Cellular cryo-tomography of nidovirus replication organelles

Wolff, G.

CHAPTER 7

General discussion

In recent years, structural biology research has changed drastically, with electron cryo-microscopy (cryo-EM) rapidly replacing the so far dominating x-ray crystallography (*68*). Nowadays, cryo-EM is able to routinely resolve molecular structures at near-atomic resolution, which can be done directly on purified, rapidly-frozen protein samples on EM grids, without the need for crystallization. This approach is enabled by single-particle analysis workflows, in which projection images of millions of copies of the same macromolecule in different orientations are acquired and subsequently computationally combined into a high-resolution 3D model (*69*). Also in the course of the COVID-19 pandemic, cryo-EM contributed important insights into the structure and function of many coronavirus proteins, shedding new light on their functions and interactions, and providing a structural basis for the developments of antiviral strategies (*308*). A prominent example is the spike protein that enables the virus to bind and enter the host cell and is a primary target for vaccine development (*309-313*). Likewise, structural studies have focused on the replication and transcription complex (RTC), the key enzymatic complex that drives coronavirus RNA synthesis (*249, 251, 252, 314, 315*). This complex is highly conserved across the coronavirus family, and thus constitutes an attractive target for antiviral drug development. The success story of cryo-EM can be primarily attributed to the developments of highly-sensitive direct electron detectors and sophisticated software that enables the extraction of high-resolution details from cryo-EM data. These technological advances also benefited the relatively small cryo-EM branch of electron cryo-tomography (cryo-ET), a method allowing the 3D analysis of larger macromolecular structures, purified organelles or cells. Cryo-EM data is inherently noisy, which is reflected in the generally limited resolution of macromolecular details in cryo-tomograms. Nonetheless, higher-resolution information can still be obtained from the data by the so-called sub-tomogram averaging approach (*71*), for which multiple subvolumes of the same macromolecule are boxed-out, aligned in 3D and averaged to increase the signal-to-noise ratio. Although the first 3D density maps obtained in this way were usually in the nanometer resolution range, more recent examples using highly sophisticated processing software prove that near-atomic resolution can be approached (*77, 316*). Cryo-ET studies have also contributed valuable insights in the molecular biology of SARS-CoV-2, e.g. by enabling the analysis of the spike on intact virus particles (*292-294*).

The major promise for the future of cryo-ET lies in its ability to determine high-resolution macromolecular structures *in situ,* i.e. without extracting them from their natural context. For the reductionist approach of the dominating structural biology methods of x-ray crystallography, nuclear magnetic resonance and single-particle cryo-EM, it is required to take cells apart and purify the component of interest in a highly homogeneous sample. However, most cellular functions are rooted in complex molecular networks rather than isolated molecular processes carried out by only one or a few actors. Purification disrupts the structural integrity of such molecular networks and weak or transient interactions are frequently lost. In contrast, cryo-ET provides a post-reductionist approach, enabling the structural analysis of macromolecules within their functional modules, and/or in their native cellular environment. This can be achieved in a relatively straightforward manner for macromolecules embedded in cellular organelles or small(er) membrane structures, such as enveloped viruses, which can be extracted and enriched for high-throughput analyses. However, sufficient electron transparency of samples, allowing their imaging by conventional transmission EM, requires thicknesses below ~300 nm, a limit that many cellular structures exceed by far. The development of focused ion beam (FIB) milling of cryogenic biological samples provides a solution that also makes the cytosolic space of larger (eukaryotic) cells

accessible for cryo-ET analysis (*66, 81*). The 100-300 nm thin cryo-lamellae generated in this manner contain unperturbed intracellular material that can be analysed directly by cryo-ET. However, this workflow is technically demanding and has a low throughput, which still limits its applicability to most biological research questions. **Chapter 3** of this thesis describes an optimized FIB milling workflow that delivers a higher yield of high-quality sample preparations. This and other very recent technical improvements (*317, 318*) are beginning to elevate the cellular cryo-ET method from a niche application to a tool accessible to a much wider research community and applicable to a larger variety of biological questions.

With the emerge of cellular cryo-ET, structural biology has entered a new era of visual proteomics (*250*). This field connects the microscopic analysis of intact cells to the structural details underlying the biochemistry of molecular interactions. We are now able to obtain holistic pictures of unperturbed molecular networks directly within the cell, which promises ground-breaking discoveries, but also poses new challenges. Cellular cryo-ET data contains a sheer wealth of information whose interpretation, however, is not straightforward. Translating the data into biologically interpretable structural information requires tailored approaches encompassing a variety of data processing steps, of which a central component is subtomogram averaging. This thesis describes the application of these revolutionary methodologies to the analysis of nidovirus replication organelles (ROs), revealing – among other features – a molecular pore complex that appears to play a crucial role in the viral replication cycle. The results presented in this thesis constitute only a first landmark in the new territory of the study of viral ROs by cellular cryo-ET, and there is no doubt that many will be added in the future.

Nidovirus replication organelles and the DMV pores

Studies of (nido)viral ROs started with their discovery in the 1960s, when the development of methods to observe thin cellular specimens by electron microscopy first allowed their direct visualization (*264, 319*). Although the connection to viral RNA synthesis was not made in those early studies, the presence of double-membrane structures comprising DMVs was found to be a characteristic feature of nidovirus-infected cells. Only around the turn of the millennium, it became clear that viral RNA synthesis might be associated with ROs, as these structures were found to systematically label for nonstructural proteins (nsps) and viral RNA, not only for nidoviruses (*51, 60*), but also for some other +RNA viruses (*274, 277, 278, 285*). The visualization of nidoviral ROs in 3D by electron tomography provided important details about the extensive membranous networks that interconnect DMVs, other virus-induced structures and the ER (*54-56, 58*). However, the improved understanding of their composition, morphology and origin raised several new questions about nidovirus RO function. A recent coronavirus study demonstrated that DMVs are the primary, if not only, site of viral RNA-synthesizing activity (*61*), as also suggested previously by the strong labelling of nidovirus-induced DMV lumen for double-stranded RNA (*55, 56, 62*). The discovery of molecular pores that span the double membranes of coronavirus-induced DMVs, described in **Chapter 4**, provides a likely pathway for viral RNA strands from the DMV lumen to the cytosol. Using an engineered recombinant virus, the largest transmembrane domain (TMD)-containing coronavirus nsp was found to be part of this DMV pore complex. Molecular pores in the ROs induced by the only distantly related arteriviruses were also detected, as described in **Chapter 5**, where it was also demonstrated that the expression of only two TMD-containing nsps is sufficient for pore formation. These findings and the nidovirus-wide conservation of three TMD-containing replicase subunits highlight their importance for viral replication (*37,*

119). Strikingly, DMVs embody several characteristics that are shared with the nucleus of eukaryotic cells: a double-membrane envelope, a polynucleotide content which is replicated and transcribed and proteinaceous pores that may selectively control the transport of macromolecules, such as RNA strands, across the two membranes. The spatial separation of transcription and translation by membranous structures thus appears as an advantageous strategy shared by eukaryotes and some viruses. Interestingly, this communality gave rise to the "viral eukaryogenesis hypothesis", which proposes that the nucleus would have originated from viruses contributing to the emergence of the eukaryotic 'super-organism', which was able to escape the limitations of prokaryotic evolution (*320-322*).

Molecular aspects of nidovirus DMV pores and comparison with other viruses

The involvement of the three TMD-containing nsps of nidoviruses in RO formation and organization is undisputed. As demonstrated in **Chapter 4**, we now know that the coronavirus TM1-containing protein (nsp3) is the core component of the molecular pore, where it forms a hexameric complex that spans the double membrane and for a major part protrudes into the cytosol. Furthermore, it was established that the N-terminal ubiquitin-like domain 1 of coronavirus nsp3 is at the top of the cytosol-facing crown structure of the DMV pore. Unfortunately, our current results do not provide direct evidence for the protein composition of the arterivirus DMV pore complex. Subtomogram averaging would provide 3D models allowing structural comparisons that may shed light into this aspect; however, this approach was unsuccessful so far, maybe reflecting some intrinsic structural heterogeneity of the complex. Nonetheless, the strong conservation of the nidovirus core replicase, including its (predicted) three TMDs (*37*), suggests that the TM1-containing subunit could play a similar role in arteriviruses and coronaviruses. This hypothesis is supported by our finding that pore complexes are present in the DMVs induced by expression of only EAV nsp2-3 (i.e. TM1 and TM2, see **Chapter 5**). The large molecular mass of the coronavirus pore suggests the presence of additional factors contributing to the complex. A likely candidate is the TM2-containing protein (nsp4 in coronaviruses, nsp3 in arteriviruses), a close interaction partner of the TM1-containing nsp and a crucial player in RO biogenesis (*44, 45, 47, 116, 125*). Although our results show that the TM3-containing subunit of arteriviruses (nsp5) is not required for the formation of the basic pore structure, it cannot be excluded that nsp5 might be recruited to the pores formed during infection. Other (host) molecular players that are (potentially) involved in nidovirus DMV formation are discussed in **Chapter 2.** Some of these factors may also contribute to pore formation, even if only indirectly, since the biogenesis of DMV-forming intermediates, such as wrapping zippered ER and the DMVs themselves, arguably is required for pore formation to occur. Similar to pore formation, the RO biogenesis is still poorly understood at the molecular level. It is conceivable, that both processes could be closely connected and mutually dependent. For instance, it was found for coronavirus DMV formation that the cleavage between nsp3 and nsp4 is essential (*116*), while the findings in **Chapter 4** suggested that nsp3 might be primarily present in the outer membrane of the DMVs, as densities resembling the large N-terminal domain of nsp3 appeared to be absent on the inner membranes. It is tempting to speculate that the TMD-containing nsps could segregate between the two DMV membranes, and that a tight regulation of the proteolytic processing of the nsps could play a crucial role in this process. The selective enwrapping of (parts of) the downstream nsps that include all RTC components could be ensured by keeping a fraction of them bound to a TM2- and/or TM3-including polyprotein recruited to the inner

membrane, which would only be proteolytically processed upon DMV closure. Additionally, the large number of functional domains present on coronavirus nsp3 alone (*199*) makes it likely that this and other nsps are involved in many, spatially separated, processes that might require different subpopulations of the same proteins.

RNA channels are not unprecedented in the realm of RNA viruses. Double-stranded RNA (dsRNA) viruses such as reoviruses are made of multi-layered icosahedral protein capsids containing viral RNA-dependent RNA polymerases (RdRp) as well as their genomic dsRNA. Since dsRNA is a strong trigger of host cell innate immune responses (*323*), most viruses seem to have developed strategies to avoid its detection. In the case of reoviruses infection, the virus particle's inner protein shell remains intact to shield the dsRNA genome and the internally occurring transcription process. The RdRps associate with RNA-exit pores present in the protein shell to guide newly-synthesized RNA strands into the cytoplasm, where they can be translated into viral proteins by host cell ribosomes or packaged into progeny virions (*324*). In the case of positive-sense RNA (+RNA) viruses, the virus genome mimics host cell mRNAs and does not require immediate shielding upon infection, but is directly released into the cytoplasm and translated into the first viral proteins. Subsequently, during genome replication and (in the case of e.g. nidoviruses) subgenomic (sg) mRNA synthesis, negative- and double-stranded RNA intermediates are generated. Their detection by innate immune sensors is thought to be prevented or delayed by their containment in virus-induced ROs (*41*). Some +RNA viruses induce ROs consisting of invaginated spherules in membranes of cellular organelles, such as ER, mitochondria or peroxisomes (*269, 270, 325, 326*) (see also **Chapter 5**). In the case of nodaviruses, these structures feature a crown-shaped molecular complex at the neck-like connection to the cytosol, which has been suggested to coordinate the transport of viral RNAs and metabolites (*172*). The nodavirus crown complex is a dodecameric assembly of 12 units of the viral protein A, a 112-kDa multidomain protein that also contains the viral RdRp (*244*). A similar complex was discovered for alphaviruses, members of the Togavirus family, in which 12 units of the non-structural protein 1 (nsP1) form a pore-complex (*245, 246*). Also here, this complex was suggested to serve as a gate at the neck of the spherular membrane invaginations that all togaviruses studied so far induce in the plasma membrane and in endosomal/lysosomal membranes (*277, 279*). The group of +RNA viruses that induces DMV-containing ROs extends beyond the *Nidovirales* order, as also hepatitis C virus-, norovirus- and enterovirus-induced ROs comprise DMVs. However, for the latter family it was recently shown that the DMVs may not contain pore-like structures in their membranes (*327*). It appears that enteroviral replication might be associated with the cytosolic side of the ROs, which comprise both single-membrane vesicles and so-called autophagy-like membranes often resembling DMVs with membrane openings. Interestingly, viral capsids seem to assemble on RO membranes, suggesting that viral RNA synthesis and encapsidation might be closely associated. Thus, it seems plausible that enteroviruses, similar to dsRNA viruses, utilize their viral capsid protein and nucleocapsid structures to shield their RNA replication from innate immune sensors, while nidoviruses and spherule-inducing viruses rely on modified host membranes for that purpose (*328*).

Viral RNA synthesis inside nidovirus ROs: where is the active RTC?

Despite the recent ultrastructural studies resulting in ground-breaking insights into the molecular architecture of +RNA virus ROs, many aspects of their formation, organization and functions remain elusive. Probably most intriguing are the whereabouts of the active viral RdRp complex and the way it integrates in the molecular playground of the viral ROs.

Clearly, it still remains to be proven that nidoviral genome replication is actually happening in the DMV lumen, as hypothesized earlier. Certainly, this scenario would offer major advantages. Replication intermediates would be hidden from cellular innate immune sensors and the replication process could be organized efficiently in a dedicated microenvironment. These features have been reasoned to be the fundamental benefits of inducing elaborate membranous RO structures. This hypothesis is supported by the finding that RTC activity is retained in purified ROs of both arteri- and coronaviruses, but lost upon detergent treatment, suggesting it is protected by membrane compartments (*163, 231*). A similar conclusion was drawn by a study combining immunofluorescence microscopy with the use of mild detergents only disrupting the plasma membrane, but not intracellular membrane compartments, which would allow the distinction between membrane-protected and cytosolic factors (*329*). However, the latter study also showed that the RdRp signal is largely dispersed in the cytosol rather than confined to putative ROs, especially at later timepoints in infection (*329*). It was hypothesized that many individual RdRp subunits are not incorporated into active RTCs and may be distributed across the cytoplasm, while active RdRps within the RTC complex might not be labelled, since the recognition of their epitopes might be blocked due to the formation of complexes with other proposed RTC components (such as nsp7, 8, 9, 13, 14, 16 (*249*)). Furthermore, as in the case of other viral nonstructural proteins, also the nidovirus RdRp is likely to be involved in other interactions not related to viral RNA synthesis (*110*). This poses a general problem in the search for active RTCs, since, for instance, antibody labelling does usually not distinguish between enzymatically active and inactive proteins. It is, however, possible, to metabolically label newly synthesized RNA and restrict this labelling to viral RNA by simultaneously inhibiting host-cellular DNA-dependent RNA polymerases. In this manner, any detected signal would imply the close proximity of active viral RTCs. Indeed, in such studies viral RNA synthesis could be detected at DMV regions in coronavirus-infected cells (*60*), while recently it was even demonstrated that DMVs are indeed the main, if not only, site of viral RNA synthesis (*61*). However, the resolution of all these labelling techniques is limited and does not allow to establish whether the signal originates inside or outside the DMV.

The cryo-ET data sets described in **Chapters 4** and **5** did not provide information about the localization of the active RTC either, as no distinct or recurring densities could be pinpointed in the DMV lumen. This raises the question of whether it is technically possible to detect putative RTC complexes in cryo-ET data in the first place. Cryo-tomograms inherently have attenuated high-resolution signals, which are especially important for the recognition of small molecular complexes directly in the tomographic data. Thus, the identification of macromolecules becomes more challenging the smaller they are, with the limit of detection currently estimated to be at around 400 kDa (*250*). This limit might be even higher for cellular cryo-ET, where the extremely crowded cytosolic space also complicates the visual distinction of individual macromolecules. For both coronaviruses and arteriviruses, the composition of the native RTC and the stoichiometry of its various (potential) subunits remain open questions. In particular the SARS-CoV-2 RdRp complex has been studied extensively, both biochemically and structurally (*40*), but all these studies are based on the use of recombinant proteins, produced in bacterial expression systems, that are allowed to assemble into protein complexes *in vitro*. The coronavirus RTC could be speculated to be quite small, e.g. an assembly of a single copy of the nsp12 RNA polymerase and nsp13 helicase in association with a few smaller subunits (like nsps 7, 8, and 9), leading to a molecular mass that may be well below 400 kDa (*249*). However, a recent study based on molecular modelling proposed much larger complexes containing six or more copies of many coronavirus nsps and having

a predicted molecular mass of more than 2 MDa (*330*). Further experimental evidence is required to validate this model and its relevance in coronavirus replication.

A compounding factor for the identification of the RTCs in cryo-tomograms is that the composition of the complex *in situ* may be different for different viral activities, like genome replication, subgenomic RNA synthesis, and mRNA capping (*40*). It is conceivable that (parts of) the RTC would at least temporarily associate to the molecular pores, ensuring the directed or even selective transport of newly-synthesized RNA strands towards the cytosol. Interestingly, small molecular densities with variable appearance were frequently found associated with the luminal side of the coronavirus DMV pores. This is also visible in the subtomogram averaging results described for the coronavirus molecular pore in **Chapter 4**, where a poorly resolved off-centred density associated with the pore near the luminal entrance of the central channel.

Encapsidation of viral RNA is likely associated with DMV pores

Both in coronavirus- and arterivirus-induced ROs, RNP(-like) structures have been found associated with or in close proximity of the cytosolic side of DMV pores, suggesting the encapsidation of viral RNA after its export. Direct molecular connections between (putative) RNP structures and DMV pores in both coronavirus- and arterivirus-induced DMVs provide further support for an involvement of the pores in RNA encapsidation. In the case of PRRSV, RNPs formed at DMV pores appear to immediately form spherical nucleocapsid shells that are morphologically identical to nucleocapsids inside virus particles. Similar observations were made in EAV-infected cells, where not only nucleocapsids but also additional RNP-related structures that so far remain ambiguous and will be discussed in more detail below, were found to be associated to pores (**Chapter 5**). For coronaviruses, the visual recognition of cytosolic RNPs was complicated by the apparent heterogeneity of these structures in the cytosol. Nonetheless, sub-tomogram averages of the pore showed diffuse densities adjacent to the top part of the cytosolic crown, which suggest frequent interactions with other molecular players (see Fig. 2C in **Chapter 4**). At the molecular level, it was demonstrated that the most N-terminal domain of coronavirus nsp3, i.e. the ubiquitin-like domain 1 (Ubl1) interacts both with N-protein (*202, 203, 331*) and single-stranded RNA (*201*).Recently, studies on SARS-CoV-2 found that Ubl1 is recruited to droplets formed by N *in vitro* (*332*), and a nuclear magnetic resonance-based study reported the transformation of N protein into a more compact conformation upon Ubl1 binding (*333*). It was hypothesized that Ubl1 could serve as a chaperone to transform N into a conformational state that is more amenable to RNA binding, similar to paramyxovirus phosphoproteins (*334*). Connections between viral RNA replication and genome encapsidation have also been found for other +RNA viruses. In the case of flaviviruses, which induce membrane invaginations similar to nodavirus ROs, budding virus particles have been observed in the direct vicinity of the membrane openings of replication spherules (*270*), suggesting virus genome packing immediately after synthesis. For enteroviruses, viral genome replication appears to be directly coupled with encapsidation by the interaction of structural proteins with the RNA replication machinery (*335*). This is in line with recent observations made in a cellular cryo-ET study (*327*), where apparent intermediates in nucleocapsid formation were observed in association with viral ROs and speculated to protect the membrane-bound replication complex. In summary, the DMV pore may be a central hub in nidovirus replication that could both functionally and spatially connect RNA synthesis inside the DMV lumen to genome encapsidation at the cytosolic face of the pore.

An interesting point of discussion in this regard is the path of sg-mRNAs, which constitute more than 90% of all viral RNA transcripts in the nidovirus-infected cell and are destined to be translated rather than packaged. Possibly also sg-mRNAs are exported by DMV pores and then diffuse into the cytoplasm until encountering a ribosome. In this regard, it would make sense if ribosomes would associate to molecular pores as well to ensure the efficient translation of sg mRNAs. Earlier studies on resin-embedded samples indeed suggested the association of ribosomes to DMV membranes (*55, 56*). However, in cellular cryo-ET data ribosomes associating to DMV membranes have not been observed so far, suggesting that previous observations may have arisen from sample preparation artifacts. Multiple factors could be involved in determining the fate of exported gRNAs and sg-mRNAs. These could include RNA elements such as gRNA-specific encapsidation signals, variable conformational states of pore complexes co-determining the type of RNA that is exported, and the concentration of cytosolic N protein. These unexplored features remain very interesting topics for future studies of nidovirus replication and DMV pore function.

Additional RNP-related structures present in the proximity of EAV-induced ROs thus far defy a straightforward structural or functional interpretation. Next to the roughly spherical nucleocapsids, RO areas frequently also contained proteinaceous sheets that can fold into tubular structures. Similar tubular sheets have been described earlier (*56*), reportedly labelled for N protein (*56, 242*), and have been found to be absent when N protein expression is knocked out during infection or in the case only N protein is overexpressed (*238*). This suggests that these represent (intermediate) ribonucleocapsid (RNP) elements requiring both N protein and viral RNA to be formed, which also aligns with their absence in EAV nsp2-3 and nsp2-7-expressing cells ((*48*), **Chapter 5**). The striking resemblance in the macromolecular patterns of the sheets and nucleocapsids strongly supports the notion that these structures are either intermediates or aberrant by-products of EAV nucleocapsid formation. Comparing the features of these sheets with a nucleocapsid of ~42 nm diameter, it appears plausible that they could be assemblies of multiple RNP elements, possibly containing several dozens of individual (viral) RNAs. A third proteinaceous structure present in regions containing EAV-induced ROs consisted of rods of about 15 nm. Since these were found to be directly associated with DMV pores, but also interconnected pores and RNP sheets or even different patches of RNP-sheets, they might represent a different, less condensed form of the same RNP. Notably, the rods consist of a relatively dense shell that appears to shield a central filament that might be RNA. The function of these structures thus far remains elusive, just like the answer to the question why they appear to be lacking in PRRSV-infected cells.

Outlook on methodological developments that could facilitate the elucidation of nidoviral RO organization and function

As outlined above, many open questions regarding the macromolecular organization of nidoviral ROs remain unanswered so far. Certainly, further improvements in cellular cryo-ET data acquisition will result in improved data quality and throughput, while advanced processing tools will help during subsequent data analysis. A main bottleneck of the cellular cryo-ET workflow is the cryo-FIB-milling step, a tedious and time-consuming process. The first software solutions for automation of the FIB-milling procedure have already become available (*83, 84, 318*), while novel hardware solutions promise the reproducible generation of high-quality cryo-lamellae (*317*). In addition, correlative electron and light cryo-microscopy (cryo-CLEM) can be utilized to guide the often difficult selection of suitable regions within the sample, which may then contain the fluorescently-tagged structure of

interest (*336, 337*). These can be rare and/or only localize in small regions of a cell, which requires a sufficiently high spatial resolution able to guide the FIB-milling process. Diffraction-limited light microscopy, with a resolution of ~400 nm, does not allow a sufficiently precise localization of the features of interest. Here, advanced super-resolution light microscopy methods promise to further close the resolution gap between light and electron microscopy (*338*). Further developments in super-resolution microscopy instrumentation (*339*) and adaptations for imaging under cryo-conditions (*340-342*) will enable a more precise and reliable localization and targeting of fluorescently tagged components in the future. Once the feature of interest is localized and contained in a cryo-lamella, the sample can be transferred into a cryo-TEM for high-resolution imaging. The tilt series acquisition on dedicated TEMs is steadily improving in throughput, with the development of advanced hardware and software solutions increasing both acquisition speed and quality. Tomogram reconstruction software packages, often originally made for the processing of resin-embedded samples, are more and more tailored and improved for the processing of low-contrast (cellular) cryo-ET data (*257, 343*). New sub-tomogram averaging strategies are starting to integrate all processing steps, from the handling of raw-data to the generation of high-resolution molecular models, in a single pipeline to allow an iterative refinement strategy that extracts as much high-resolution information as possible from the datasets (*77, 316, 344*). However, to be able to perform high-resolution subtomogram averaging, a first identification of the macromolecule of interest is pivotal. Cryo-CLEM, as discussed above, could already provide guidance during this step. However, the unambiguous identification of (relatively small) macromolecules can be difficult in the inherently noisy cryo-ET data and especially in crowded environments like the cytosol. Here, sophisticated denoising routines incorporating deep-learning strategies promise an easier visual identification of the macromolecule of interest (*345, 346*). In case (parts of) a macromolecular structure are already structurally characterized, recognition algorithms based on template matching (*347, 348*), deep learning (*349, 350*) or artificial intelligence could be applied. These are currently under development and could facilitate the automated mapping of particles for the characterization of their spatial organization or for the generation of high-resolution structures by sub-tomogram averaging. These approaches may provide detailed insights into protein structure and function, as long as sufficient resolution is obtained to allow the annotation of protein domains or even the *de novo* mapping of peptide chains. In some cases, the nature of the sample may prevent a straightforward analysis resulting in sufficient molecular detail, which ultimately depends on the availability of large copy numbers (i.e. thousands to millions) of the macromolecule of interest. This, for example, is the case for the coronavirus DMV pores, of which only a handful are found in each cryo-tomogram, with the number decreasing for higher-resolution data comprising a smaller field of view. Here, alternative routes to the cellular cryo-ET workflow that increase the yield of the macromolecules of interest in the EM sample have to be developed. In the case of nidoviral ROs, a purification procedure that would yield a highly concentrated DMV preparation would be favourable. Similar purifications that included differential- or gradient purification protocols have already been performed for nodavirus-induced ROs in mitochondrial membranes (*244, 269*). Such systems have the advantage of offering an easily manipulatable environment for functional analyses, which could be supplemented with therapeutic compounds and viral- or host factors presumed to be involved in underlying molecular processes. Interdisciplinary approaches such as lipidomics (*232, 351*) or proteomics (*352*) could be utilized to study the molecular composition of these systems under different conditions. The comparison of nsp-expression systems in such analyses could provide valuable insights into the interplay of

individual RO components with host-cellular factors.

As demonstrated for the coronavirus DMV pore (see **chapter 4**), where a hitherto unseen molecular complex demands the *de-novo* identification of its components, the use of labels compatible with the preservation of the native environment can be crucial. Ideally, the unperturbed sample would contain an electron dense tag that is recognizable in cryo-ET imaging to guide the search for the (maybe unidentified) structure of interest. However, labels meeting all these requirements are not yet available. A proposed solution in this direction is a protein tag that accumulates soluble gold salts, which would form clusters that are visible in the EM data (*353*). However, establishing this method for cryo-EM samples seems difficult, which might in part be due to the high toxicity of the gold salts that would have to be added to the culture media. Alternative strategies imply compromises, for example by disrupting cells for the isolation of the structure of interest in order to make these accessible for electron-dense labels, such as gold-beads. Subsequently, established immune-labelling EM protocols, relying on the binding of gold-beads coupled to antibodies, can be easily adapted for their use in such cryo-ET analyses (*172*). Although this strategy allows a straightforward detection of proteins and would be particularly suited for the screening of different candidates, the relatively large antibodies make the localization imprecise. A more precise tagging could be achieved by the use of small protein tags fused to the macromolecule of interest that can directly bind to functionalized gold beads (*244*). However, the heavily electron-scattering metal beads can produce artefacts during reconstruction and might obscure the sample (*354*). This can be avoided by the use of DNA-origami signposts (*355*), where a signpost-shaped DNA-assembly points directly at the molecule of interest and could also be used to guide the subtomogram averaging process. Such strategies could also be applied on macromolecules readily accessible on the surface of plasma membrane, however, for intracellular structures that are difficult to extract, alternative (*in situ*) approaches are required. The addition or deletion of a relatively large molecular mass, i.e. recognizable in sub-tomogram averages, can be used to reveal the identity and orientation of macromolecular components, as demonstrated in **Chapter 4**. In addition, an unambiguous identification of the structure of interest relies on subtomogram, and cannot be done directly in the tomograms. Current novel developments include protein tags, that would be highly sensitive to electron radiation. These tags would emit gas once irradiated, leading to a bubble that is trapped within the specimen and visible in the respective cryo-tomogram (*356*). Clearly, this and other future technical developments will facilitate addressing many of the open questions about the molecular playground of +RNA virus ROs and may – for example – lead to exciting new insights into the role of the DMV pore in the nidovirus replication cycle.

GENERAL DISCUSSION