Beyond structures of highly symmetric purified viral capsids by cryo-EM

Robert Stass¹,*, Serban L. Ilca¹,*, Juha T. Huiskonen¹,²,#

¹ Division of Structural Biology, Wellcome Centre for Human Genetics, Roosevelt Drive, University of Oxford, OX3 7BN, Oxford, UK
² Helsinki Institute of Life Science and Molecular and Integrative Biosciences Research Program, Faculty of Biological and Environmental Sciences, Viikinkaari 1, University of Helsinki, 00014 Helsinki, Finland
* These authors contributed equally
# Correspondence: juha.huiskonen@helsinki.fi

Abstract

Cryogenic transmission electron microscopy (cryo-EM) is widely used to determine high-resolution structures of symmetric virus capsids. The method holds promise for extending studies beyond purified capsids and their symmetric protein shells. The non-symmetric genome component has been addressed in dsRNA cypoviruses and ssRNA bacteriophages Qβ and MS2. The structure of human herpes simplex virus type 1 capsids has been determined within intact virions to resolve capsid–tegument interactions. Electron tomography under cryogenic conditions (cryo-ET), has allowed resolving an early membrane fusion intermediate of Rift Valley fever virus. Antibody-affinity based sample grids allow capturing of virions directly from cell cultures or even clinical samples. These and other emerging methods will support studies to address viral entry, assembly and neutralization processes at increasingly high resolutions and native conditions.

Highlights

• Single-particle averaging has evolved to determine high resolution structures of large icosahedral viruses and viral genomes.
• Sub-tomogram averaging continues to produce high resolution structures of viral surface proteins and nucleocapsids.
• Tomography of viruses has provided insights into dynamic viral entry processes.

Introduction

Cryogenic transmission electron microscopy (cryo-EM) has matured in recent years into a versatile, high-resolution technique in structural biology. This is largely due to the development of direct detectors that now allow recording data with better quality than ever before. In structural virology, cryo-EM is becoming increasingly useful in determining structures of virus capsids with helical [1-4] and icosahedral [5-14] symmetry. Due to the ordered nature of these assemblies and their structural homogeneity it is possible to reach resolutions better than 3 Å by combining data from several hundreds, often thousands of capsids [5]. At this resolution it is possible to recognise density for some amino-acid side chains, allowing de novo structure determination of previously uncharacterized viruses [10]. Electron tomography of
vitrified specimens (cryo-ET) can determine the structures of heterogeneous viral particles with pleomorphic shapes [15-18]. Sub-tomogram averaging (averaging of repeating structures from tomographic reconstructions, or tomograms) can be further exploited to resolve higher resolution detail of any repeating structures, such as the viral surface glycoproteins [15,16,18] and internal nucleocapsids [18-21]. Different approaches to study virus structures by cryo-EM discussed in this review are illustrated in Figure 1. We will first exemplify recent high-resolution structures of large viral capsids with icosahedral symmetry as well as the first structures of viral genomes. We will then move on to review current examples of viral surface and nucleocapsid structures determined by sub-tomogram averaging. Finally, we will highlight studies that have started to move the structural virology field beyond structures of isolated purified virions, with the aim of deciphering the infection and assembly pathways within the cellular context.

Main text

High resolution structures of large viruses from single particle averaging

Viral capsids are often ideal targets for high-resolution structure determination due to their high symmetry, relatively large size compared to many other protein complexes, and structural rigidity. Currently the highest resolution structures of icosahedrally symmetric virus capsids determined by single particle averaging are for the capsids of rhinovirus (~32 nm in diameter, 2.3 Å resolution) [13] and Tulane virus (~40 nm in diameter; 2.6 Å resolution) [5]. The much larger size of many virus capsids, however, poses several challenges and limitations as well. The thickness of the sample limits the resolution in cryo-EM reconstructions, partially due to the defocus gradient across the specimen and the curvature of the Ewald sphere [22]. Another practical limitation is that only a few large virus particles will fit in each field of view, necessitating longer data collection sessions. Furthermore, large virus capsids may be more prone to deformations than small capsids. Finally, large reconstruction box sizes still pose a challenge for current image processing software tools and either modifications to these or computers with more memory are required to allow refinement and reconstruction of very large particles.

Recently, the very large capsid structures (~130 nm in diameter) of human cytomegalovirus (HCMV) [23], Kaposi’s sarcoma-associated herpesvirus (KSHV) [14], herpes simplex virus type 1 (HSV-1) [24] and herpes simplex virus type 2 (HSV-2) [25] have been determined to resolutions of 3.9 Å, 4.2 Å, 4.2 Å, and 3.1 Å, respectively. In the KSHV and HSV-1 studies, the capsid structure was resolved from images of intact virions (~200 nm in diameter). This further increased the sample thickness, but in the case of HSV-1 allowed solving the capsid-associated tegument complexes (Figure 2A) [24]. A relatively low magnification and the super-resolution capability of a direct-electron detector were used to acquire ~10,000 movies to yield a sufficient number of particles (~40,000) required to reach a high enough resolution for atomic model building. To facilitate the reconstruction of these large viruses in a box with a size of up to 1,440×1,440×1,440 voxels, the software was modified to further optimize memory usage [14]. For both HSV-1 and HSV-2, a sub-particle based approach was employed to account for the local defocus and structural deformation of capsid subunits leading to resolution improvements of 3.9 Å to 3.5 Å, and ~4 Å to 3.1 Å, respectively [24-26]. We expect similar studies to become more
common due to increased accessibility to high-end microscopes allowing automated
data acquisition over several days, new detectors with a larger field of view that will
allow capturing more particles in one image, and improvements in reconstruction
algorithms.

**Asymmetric reconstruction of viral genomes**

The genomes of viruses with icosahedrally symmetric capsids remain challenging
targets for cryo-EM structure determination. As the nucleic acid is asymmetric and
possibly flexible, its structure gets incorrectly averaged when icosahedral symmetry is
applied. Recent studies on the genomes of both single-stranded and double-stranded
RNA viruses have overcome this limitation by asymmetric reconstruction and novel
data analysis strategies. Aided by their relatively rigid genome and by the presence of
a single maturation protein copy to break the icosahedral symmetry of the capsid, the
structures of the ssRNA bacteriophages Qβ [27] and MS2 [28-30] have been
determined. A graph-theoretical analysis of MS2 cryo-ET data allowed proposing a
model for the asymmetric organization of a ssRNA viral genome [28]. For both Qβ
and MS2, standard asymmetric refinements of single particle data were also
performed recently [27,29,30]. The latest study on MS2 was based on a dataset of
more than 300,000 particles, revealed multiple conformations with highly localized
differences, and allowed building of atomic models for more than 80% of the entire
genome sequence (Figure 2B) [30].

In the case of dsRNA viral genomes, only the genome structure of cypovirus, in both
transcribing and non-transcribing states, has been determined [31,32]. Both studies
relied on subtracting capsid contributions from the images followed by either a
standard asymmetric refinement with manual selection of particles to be included in
the final map [31] or a refinement based on relaxing symmetry [32]. These studies
revealed a non-spooled organization of the genome following pseudo-D3 symmetry.
Ten RNA-dependent RNA polymerases (RdRPs) occupy specific locations, leaving
two out of the twelve possible positions vacant. Notably, the RdRPs were resolved up
to 3.3-Å resolution and were shown to be in different conformations in the
transcribing and non-transcribing states. The individual RNA strands were readily
discernible and overall adopted an almost identical conformation in the two states,
with differences only in the regions directly interacting with the RdRPs.

**In situ structures of viral envelope glycoproteins and nucleocapsids**

Many enveloped virions are pleomorphic which prevents structural determination by
cryo-EM and conventional single particle averaging. By using tomography and sub-
tomogram averaging, the surface glycoprotein spikes have recently been mapped in
Lassa fever virus [16], Tula hantavirus [15], rubella virus [18], and a prototypic
foamy virus [33]. In the case of rubella virus, the spikes were found to be organised in
a helical arrangement and the structures of the individual subunits enabled fitting of
atomic models. In the latter example, single particle and sub-tomogram averaging
were combined in an unconventional way: a locally ordered patch of foamy virus
surface glycoprotein trimers, first derived by standard sub-tomogram averaging, was
used as a reference in a subsequent template-based particle picking and single particle
refinement to obtain a resolution of ~9 Å [33]. Different variations of this
combination are expected to be beneficial in analysing locally ordered lattices of
glycoproteins on the surface of enveloped viruses.
Many pleomorphic viruses contain internal structures that are also suitable for cryo-EM structure determination. Sub-tomogram averaging has been used to determine the structure of the helical nucleocapsids within Ebola (Figure 2C) and Marburg virus virions [21]. The same approach has been used to determine the structure of the mature HIV-1 cone shaped capsid revealing its assembly mechanism from pentamers and hexamers [20]. The structure of the immature HIV-1 capsid from virus-like particles in the presence of a maturation inhibitor has been determined to 3.9-Å resolution from a large dataset consisting of ~130,000 sub-tomograms [19]. This resolution, previously attainable only by single-particle averaging, was reached by applying frame-based motion filtering and exposure filtering to account for sample movements at each tilt and for the large electron dose of tomographic data collection, respectively. This structure suggested that maturation inhibitors work by stabilising the immature lattice, preventing proteolytic cleavage. The resolution of this structure was later improved to 3.4-Å resolution by using 3D contrast transfer function correction during reconstruction of the tomographic volume [34].

**Tomography of virus entry intermediates**

Cryo-ET can be a powerful method for examining dynamic viral processes such as the entry of enveloped viruses into cells. This process involves fusion of the viral membrane with the host cell membrane, a process catalyzed by viral surface glycoproteins called fusion proteins grouped in classes I–III. A number of publications have studied membrane fusion of influenza A virus (class I) to liposomes [35-39]. In one such study, tomograms calculated over a time course of 30 minutes revealed multiple stages of the fusion process, including membrane contact, dimple formation, hemifusion and pore formation [36]. A different study examined the formation of the hemifusion state by employing a Volta phase plate to improve the signal-to-noise ratio of the tomograms [37]. Two independent mechanisms were proposed for fusion pore formation: a rupture-insertion pathway and a hemifusion-stalk pathway, with the latter suggested to be more relevant for virus infection. Sub-tomogram averaging has also been used to study the orientation of full-length HSV-1 fusion protein gB (class III) in pre-fusion and post-fusion conformation on vesicles [40,41]. These two studies supported two alternative orientations of gB, the former placing the fusion loops membrane distal and the latter placing them membrane proximal. Finally, the fusion process has been studied in Rift Valley fever virus, a large enveloped virus with a class II fusion protein [42]. Virus–liposome contact sites resolved 30 seconds after low pH treatment by tomography and sub-tomogram averaging revealed a putative early fusion intermediate, in which the viral glycoprotein spikes were inserted into the liposome membrane.

**Towards more native structural virology**

The structural biologist’s perspective of the virosphere is biased by those viruses that can be isolated and purified in large quantities for structural studies. In addition, purification and cell culture conditions themselves may perturb the virion structure. For example, respiratory syncytial virus particles have been observed to convert from filamentous to noninfectious spherical particles over time at ambient or cell culture temperatures [43]. Novel cryo-EM grids based on antibody-affinity capture may provide means for gently purifying many types of virions directly from cell culture supernatants or even clinical and environmental samples. Such grids have already been used to analyse the structures of both enveloped [44] and non-enveloped viruses [5] by cryo-EM. In the latter case, a 2.6-Å resolution single particle reconstruction of Tulane virus was achieved from 3 µL of sample at concentration as low as ~2 µg/mL.
Cryo-ET allows studying virus–host cell interactions, such as entry and budding, in the native cellular context. By using this approach it has been shown that influenza A virus buds from its host cells in filamentous form [45]. This form has been suggested to be clinically more relevant than the spherical form, produced by some laboratory-adapted strains. In another example, murine leukemia virus-like particles were imaged interacting with a host cell plasma membrane, and the envelope protein was suggested to undergo a large-scale conformational change upon binding to a cellular receptor [17]. A tomography study of cyanophage P-SSP7 entry into its host marine cyanobacterium Prochlorococcus demonstrated how the phages first adsorb to the surface with their tails held away from the cell and later “stand up” with their tail perpendicular to the surface ready for genome injection (Figure 2D) [46]. Regarding intracellular events, vesicle-coated transport of an adenovirus along a microtubule has been visualized in human epithelial cells recently [47]. This study was restricted to imaging viruses at the cell periphery where the sample is thin enough for electrons to penetrate. Thinning of cells by focused ion beam milling is expected to enable imaging of viruses throughout the whole cell. This approach has recently been used to image egress of herpesvirus capsids through the nuclear membrane [48]. These examples give glimpses of the future potential of in situ structural virology. Use of other emerging technologies, such as the Volta phase plate [37], are expected to contribute to this success.

**Conclusions**

The remaining challenges in the structural biology of large icosahedral symmetric viruses lie in accounting for the defocus gradient across the specimen, correcting for global deviations from perfect spherical shape, sorting out local variations in the subunit conformations and correctly dealing with deviations from the icosahedral symmetry. Efficient solutions to each of these are required in order to resolve all of the ordered components of the virus, including the viral genome. On the sample preparation side, there is a need for addressing structures under conditions as native as possible, for instance by using on-the-grid purification based on antibodies. Focused ion beam milling is an emerging method in cellular structural biology that is also expected to have a profound impact in structural virology, as the method will allow following viral entry and assembly intermediates in cellula. Taken together, these improvements will allow descriptions of virus–host cell interactions during viral infection with improved structural preservation and resolution.

**Acknowledgements**

This work was funded by the Helsinki Institute of Life Science HiLIFE, University of Helsinki, European Research Council under the European Union’s Horizon 2020 research and innovation programme (649053 to J.T.H.), and Wellcome Trust Four-Year PhD Studentship (109135/Z/15/A to S.L.I.).

**Annotated References**

* [14] This paper on Kaposi’s sarcoma-associated herpesvirus reports many technological advances that were required to reconstruct the capsid of this large virus to 4.2-Å resolution. This structure allowed identifying structural motifs that stabilize the capsid and that that could be targeted by antivirals.
This study is one of the few examples where sub-tomogram averaging has been applied on viruses binding to the host cell. This approach suggested a large-scale structural rearrangements in the retroviral Env protein upon receptor binding.

This study on HIV-1 capsid demonstrated that sub-tomogram averaging can reach the same resolution as current single particle averaging methods. The structure suggested that maturation inhibitors work by stabilising the immature Gag lattice and thus blocking proteolytic cleavage.

Structures of the recombinant nucleocapsids as well as nucleocapsids from intact Ebolavirus virions were reconstructed by sub-tomogram averaging. Models created by flexible fitting of crystallographic structures suggested a mechanism for RNA encapsidation.

The highest resolution structure of a HSV-2 capsid, determined here at 3.1-Å resolution, was made possible by computationally compensating for flexibility in the viral capsid in addition to correcting for the defocus gradient across it. This study provides the most accurate description of protein–protein interactions in herpesviruses to date.

These authors were able to trace the ssRNA genome backbone and to build atomic models for sixteen RNA stem-loops in an asymmetric virus structure for the first time. The mapped interactions of the genome with the maturation and capsid proteins sheds light on the assembly and genome delivery mechanism.

The Env surface glycoprotein structure from a prototypic foamy virus, a member of the Retroviridae family, was reconstructed to 9-Å resolution by a combination of sub-tomogram and single particle averaging. This paper is one of the first examples of such a combination of methods used to address the structure of viral surface proteins.

This paper reports the highest resolution structure determined by sub-tomogram averaging to date. Resolution of 3.4 Å was reached by introducing 3D contrast transfer correction in tomographic reconstruction.

This study on influenza virus fusion to liposomes is a great example of how cryo-ET can be used to shed light on dynamic processes involving membranes.

In another study two different types of membrane influenza virus fusion pathways could be imaged by using Volta phase plate technology.

These authors used localized reconstruction to compensate for flexibility in large, enveloped Rift Valley fever virus particles and sub-tomogram averaging to resolve the interaction between its class II fusion protein and liposomes. This study is the first to report the structure of a viral fusion protein in an intermediate state.

The adsorption of a marine cyanophage to its host cell was imaged by cryo-EM and tomography. Subtomogram averaging and classification led to insights on the role of the phage tail fibres in this process.
References


Figure 1. **Different approaches to study virus structures by cryo-EM.** This schematic figure illustrates three approaches for studying virus structures discussed in this review (gray boxes). From left to right, the virus purification/isolation strategy, the type of sample on the electron microscopy sample grid, methods applied (single-particle averaging, SPA; and sub-tomogram averaging, STA) and the types of structures resulting from each approach are illustrated. The trend towards more native structural virology is indicated (blue arrow).

<table>
<thead>
<tr>
<th>Purification strategy</th>
<th>Sample on the grid</th>
<th>Methods</th>
<th>Resulting structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purified virions</td>
<td>Virions</td>
<td>Cryo-EM &amp; STA</td>
<td>Symmetric capsids</td>
</tr>
<tr>
<td>Large-scale cell culture and purification</td>
<td></td>
<td></td>
<td>Surface spikes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Genomes</td>
</tr>
<tr>
<td>Captured virions</td>
<td>Virions free of</td>
<td>Cryo-EM &amp; STA</td>
<td>Native structures of virions and their components</td>
</tr>
<tr>
<td>Small-scale cell culture or clinical/</td>
<td>purification artifacts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>environmental samples</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Virions in situ</td>
<td>Virus infected cells</td>
<td>Cellular Cryo-ET</td>
<td>Virus–host cell interactions in situ</td>
</tr>
<tr>
<td>Small-scale cell culture</td>
<td></td>
<td>&amp; STA</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2. Recent examples of viral structures studied by cryo-EM and cryo-ET. (A) A close-up of the vertex region of herpes simplex virus 1 capsid, reconstructed by localized reconstruction from single particle cryo-EM data. The inset shows one capsid-associated tegument complex (CATC). Different components are colored and labelled. The bi-lobed head region of the CATC is indicated with a box. Figure reproduced with permission from [24]. (B) A cut-open view of MS2 bacteriophage density, reconstructed by standard asymmetric refinement of single particle cryo-EM data. The viral capsid is radially depth colored, the maturation protein is magenta and the single-stranded RNA genome is blue. Examples of major and minor grooves resolved in in double-stranded regions of the genome are indicated. Figure reproduced with permission from [30]. (C) The structure of Ebola virus nucleoprotein solved by sub-tomogram averaging. N-terminal lobe in green, orange, and cyan; C-terminal lobe in blue, pink, purple, red, and brown. Putative RNA density is in yellow. The inset shows a model of the nucleocapsid, derived by placing the nucleoprotein structure back in its original positions in a tomogram is shown in gray. One subunit is in pink. Scale bar is 2 nm. Figure reproduced with permission from [21]. (D) A tomogram of cyanophage P-SSP7 particles (red) adhering to its host marine cyanobacterium reconstructed from cryo-ET data collected using a Volta phase plate. The cell wall is orange, the plasma membrane is light yellow, the thylakoid membrane is green, carboxysomes are cyan, the polyphosphate body is blue, and cytoplasmic granules (likely to correspond mainly to ribosomes) are light purple. Scale bar is 200 nm. Figure reproduced from [46].