See also: Nidovirales.

Further Reading

Britton P and Cavanagh D (2007) Avian coronavirus diseases and infectious bronchitis vaccine development. In: Thiel V (ed.) *Coronaviruses: Molecular and Cellular Biology* pp. 161–181. Norfolk, UK: Caister Academic Press.

Britton P and Cavanagh D (2007) Nidovirus genome organization and expression mechanisms. In: Perlman S Gallagher T, and Snijder EJ (eds.) *Nidoviruses*, pp. 29–46. Washington, DC: ASM Press.

Cavanagh D (2003) SARS vaccine development: Experiences of vaccination against avian infectious bronchitis coronavirus. *Avian Pathology* 32: 567–582.

Cavanagh D (2005) *Coronaviridae*: A review of coronaviruses and toroviruses. In: Schmidt A and Wolff MH (eds.) *Coronaviruses with Special Emphasis on First Insights Concerning SARS*, pp. 1–54. Basel: Birkhäuser.

Cavanagh D (2005) Coronaviruses in poultry and other birds. *Avian Pathology* 34: 439–448.

Enjuanes L, Almazán F, Sola I, and Zuñiga S (2006) Biochemical aspects of coronavirus replication: A virus–host interaction. *Annual Reviews in Microbiology* 60: 211–230.

González JM, Gomez-Puertas P, Cavanagh D, Gorbalenya AE, and Enjuanes L (2003) A comparative sequence analysis to revise the current taxonomy of the family *Coronaviridae*. *Archives of Virology* 148: 2207–2235.

Masters PS (2006) The molecular biology of coronaviruses. *Advances in Virus Research* 66: 193–292.

Siddell S, Ziebuhr J, and Snijder E (2005) Coronaviruses, toroviruses and arteriviruses. In: Mahy BWJ and ter Meulen V (eds.) *Topley and Wilson's Microbiology and Microbial Infections, Virology*, pp. 823–856. London: Hodder Arnold.

Coronaviruses: Molecular Biology

S C Baker, Loyola University of Chicago, Maywood, IL, USA

Glossary

Cell tropism Process that determines which cells can be infected by a virus. Factors such as receptor express can influence the cell type that can be infected.

Discontinuous transcription Process by which the coronavirus leader sequence and body sequence are joined to generate subgenomic RNAs.

Double membrane vesicles (DMVs) Vesicles that are generated during coronavirus replication when viral replicase proteins sequester host cell membranes. These vesicles are the site of coronavirus RNA synthesis.

Transcriptional regulatory sequences (TRSs) Sequences that are recognized by the coronavirus transcription complex to generate leader-containing subgenomic RNAs.

Introduction

Coronaviruses (CoVs) were first identified during the 1960s by using electron microscopy to visualize the distinctive spike glycoprotein projections on the surface of enveloped virus particles. It was quickly recognized that CoV infections are quite common, and that they are responsible for seasonal or local epidemics of respiratory and gastrointestinal disease in a variety of animals. CoVs have been named according to the species from which they were isolated and the disease associated with the viral infection. Avian infectious bronchitis virus (IBV) infects chickens, causing respiratory infection, decreased egg production, and mortality in young birds. Bovine coronavirus (BCoV) causes respiratory and gastrointestinal disease in cattle. Porcine transmissible gastroenteritis virus (TGEV) and porcine epidemic diarrhea virus (PEDV) cause gastroenteritis in pigs. These CoV infections can be fatal in young animals. Feline infectious peritonitis virus (FIPV) and canine coronavirus (CCoV) can cause severe disease in cats and dogs. Depending on the strain of the virus and the site of infection, the murine CoV mouse hepatitis virus (MHV) can cause hepatitis or a demyelinating disease similar to multiple sclerosis. CoVs also infect humans. Human coronaviruses (HCoVs) 229e and OC43 are detected worldwide and are estimated to be responsible for 5–30% of common colds and mild gastroenteritis. Interestingly, HCoV-OC43 and BCoV share considerable sequence similarity, indicating a likely transmission across species (either from cows to humans or vice versa) and then adaptation of the virus to its host. In contrast to the relatively mild infections caused by HCoV-229e and HCoV-OC43, the CoV responsible for severe acute respiratory syndrome (SARS-CoV) causes atypical pneumonia with a 10% mortality rate. Two additional HCoVs, HCoV-NL63 and HCoV-HKU1, have been recently identified using molecular methods and are associated with upper and lower respiratory tract infections in children, and elderly

and immunosuppressed patients. CoVs are grouped according to sequence similarity. CoVs that infect mammals are assigned to group 1 and group 2, whereas CoVs that infect birds are in group 3.

To date, the most infamous example of zoonotic transmission of a CoV is the outbreak of SARS in 2002–03. We now know that the outbreak started with cases of atypical pneumonia in the Guangdong Province in southern China in the fall of 2002. The infection was spread to tourists visiting Hong Kong in February, 2003, resulting in the dissemination of the outbreak to Hong Kong, Vietnam, Singapore, and Toronto, Canada. After attempting to treat cases of atypical pneumonia in Vietnam and acquiring the infection himself, Dr. Carlo Urbani alerted the World Health Organization (WHO) that this disease of unknown origin may be a threat to public health. The WHO rapidly organized an international effort to identify the cause of the outbreak, and within months a novel CoV was isolated from SARS patients and identified as the causative agent. Sequence analysis revealed that the virus was related to, but distinct from, all known CoVs. This led to an intensive search for an animal reservoir for this novel CoV. Initially, the masked palm civet and raccoon dog were implicated in the chain of transmission, since a SARS-CoV-like virus could be isolated from some animals found in wild animal markets in China. However, SARS-CoV-like viruses were never detected in animals captured from the wild, indicating that the civets may

have only served as an intermediate host in the chain of transmission. Further investigation revealed that the likely reservoir for SARS-CoV is the Chinese horseshoe bat (*Rhinolophus* spp.), which is endemically infected with a virus, named bat-SARS-CoV, that is closely related to SARS-CoV. The existence of an animal reservoir presents the possibility of re-emergence of this significant human pathogen. By improving our understanding of the molecular aspects of CoV replication and pathogenesis, we may facilitate development of appropriate antiviral agents and vaccines to control and prevent diseases caused by known and potentially emerging CoV infections.

Molecular Features of CoVs

CoV virions (**Figure 1(a)**) are composed of a large RNA genome, which combines with the viral nucleocapsid protein (N) to form a helical nucleocapsid, and a host cell-derived lipid envelope which is studded with virus-specific proteins including the membrane (M) glycoprotein, the envelope (E) protein, and the spike (S) glycoprotein. CoV particles vary somewhat in size, but average about 100 nm in diameter. The genomic RNA (gRNA) inside the virion, which ranges in size from 27 to 32 kb for different CoVs, is the largest viral RNA identified to date. CoV gRNAs have a broadly conserved structure which is illustrated by the SARS-CoV genome shown in **Figure 1(b)**. The gRNA is
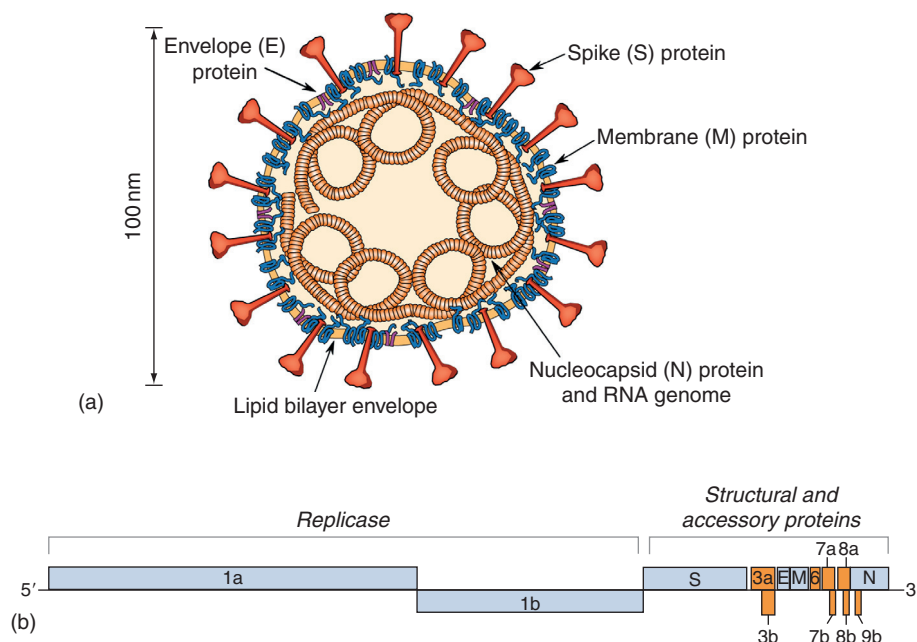

Figure 1 CoV virion and the genome of SARS-CoV. (a) Schematic diagram of a CoV virion with the minimal set of four structural proteins required for efficient assembly of the infectious virus particles: S, spike glycoprotein; M, membrane glycoprotein; E, envelope protein; and N, nucleocapsid phosphoprotein which encapsidates the positive-strand RNA genome. (b) Schematic diagram of the gRNA of SARS-CoV. Translation of the first two open reading frames (ORF1a and ORF1b) generates the replicase polyprotein. ORFs encoding viral structural and accessory (orange) ORFs are indicated at the 3′ end of the genome. (a) Reprinted from Masters PS (2006) The molecular biology of coronaviruses. *Advances in Virus Research* 66: 193–292, with permission from Elsevier.

capped at the 5′ end, with a short leader sequence followed by two long open reading frames (ORFs) encoding the replicase polyprotein. The remaining part of the genome encodes the viral structural and so-called accessory proteins. The structural protein genes are always found in the order S–E–M–N, but accessory protein genes may be interspersed at various sites between the structural genes. SARS-CoV has the most complex genome yet identified, with eight ORFs encoding accessory proteins. The expression of these ORFs is not required for viral replication, but they may play a role in the pathogenesis of SARS. In addition, the products of accessory genes may be incorporated into the virus particle, potentially altering the tropism or enhancing infectivity. For SARS-CoV, the proteins encoded in ORFs 3a, 6, 7a, and 7b have been shown to be incorporated in virus particles, but the exact role of these proteins in enhancing virulence is not yet clear.

The features of CoV structural proteins are shown in Figure 2. For each structural protein, a schematic diagram of the predicted structure of the protein is shown on the left and a linear display of the features is shown on the right. The CoV spike glycoprotein is essential for attachment of the virus to the host cell receptor and fusion of the virus envelope with the host cell membrane. CoV spike glycoproteins assemble as trimers with a short cytoplasmic tail and hydrophobic transmembrane domain anchoring the protein into the membrane. The spike glycoprotein is divided into the S1 and S2 regions, which are sometimes cleaved into separate proteins by cellular proteases during the maturation and assembly of virus particles. S1 contains the receptor-binding domain (RBD) and has been shown to provide the specificity of attachment for CoV particles. The cellular receptors and corresponding RBDs in S1 have been identified for several CoVs. MHV binds to murine carcinoembryonic
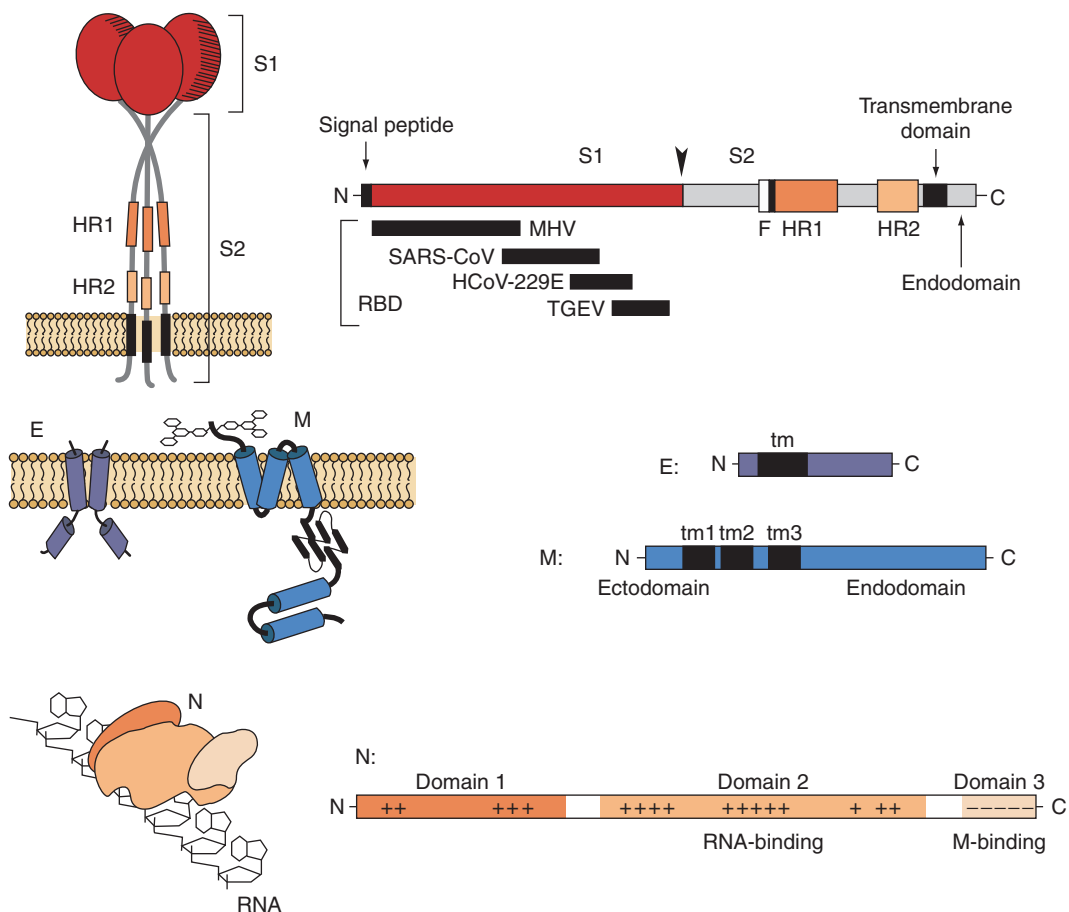

Figure 2 Diagrammatic representation of the spike trimer assembled on membranes, with the S1 receptor binding domain (RBD) and S2 fusion domain indicated. The linear map of spike indicates the location of the RBDs for three CoVs, and the relative location of the heptad repeat domains 1 and 2 (HR1 and HR2) which mediate the conformational changes required to present the fusion peptide (F) to cellular membranes. The membrane (M), envelope (E), and nucleocapsid (N) proteins represented in association with membranes or viral RNA. The linear map of each protein highlights the transmembrane domains of M and E and the RNA-binding and M protein-binding domains of N. Domains 1 and 2 of N are rich in arginine and lysine (indicated by +). Reprinted from Masters PS (2006) The molecular biology of coronaviruses. *Advances in Virus Research* 66: 193–292, with permission from Elsevier.

antigen-related cell adhesion molecules (mCEACAM1 and MCEACAM2); TGEV, FIPV, and HCoV-229e bind to species-specific versions of aminopeptidase N. Interestingly, both HCoV-NL63 and SARS-CoV have been shown to bind to human angiotensin-converting enzyme 2 (ACE2). ACE2 is expressed in both the respiratory and gastrointestinal tracts, consistent with virus replication at both these sites.

Once the S1 portion of the spike has engaged the host cell receptor, the protein undergoes a dramatic conformational change to promote fusion with the host cell membrane. Depending on the virus strain, this can occur at the plasma membrane on the surface of the cell, or in acidified endosomes after receptor-mediated endocytosis. The critical elements in the conformational change are the heptad repeats, HR1 and HR2, and the fusion peptide, F. After engaging the receptor, there is a dissociation of S1 which likely triggers the rearrangement of S2 so that HR1 and HR2 are brought together to form an antiparallel, six-helix bundle. This new conformation brings together the viral and host cell membranes and promotes the fusion of the lipid bilayers and introduction of the nucleocapsid into the cytoplasm. During infection, the spike glycoprotein is also present on the surface of the infected cell where it may (depending on the virus strain) promote fusion with neighboring cells and syncytia formation. The spike glycoprotein is also the major antigen to which neutralizing antibodies develop. The spike protein is a target for development of therapeutics for treatment of CoV infections. Monoclonal antibodies directed against the spike neutralize the virus by blocking binding to the receptor; synthetic peptides that block HR1-HR2 bundle formation have also been shown to block CoV infection.

The membrane (M) and envelope (E) proteins are essential for the efficient assembly of CoV particles. M is a triple-membrane-spanning protein that is the most abundant viral structural protein in the CoV virion. The ectodomain of M is generally glycosylated, and is followed by three transmembrane domains and an endodomain which is important for interaction with the nucleocapsid protein and packaging of the viral genome. The E protein is present in low copy numbers in the virion, but is important for efficient assembly. In the absence of E protein, few or no infectious virus particles are produced. The exact role of the E protein in the assembly of virus particles is still unknown, but recent studies suggest that E may act as an ion channel. The nucleocapsid protein (N) is an RNA-binding protein and associates with the CoV gRNA to assemble ribonucleoprotein complexes. The N protein is phosphorylated, predominantly at serine residues, but the role of phosphorylation is currently unknown. The N protein has three conserved domains, each separated by highly variable spacer elements. Domains 1 and 2 are rich in arginine and lysine residues, which is typical of many RNA-binding proteins. Domain 3 is essential for interaction

with the M protein and assembly of infectious virus particles. The N protein has been shown to be an important cofactor in CoV RNA synthesis and is proposed to act as an RNA chaperone to promote template switching, as described below.

Replication and Transcription of CoV RNA

The replication and transcription of CoV RNA takes place in the cytoplasm of infected cells (**Figure 3**). The CoV virion attaches to the host cell receptor via the spike glycoprotein and, depending on the virus strain, the spike mediates fusion directly with the plasma membrane or the virus undergoes receptor-mediated endocytosis and spike-mediated fusion with endosomal membranes to release the viral gRNA into the cytoplasm. Once the positive-strand RNA genome is released, it acts as a messenger RNA (mRNA) and the 5′ end (ORF1a and ORF1b) is translated by ribosomes to generate the viral RNA-dependent RNA polymerase polyprotein, termed the viral replicase. Translation of ORF1b is dependent on ribosomal frameshifting, which is facilitated by a slippery sequence and RNA pseudoknot structure present in all CoV gRNAs. The replicase polyprotein is processed by replicase-encoded proteases (papain-like proteases and a poliovirus 3C-like protease) to generate 16 mature replicase products. These viral replicase proteins sequester host cell membranes to generate distinctive double-membrane vesicles (DMVs) that have been shown to be the site of CoV RNA synthesis. The replicase complex on the DMVs then mediates the replication of the positive-strand RNA genome to generate full-length and subgenomic negative-strand RNAs, and the subsequent production of positive-strand gRNAs and sgmRNAs. The sgmRNAs are translated to generate viral structural and accessory proteins, and virus particles assemble with positive-strand gRNA in the endoplasmic reticulum-Golgi intermediate compartment (ERGIC) and bud into vesicles, with subsequent release from the cell. Depending on the virus strain, this replication can be robust and result in destruction of the host cell or a low-level, persistent infection that can be maintained in cultured cells or infected animals.

A hallmark of CoV transcription is the generation of a nested set of mRNAs, with each mRNA having the identical 'leader' sequence of approximately 65–90 nt at the 5′ end (**Figure 4(a)**). The leader sequence is encoded only once at the 5′ end of the gRNA. Each subgenomic mRNA (sgmRNA) has the identical leader sequence fused to the 5′ end of the body sequence. How are the leader-containing mRNAs generated during CoV transcription? Current evidence supports a model of discontinuous transcription, whereby the replicase complex switches templates during the synthesis of negative-strand RNA
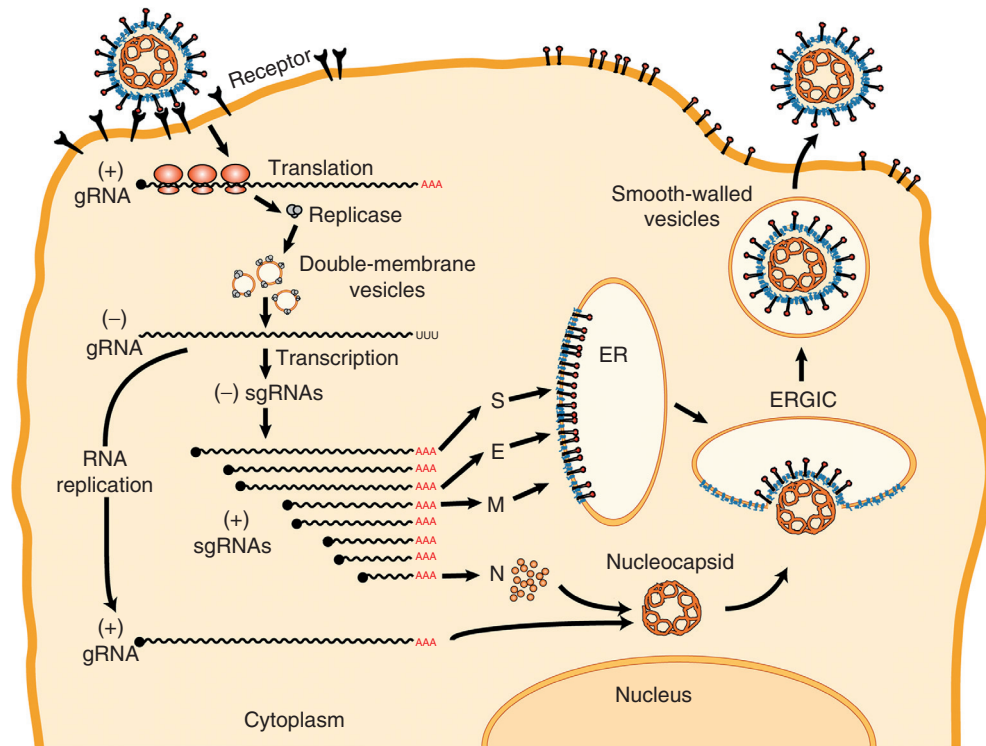
Figure 3 Replication cycle of CoVs. The spike glycoprotein on the virus particle interacts with host cell receptors to mediate fusion of the virus and host cell membranes and release of the positive-strand RNA genome into the cytoplasm. The 5′-proximal open reading frames (ORF1a and ORF1b) are translated from the gRNA to generate the replicase polyprotein. The replicase polyprotein is processed by viral proteases into 16 nonstructural proteins which assemble with membranes to generate double-membrane vesicles (DMVs) where RNA synthesis takes place. A nested set of 3′ co-terminal subgenomic (sg) RNAs is generated by a discontinuous transcription process. The sgRNAs are translated to generate the viral structural and accessory proteins. Viral gRNA is replicated and associates with nucleocapsid protein and viral structural proteins in the endoplasmic reticulum-Golgi intermediate compartment (ERGIC), where virus particles bud into vesicles before transport and release from the cell. Reprinted from Masters PS (2006) The molecular biology of coronaviruses. *Advances in Virus Research* 66: 193–292, with permission from Elsevier.

(**Figure 4(b)**). The key sequence element in this process is the transcriptional regulatory sequence (TRS). The TRS is a sequence of approximately 6–9 nt (5′-ACGAAC-3′ for SARS-CoV) which is found at the end of the leader sequence and at each intergenic region (the sites between the open reading frames encoding the viral structural and accessory proteins). Site-directed mutagenesis and deletion analysis has revealed the critical role of the TRS in mediating transcription of sgmRNAs. Deletion of any intergenic TRS results in loss of production of the corresponding sgmRNA. In addition, the CoV leader TRS and the intergenic TRS sequences must be identical for optimal production of the sgmRNAs. A three-step working model for template switching during negative-strand RNA synthesis has been proposed to describe the process for the generation of CoV leader-containing sgmRNAs (**Figure 4(b)**). In this process, the 5′ end and 3′ end of the gRNA form a complex with host cell factors and the viral replication complex. The 3′ end of the positive strand is used as the template for the initiation of transcription of negative-strand RNA. Negative-strand RNA synthesis continues up to the point of the TRS. At each TRS,

the viral replicase may either read through the sequence to generate a longer template, or switch templates to copy the leader sequence. The template switch allows the generation of a leader-containing sgmRNA. In this model, alignment of the leader TRS, the newly synthesized negative-strand RNA, and the genomic TRS is critical for the template switching to occur. Disruption of the complex, or loss of base-pairing within the complex, will result in the loss of production of that sgmRNA. Further studies of the CoV replication complex may yield new insights into the role of the viral helicase and endoribonuclease in the generation of the leader-containing CoV RNAs.

Another hallmark of CoV replication is high-frequency RNA recombination. RNA recombination occurs when a partially synthesized viral RNA dissociates from one template and hybridizes to similar sequences present in a second template. Viral RNA synthesis continues and generates a progeny virus with sequences from two different parental genomes. This RNA recombination event is termed copy-choice recombination. Copy-choice RNA recombination can be demonstrated experimentally when two closely related CoV strains (such as MHV-JHM and
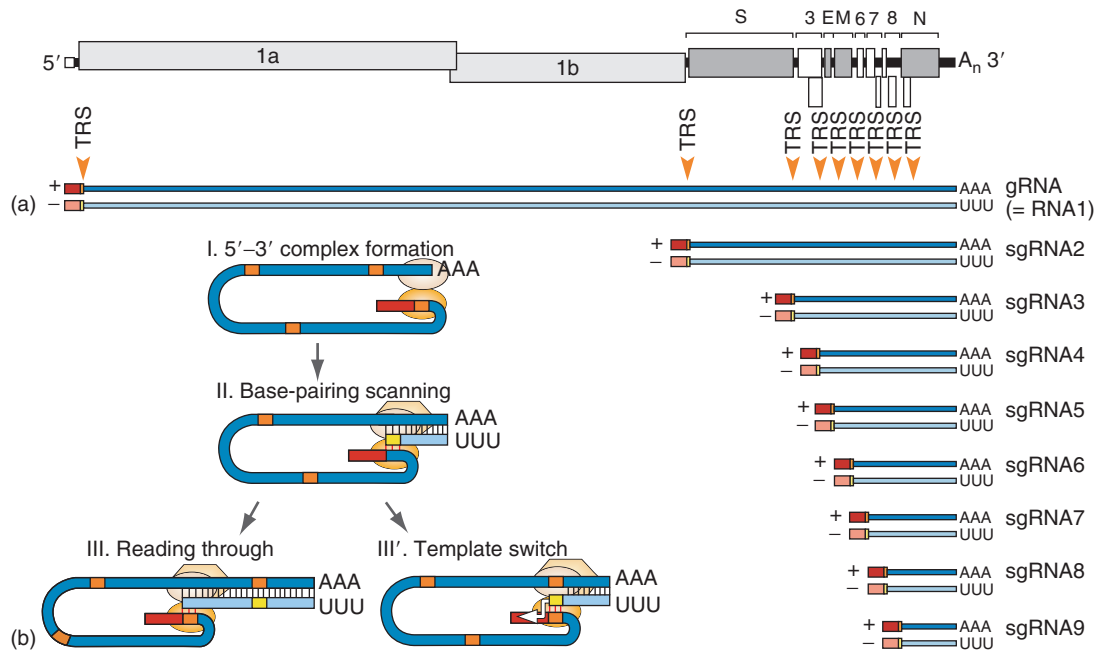
Figure 4 Model of SARS-CoV gRNA and sgRNAs, and a working model of discontinuous transcription. (a) Diagram of gRNA and the nested set of sgRNAs of SARS-CoV. The 5′ leader sequence, the transcriptional regulatory sequences (TRSs), and the positive- and negative-sense sgRNAs are indicated. (b) A working model of CoV discontinuous transcription. I. 5′–3′ complex formation. Binding of viral and cellular proteins to the 5′ and 3′ ends of the CoV gRNA is represented by ellipsoids. The leader sequence is indicated in red, the TRS sites are in orange. II. Base-pair scanning step. Minus-strand RNA (light blue) is synthesized from the positive-strand template by the viral transcription complex (hexagon). At the TRS site, base-pairing may occur between the template, the nascent negative-strand RNA, and the leader TRS sequence (dotted lines). III. The synthesis of negative-strand RNA can continue to make a longer sgRNA III, or a template switch can take place III′ to generate a leader-containing subgenomic negative-strand RNA, which could then serve as the template for leader-containing positive-strand sgRNAs. Modified from Enjuanes L, Almazán F, Sola I, and Zunia S (2006) Biochemical aspects of coronavirus replication: A virus–host interaction. *Annual Reviews in Microbiology* 60: 211–230.

MHV-A59) are used to coinfect cells. Recombinant viruses with cross-over sites throughout the genome can be isolated, although sequences within the spike glycoprotein may be a 'hot spot' for recombination due to the presence of RNA secondary structures that may promote dissociation and reassociation of RNA. It has been proposed that copy-choice recombination is also the mechanism by which many CoVs have acquired accessory genes, and it has been exploited experimentally for the deletion or insertion of specific sequences in CoV genomes to assess their role in virus replication and pathogenesis.

CoV Accessory Proteins

Sequence analysis of CoVs isolated from species ranging from birds to humans has revealed that all CoVs encode a core canonical set of genes, replicase (rep), spike (S), envelope (E), membrane (M), and nucleocapsid (N), and additional, so-called accessory genes (Table 1). The canonical genes are always found in the same order in the genome: rep-S-E-M-N. Reverse genetic studies (see below) have shown that this is the minimal set of

Table 1 Coronavirus **canonical** and *accessory* proteins

Virus	Proteins: canonical (rep-S-E-M-N) and accessory
Group 1	
TGEV	**rep-S**-*3a,3b*-**E-M-N**-*7*
FIPV	**rep-S**-*3a,3b,3c*-**E-M-N**-*7a,7b*
HCoV-229E	**rep-S**-*4a,4b*-**E-M-N**
PEDV	**rep-S**-*3*-**E-M-N**
HCoV-NL63	**rep-S**-*3*-**E-M-N**
Group 2	
MHV	**rep**-*2a, HE*-**S**-*4-5a,***E-M-N**,*7b*
BCoV	**rep**-*2a, HE*-**S**-*4a,4b-5,***E-M-N**,*7b*
HCoV-OC43	**rep**-*2a, HE*-**S**-*5,***E-M-N**,*7b*
HCoV-HKU1	**rep**-*HE*-**S**-*4*-**E-M-N**,*7b*
SARS-CoV	**rep-S**-*3a,3b*-**E-M**-*6-7a,7b-8a,8b*-**N**,*9b*
Bat-SARS-CoV	**rep-S**-*3*-**E-M**-*6-7a,7b-8*-**N**,*9b*
Group 3	
Avian IBV	**rep-S**-*3a,3b,3c*-**E-M**-*5a,5b*-**N**

genes required for efficient replication and assembly of infectious CoV particles. However, the genomes of all CoVs sequenced to date encode from one to eight additional ORFs, which code for accessory proteins. As the

name implies, these accessory proteins are not required for CoV replication in tissue culture cell lines, but they may play important roles in tropism and pathogenesis *in vivo*. How were these additional genes acquired? Current evidence indicates that these additional sequences may have been acquired by RNA recombination events between co-infecting viruses. For example, the hemagglutinin-esterase (HE) glycoprotein present in four different CoVs (MHV, BCoV, HCoV-OC43, and HCoV-HKU-1) was likely acquired by recombination of an ancestral CoV with the HE glycoprotein gene of influenza C. Interestingly, the expression of the HE gene has no effect on replication of the virus in cultured cell lines, but has been shown to enhance virulence in infected animals. Other CoV accessory genes may have been acquired through recombination with host cell mRNA or other viral mRNAs. The specific role of the accessory proteins in CoV replication and pathogenesis is under investigation. For SARS-CoV, accessory protein 6 has been implicated as an important factor in viral pathogenesis. Researchers have shown that mice infected with murine CoV expressing SARS-CoV protein 6 rapidly succumb to the infection, indicating that the protein 6 enhances virulence. In addition, recent studies suggest that SARS-CoV accessory proteins may play a role in blocking host cell innate immune responses, which may enhance viral replication and virulence. Other accessory proteins, such as SARS-CoV 3a and 7a, have been shown to be packaged into virus particles, where they may enhance infectivity or alter cell tropism. Future studies will be aimed at elucidating how CoV accessory proteins may modulate the virulence of CoV infection.

Manipulating CoV Genomes Using RNA Recombination and Reverse Genetics

Genetic manipulation of CoV sequences is challenging because of the large size (27–32 kbp) of the RNA genomes. However, two approaches have been developed to allow researchers to introduce mutations, deletions, and reporter genes into CoV genomes. These approaches are (1) targeted RNA recombination and (2) reverse genetics using infectious cDNA constructs of CoV. The first approach exploits high-frequency copy-choice recombination to introduce mutations of interest into the 3' end of the CoV gRNA. In the first step of targeted RNA recombination, a cDNA clone encoding the region from the spike glycoprotein to the 3' end of the RNA is generated. These sequences can be easily manipulated in the laboratory to introduce mutations or deletions, or for the insertion of reporter or accessory genes, into the plasmid DNA. Next, RNA is transcribed from the plasmid DNA and the RNA is transfected into cells coinfected with the CoV of interest. RNA recombination occurs between

the replicating CoV and the transfected substrate RNA, and viruses with the 3' end sequences derived from the transfected substrate RNA will be generated. The recombinant viruses are generated by high-frequency copy-choice recombination, but the challenge is to sort or select for the recombinant virus of interest from the background of wild-type virus. To facilitate selection of recombinant viruses, Masters and Rottier introduced the idea of host range-based selection. They devised a clever plan to use a mouse hepatitis virus (MHV) that encodes the spike glycoprotein from a feline CoV as the target for their recombination experiments. This feline-MHV, termed fMHV, will infect only feline cell lines. Substrate RNAs that encode the MHV spike and mutations of interest in the 3' region of the genome can be transfected into feline cells infected with fMHV, and progeny virus can be collected from the supernatant and subsequently selected for the ability to infect murine cell lines. Recombinant CoVs that have incorporated the MHV spike gene sequence (and the downstream substrate RNA with mutations of interest) can be selected for growth on murine cells, thus allowing for the rapid isolation of the recombinant virus of interest. This host range-based selection step is now widely used by virologists to generate recombinant viruses with specific alterations in the 3' end of the CoV genome.

The second approach for manipulating CoV sequences, generating infectious cDNA constructs of CoV, has been developed in several laboratories. Full-length CoV sequences have been cloned and expressed using bacterial artificial chromosomes (BACs), vaccinia virus vectors, and from an assembled set of cDNA clones representing the entire CoV genome. The generation of a full-length cDNA and subsequent generation of a full-length CoV gRNA allows for reverse genetic analysis of CoV sequences. Successful reverse genetics systems are now in place to study the replication and pathogenesis of SARS-CoV, MHV, HCoV-229e, and IBV. These reverse genetics systems have allowed researchers to introduce mutations into the replicase gene and identify sites that are critical for enzymatic activities of many replicase products such as the helicase, endoribonuclease, and the papain-like proteases. Reverse genetic approaches are also being used to investigate the role of the TRSs in controlling the synthesis of CoV mRNAs. Interestingly, the SARS-CoV genome can be 're-wired' using a novel, nonanonical TRS sequence, which must be present at both the ends of the leader sequence and at each intergenic junction. This 're-wired' SARS-CoV may be useful for generating a live-attenuated SARS-CoV vaccine. An important feature of this 're-wired' virus is that it would be nonviable if it recombined with wild-type virus, since the leader TRS and downstream TRS would no longer match in a recombinant virus. The development of reverse genetics systems for CoVs has opened

the door to investigate how replicase gene products function in the complex mechanism of CoV discontinuous transcription, and provides new opportunities to generate novel CoVs as potential live-attenuated or killed virus vaccines to reduce or prevent CoV infections in humans and animals.

Vaccines and Antiviral Drug Development

Because of the economic importance of CoV infection to livestock and domestic animals, a variety of live-attenuated and killed CoV vaccines have been tested in animals. Vaccines have been developed against IBV, TGEV, CCoV, and FIPV. However, these vaccines do not seem to provide complete protection from wild-type virus infection. In some cases, the wild-type CoV rapidly evolves to escape neutralization by vaccine-induced antibodies. In studies of vaccinated chickens, a live-attenuated IBV vaccine has been shown to undergo RNA recombination with wild-type virus to generate vaccine escape mutants. Killed virus vaccines may also be problematic for some CoV infections. Vaccination of cats with a killed FIPV vaccine has been shown to exacerbate disease when cats are challenged with wild-type virus. Therefore, extensive studies will be required to carefully evaluate candidate vaccines for SARS-CoV. A variety of approaches are currently under investigation for developing a SARS-CoV vaccine, including analysis of killed virus vaccines, live-attenuated virus vaccines, DNA immunization, and viral vector vaccines (such as modified vaccine virus Ankara, canarypox, alphavirus, and adenovirus vectors). The development of improved animal models for SARS will be essential for evaluating SARS-CoV candidate vaccines. Transgenic mice expressing human ACE-2 may be an appropriate small animal model. Initial studies suggest that Syrian hamsters and ferrets develop pneumonia and lung pathology similar to that seen in humans after infection with SARS-CoV, and therefore may be appropriate animal models for viral pathogenesis. CoV vaccine studies will benefit from an improved understanding of conserved viral epitopes that can be targeted for vaccine development.

The use of neutralizing monoclonal antibodies directed against the SARS-CoV spike glycoprotein is another approach that may provide protection from severe disease. The success in the development and use of humanized monoclonal antibodies against respiratory syncytial virus (family *Paramyxoviridae*) to protect infants from severe disease indicates that this approach is certainly worth investigating. Preliminary studies have indicated that patient convalescent serum and monoclonal antibodies directed against the SARS-CoV spike glycoprotein efficiently neutralize infectious virus. Further studies are essential to evaluate any concerns about potential antibody-mediated enhancement of disease and to determine if neutralization escape mutants arise rapidly after challenge with infectious virus. Studies evaluating monoclonal antibodies directed against a variety of structural proteins, and monoclonal antibodies directed against conserved sites in the spike glycoprotein will provide important information on the efficacy of passive immunity to protect against SARS.

Currently, there are no antiviral drugs approved for use against any human CoV infection. With the potential for the emergence or re-emergence of pathogenic CoV from animal reservoirs, there is considerable interest in identifying potential therapeutic targets and developing antiviral drugs that will block viral replication and reduce the severity of CoV infections in humans. Two promising targets for antiviral drug development are the SARS-CoV protease domains, the papain-like protease (PL^{pro}) and the 3C-like protease (3CL^{pro}, also termed the main protease, M^{pro}) (**Figure 5**). These two protease domains
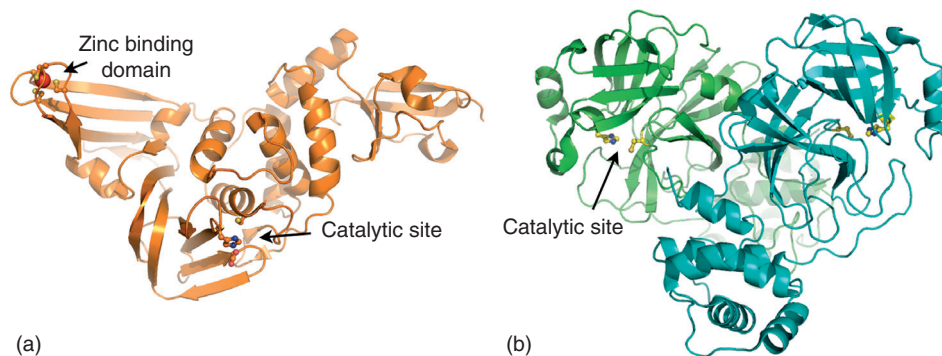

Figure 5   CoV proteases are targets for antiviral drug development; X-ray structures of the two SARS-CoV protease domains encoded in the replicase polyprotein. (a) The SARS-CoV papain-like protease (PL^{pro}) with catalytic triad cysteine, histidine, and aspartic acid residues, and zinc-binding domain indicated. (b) The 3C-like protease (3CL^{pro}, also termed main protease, M^{pro}) dimer with catalytic cysteine and histidine residues indicated.

are encoded within the replicase polyprotein gene and protease activity is required to generate the 16 replicase nonstructural proteins (nsp1–nsp16) that assemble to generate the viral replication complex. The crystal structure of the 3CL^{pro} was determined first from TGEV and then from SARS-CoV. Rational drug design, much of which was based on our knowledge of inhibitors directed against the rhinovirus 3C protease, has provided promising lead compounds for 3CL^{pro} antiviral drug development. Interestingly, these candidate antivirals have been shown to inhibit the replication of SARS-CoV and other group 2 CoVs such as MHV, and the less related group 1 CoV, HCoV-229e. This indicates that the active site of 3CL^{pro} is highly conserved among CoVs and that antiviral drugs developed against SARS-CoV 3CL^{pro} may also be useful for inhibiting the replication of more common human CoVs such as HCoV-229e, HCoV-OC43, HCoV-NL63, and HCoV-HKU1. Further studies are needed to determine if these inhibitors can be developed into clinically useful antiviral agents.

Analysis of SARS-CoV papain-like protease led to the surprising discovery that this protease is also a viral deubiquitinating (DUB) enzyme. The SARS-CoV PL^{pro} was shown to be required for processing the amino-terminal end of the replicase polyprotein and to recognize conserved cleavage site (-LXGG). The LXGG cleavage site is also the site recognized by cellular DUBs to remove polyubiquitin chains from proteins targeted for degradation by proteasomes. Analysis of the X-ray structure of the SARS-CoV PL^{pro} has revealed that it has structural similarity to known cellular DUBs. These studies suggest that CoV papain-like proteases have evolved to have both proteolytic processing and DUB activity. The DUB activity may be important in preventing ubiquitin-mediated degradation of viral proteins, or the DUB activity may be important in subverting host cell pathways to enhance viral replication. PL^{pro} inhibitors are now being developed using structural information and by performing high-throughput screening of small molecule libraries to identify lead compounds. Additional CoV replicase proteins, particularly the RNA-dependent RNA polymerase, helicase, and endoribonuclease, are also being targeted for antiviral drug development.

Future Perspectives

The development of targeted RNA recombination and reverse genetics systems for CoVs has provided new opportunities to address important questions concerning the mechanisms of CoV replication and virulence, and to design novel CoV vaccines. In the future, improved small animal models for testing vaccines and antivirals, and the availability of additional X-ray crystallographic structure information for rational drug design will be critical for further progress toward development of effective vaccines and antiviral drugs that can prevent or reduce diseases caused by CoVs.

See also: Coronaviruses: General Features; Nidovirales; Severe Acute Respiratory Syndrome (SARS); Torovirus.

Further Reading

Baker SC and Denison M (2007) Cell biology of nidovirus replication complexes. In: Perlman S, Gallagher T, and Snijder E (eds.) *The Nidoviruses*. Washington, DC: ASM Press.

Baric RS and Sims AC (2005) Development of mouse hepatitis virus and SARS-CoV infectious cDNA constructs. *Current Topics in Microbiology and Immunology* 287: 229–252.

Enjuanes L, Almazán F, Sola I, and Zuñiga S (2006) Biochemical aspects of coronavirus replication: A virus–host interaction. *Annual Reviews in Microbiology* 60: 211–230.

Lau YL and Peiris JS (2005) Pathogenesis of severe acute respiratory syndrome. *Current Opinion in Immunology* 17: 404–410.

Li W, Wong SK, Li F, *et al.* (2006) Animal origins of the severe acute respiratory syndrome coronavirus: Insights from ACE2-S-protein interactions. *Journal of Virology* 80: 4211–4219.

Masters PS (2006) The molecular biology of coronaviruses. *Advances in Virus Research* 66: 193–292.

Masters PS and Rottier PJM (2005) Coronavirus reverse genetics by targeted RNA recombination. *Current Topics in Microbiology and Immunology* 287: 133–160.

Perlman S and Dandekar AA (2005) Immunopathogenesis of coronavirus infections: Implications for SARS. *Nature Reviews Immunology* 5: 917–927.

Ratia K, Saikatendu K, Santarsiero BD, *et al.* (2006) Severe acute respiratory syndrome coronavirus papain-like protease: Structure of a viral deubiquitinating enzyme. *Proceedings of the National Academy of Sciences, USA* 103: 5717–5722.

Shi ST and Lai MMC (2005) Viral and cellular proteins involved in coronavirus replication. *Current Topics in Microbiology and Immunology* 287: 95–132.

Stadler K, Masignani V, Eickmann M, *et al.* (2003) SARS – Beginning to understand a new virus. *Nature Reviews Microbiology* 1: 209–218.

Thiel V and Siddell S (2005) Reverse genetics of coronaviruses using vaccinia virus vectors. *Current Topics in Microbiology and Immunology* 287: 199–228.

Wang L, Shi Z, Zhang S, Field H, Daszak P, and Eaton BT (2006) Review of bats and SARS. *Emerging Infectious Diseases* 12: 1834–1840.

Yang H, Xie W, Xue X, *et al.* (2005) Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biology* 3: 1742–1751.

Yount B, Roberts RS, Lindesmith L, and Baric RS (2006) Rewiring the severe acute respiratory syndrome coronavirus (SARS-CoV) transcription circuit: Engineering a recombinantion-Resistant genome. *Proceedings of the National Academy of Sciences, USA* 103: 12546–12551.

Relevant Websites

http://patric.vbi.vt.edu – Coronavirus bioinformatics resource, PATRIC (PathoSystems Resource Integration Center).

http://www.cdc.gov – Severe acute respiratory syndrome (SARS) resource, Centers for Disease Control and Prevention.