Cellular cryo-tomography of nidovirus replication organelles

Wolff, G.

CHAPTER 1

General introduction

The order Nidovirales

The *Nidovirales* is an order of positive-sense single-stranded RNA (+RNA) viruses that includes a large number of animal and human pathogens. Established in 1996 by the International Committee of Taxonomy of Viruses (ICTV), this order initially united two families of related viruses, namely the families of *Arteriviridae* and *Coronaviridae* (*1-3*). Over the years, mainly due to accelerated virus discovery driven by improvements in viral metagenomics, more virus taxa and species were added, resulting in an order currently containing 14 different families that include over a hundred virus species (*4*). A distinct nidovirus feature is embedded in the order's name, where "nido", originating from the Latin *nidus,* means "nest", and refers to the nested set of subgenomic (sg) mRNAs that nidoviruses produce during replication (*5*). These sg-mRNAs share identical 5' and/or 3' terminal sequences with the genomic (g) RNA and encode viral structural proteins and some of the so-called accessory proteins. The nidovirus gRNA is non-segmented and linear, and its length ranges from 12 to 41 kb (*6*).

The largest and best-studied nidovirus family is the *Coronaviridae*, which has become notorious for viruses causing potentially lethal diseases in humans. Amongst the human coronaviruses, four are endemic annually circulating viruses (namely the human coronaviruses (HCoVs) 229E, NL63, OC43 and HKU1) causing common colds that do not pose major health threats. During the last two decades, however, three previously unknown zoonotic coronaviurses able to infect humans emerged. These agents presumably were transmitted from (intermediate) animal hosts such as bats, civet cats or camels, and can cause potentially lethal respiratory disease in humans. First, there was the severe acute respiratory syndrome coronavirus (SARS-CoV) responsible for an outbreak in southeast Asia in the winter of 2002-2003 (*7*), which was successfully contained by strict monitoring and quarantine measures. Since 2012, and probably also earlier, smaller outbreaks of the Middle East respiratory syndrome coronavirus (MERS-CoV) have occurred in the Middle East (*8*). Clusters of MERS-CoV cases in humans mainly originate from transmission from dromedary camels, which can be followed by generally inefficient local spread, often in household or hospital settings. And finally, emerging late 2019 in Wuhan, China, SARS-CoV-2 is the causative agent of the on-going COVID-19 pandemic (*9*) that so far resulted in more than 6 million recorded fatalities worldwide (as of 15th of March 2022). SARS-CoV-2 and especially its variants that have further adapted to the human host (e.g. the delta (*10, 11*) and omicron variants (*12, 13*) are easily transmitted between humans and will likely become endemic in the near future (*14*). The COVID-19 pandemic and its immense socioeconomical impact resulted in a boost of worldwide efforts to develop antiviral therapies and vaccines. By now several vaccines have been registered, authorized (*15*), and administered worldwide to help reduce the spread and impact of SARS-CoV-2 infections. Additionally, the first antiviral drugs targeting human coronaviruses have recently been approved for the treatment of COVID-19 patients (*16, 17*), including also orally administered drugs (*18*). Certainly, zoonotic (corona)virus spillovers will continue to pose a constant threat to our health, society, and economy. Thus, our preparedness for zoonotic coronaviruses should be improved, including an enhanced understanding of their basic molecular biological properties, transmission and evolution. Together with the availability of improved vaccines and therapeutics, the detection and monitoring of potential progenitor viruses and early transmission events (*19*), this may enable us to rapidly contain or even prevent novel coronavirus outbreaks.

The second largest family in the *Nidovirales* order are the *Arteriviridae*, which consists of mammalian pathogens that, for example, infect horses, pigs, mice and monkeys. A

considerable cluster of arteriviruses infect non-human primates, however, so far, no cross-species transmission to humans has been reported. Some arteriviruses have a major impact on domestic animal breeding and food industry (e.g. porcine reproductive and respiratory syndrome virus [PRRSV] and alphaarterivirus equid, formerly equine arteritis virus [EAV], referred to as EAV for simplicity hereafter). Outbreaks of PRRSV that resulted in reproductive losses, increased pneumonia, and reduced pig growth have been recognized since the 1980s (*20*). The economic impact of PRRSV in the swine industry is significant, with yearly costs of several hundreds of million dollars estimated for the US alone in a study published in 2013 (*21*). Several vaccines against PRRSV are available, however, these can only partially reduce the disease severity against the circulating heterologous field strains (*22*). While EAV infection in horses was first recognized in the 1950s (*23*), an increase in worldwide travel and international horse riding events contributed to its increased economic impact especially during the last decades. Since the establishment of the carrier state in stallions used for breeding is a major driver for virus spread, preventive efforts are made to reduce this impact by the vaccination of breeding stallions and the organized testing of shipped horse semen (*24*).

Nidovirus replication cycle

The replicative cycle of nidoviruses can be divided in several steps: virus entry and genome release, genome translation, genome replication and sg-mRNA synthesis, abundant synthesis of viral nonstructural and structural proteins, and progeny virus assembly and release (see Fig. 1). Nidoviruses are enveloped and attach to host cell receptors exposed on the plasma membrane via binding of viral glycoproteins embedded in the viral envelope. In the case of coronaviruses, a spike glycoprotein trimer is responsible for virus particle attachment and membrane fusion (*25, 26*). Arterivirus particles, on the other hand, contain four different glycoproteins: GP2, GP3 and GP4 form a heterotrimer (minor GP-complex), while GP5 together with the unglycosylated membrane protein (M) forms GP5-M heterodimers (*27*). After attachment, arterivirus particles appear to enter the host cell by receptor-mediated endocytosis (*28, 29*). Also coronaviruses can rely on endocytosis and internal spike-mediated fusion, but virus particles may also fuse directly with the plasma membrane to release the viral genome into the cytoplasm. In fact, both entry pathways may be used by the same virus, as it is the case for coronaviruses like the SARS-CoVs (*30, 31*). For the SARS-CoV-2 omicron variant, a major shift from plasma membrane fusion (as primarily used by the delta variant) to internal fusion has been reported, and was postulated to be linked to the strikingly different pathogenic properties of this variant (*32, 33*).

Upon release of the nidoviral genome into the cytoplasm, the 5'-proximal two thirds of the genome, which contain the two large replicase open reading frames (ORFs) ORF1a and ORF1b, can directly be translated by host-cell ribosomes. ORF1a and ORF1b briefly overlap, while a -1 programmed ribosomal frameshift (RFS) at the end of ORF1a directs a fraction of the ribosomes into ORF1b (*34, 35*). Consequently, the full-length replicase polyprotein pp1ab will be expressed in lower quantities than polyprotein pp1a, which leads to a lower abundance of the proteins encoded in ORF1b. The two polyproteins are proteolytically processed into up to 16 so-called non-structural proteins (nsps) (*36, 37*). This is accomplished by a combination of viral proteases that are embedded in the ORF1a-encoded polyprotein. Most cleavage sites are processed by the picornavirus 3C-like or main protease (M^{pro}) that is conserved across all nidoviruses. In addition, one or two papain-like proteases (PL^{pro}s) exist in coronaviruses, while for arteriviruses up to four different PL^{pro}s can be present (*38*). Most nsps are involved in the

formation of a replication and transcription complex (RTC) (*39*), with its core component being the subunit containing the RNA-dependent RNA polymerase (RdRp), which is nsp9 or nsp12 in arteri- or coronaviruses, respectively. This enzymatic complex carries out all viral RNA synthesis steps, such as the synthesis of negative-sense RNA intermediates, genome replication, the transcription of sg-mRNAs, and the 5' capping of genome and sg-mRNAs (*40*). Nidovirus genome replication is associated with modified host cell membranes termed viral replication organelles (ROs), which will be (*41*). The nidovirus ROs will be discussed in further detail in the next section. For virion assembly, newly synthesized full-length gRNA is encapsidated by the cytosolic N protein and nucleocapsid then buds from membranes of the exocytotic pathway in which the viral envelope (glyco-)proteins have accumulated. For coronaviruses, the primary budding sites were found to be membranes of the ER-to-Golgi intermediate compartment (ERGIC) (*42*). The fully assembled progeny virions are then

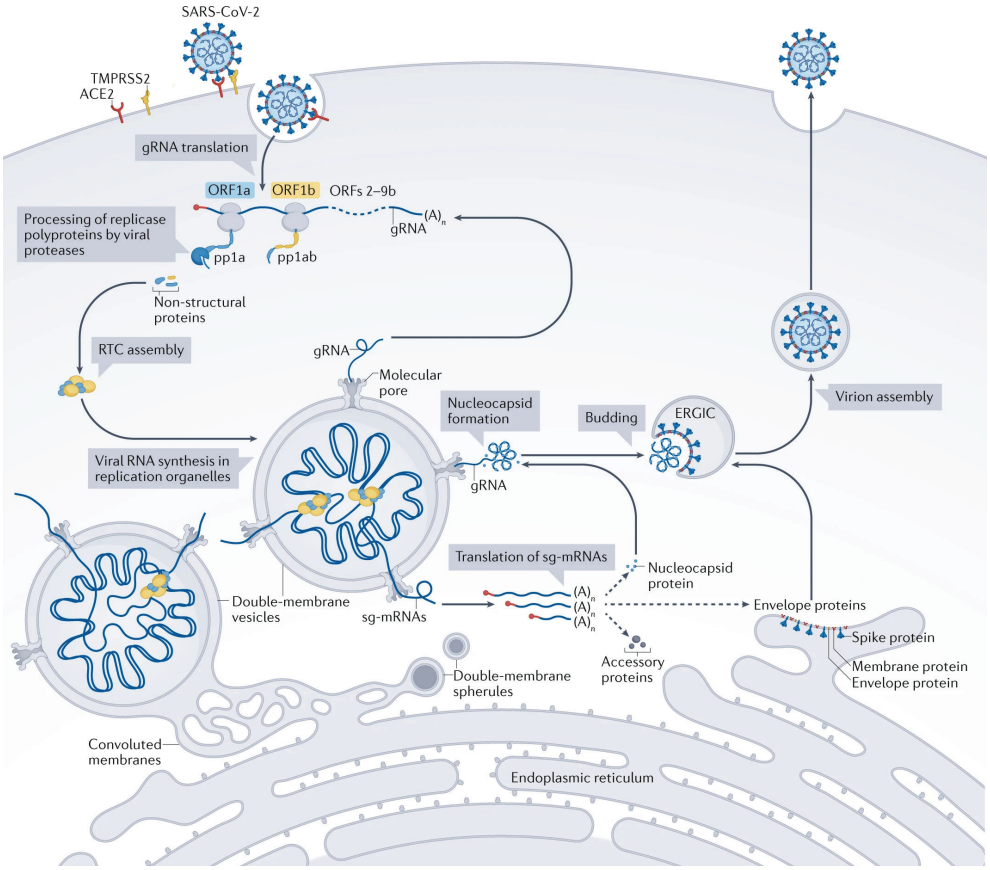

Figure 1. The nidovirus replication cycle exemplified by SARS-CoV-2 infection. Extracellular virus particles (top left) bind to host receptors in the plasma membrane to initiate cell entry. In the cell, the viral genome is released into the cytoplasm and translated by host-cell ribosomes into large polyproteins that are subsequently proteolytically cleaved by viral proteases into non-structural proteins (nsps). Some nsps assemble into replication and transcription complexes (RTCs), while others induce the formation of replication organelles dominated by double-membrane vesicles, at which viral RNA synthesis occurs. Newly synthesized viral RNAs comprise sub-genomic mRNAs (sg-mRNAs) that are translated into viral structural proteins and accessory proteins, and genomic RNA (gRNA) that can either be translated into nsps or encapsidated and packed into budding virus particles. Finally, fully assembled virions are released via exocytotic or lysosomal trafficking pathways. The figure was adapted with permission from (*40*).

secreted by exocytosis or might follow alternative routes to further spread infection, like the lysosomal trafficking pathway (*43*).

The nidovirus replication organelles

The spatiotemporal organization of the viral genome replication in nidoviruses is the focus of this thesis. All +RNA viruses of eukaryotes studied so far induce the remodeling of (different) host cell membranes to form ROs that support viral RNA synthesis. These ROs are thought to compartmentalize viral RNA replication to create micro-environments that concentrate viral and host factors relevant for efficient RNA synthesis, and to shield the process from host cell innate immune sensors. Based on their RO morphology, +RNA viruses can roughly be divided in two groups: (I) viruses inducing membrane invaginations called spherules and (II) viruses inducing more elaborate membrane structures, including double membrane vesicles (DMVs). ROs induced by nidoviruses belong to the DMV-inducing group, whose structural and functional aspects are reviewed in greater detail in **Chapter 2**. In the case of nidoviruses, three nsps in particular are assumed to be involved in RO formation, namely nsp3, nsp4 and nsp6 in coronaviruses (*44-46*), which serve similar functions as the corresponding nsp2, nsp3 and nsp5 in arteriviruses (*47, 48*). These three nsps contain trans-membrane domains (TMD) and are thought to hijack ER membranes for RO formation (*49-51*). While DMVs are the predominant feature of nidoviral ROs, additional membranous elements can be induced, especially by coronaviruses, such as zippered ER, convoluted membranes, or double-membrane spherules (see also Fig. 2) (*52-55*). About 15 years ago, electron tomography (ET) started to be applied in the field, which significantly improved our understanding of the three-dimensional (3D) organization of viral ROs. It became apparent that nidovirus-induced DMVs as well as other RO elements are part of a large reticulovesicular network of interconnected (double-) membrane structures (*48, 54-59*). These studies also characterized DMVs as fully sealed compartments, as they lacked the membrane openings towards the cytosol that are an integral feature of the spherular type of ROs. Interestingly, newly synthesized viral RNA was found to be associated with DMVs (*60, 61*), but not with other coronavirus-induced RO structures (*61*). Immunoelectron microscopy (IEM) provided evidence that viral RNA synthesis could occur inside DMVs, since the DMV lumen strongly labelled for double-stranded RNA (dsRNA) (*55, 56, 62*), a viral RNA synthesis intermediate not found in uninfected cells. However, these findings posed a topological conundrum as viral +RNA synthesized inside the DMVs would have to reach the cytoplasm for either translation into viral proteins or packaging into progeny virus particles. Thus, if viral RNA synthesis would indeed occur within sealed DMVs, the transfer of newly made viral RNA across the double-membranes would appear to be not straightforward. One scenario would be that the primary synthesis of full-length viral genomes and sg-mRNAs occurs prior to or during DMV formation, a process difficult to study possibly due to its rapidness. In this case, sealed DMVs would mainly serve as hiding places for excess replication intermediates such as dsRNA. Alternatively, DMVs could contain molecular complexes with a channel that would facilitate the exchange of metabolites and RNA strands between the DMV lumen and the cytosol. The detection of such complexes might not be possible by conventional EM methods, like those used in the studies described so far. For such EM methods, the samples are fixed, heavy-metal stained, resin embedded, and sectioned for analysis in a transmission EM (TEM). The resulting images have excellent contrast and are well suited for the analysis of membranous structures like cell organelles or viral ROs. However, such sample preparation protocols can also introduce artefacts (*63, 64*) and the biological content is visualized only indirectly by the

heavy metal staining accumulating at certain cellular components. The directly visualization of molecular complexes has become possible with the emergence of cryo-EM, a technique imaging only native biological samples in a frozen-hydrated state (*65*). For this method, biological material is rapidly frozen (vitrified) in its aqueous environment and can directly be analyzed at macromolecular resolution. The main goal of this thesis was to study the macromolecular elements of nidoviral ROs by cellular cryo-ET, a newly developed method allowing the visualization of the cellular ultrastructure in its native state (*66*).

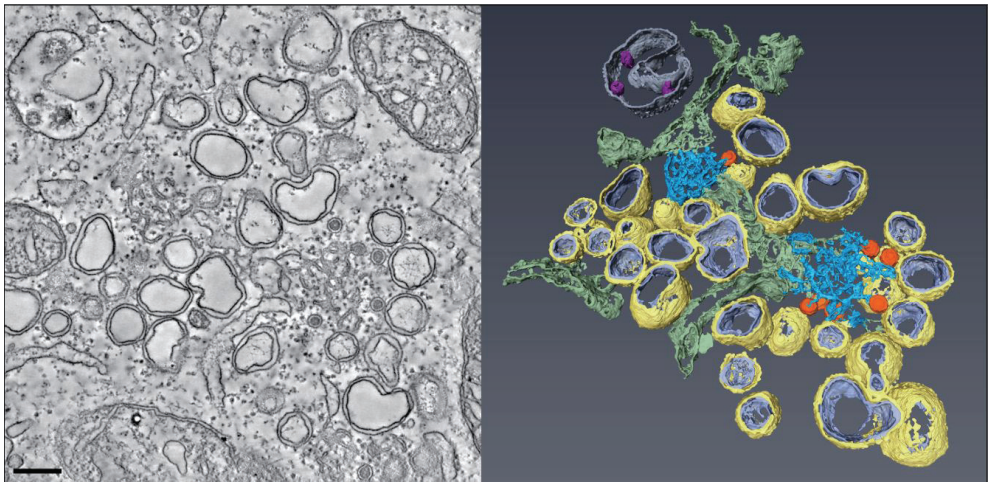

Figure 2. 3D morphology of nidovirus ROs exemplified by tomographic data of MERS-CoV-induced replication organelles (ROs). A tomographic slice (left) and the corresponding surface-rendered model (right) of MERS-CoV-induced ROs in Huh7 cells. Different types of double-membrane modifications found in ROs are depicted, such as convoluted membranes (blue), double-membrane spherules (orange), and double-membrane vesicles (yellow and purple, outer and inner membranes, respectively). Next to the RO elements, also ER membranes (green) and a vesicle (silver) that contains virus particles (dark pink) are shown. Scalebar 200 nm. The figure was adapted from (*61*).

Cellular cryo-tomography

Electron microscopy has accompanied virus research ever since it empowered the visualization of the first virus particles in the 1930's (*67*). In a transmission electron microscope (TEM) an electron beam is transmitted through samples that are not thicker than a few hundred nanometers to produce a 2D projection image that can be acquired by a detector. With the development of cryo-EM in the 1980's, it became possible to directly visualize molecular structures in the absence of heavy-metal staining. Technical developments, such as automation, dedicated image-processing software and, most critically, direct electron detectors were critical to the extraordinary success and fast growth in recent years of cryo-EM, which is expected to become the leading method in the field of structural biology (*68*). Cryo-EM samples are usually plunge-frozen in a liquid cryogen, a process so rapid that the water transforms into a solid amorphous (i.e. vitreous) state without the formation of ice crystals (*65*). In this manner, macromolecules can retain their native state. However, the lack of stain together with the strict low electron dose-imaging conditions that are required to prevent electron beam-induced radiation damage results in low-contrast images. Consequently, the resulting data does not allow the direct interpretation of high-resolution details. For this step, single particle analysis (SPA), the approach that now yields the overwhelming majority of published cryo-EM structures, can be employed to increase the signal-to-noise ratio (SNR)

of the molecule of interest. In this case, large numbers of images of a sample containing the purified macromolecule are acquired, resulting in thousands to millions of views of the (in principle) same structure in different orientations. These views can then be computationally combined in high-resolution 3D models that nowadays routinely achieve near-atomic resolution for sufficiently homogenous samples (*69*).

Electron tomography (ET) is an alternative approach to obtain 3D information on EM samples. By gradually tilting the specimen inside the TEM and taking projection images over a large tilt range, one can generate a tilt series containing different views of the specimen, which can then be computationally combined into a 3D volume (for a review see (*70*)). A major advantage of this method is that it can readily provide valuable 3D information of the imaged samples, as already demonstrated for conventional resin-embedded EM samples of +RNA virus-induced ROs (see Fig. 2, **Chapter 2** and **Chapter 6**). This method can also be applied to frozen-hydrated samples; however, the inherent noisiness of cryo-EM data also limits the interpretation of structural details in cryo-ET. Nevertheless, in case a certain structure is repeatedly found in the cryo-tomograms, these individual reconstructions can be cropped out, aligned in 3D and averaged. This approach, known as subtomogram averaging (STA), follows a similar principle as SPA and can be employed to increase the final resolution of the structure of the macromolecule of interest (*71*). STA can nowadays achieve sub-nanometer resolution, providing structural detail that allows the assignment of individual protein domains (*72-74*). Ongoing methodological developments slowly push the limits of resolved molecular detail by STA towards similar near-atomic resolution ranges as achieved with SPA (*75-77*).

Cryo-ET can be easily applied to thin samples, like purified proteins or virus particles. However, the imaging of cellular content is not straightforward, as most cells by far exceed the maximum sample thickness of ~300 nm suitable for ET. To overcome this, a method based on cryo-ultramicrotomy, often referred to as CEMOVIS, was originally developed (*78, 79*). Frozen-hydrated specimens are sectioned with a diamond-knife while being kept under cryogenic conditions, resulting in thin cellular frozen-hydrated slices that can then be imaged in a TEM. Unfortunately, this process not only is extremely delicate and requires extensive experience and skills, but the mechanical sectioning also introduces artefacts in the sample (e.g. compression) that can introduce significant distortions (*80*).

During the past decade, the method of focused ion beam (FIB)-milling has emerged for the preparation of thin slices from vitreous cellular samples in an artefact-free manner (*66, 81*). Here, the intracellular content is exposed, enabling cryo-ET studies of native macromolecules within their intact cellular environment. In the course of the cellular cryo-ET workflow (see also Fig. 3), the vitrified cells are transferred to a scanning electron microscope (SEM) equipped with a FIB and a cryo-stage that is able to maintain the sample at cryogenic temperatures. Within the cryo-FIB-SEM device, the FIB is aimed on the vitrified cells at a shallow angle and removes (biological) material above and below a predefined region of interest. In this manner, a 100 to 300 nm thin cryo-lamella is generated that contains the intracellular features of interest accessible for high-resolution cellular cryo-ET (*82*). The milling of fragile cryo-lamellae is a time-consuming process that requires careful attention and experience. Automation (of parts) of this process has been developed recently to reduce the effective user time (*83, 84*). Nonetheless, the fabrication of the thinned lamellae that allow the acquisition of high-quality cryo-ET data remains difficult and the yields of samples of sufficient quality are low. In **Chapter 3** of this thesis, an optimization of this workflow that allowed us to more reproducibly generate high-quality cryo-lamellae is outlined.

Successful application of the cellular cryo-ET workflow can reveal unprecedented details of macromolecular complexes. Therefore, this methodology bridges the gap between the fields of cell and structural biology, enabling the annotation of functional macromolecules directly within their cellular context. An early groundbreaking study using this approach revealed the spatial organization of the cytoplasmic translation machinery as well as the dynamics of nuclear pore complexes embedded in the nuclear envelope of HeLa cells (*86*). Subsequently, various studies provided novel insights into intracellular processes, such as the gradient organization of ribosome biogenesis within nucleoli (*87*), the molecular basis of sarcomeres in skeletal muscles (*88*), the native structure of the COPI coat (*89*), the dilation and constriction of nuclear pore complexes (*90*), or how HIV capsids can transit these to reach the nucleus of an infected cell (*91*).

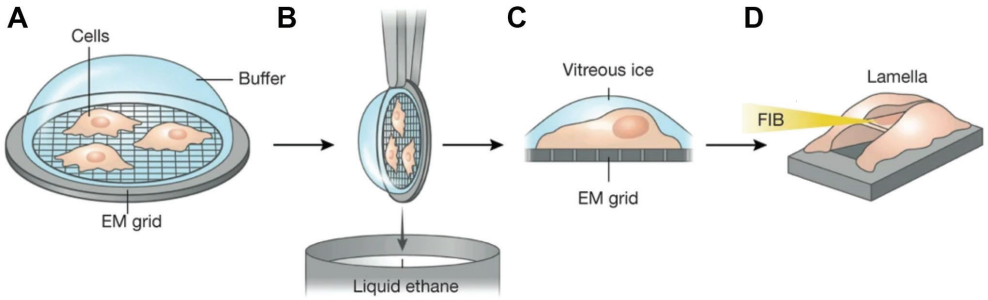

Figure 3. Schematic outline of the sample preparation workflow for cellular cryo-ET specimen. (A) Cells maintained in culture medium, are grown or pipetted on EM grids and transferred to a plunge-freezing device. After the removal of excess of liquid from the grids (B) they are immediately plunge-frozen in a liquid cryogen, such as liquid ethane. (C) Subsequently, the vitrified samples are kept under cryogenic temperatures and transferred to a cryo-FIB-SEM, where (D) a region of interest is selected and FIB-milling tangentially removes material above and below that region. The resulting cryo-lamella is thinner than 300 nm, which allows high-resolution imaging in a TEM. The figure was adapted with permission from (*85*).

Thesis outline

The focus of this thesis is the structure and function of coronavirus- and arterivirus-induced ROs, as observed in their native state by means of cellular cryo-ET. Novel structural details unveiled in this work provide strong evidence for a defined pathway of newly synthesized viral RNA, from the DMV interior to the cytosol. **Chapter 2** reviews structural and functional aspects of DMVs and related RO structures in the replication of +RNA viruses. The different (potential) paths of DMV biogenesis, the involvement of viral factors as well as host proteins and lipids are discussed, in addition to the central role of DMVs in viral RNA synthesis. **Chapter 3** describes some of the technical hurdles that cellular cryo-ET and especially the FIB-milling of cryo-lamellae can pose, and how an improved protocol was designed to increase the throughput of this method when applied to vitrified (infected) eukaryotic cells. **Chapter 4** provides the first insights into the native structure of coronavirus-induced ROs. This led to the discovery of a hitherto unknown and unique molecular pore spanning the two membranes of the DMVs that are induced in cells infected with either murine hepatitis virus (MHV), a model mouse coronavirus, or SARS-CoV-2. This hexameric complex contains a central channel wide enough to allow the transit of RNA strands and thus suggests a path for viral RNA travelling from the DMV lumen to the cytosol. Furthermore, we found that the core of the pore complex is formed by six copies of coronavirus nsp3, a very large replicase subunit of about 2000 amino acids. Our results supported a model in which the DMV pore

complex could coordinate viral RNA synthesis with RNA export and encapsidation. In **Chapter 5,** cellular cryo-ET of arterivirus-induced ROs extends the existence of DMV molecular pores across the order *Nidovirales*, highlighting its conserved nature and (presumed) functional importance. The significantly smaller pore complexes detected in PRRSV- and EAV-induced DMVs show associations with distinct nucleocapsid structures similar to those found in complete virus particles, providing further evidence for a vRNA pathway from the pore to virus particle assembly. Additionally, EAV-induced molecular pores in DMVs were found to also assemble outside the context of the infected cell, upon DMV formation induced by the expression of only two TMD-containing nsps. **Chapter 6** discusses technical developments in EM that contributed to our enhanced understanding of viral ROs. A broad spectrum of EM techniques applied to +RNA virus-infected cells and the resulting insights are reviewed, concluding with a description of the obvious differences and striking similarities between the protein complexes found in the two morphologically distinct +RNA viral RO types, namely DMVs and invaginated spherules. In **Chapter 7** the broader context of pore-like complexes in viral ROs and beyond is discussed on the basis of the previous chapters and recent literature. Prospects of emerging schemes in the molecular organization of nidovirus and other +RNA induced ROs are sketched and it is outlined how novel EM tools could facilitate their further characterization.