Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Coronaviruses: General Features (Coronaviridae)

Paul Britton, The Pirbright Institute, Pirbright, United Kingdom

© 2021 Published by Elsevier Ltd.

This is an update of D. Cavanagh, P. Britton, Coronaviruses: General Features, In Encyclopedia of Virology (Third Edition), edited by Brian W.J. Mahy, Marc H.V. Van Regenmortel, Elsevier Ltd., 2008, doi:10.1016/B978-012374410-4.00370-8.

Glossary

Infectious clone A full-length DNA copy of an RNA virus genome from which full-length viral RNA can be generated, leading to production of infectious virus.

Nidovirales (nidoviruses) An order comprising positive-sense RNA coronaviruses, toroviruses, arteriviruses, and roniviruses that have a common genome organization and expression, similar replication/transcription strategies, and form a nested set of 3’ co-terminal subgenomic mRNAs (nidus, Latin for nest).

Ribosomal frameshifting Movement (shift) backward by one nucleotide of a ribosome that is on an RNA, caused by particular RNA structures and sequences. Subsequent continuation of the progress of the ribosome is in a different open reading frame.

Introduction

Coronaviruses are known to cause disease in humans, other mammals, and birds. They cause major economic loss, sometimes associated with high mortality, in neonates of some domestic species (e.g., chickens, pigs). In humans, they are responsible for respiratory and enteric diseases. Coronaviruses do not necessarily observe species barriers, as illustrated most graphically by the recent spread of three zoonotic viruses in humans, severe acute respiratory syndrome (SARS) coronavirus, Middle East respiratory syndrome (MERS) coronavirus and in 2019 SARS-CoV-2 responsible for the disease Covid-19 causing a global Pandemic infection. All of which appear to have originated from bats via an intermediate wild animal species before infecting humans, with lethal consequences. As a group, coronaviruses are not limited to particular organs; target tissues include the nervous system, immune system, kidney, and reproductive tract in addition to many parts of the respiratory and enteric systems. A great advance in recent years has been the development of systems (‘infectious clones’) for modifying the genomes of coronaviruses to study all aspects of coronavirus replication, and for the development of new vaccines.

Taxonomy and Classification

Coronaviruses are part of the Order Nidovirales, which is divided into five Suborders; the Arterivirinae, which contains the Family Arteriviridae; the Coronavirinae, which contains the Family Coronaviridae; the Toruvirinae, which contains the Family Toruviridae; the Mesonvirinae, which contains the Family Mesonviridae; and the Ronuvirinae, which contains the Family Ronuviridae. The viruses generically known as coronaviruses fall within the Coronavirinae Family which are divided into four genera, the Alphacoronavirus, Betacoronavirus, Gammacoronavirus and Deltacoronavirus. Members of the Nidovirales Order have a similar genome organization and produce a nested set of subgenomic mRNAs and contain related specialized enzymes that are involved in replication of the RNA. Coronaviruses have been placed into different genera, initially on the basis of serological relationships which has subsequently been refined by gene sequencing.

Virion Properties

Virions have a buoyant density of approximately 1.18 g ml⁻¹ in sucrose. Being enveloped viruses (Fig. 1(a)), they are destroyed by organic solvents such as ether and chloroform.

Virion Structure and Composition

All coronaviruses have four structural proteins in common (Fig. 1(b)): a large surface glycoprotein (S; c. 1150–1450 amino acids); a small envelope protein (E; c. 100 amino acids, present in very small amounts in virions); integral membrane glycoprotein (M; c. 250 amino acids); and a phosphorylated nucleocapsid protein (N; c. 500 amino acids). Some Betacoronaviruses have an additional structural glycoprotein, the hemagglutinin-esterase protein (HE; c. 425 amino acids). This is not essential for replication in vitro and may affect tropism in vivo.
Virions are c. 120 nm in diameter, although they can be up to twice that size, and the ring of S protein spikes is approximately 20 nm deep. When present, the HE protein forms a layer 5–10 nm deep. In some species, the S protein is cleaved into two subunits, the N-terminal S1 fragment being slightly smaller than the C-terminal S2 sequence. The S protein is anchored in the envelope by a transmembrane region near the C-terminus of S2. The functional S protein is highly glycosylated and exists as a trimer. The bulbous outer part of the mature S protein is formed largely by S1 while the stalk is formed largely by S2, having a coiled-coil structure. S1 is the most variable part of the S protein; some serotypes of the avian coronavirus, infectious bronchitis virus (IBV) differ from one another by 40% of S1 amino acids. S1 is the major inducer of protective immune responses. Variation in the S1 protein enables one strain of virus to avoid immunity induced by another strain of the same species.

The M glycoprotein is the most abundant protein in virions. In most cases, only a small part (~20 amino acids) at the N-terminus protrudes at the surface of the virus. There are three membrane-spanning segments and the C-terminal half of the M protein is within the lumen of the virus. In transmissible gastroenteritis virus (TGEV), a proportion of M molecules have four membrane-spanning segments, resulting in the C-terminus also being exposed on the outer surface of the virus (M’ in Fig. 1(b)). The E protein is anchored in the membrane by a sequence near its N-terminus.

**Fig. 1** (a) Electron micrograph of an IBV virion, showing the bulbous S protein. (b) Diagrammatic representation of the composition and structure of a coronavirus virion: S, spike glycoprotein; M, M’, integral membrane glycoprotein; E, small envelope protein; N, nucleocapsid protein; NC, nucleocapsid (nucleoprotein) comprising the RNA genome and N protein. Cryoelectron microscopy of TGEV has indicated a core structure comprising the NC and the M protein. Two forms of M protein (M, M’) have been observed for TGEV (see main text). The coronavirus membrane proteins, S, E, M, and M’, are inserted into a lipid bilayer (MEM) derived from internal cell membranes. (b) Reproduced from González, J.M., Gomez-Puertas, P., Cavanagh, D., Gorbatenya, A.E., Enjuanes, L., 2003. A comparative sequence analysis to revise the current taxonomy of the family. Coronaviridae. Archives of Virology 148, 2207–2235, with permission from Springer-Verlag.
Coronaviruses have the largest known RNA genomes, which comprise 28–32 kb of positive sense, single-stranded RNA. The overall genome organization being 5′ UTR–polymerase gene–structural protein genes–3′ UTR, where the UTRs are untranslated regions (Fig. 2). The first 60–90 nucleotides at the 5′ end form a leader sequence. The structural protein genes are in the same order in all coronaviruses: (HE)–S–E–M–N. Interspersed among these genes are one or more other genes, often referred to as accessory genes as they have been demonstrated to be non-essential for replication in several coronaviruses using reverse genetic systems. The accessory genes encode small proteins of mainly unknown function. Some of these genes encode two or three proteins. In some cases (e.g., gene 3 of IBV and gene 5 of murine hepatitis virus (MHV)), translation of the third and second open reading frame (ORF), respectively, is affected by the preceding ORFs acting as internal ribosome entry sites. The proteins encoded by these small ORFs are mostly not required for replication in vitro; some of them might function as antagonists of innate immune responses.

Following entry into a cell and the release of the virus ribonucleoprotein (genome surrounded by the N protein) into the cytoplasm, ribosomes translate gene 1, which is approximately 20 kb, into two polyproteins (pp1a and pp1ab). These are cleaved by gene 1-encoded proteases, to generate 15 or 16 proteins (Fig. 3). Translation of gene 1 results in two polyproteins, translation of the ORF 1a region results in pp1a and translation of the ORF 1b region results in pp1ab, the latter involves ribosomal frameshifting, which has two elements, a slippery site followed by an RNA pseudoknot. At the slippery site (UUIUAAC in IBV), the ribosome slips one nucleotide backward and then moves forward, this time in a −1 frame compared with translation ORF 1a, resulting in the synthesis polyprotein 1ab, in which the proteins encoded by ORF 1b are in effect fused with the proteins from ORF 1a.

Proteins, including the RNA-dependent RNA polymerase (RdRp), from gene 1 associate to form the replicase complex, which is membrane associated. Coronavirus subgenomic mRNAs are generated by a discontinuous process. At the beginning of each gene is a common sequence (CUUAACAA in the case of IBV) called a transcription regulatory sequence (TRS). When the polymerase producing the nascent negative sense RNA reaches a TRS, RNA synthesis is attenuated, followed by continuation at the 5′ end of genomic RNA. This results in the addition of a negative copy of the leader sequence to the negative-sense RNA, resulting in a negative-sense copy of an sg mRNA. Of course, progress of the polymerase is not always halted at a TRS. Rather, it sometimes continues, producing a nested set of negative-sense sg mRNAs. These negative-sense copies of the sg mRNAs are the templates for the generation of the positive-sense sg mRNAs (Fig. 3). The amount of each sg mRNA does not necessarily decrease in a linear fashion; the efficiency of termination by a TRS is dependent on adjacent sequences, which are different for each gene. The leader sequence is found at the very 5′ end of the genomic RNA and at the 5′ ends of each sg mRNA.
Replication Cycle

The N-terminal (S1) part of the S protein mediates attachment to cells. It is a determinant of host species specificity and, in some cases, pathogenicity, by determining susceptible cell range (tissue tropism) within a host. The C-terminal S2 part triggers fusion of the virus envelope with cell membranes (plasma membrane or endosomal membranes), which can occur at neutral or slightly acidic pH, depending on species or even strain. The virus glycoproteins (S, M, and HE, when present) are synthesized at the endoplasmic reticulum. Both subunits of the S protein are multiply glycosylated, while the M protein has one or two glycans close to its N-terminus. Interestingly, glycosylation of the M protein can be either N- or O-linked, depending on the type of coronavirus, although experiments using reverse genetics showed that conversion of an O-linked glycosylated M protein to an N-linked version had no effect on virus growth.

Early and late in infection, formation of virus particles can occur in the endoplasmic reticulum–Golgi intermediate compartment (ERGIC) and endoplasmic reticulum, but most assembly occurs in the Golgi membranes. The M protein is not transported to the plasma membrane; its location at internal membranes determines the sites of virus particle formation. It interacts with the N protein (as part of the RNP) and C-terminal part of the S protein, retaining some, though not all, of the S protein at internal membranes. The E protein is essential for virus particle formation, though it is not known how it functions. It has a sequence that determines its accumulation at internal membranes, and its interaction with the M protein. The latter’s interaction with the N protein enables the formation of virus particles with spikes.

Genome Replication and Recombination

Following infection of a susceptible cell, the coronavirus genomic RNA is released from the virion into the cytoplasm and immediately recognized as an mRNA for the translation of the replicase pp1a and pp1ab proteins. These proteins are cleaved by ORF1a-encoded proteases, after which they become part of replicase complexes for the synthesis of either complete negative-sense copies of the genomic RNA or negative-sense copies of the sg mRNAs. The negative-sense RNAs are used as templates for the synthesis of genomic RNA and sg mRNAs (Fig. 2). RNA synthesis appears to take place in replication complexes associated with rearranged host membranes generated by several pp1a derived proteins. Following synthesis of the sg mRNAs, the structural proteins are produced for the assembly and encapsidation of the de novo-synthesized genomic RNA, resulting in the release of new infectious coronavirus virions. The release of new virions starts 4–7 h after the initial infection. As indicated above, the synthesis of the sg mRNAs is the result of a discontinuous process in which the synthesis of a negative-sense copy of an sg mRNA is completed.
by the addition of the negative-sense leader sequence by a recombination mechanism. If a cell is infected with two related coronaviruses, the polymerase may swap between two RNA templates, in a similar way to addition of the leader sequence. This ‘copy-choice’ mechanism of genetic recombination results in a chimeric RNA. Such RNAs may give rise to new viruses with modified genomes with a capacity to infect a different cell and, in some cases, new host species.

**Diseases and Host Range**

Probably all coronaviruses replicate in epithelial cells of the respiratory and/or enteric tracts, though not necessarily producing clinical damage at those sites. The avian coronavirus (a Gamma-coronavirus), IBV, not only causes respiratory disease but can also damage gonads in both females and males, and causes serious kidney disease (dependent on the strain of virus, and to some extent on the breed of chicken). IBV is able to replicate at virtually every epithelial surface in the host. Some coronaviruses have their most profound effect in the alimentary tract (e.g., the porcine coronavirus transmissible gastroenteritis virus (TGEV) causes ≥90% mortality in neonatal pigs). Human coronaviruses are known to be associated with enteric and respiratory diseases (e.g., diarrhea), in addition to respiratory disease. SARS-CoV was also associated with diarrhea in humans, in addition to serious lung disease. Other coronaviruses, for example, MHV and porcine HEV, spread to cells of the central nervous system, producing disease, for example, acute or chronic demyelination in the case of MHV.

Coronavirus replication and disease are not necessarily restricted to a single host species. Canine enteric CoV and feline CoV can replicate and cause disease in pigs; these two viruses have proteins with very high amino acid identity to those of porcine TGEV. Canine respiratory CoV has proteins, including the S protein (which is the attachment protein and a determinant of host range), with very high amino acid identity (≥75%) to other group 2 viruses Hu CoV-OC43 and BCoV. This raises the possibility of co-infection in these hosts. There is evidence that pheasant CoV can infect chickens, and IBV infect teal (a duck), though without causing disease. The most dramatic demonstration that coronaviruses can have a wide host range was provided by SARS-CoV. This appears to have its origin in bats, was transferred to various other species (e.g., civet cat) that were captured for trade, and then caused lethal disease in humans. A similar zoonotic pathway has been hypothesized for the infection of humans by SARS-CoV-2, again with an origin in bats but via, as yet unknown, intermediate host or hosts, though generally accepted via a wild animal in a live animal market.

Persistent infections in vivo are well known for MHV, and less well known for other coronaviruses (e.g., IBV). Following infection of very young chickens, IBV is re-excreted when hens start to lay eggs. The trigger for release is probably the stress of coming into lay.

The S protein is a determinant of both tissue tropism within a host and host range. This has been elegantly demonstrated by genetic manipulation of the genome of MHV, which is unable to attach to feline cells. Replacement of the MHV S protein gene with that of CoV from feline coronavirus resulted in a recombinant virus that was able to attach, and subsequently replicate in, feline cells. However, other proteins can also affect pathogenicity. Research with genetically modified coronaviruses, using targeted recombination or ‘infected clones’, has shown that modifications to proteins encoded in ORF1 and the small genes interspersed among the structural protein genes, result in attenuation of pathogenicity. Although the roles of these ‘accessory proteins’ are not known, this may offer a route to the development of a new generation of live vaccines. Currently, the most widely used prophylactics for control of IBV in chickens include killed vaccines and live vaccines attenuated by passage in embryonated eggs. However, disease control is complicated by extensive variation in the S1 protein which is the inducer of protective immunity.

**See also:** Coronaviruses: Molecular Biology (Coronaviridae), Enveloped, Positive-Strand RNA Viruses (Nidovirales). Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) (Coronaviridae). Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) (Coronaviridae)

**Further Reading**

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

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

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., Wolff, M.H. (Eds.), Coronaviruses with Special Emphasis on First Insights Concerning SARS. Basel: Birkhäuser, pp. 1–54.

Cavanagh, D., 2005. Coronaviruses in poultry and other birds. Avian Pathology 34, 439–446.

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

González, J.M., Gomez-Puertas, P., Cavanagh, D., Gorbalenya, A.E., Enjuanes, L., 2003. A comparative sequence analysis to revise the current taxonomy of the family Coronaviridae: Archives of Virology 148, 2207–2235.

Masters, P.S., 2006. The molecular biology of coronaviruses. Advances in Virus Research 66, 193–292.

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