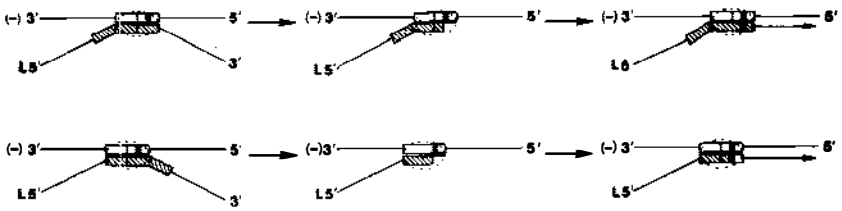


Figure 3 Possible mechanisms for the generation of heterogeneity of mRNAs. Each open box represents a UCUAA pentanucleotide. The darkened boxes represent the nucleotide mismatches between the leader and template RNAs.

directly involved in leader RNA binding. In all of these RNA subspecies, most of the mismatches between the leader and the intergenic sequences are corrected. However, a few of them are not. Again, the rule of the mismatch repair is not known.

Transcriptional Regulation of Optional mRNA Species

In addition to the major mRNA species observed in all of the coronaviruses, several additional mRNA species have been detected in some strains of MHVs. The most dramatic example is the transcription of a mRNA 2-1 (old term, mRNA 2b) with a size between that of mRNAs 2 and 3 (98). This mRNA corresponds to a functional gene HE but is transcribed only in a few MHV strains. For instance, it is transcribed only in JHM strains containing two UCUAA repeats in the leader sequence at the 5' end of genome, but not by the JHM strain with three UCUAA repeats (75). The transcriptional initiation signal of this gene is UUUAAUAAAC, which differs slightly from the consensus sequence (UCUAAAC). This differential transcription by the leader RNA with either two or three UCUAA repeats is not seen with other intergenic regions. Thus, obviously, this phenomenon is related to the peculiar sequence at the transcription initiation site of this mRNA. Sequence analysis shows that the leader RNA forms the most stable hybrid with the sequences upstream and downstream of this consensus intergenic sequence, rather than with the consensus sequence itself (Figure 4). The loop formed by this kind of leader binding might hinder transcription. Perhaps the loop created by three repeats is too big to allow the transcription to proceed. This differential transcription has also been noted with several recombinant MHVs (see below). The recombinant viruses with A59-derived leader sequence, which contains two UCUAA repeats, and JHM-derived gene 2b (HE) sequence transcribe mRNA 2-1, but A59 virus itself does not (75). The intergenic sequence of gene 2b of A59 contains a nucleotide change (UCUAAAC → UCUGAAC) (Figure 4). This nucleotide change may account for the failure of A59 to transcribe this mRNA. These results indicate the importance of both leader sequence and intergenic sequence in mRNA transcription.

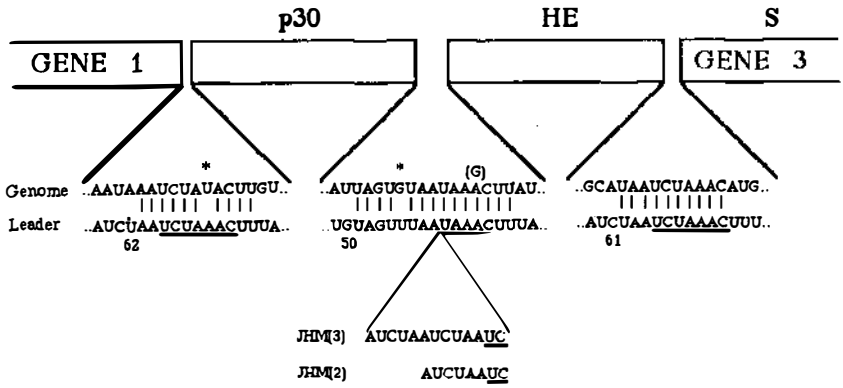


Figure 4 The sequences at the transcription initiation site for gene HE and the leader of MHV. JHM(2) is the JHM strain containing two UCUAA repeats while JHM(3) contains three UCUAA repeats. (G) is the mutated nucleotide in A59 RNA. Modified from Ref. 98 with permission.

Several other mRNAs (e. g. mRNA 3a) are also differentially expressed (72, 75), depending on the sequence of both the leader RNA and intergenic sequences, although their transcriptional initiation sites have not been precisely determined. In addition, several minor RNA species are frequently observed in coronavirus-infected cells. These RNAs may be transcribed from the sequences that have only partial homology to the consensus sequence, thus representing suboptimal transcriptional initiation signals. The transcriptional regulation and functional significance of these RNAs are currently unknown.

These studies suggest that the 3' end of the leader RNA, particularly the region around the UCUAA repeats, is important for leader-primed transcription. This conclusion is further supported by a study using the in vitro transcription system, which showed that only the leader RNA containing the UCUAA repeats could participate in mRNA transcription in trans (S. Baker & M. M. C. Lai, unpublished observations). The smaller leader RNA lacking this repeat could not be used for transcription in vitro. Thus, if the template RNA used is subgenomic, this UCUAA repeat sequence must be present in these subgenomic templates.

NEGATIVE-STRANDED RNA SYNTHESIS

Coronavirus RNA synthesis is clearly mediated through a negative-stranded RNA template, which constitutes only a very small percentage (1–2%) of the intracellular virus-specific RNAs (89, 94). The synthesis of this negative-stranded RNA from the incoming virion RNA by the virus-encoded RNA

polymerase precedes the synthesis of mRNAs. Indeed, the negative-stranded RNA template of coronavirus could be detected as early as two hours after infection and peaked at five to six hours after infection (89, 94). Apparently, the majority of the negative-stranded RNA is synthesized early in the infection; however, a small amount of it may continue to be synthesized until late in the infection during the peak of mRNA synthesis (94). The negative-stranded RNA is probably utilized for multiple rounds of mRNA synthesis because the positive-stranded RNA outnumbers the negative-stranded RNA (89, 94). Therefore, the negative-stranded RNA is stable. This stability is related to the fact that all of the negative-stranded RNA is present as double-strand RNA; no free negative-stranded RNA can be detected in coronavirus-infected cells (94). Whether the negative-stranded RNA synthesized late in the infection contributes much to the mRNA synthesis is not clear, since the UV transcription study, in which the infected cell RNA was irradiated at six h after infection, suggested that the template RNA used for mRNA synthesis is made before, but not during, mRNA synthesis (42).

The negative-strand template RNA was initially thought to consist of only the full-length genomic-sized species (58). More recently, however, subgenomic negative-stranded RNAs and replicative form (RF) RNAs corresponding to the sizes of the subgenomic mRNAs were detected (94a, 97). These subgenomic negative-stranded RNAs have the same relative ratio among them as that among the subgenomic mRNAs. They could be generated either by termination at the intergenic sequences during negative-stranded RNA synthesis, or, alternatively, by replication of the subgenomic mRNAs. If the former is the case, then the intergenic sequences (UCUAAAC) will serve not only as the leader RNA binding sites during positive-stranded RNA synthesis but also the termination signals of the negative-stranded RNA synthesis. At the present time, the precise structure of the template RNAs used for subgenomic mRNA synthesis has not been resolved.

The mechanism of initiation of the negative-stranded RNA synthesis is not clear. The negative-stranded RNA contains a poly(U) tract (J. Leibowitz, personal communication); thus, it is most likely initiated from the poly(A) at the 3' end of the genomic RNA. Interestingly, at approximately 80 nucleotides from the 3' end of the genomic RNA is a stretch of sequences (CGAAGAGC) that is conserved among all of the coronaviruses sequenced so far (62, 96). This conserved sequence may be the recognition signal for the negative-stranded RNA synthesis. Boursnell et al (11) pointed out that the 5' end sequence and the complementary sequence of the 3' end of IBV genomic RNA are homologous, suggesting that similar sequences regulate the replication of positive and negative sense RNA. However, this sequence homology was not seen in MHV RNA.

REPLICATION OF GENOMIC RNA

The genomic-sized RNA in coronavirus-infected cells consists of two populations: one serves as mRNA 1, which is used for the translation of the 5' most ORF (ORFs 1a and 1b), and the other is the RNA that is eventually packaged into the virion. Whether these two RNA populations have any structural differences is not clear. Presumably early in infection, most of the genomic-sized RNA serves as the mRNA for the synthesis of RNA polymerases required for RNA transcription; later in infections, most of these RNAs are packaged into virions. Although one cannot distinguish between these two RNA populations, in BCV-infected cells, the relative ratio of the genomic RNA to subgenomic RNAs increases later in infection (44), suggesting that mRNA synthesis (transcription) shifts to genomic RNA synthesis (replication). Furthermore, late in the infection, most (95%) of the genomic-sized RNA is associated with nucleocapsids, while the remaining portion (5%) is associated with polysomes (91, 106). This result also suggests a functional distinction between virion genomic RNA and the genomic-sized mRNAs. Whether the replication and transcription of coronavirus RNAs are carried out by different mechanisms of RNA synthesis is unknown. These processes probably occur on different RNA synthetic machineries because the studies of RNA polymerases from infected cells revealed the presence of two separate membrane complexes containing polymerase activities (15, 31). One of these complexes synthesizes all seven RNAs, while the other synthesizes mainly genomic-sized RNA (15). Thus, the former is likely a transcription complex while the latter is a replication complex, both of which are separable from each other.

One of the key questions related to coronavirus RNA replication asks whether the replication of RNA also involves the free leader RNA species, as in subgenomic RNA transcription. In other words, is the synthesis of genomic RNA, which is a faithful continuous copy of the virion genomic RNA, discontinuous? No direct answer to this question is currently available. However, the studies of the replication of defective-interfering (DI) RNAs suggest a possible answer. When an MHV DI RNA was transfected into cells, its leader RNA readily changed into that of the helper virus, suggesting that the free leader RNA from the helper virus is utilized for DI RNA replication (76). Because the replication of DI RNA is generally thought to be similar to that of the wild-type virus RNA, albeit much faster, the replication of coronavirus genomic RNA likely also utilizes a free leader RNA species and proceeds in a discontinuous manner. This possibility is also suggested from the heterogeneity of both the wild-type genomic RNA and DI RNA in coronavirus-infected cells (56, 75). As discussed above, the 5' end leader sequences of the genomic RNA as well as the DI RNA in different MHVs

contain several UCUAA repeats. The uncloned populations of MHVs contain a mixture of genomic RNAs with two, three, or four UCUAA repeats (75). Furthermore, starting with a pure population of MHV, the repeat number in the genomic RNA rapidly becomes mixed during serial passages in tissue culture. Likewise, DI RNAs also contain a mixed subpopulation of RNAs with different repeat numbers (56). These findings are most consistent with the interpretation that a free leader RNA is utilized for the replication of genomic RNA and DI RNA, in the same way as subgenomic mRNA transcription. The imprecise binding of the free leader RNA during RNA replication could result in the heterogeneity of the genomic RNA. The genomic RNA with a lower number of UCUAA repeats seems to have an advantage in replication, as these RNAs become the predominant RNA species during serial virus passages (75).

The replication of coronavirus genomic RNA has an additional unique feature: the replication of the rest of the genome, other than the leader sequence, may proceed in a discontinuous and nonprocessive manner (7). The crux of this mechanism is that the RNA synthesis pauses at sites of extensive secondary structure of the RNA template; the incomplete RNA products dissociate from the template because of a nonprocessive nature of the coronavirus RNA polymerases. These dissociated RNA products then rebind to the RNA template to continue RNA synthesis. This discontinuous, nonprocessive mechanism of RNA synthesis was suggested from the demonstration of high-frequency RNA recombination between coronaviruses (74). The recombination frequency could be as high as 10–20%. Such a high frequency RNA recombination is reminiscent of RNA reassortment among viruses containing segmented RNA genomes, such as influenza virus or reovirus. Therefore, segmented RNA intermediates are likely present during RNA replication. These RNA intermediates can bind to the RNA template of a different virus during mixed infections. Indeed, leader-containing RNA intermediates of different sizes have been detected in MHV-infected cells (7). The termination points of these RNA species correspond to the sites of extensive secondary structures in the MHV genomic RNA (7). These RNA intermediates were separated from the template RNA (7) and could thus potentially participate in RNA replication by rebinding to the template RNA. Because the recombinants with multiple crossovers could be easily isolated (45–47, 72), the dissociation and reassociation of RNA intermediates probably occurs frequently during RNA replication.

The mechanism of switching from the synthesis of mRNAs to the replication of genomic RNA is not clear. The genomic RNA synthesis and encapsidation are coupled in MHV-infected cells (89). Since the encapsidation requires N protein, the accumulation of the N protein could possibly tip the balance of RNA synthesis from transcription toward RNA replication.

Conceivably the synthesis of mRNAs during infection leads to the synthesis of viral structural and nonstructural proteins. Once enough N proteins are made, the replication of the genomic RNA could be facilitated by the binding of N proteins to the nascent RNA products. This mechanism of switching from transcription to replication has also been demonstrated for vesicular stomatitis virus (9), which is a negative-stranded RNA virus and contains a helical nucleocapsid. Because coronavirus also has a helical nucleocapsid, a similar mechanism may regulate coronavirus RNA replication. The binding of the N protein to the RNA has been demonstrated in vitro (4, 112). The binding sites appear to reside in the leader RNA (112). Curiously the N proteins also bind to all of the subgenomic mRNA species of coronaviruses (4). The significance of such a binding is not clear. The N protein binding to the virion RNA demonstrated in vitro does not directly lead to the packaging of RNA into virus particles. Despite the binding of N protein to all of the mRNAs, only the genomic sized RNA is packaged into mature virion. Furthermore, many DI RNA containing the leader RNA are not packaged efficiently (73). Recent studies using MHV DI RNAs suggest that the RNA packaging signal may reside in the 3' end of gene 1 (S. Makino, K. Yokomori, & M. M. C. Lai, unpublished observation). Thus, the encapsidation by N proteins is necessary but not sufficient for RNA packaging into mature virions. This binding may represent a regulatory mechanism for RNA synthesis rather than a virion assembly process. Additional events are required for virus packaging of RNA.

ENZYMOLOGY OF RNA SYNTHESIS

Coronaviruses do not carry RNA polymerases in the virion. Therefore, the RNA polymerases used for RNA transcription have to be synthesized de novo from the incoming virion genomic RNA initially and then from the newly synthesized mRNAs later in the infection. Indeed, the inhibitors of protein synthesis such as cycloheximide, when added to the infected cells immediately after infection, inhibit RNA synthesis (71, 94). The dependency of RNA synthesis on protein synthesis continues throughout the viral replication cycle. Thus, either RNA polymerases or some cofactors have to be continuously synthesized.

Two types of RNA-dependent RNA polymerase activities specific for coronaviruses have been detected in the lysates of several coronavirus-infected cells (14, 26, 31, 63, 71). One is detected two to three hours after infection by MHV and disappeared by four or five hours postinfection (14). The other peaked at seven to eight hours postinfection. These two activities have different enzymatic properties, including pH optimum and cationic requirements (14). The detection of these enzymatic activities depends on

active protein synthesis (71). The first or early polymerase is most likely responsible for negative-stranded RNA synthesis, while the second or late polymerase is responsible for the positive-stranded RNA synthesis (15). The latter activity was associated with two separable membrane complexes, one of which is probably responsible for the synthesis of genomic RNA and the other for both genomic and subgenomic mRNAs (15). Whether the protein components of these two complexes are different has not been shown. Since the enzymatic properties of the early and late RNA polymerases are different (14), and the negative-stranded and positive-stranded RNA synthesis have different sensitivities to inhibition by cycloheximide (94), the negative- and positive-stranded RNAs are most likely synthesized by different RNA polymerases or, at least, by the same polymerases modified differently by other viral or cellular proteins.

The proteins responsible for the polymerase activities have not been detected. Theoretical considerations suggest that the 5'-most gene (gene 1) of the coronavirus genome encodes RNA polymerases. This is based on the reasoning that only the 5'-most gene of the virion genomic RNA can be translated into proteins, and that the synthesis of RNA polymerases is required before any other macromolecular synthesis can take place. The size of this gene ranges from 18 kb for IBV (11) to 23 kb for MHV (2, 86) and accounts for more than two thirds of the entire viral genome. The IBV gene 1 contains 2 overlapping ORFs, which can potentially encode 441 kd and 300 kd proteins, respectively (11). There is a pseudoknot-forming sequence around the junction between the two ORFs (17). These two ORFs can be translated into a continuous protein by a ribosomal frameshifting mechanism by virtue of this pseudoknot structure (16, 17). None of the protein products have been detected in the coronavirus-infected cells, however. The large size of this potential gene product suggests that it contains several different functional domains or that it is processed into several different proteins. In vitro translation of the genomic RNA has yielded a protein of 250 kd (29), which represents only a portion of this potential gene product. Thus, it may be a cleavage product of the primary translation product. Furthermore, the N-terminus of this protein was cleaved into a 28 kd protein during in vitro translation and in the infected cells (29, 30). A coupled transcription-translation of MHV cDNA has shown that this cleavage may be carried out by an autoprotease activity inherent in the protein itself (3). The protease activity is localized somewhere downstream of the cleavage site (3). This is the first detected enzymatic activity and processing event associated with the coronavirus RNA polymerase. Sequence analysis of this gene has suggested that this protein may contain protease and helicase activities in addition to the polymerase activities (38). These activities have yet to be demonstrated biochemically.

Temperature-sensitive (ts) mutants defective in RNA synthesis at the nonpermissive temperature have contributed additional insights into the properties of this protein. The *ts* mutants of MHV have been grouped into seven complementation groups, six of which have RNA (−) phenotypes (64). The gene functions affected in these *ts* mutants have not been identified. More recent studies by Baric et al have identified 4 RNA$^-$ complementation groups, which are defective in the synthesis of genomic RNA, leader RNA, mRNAs, and negative-stranded RNA, respectively. The ts lesions of these mutants appear to map within the gene 1. Therefore, this gene product appears to have several different functional domains separately responsible for the synthesis of different RNA species. Whether these functional domains represent different RNA polymerases or represent different factors modifying a core polymerase is not clear. It would be interesting to know whether these functional domains are indeed separated into individual proteins by processing of a polyprotein.

The function and enzymology of polymerases can best be studied in an in vitro transcription and replication system. Several in vitro transcription systems have been described for MHV and TGEV (14, 26, 31, 63, 71). These systems can synthesize genomic RNA, and, in some cases, subgenomic RNAs using endogenous viral RNAs as templates. They do not appear to be able to initiate RNA synthesis. The proteins required for the RNA synthesis in these systems have not been identified. However, the antibody against the nucleocapsid (N) protein can inhibit RNA synthesis (26). Thus, in addition to the RNA polymerases, the N protein may also play a role in RNA synthesis. Recently, an RNA transcription and replication system has been developed using lysolecithin permeabilization of the MHV-infected cells (S. C. Baker & M. M. C. Lai, unpublished observation). This system can utilize exogeneously supplied leader RNA as a primer for RNA synthesis. The leader RNA can apparently be trimmed at various points and incorporated into newly synthesized RNA. This may represent the specific nuclease activity postulated in the leader-primed transcription model.

HIGH-FREQUENCY RNA RECOMBINATION

One unique phenomenon associated with coronaviruses, which may result from the peculiar mechanism of coronavirus RNA synthesis, is the high-frequency RNA recombination (74). Although coronaviruses contain a single piece of RNA genome, they can undergo genetic crossovers at a frequency reminiscent of the RNA reassortment in viruses containing segmented genomes. The high frequency of recombination has been demonstrated not only by the large number of recombinants among the progeny viruses in a mixed infection (74), but also by the large number of crossovers in the

genome of recombinants during a single round of infection (45–47, 72). Most remarkably, when two selection markers (e. g. temperature sensitivity in virus replication and resistance to neutralization by monoclonal antibodies) were used, the recombinants obtained contained not only a crossover between the two selection markers, but also crossovers at regions outside the selection markers (45, 46, 72). This finding indicates that the recombination frequency is so high that no selection pressure was needed for the isolation of recombinants. A genetic map has recently been established for various ts mutants of MHV based on the recombination frequencies (R. S. Baric, personal communication). A calculation of recombination frequencies suggests that the entire genome of MHV will have a recombination frequency of roughly 25%. This number is extremely high, suggesting that freely segregated RNA segments must be present during MHV RNA synthesis. Baric et al (7) presented the biochemical evidence for the existence of such RNA segments.

RNA recombination among coronaviruses have been demonstrated not only in tissue culture, but also in infected animals (45) and in natural infections. Sequence analysis of several natural isolates of IBV suggested that different parts of the genome have different degrees of homology among IBV strains (51). This finding suggests that they may have arisen from recombination among different IBV strains. Furthermore, the gene order of IBV is different from that of other coronaviruses (Figure 1). For instance, a gene is inserted between M and N genes. Some of these insertions and rearrangements may have been derived from RNA recombination. Most remarkably, some, but not all, coronaviruses contain an HE gene related to the hemagglutinin protein of influenza C virus (69). Functionally, it also preserves the esterase activity of the HN protein of influenza C virus (122, 123, 129). This gene in coronaviruses may have been derived by recombination between coronaviruses and influenza C virus.

The possible sites of recombination between different MHV strains appear to be distributed over the entire genome. However, preferred recombination sites exist which probably correspond to sites of secondary structure of RNA. These sites may promote transcription pausing, and generate incomplete RNA intermediates, which lead to RNA recombination. Recently, a recombination hot spot was identified near a hypervariable region, which undergoes very frequent deletions in various MHV strains (3a). Thus, the same RNA secondary structure may promote both recombination and deletions.

This very same mechanism may also be responsible for the generation of DI RNAs, which have been observed in MHV-infected cells (73, 77, 81). These DI RNAs are composed of several discontiguous parts of the wild-type viral genomic RNA (78). They might have been derived by a nonhomologous recombination caused by the binding of pausing RNA intermediates to a site downstream of the original pausing sites, thus deleting various internal por-

tions of the genome. This nonhomologous recombination likely occurs between different RNA molecules of the same virus because DI RNAs are known to be generated only during high multiplicity of infection. A high degree of potential secondary structure formation has been noted at the RNA fusion sites of a DI RNA (78).

The occurrence of high-frequency RNA recombination is unusual among other RNA viruses. Only picornavirus (48) and brome mosaic virus (22) reportedly undergo recombination, at a low frequency. The high frequency of coronavirus recombination is probably caused by the nonprocessive nature of the coronavirus RNA polymerase. This high-frequency recombination may play a significant role in coronavirus biology. For instance, RNA recombination, along with the postulated proofreading activity of the coronavirus RNA polymerase, may allow the viral RNA to overcome the deleterious effects of errors in RNA synthesis, which are expected to occur in almost every RNA molecule because of its large size. Thus, the functional integrity of the large RNA genome of coronaviruses may depend on the ability of the coronavirus to recombine.

PERSPECTIVES

Coronavirus RNA synthesis and gene expression clearly involve very unique mechanisms. The leader-primed transcription and high-frequency RNA recombination are features not observed in any other systems. Several unanswered questions currently remain. Chief among them are:

1. What are the properties of RNA-dependent RNA polymerases? So far, the proteins responsible for the polymerase activites have not been identified. The sequence information of the coronavirus RNA genome suggests that the RNA polymerase may be derived from the processing of a large precursor protein. Biochemical characterization of the structure and function of these proteins is essential for further understanding of the coronavirus RNA synthesis. These polymerases are expected to have a nonprocessive enzymatic activity. Furthermore, they may carry a nuclease activity that can specifically recognize mismatched sequences during coronavirus RNA synthesis. In other words, they may have a proofreading activity to correct sequence errors during RNA synthesis. This activity effectively reduces the error frequency of RNA synthesis, thus ensuring that the coronavirus RNA genomes, which have an extremely large size, can be faithfully synthesized. If proved, coronavirus polymerase will be the first RNA polymerase with a possible proofreading activity.
2. What is the sequence requirement of the leader-primed transcription? The previous studies of mRNA transcription in coronavirus-infected cells have

revealed insights into the sequence requirements of mRNA synthesis. Further understanding will likely come from site-specific mutagenesis of the leader RNA and the template RNA. These studies will require an in vitro transcription and replication system that permits initiation of RNA synthesis and utilizes exogenous leader RNA and template RNA. Additionally, the structure of the negative-strand RNA templates needs to be determined.

3. What are the functions of the nonstructural proteins other than RNA polymerases? Practically no information points to the functions of these proteins at the present time. They may modify RNA polymerases or participate in virus assembly. The isolation and characterization of temperature-sensitive mutants affecting these proteins should help identify their functions. The mechanism of translation of these proteins is also an unresolved issue. Several of these proteins are made from overlapping ORFs present in the same gene. The mechanism of translation of downstream ORFs is not clear.
4. What is the mechanism of RNA recombination? What is the RNA structure responsible for recombination? And what is the functional significance of RNA recombination in coronavirus biology? Conceivably, RNA recombination provides a mechanism to remove errors during coronavirus RNA synthesis, thus ensuring the integrity of the coronavirus RNA genome. But it could also lead to the divergence of virus strains and have impact on viral pathogenesis.

The answers to these and other questions will undoubtedly help unravel the unique features of coronavirus RNA synthesis and gene expression.

Literature Cited

1. Armstrong, J., Niemann, H., Smeekens, S., Rottier, P., Warren, G. 1984. Sequence and topology of a model intracellular membrane protein, E1 glycoprotein, from a coronavirus. *Nature* 308:751–52
2. Baker, S. C., La Monica, N., Shieh, C. K., Lai, M. M. C. 1989. Murine coronavirus gene 1 polyprotein contains an autoproteolytic activity. In *Pathogenesis and Molecular Biology of Coronavirus,* ed. D. Cavanagh, T. D. K. Brown, New York: Plenum
3. Baker, S. C., Shieh, C. K., Soe, L. H., Chang, M.-F., Vannier, D. M., Lai, M. M. C. 1989. Identification of a domain required for the autoproteolytic cleavage of murine coronavirus gene A polyprotein. *J. Virol.* 63:3693–99

3a. Banner, L., Keck, J. G. Lai, M. M. C. 1990. A clustering of RNA recombination sites adjacent to a hypervariable region of the peplomer gene of murine coronavirus. *Virology.* 175:548–55

4. Baric, R. S., Nelson, G. W., Fleming, J. O., Deans, R. J., Keck, J. G., et al. 1988. Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. *J. Virol.* 62:4280–87
5. Baric, R. S., Shieh, C. K., Stohlman, S. A., Lai, M. M. C. 1987. Analysis of intracellular small RNAs of mouse hepatitis virus: evidence for discontinuous transcription. *Virology.* 156:342–54
6. Baric, R. S., Stohlman, S. A., Lai, M. M. C. 1983. Characterization of replicative intermediate RNA of mouse hepati-

tis virus: presence of leader RNA sequences on nascent chains. *J. Virol.* 48: 633–40
7. Baric, R. S., Stohlman, S. A., Razavi, M. K., Lai, M. M. C. 1985. Characterization of leader-related small RNAs in coronavirus-infected cells: further evidence for leader-primed mechanism of transcription. *Virus Res.* 3:19–33
8. Binns, M. M., Boursnell, M. E. G., Cavanagh, D., Pappin, D. J. C., and Brown, T. D. K. 1985. Cloning and sequencing of the gene encoding the spike protein of the coronavirus IBV. *J. Gen. Virol.* 66:719–26
9. Blumberg, B. M., Leppert, M., Kolakofsky, D. 1981. Interaction of VSV leader RNA and nucleocapsid protein may control VSV genome replication. *Cell* 23:837–45
10. Boursnell, M. E. G., Binns, M. M., Brown, T. D. K. 1985. Sequencing of coronavirus IBV genomic RNA: three open reading frames in the 5' "unique" region of mRNA D. *J. Gen. Virol.* 66:2253–58
11. Boursnell, M. E. G., Brown, T. D. K., Foulds, I. J., Green, P. F., Tomley, F. M., Binns, M. M. 1987. Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *J. Gen. Virol.* 68:57–77
12. Boyle, J. F., Weismiller, D. G., Holmes, K. V. 1987. Genetic resistance to mouse hepatitis virus correlates with absence of virus-binding activity on target tissues. *J. Virol.* 61:185–89
13. Brayton, P. R., Ganges, R. G., Stohlman, S. A. 1981. Host cell nuclear function and murine hepatitis virus replication. *J. Gen. Virol.* 56:457–60
14. Brayton, P. R., Lai, M. M. C., Patton, C. D., Stohlman, S. A. 1982. Characterization of two RNA polymerase activities induced by mouse hepatitis virus. *J. Virol.* 42:847–53
15. Brayton, P. R., Stohlman, S. A., Lai, M. M. C. 1984. Further characterization of mouse hepatitis virus RNA-dependent RNA polymerase. *Virology.* 133:197–201
16. Brierley, I., Boursnell, M. E. G., Binns, M. M., Bilimoria, B., Blok, V. C., et al. 1987. An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J.* 6:3779–85
17. Brierley, I., Digard, P., Inglis, S. C. 1989. Characterization of an efficient coronavirus ribosomal frame-shifting signal: requirement for an RNA pseudoknot. *Cell* 57:537–47
18. Brown, T. D. K., Boursnell, M. E. G. 1984. Avian infectious bronchitis virus genomic RNA contains sequence homologies at the intergenic boundaries. *Virus Res.* 1:15–24
19. Brown, T. D. K., Boursnell, M. E. G., Binns, M. M. 1984. A leader sequence is present on mRNA A of avian infectious bronchitis virus. *J. Gen. Virol.* 65:1437–42
20. Budzilowicz, C. J., Weiss, S. R. 1987. *In vitro* synthesis of two polypeptides from a nonstructural gene of coronavirus mouse hepatitis virus strain A59. *Virology* 157:509–15
21. Budzilowicz, C. J., Wilczynski, S. P., Weiss, S. R. 1985. Three intergenic regions of coronavirus mouse hepatitis virus strain A59 genome RNA contain a common nucleotide sequence that is homologous to the 3' end of the viral mRNA leader sequence. *J. Virol.* 53:834–40
22. Bujarski, J. J., Kaesberg, P. 1986. Genetic recombination between RNA components of a multipartite plant virus. *Nature* 321:528–31
23. Cavanagh, D. 1983. Coronavirus IBV: further evidence that the surface projections are associated with two glycopolypeptides. *J. Gen. Virol.* 64:1787–91
23a. Cavanagh, D., Brian, D. A., Enjuanes, L., Holmes, K. V., Lai, M. M. C., et al. 1990. Recommendations of the coronavirus study group for the nomenclature of the structural proteins, mRNAs, and genes of coronaviruses. *Virology.* In press
24. Cavanagh, D., Davis, P. J., Pappin, D. J. C. 1986a. Coronavirus IBV glycopolypeptides: locational studies using proteases and saponin, a membrane permeabilizer. *Virus Res.* 4:145–56
25. Collins, A. R., Knobler, R. L., Powell, H., Buchmeier, M. J. 1982. Monoclonal antibodies to murine hepatitis virus-4 (strain JHM) define the viral glycoprotein responsible for attachment and cell-cell fusion. *Virology* 119:358–71
26. Compton, S. R., Rogers, D. B., Holmes, K. V., Fertsch, D., Remenick, J., McGowan, J. J. 1987. *In vitro* replication of mouse hepatitis virus strain A59. *J. Virol.* 61:1814–20
27. Dea, S., Tijssen, P. 1988. Identification of the structural proteins of turkey enteric coronavirus. *Arch. Virol.* 99:173–86
28. De Groot, R. J., Luytjes, W., Horzinek, M. C., van der Zeijst, B. A. M., Spaan,

W. J. M., Lenstra, J. A. 1987. Evidence for a coiled-coil structure in the spike proteins of coronaviruses. *J. Mol. Biol.* 196:963–66
29. Denison, M. R., Perlman, S. 1986. Translation and processing of mouse hepatitis virus virion RNA in a cell-free system. *J. Virol.* 60:12–18
30. Denison, M., Perlman, S. 1987. Identification of putative polymerase gene product in cells infected with murine coronavirus A59. *Virology* 157:565–68
31. Dennis, D. E., Brian, D. A. 1982. RNA-dependent RNA polymerase activity in coronavirus-infected cells. *J. Virol.* 42:153–64
32. Deregt, D., Babiuk, L. A. 1987. Monoclonal antibodies to bovine coronavirus: characteristics and topographical mapping of neutralizing epitopes on the E2 and E3 glycoproteins. *Virology* 161:410–20
33. Deregt, D., Sabara, M., Babiuk, L. A. 1987. Structural proteins of bovine coronavirus and their intracellular processing. *J. Gen. Virol.* 68:2863–77
34. Dubois-Dalcq, M. E., Doller, E. W., Haspel, M. V., Holmes, K. V. 1982. Cell tropism and expression of mouse hepatitis viruses (MHV) in mouse spinal cord cultures. *Virology* 119:317–31
35. Evans, M. R., Simpson, R. W. 1980. The coronavirus avian infectious bronchitis virus requires the cell nucleus and host transcriptional factors. *Virology* 105:582–91
36. Fields, B. N. 1981. Genetics of reovirus. *Curr. Top. Microbiol. Immunol.* 91:1–24
37. Garwes, D. J., Reynolds, D. J. 1981. The polypeptide structure of canine coronavirus and its relationship to porcine transmissible gastroenteritis virus. *J. Gen. Virol.* 52:153–57
38. Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., Blinov, V. M. 1989. Coronavirus genome: prediction of putative functional domains in the nonstructural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Res.* 17:4847–61
39. Hogue, B. G., Brian, D. A. 1986. Structural proteins of human respiratory coronavirus OC43. *Virus Res.* 5:131–44
40. Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., VandePol, S. 1982. Rapid evolution of RNA genomes. *Science* 215:1577–85
41. Holmes, E. V., Doller, E. W., Sturman, L. S. 1981. Tunicamycin-resistant glycosylation of coronavirus glycoprotein: demonstration of a novel type of viral glycoprotein. *Virol.* 115:334–44
42. Jacobs, L., Spaan, W. J. M., Horzinck, M. C., van der Zeijst, B. A. M. 1981. The synthesis of the subgenomic mRNAs of mouse hepatitis virus is initiated independently: evidence from UV transcription mapping. *J. Virol.* 39:401–6
43. Jacobs, L., van der Zeijst, B. A. M., Horzinek, M. C. 1986. Characterization and translation of transmissible gastroenteritis virus mRNAs. *J. Virol.* 57:1010–15
44. Keck, J. G., Hogue, B. G., Brian, D. A., Lai, M. M. C. 1988. Temporal regulation of bovine coronavirus RNA synthesis. *Virus Res.* 9:343–56
45. Keck, J. G., Matsushima, G. K., Makino, S., Fleming, J. O., Vannier, D. M., Stohlman, S. A., Lai, M. M. C. 1988. In vivo RNA-RNA recombination of coronavirus in mouse brain. *J. Virol.* 62:1810–13
46. Keck, J. G., Soe, L. H., Makino, S., Stohlman, S. A., Lai, M. M. C. 1988. RNA recombination of murine coronaviruses: recombination between fusion-positive MHV-A59 and fusion-negative MHV-2. *J. Virol.* 62:1989–98
47. Keck, J. G., Stohlman, S. A., Soe, L. H., Makino, S., Lai, M. M. C. 1987. Multiple recombination sites at the 5'-end of murine coronavirus RNA. *Virology* 156:331–41
48. King, A. M. Q., McCahon, D., Slade, W. R., Newman, J. W. I. 1982. Recombination in RNA. *Cell* 29:921–28
49. King, B., Potts, B. J., Brian, D. A. 1985. Bovine coronavirus hemagglutinin protein. *Virus Res.* 2:53–59
50. Krzystyniak, K., Dupuy, J. M. 1984. Entry of mouse hepatitis virus 3 into cells. *J. Gen. Virol.* 65:227–31
51. Kusters, J. G., Niesters, H. G. M., Lenstra, J. A., Horzinek, M. C., van der Zeijst, B. A. M. 1989. Phylogeny of antigenic variants of avian coronavirus IBV. *Virology* 169:217–21
52. Lai, M. M. C. 1986. Coronavirus leader-RNA-primed transcription: an alternative mechanism to RNA splicing. *BioEssays* 5:257–60
53. Lai, M. M. C. 1987. Replication of coronavirus RNA. In *RNA Genetics*, ed. E. Domingo, J. J. Holland, P. Ahlquist, 1:115–36. Boca Raton, Fl: CRC
54. Lai, M. M. C., Baric, R. S., Brayton, P. R., Stohlman, S. A. 1984. Characterization of leader RNA sequences on the virion and mRNAs

of mouse hepatitis virus, a cytoplasmic RNA virus. *Proc. Natl. Acad. Sci. USA* 81:3626–30
55. Lai, M. M. C., Brayton, P. R., Armen, R. C., Patton, C. D., Pugh, C., 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.* 39:823–34
56. Lai, M. M. C., Makino, S., Soe, L. H., Shieh, C.-K., Keck, J. G., Fleming, J. O. 1987. Coronavirus: a jumping RNA transcription. *Cold Spring Harbor Symp.* 52:359–65
57. Lai, M. M. C., Patton, C. D., Baric, R. S., Stohlman, S. A. 1983. Presence of leader sequences in the mRNA of mouse hepatitis virus. *J. Virol.* 46:1027–33
58. Lai, M. M. C., Patton, C. D., Stohlman, S. A. 1982. Replication of mouse hepatitis virus: negative-stranded RNA and replicative form RNA are of genome length. *J. Virol.* 44:487–92
59. Lai, M. M. C., Patton, C. D., Stohlman, S. A. 1982. Further characterization of mouse hepatitis virus: presence of common 5' -end nucleotides. *J. Virol.* 41:557–65
60. Lai, M. M. C., Stohlman, S. A. 1978. The RNA of mouse hepatitis virus. *J. Virol.* 26:236–42
61. Lai, M. M. C., Stohlman, S. A. 1981. Comparative analysis of RNA genomes of mouse hepatitis viruses. *J. Virol.* 38:661–70
62. Lapps, W., Hogue, B. G., Brian, D. A. 1987. Sequence analysis of the bovine coronavirus nucleocapsid and matrix protein genes. *Virology* 157:47–57
63. Leibowitz, J. L., De Vries, J. R. 1988. Synthesis of virus-specific RNA in permeabilized murine coronavirus-infected cells. *Virology* 166:66–75
64. Leibowitz, J. L., De Vries, J. R., Haspel, M. V. 1982. Genetic analysis of murine hepatitis virus strain JHM. *J. Virol.* 42:1080–87
65. Leibowitz, J. L., Perlman, S., Weinstock, G., De Vries, J. R., Budzilowicz, C., et al. 1988. Detection of a murine coronavirus nonstructural protein encoded in a downstream open reading frame. *Virology* 164:156–64
66. Leibowitz, J. L., Weiss, S. R., Paavola, E., Bond, C. W. 1982. Cell-free translation of murine coronavirus RNA. *J. Virol.* 43:905–13
67. Leibowitz, J. L., Wilhemsen, K. C., Bond, C. W. 1981. The virus-specific intracellular RNA species of two murine coronaviruses: MHV-A59 and MHV-JHM. *Virology* 114:39–51
68. Lomniczi, B., Kennedy, I. 1977. Genome of infectious bronchitis virus. *J. Virol.* 24:99–107
69. Luytjes, W., Bredenbeek, P. J., Noten, A. F., Horzinek, M. C., Spaan, W. J. 1988. Sequence of mouse hepatitis virus A59 mRNA 2: indications for RNA-recombination between coronavirus and influenza C virus: *Virology* 166:415–22
70. Luytjes, W., Sturman, L. S., Bredenbeek, P. J., Charite, J., van der Zeijst, B. A., Horzinek, M. C., Spaan, W. J. 1987. Primary structure of the glycoprotein E2 of coronavirus MHV-A59 and identification of the trypsin cleavage site. *Virology* 161:479–87
71. Mahy, B. W. J., Siddell, S., Wege, H., ter Meulen, V. 1983. RNA-dependent RNA polymerase activity in murine coronavirus-infected cells. *J. Gen. Virol.* 64:103–11
72. Makino, S., Fleming, J. O., Keck, J. G., Stohlman, S. A., 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. USA* 84:8567–71
73. Makino, S., Fujioka, N., Fujiwara, K. 1985. Structure of the intracellular defective viral RNAs of defective interfering particles of mouse hepatitis virus. *J. Virol.* 54:329–36
74. Makino, S., Keck, J. G., Stohlman, S. A., Lai, M. M. C. 1986. High-frequency RNA recombination of murine coronaviruses. *J. Virol.* 57:729–37
75. Makino, S., Lai, M. M. C. 1989. Evolution of the 5'-end of genomic RNA of murine coronaviruses. *Virology* 169:229–32
76. Makino, S., Lai, M. M. C. 1989. High-frequency leader sequence switching during coronavirus defective RNA replication. *J. Virol.* 63:5285–92
77. Makino, S., Shieh, C.-K., Keck, J. G., Lai, M. M. C. 1988. Defective interfering particles of murine coronavirus: mechanism of transcription of defective viral RNA. *Virology* 163:104–11
78. Makino, S., Shieh, C.-K., Soe, L. H., Baker, S. C., Lai, M. M. C. 1988. Primary structure and translation of a defective interfering RNA of murine coronavirus. *Virology* 166:550–60
79. Makino, S., Soe, L. H., Shieh, C.-K., Lai, M. M. C. 1988. Discontinuous transcription of coronavirus mRNAs

generates a variable number of pentanucleotide repeats at the leader fusion site. *J. Virol.* 62:3870–73
80. Makino, S., Stohlman, S. A., Lai, M. M. C. 1986. Leader sequences of murine coronavirus mRNAs can be freely reassorted: evidence for the role of free leader RNA in transcription. *Proc. Natl. Acad. Sci. USA* 83:4204–8
81. Makino, S., Taguchi, F., Fujiwara, K. 1984. Defective interfering particles of mouse hepatitis virus. *Virology* 133:9–17
82. Massalski, A., Coulter-Mackie, M., Dales, S. 1981. Assembly of mouse hepatitis virus strain JHM. In *Biochemistry and Biology of Coronaviruses,* ed. V. ter Meulen, S. Siddell, H. Wege, pp. 111–18. New York: Plenum.
83. Mizzen, L., Hilton, A., Cheley, S., Anderson, R. 1985. Attenuation of murine coronavirus infection by ammonium chloride. *Virology* 142:378–88
84. Niemann, H., Geyer, R., Klenk, H. D., Linder, D., Stirm, S., Wirth, M. 1984. The carbohydrates of mouse hepatitis virus (MHV) A59: structures of the O-glycosidically linked oligosaccharides of glycoprotein E1. *EMBO J.* 3:665–70
85. Oshiro, L. S. 1973. Coronaviruses. In *Ultrastructure of Animal Viruses and Bacteriophages: An Atlas,* ed. A. J. Dalton, F. Haguenau, pp. 331–43. New York: Academic
86. Pachuk, C. J., Bredenbeek, P. J., Zoltick, P. W., Spaan, W. J. M., Weiss, S. R. 1989. Molecular cloning of the gene encoding the putative polymerase of mouse hepatitis coronavirus strain A59. *Virology* 171:141–48
87. Palese, P. 1977. The genes of influenza virus. *Cell* 10:1–10
88. Parker, M. D., Cox, G. J., Deregt, D., Fitzpatrick, D. R., Babiuk, L. A. 1989. Cloning and *in vitro* expression of the gene for the E3 haemagglutinin glycoprotein of bovine coronavirus. *J. Gen. Virol.* 70:155–64
89. Perlman, S., Ries, D., Bolger, E., Chang, L. J., Stoltzfus, C. M. 1986. MHV nucleocapsid synthesis in the presence of cycloheximide and accumulation of negative strand MHV RNA. *Virus Res.* 6:261–72
90. Rasschaert, D., Laude, H. 1987. The predicted primary structure of the peplomer protein E2 of the porcine coronavirus transmissible gastroenteritis virus *J. Gen. Virol.* 68:1833–90
91. Robb, J., Bond, C. W. 1979. Pathogenic murine coronaviruses. I. Characterization of biological behavior *in vitro* and virus-specific intracellular RNA of strongly neurotropic JHMV and weakly neurotropic A59 viruses. *Virology* 94:352–70
92. Rottier, P., Brandenburg, D., Armstrong, J., van der Zeijst, B., Warren, G. 1984. Assembly *in vitro* of a spanning membrane protein of the endoplasmic reticulum: the El glycoprotein of coronavirus mouse hepatitis virus A59. *Proc. Natl. Acad. Sci. USA* 81:1421–25
93. Rottier, P. J., Welling, G. W., Welling-Wester, S., Niesters, H. G., Lenstra, J. A., van der Zeijst, B. A. M. 1986. Predicted membrane topology of the coronavirus protein E1. *Biochemistry* 25: 1335–39
94. Sawicki, S. G., Sawicki, D. L. 1986. Coronavirus minus-strand RNA synthesis and effect of cycloheximide on coronavirus RNA synthesis. *J. Virol.* 57: 328–34
94a. Sawicki, S. G., Sawicki, S. G 1990. Coronavirus transcription: Subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J. Virol.* 64:1050–56
95. Schmidt, L., Skinner, M., Siddell, S. 1987. Nucleotide sequence of the gene encoding the surface projection glycoprotein of coronavirus MHV-JHM. *J. Gen. Virol.* 68:47–56
96. Schreiber, S. S., Kamahora, T., Lai, M. M. C. 1989. Sequence analysis of the nucleocapsid protein gene of human coronavirus 229E. *Virology* 169:142–51
97. Sethna, P. B., Hung, S.-L., Brian, D. A. 1989. Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. *Proc. Natl. Acad. Sci. USA* 86:5626–30
98. Shieh, C.-K., Lee, H. J., Yokomori, K., Makino, S., Lai, M. M. C. 1989. Identification of a new transcriptional initiation site and the corresponding functional gene 2b in murine coronavirus RNA genome. *J. Virol.* 63:3729–36
99. Shieh, C.-K., Soe, L. H., Makino, S., Chang, M. F., Stohlman, S. A., Lai, M. M. C. 1987. The 5'-end sequence of the murine coronavirus genome: implications for multiple fusion sites in leader-primed transcription. *Virology* 156:321–30
100. Siddell, S. 1983. Coronavirus JHM: coding assignments of subgenomic mRNAs. *J. Gen. Virol.* 64:113–25

101. Skinner, M. A., Ebner, D., Siddell, S. G. 1985. Coronavirus MHV-JHM mRNA 5 has a sequence arrangement which potentially allows translation of a second, downstream open reading frame. *J. Gen. Virol.* 66:581–92
102. Smith, A. R., Boursnell, M. E. G., Binns, M. M., Brown, T. D. K., Inglis, S. C. 1987. Identification of a new gene product encoded by mRNA D of infectious bronchitis virus. In *Coronaviruses* ed. M. M. C. Lai, S. A. Stohlman, pp. 47–54. New York: Plenum
103. Soe, L. H., Shieh, C.-K., Baker, S. C., Chang, M.-F., Lai, M. M. C. 1987. Sequence and translation of the murine coronavirus 5'-end genomic RNA reveals the N-terminal structure of the putative RNA polymerase. *J. Virol.* 61:3968–76
104. Spaan, W., Cavanagh, D., Horzinek, M. C. 1988. Coronavirus: structure and genome expression. *J. Gen. Virol.* 69: 2939–52
105. Spaan, W., Delius, H., Skinner, M., Armstrong, J., Rottier, P., et al. 1983. Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO J.* 2:1839–44
106. Spaan, W. J. M., Rottier, P. J. M., Horzinek, M. C., van der Zeijst, B. A. M. 1981. Isolation and identification of virus-specific mRNAs in cells infected with mouse hepatitis virus (MHV-A59). *Virology* 108:424–34
107. Stern, D. F., Kennedy, S. I. T. 1980. Coronavirus multiplication strategy. I. Identification and characterization of virus-specified RNA. *J. Virol.* 34:665–74
108. Stern, D. F., Sefton, B. M. 1982. Coronavirus proteins: structure and function of the oligosaccharides of the avian infectious bronchitis virus glycoproteins. *J. Virol.* 44:804–12
109. Stern, D. F., Sefton, B. M. 1982. Coronavirus proteins: biogenesis of avian infectious bronchitis virus virion proteins. *J. Virol.* 44:794–803
110. Stern, D. F., Sefton, B. M. 1982. Synthesis of coronavirus mRNAs: kinetics of inactivation of IBV RNA synthesis by UV light. *J. Virol.* 42:755–59
111. Stern, D. F., Sefton, B. M. 1984. Coronavirus multiplication: locations of genes for virion proteins on the avian infectious bronchitis virus genome. *J. Virol.* 50:22–29
112. Stohlman, J. 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.* 62:4288–95
113. Stohlman, S. A., Lai, M. M. C. 1979. Phosphoproteins of murine hepatitis viruses. *J. Virol.* 32:672–75
114. Sturman, L., Holmes, K. 1985. The novel glycoprotein of coronaviruses. *Trends Biochem. Sci.* 10:17–20
115. Sturman, L. S., Holmes, K. V. 1977. Characterization of a coronavirus. II. Glycoproteins of the viral envelope: tryptic peptide analysis. *Virology* 77:650–60
116. Sturman, L. S., Holmes, K. V. 1983. The molecular biology of coronaviruses. *Adv. Virus Res.* 28:35–112
117. Sturman, L. S., Holmes, K. V., Behnke, J. 1980. Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. *J. Virol.* 33:449–62
118. Sturman, L. S., Ricard, C. S., Holmes, K. V. 1985. Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: activation of cell-fusing activity of virions by trypsin and separation of two different 90K cleavage fragments. *J. Virol.* 56:904–11
119. Sutton, R. E., Boothroyd, J. C. 1986. Evidence for *trans* splicing in trypanosomes. *Cell* 47:527–35
120. Tooze, J., Tooze, S., Warren, G. 1984. Replication of coronavirus MHV-A59 in sac-cells: determination of the first site of budding of progeny virions. *Eur. J. Cell Biol.* 33:281–93
121. Vlasak, R., Krystal, M., Nacht, R., Palese, P. 1987. The influenza C virus glycoprotein (HE) exhibits receptor binding (hemagglutinin) and receptor destroying (esterase) activities. *Virology* 160:419–25
122. Vlasak, R., Luytjes, W., Leider, W., Spaan, W., Palese, P. 1988. The E3 protein of bovine coronavirus is a receptor-destroying enzyme with acetylesterase activity. *J. Virol.* 52:4686–90
123. Vlasak, R., Luytjes, W., Spaan, W., Palese, P. 1988. Human and bovine coronaviruses recognize sialic acid–containing receptors similar to those of influenza C viruses. *Proc. Natl. Acad. Sci. USA* 85:4526–29
124. Wege, H., Dörries, R., Wege, H. 1984. Hybridoma antibodies to the murine coronavirus JHM: characterization of epitopes on the peplomer protein (E2). *J. Gen. Virol.* 65:1931–42
125. Wege, H., Muller, A., ter Meulen, V. 1978. Genomic RNA of the murine

coronavirus JHM. *J. Gen. Virol.* 41: 217–27
126. Wege, H., Siddell, S., ter Meulen, V. 1982. The biology and pathogenesis of coronaviruses. *Curr. Top. Microbiol. Immunol.* 99:165–200
127. Wilhelmsen, K. C., Leibowitz, J. L., Bond, C.W., Robb, J. A. 1981. The replication of murine coronavirus in enucleated cells. *Virology* 110:225–30
128. Yogo, Y., Hirano, N., Hino, S., Shibuta, H., Matumoto, M. 1977. Polyadenylate in the virion RNA of mouse hepatitis virus. *J. Biochem.* 82:1103–8
129. Yokomori, K., La Monica, N., Makino, S., Shieh, C.-K., Lai, M. M. C. 1989. Biosynthesis, structure, and biological activities of envelope protein gp65 of murine coronavirus. *Virology* 173:683–91
130. Kapke, P. A., Brian, D. A. 1986. Sequence analysis of the porcine transmissible gastroenteritis coronavirus nucleocapsid protein gene. *Virology* 151:41–49