Computational approaches for the analysis of cryo-electron tomograms

Inaugural dissertation

for the attainment of the title of doctor in the Faculty of Mathematics and Natural Sciences at the Heinrich Heine University Düsseldorf

presented by

David Kartte
from Bonn, Germany

Düsseldorf, October 2025

from the institute for Biology
at the Heinrich Heine University Düsseldorf
and the Ernst Ruska-Centre for Microscopy and Spectroscopy with Electrons:
Structural Biology (ER-C-3)
at the Forschungszentrum Jülich

Published with permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Reviewer: Prof. Dr. Carsten Sachse

Second reviewer: Prof. Dr. Gunnar Schröder

Date of the oral examination: 29.01.2026

Eidesstattliche Erklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

_____

Ort, Datum

_____

David Kartte

Acknowledgements

First, I would like to thank Prof. Dr. Carsten Sachse for his exceptional mentorship and continuous support during my PhD. He provided me with the environment and freedom to develop and test my ideas, and his guidance and critical feedback were indispensable. I am also grateful for the numerous opportunities for personal and professional growth he provided, from enabling my participation in stimulating conferences and workshops to facilitating connections and discussions with interesting and inspiring researchers.

I would like to thank Prof. Dr. Ralf Bartenschlager and Dr. Nils-Alexander Lakomek for providing valuable collaborative opportunities, their insightful scientific advice, and for their dedicated service on my thesis advisory committee. I would like to extend my gratitude to all members of the COVIPA project for providing numerous learning opportunities, fostering stimulating scientific discussions, and creating such an enjoyable collaborative atmosphere during our meetings. I am also grateful to Prof. Dr. Gunnar Schröder for kindly agreeing to serve as my second reviewer.

I would like to thank Dr. Benedikt Junglas, who was an excellent and insightful collaboration partner. Thank you for the opportunity to take part in your fascinating project. I would also like to thank Dr. Sabrina Berkamp for always taking the time to provide advice and for the valuable feedback on my writing.

I would like to thank Kevin Boga for the companionship, the excellent data and the opportunity for collaboration. I also greatly appreciate your thoughtful feedback on countless manuscripts and presentations. Thank you to Philipp Schönnenbeck for the many discussions about software, programming, bugs and errors. I am also thankful for your help setting things up at the beginning of my PhD and for your

reliability and competence when we were teaching. Thank you to Esther Hudina, I very much enjoyed working with you and I will miss our shared lunches. Thank you to Dr. Daniel Mann for helping to provide the computational resources, and especially for your early help with software and Jureca. In that context, I am also grateful to Chaitali Suhas Bagwe for the IT support. Thank you to Ursula Kessy for the administrative assistance. I would also like to thank everyone in our institute for providing such a friendly and supportive working environment, especially Aikaterini Filopoulou, Tom Goetze, Dr. Thomas Heidler, Dr. Lisa Jungbluth, Dr. Alexandros Katrinidis, Ahad Ali Kazmi, Dr. Claire Ortmann de Percin Northumberland, Ilona Ritter, Dr. Saba Shazad, Muratha Sottatipreedawong, Pia Sundermeyer, Jun. Prof. Dr. Irene Vercellino and Dr. Iris von der Hocht.

Finally, I would like to express my deep gratitude to my parents, my brothers, and my girlfriend. None of this would have been possible without their companionship and support.

Summary

Cryo-electron tomography provides three-dimensional structural information of biological specimen, allowing the study of cellular mechanisms in their native cellular context on the nanometer scale. New techniques allow for faster data generation and image processing, enabling the generation of large tomogram datasets and thus facilitating thorough ultrastructural analysis. However, data quality control is time consuming to conduct manually and requires expert knowledge, and quantifying the characteristics of interest is tedious work and subject to bias when conducted manually. This thesis aims at enhancing the capabilities of tomogram analysis by providing a tool for tomography data quality analysis and exploring a framework for automated ultrastructural analysis and quantification of characteristics of interest.

The newly developed method for local quality assessment termed RESOLVE (**RESO**lution estimation in **V**arying **E**nvironments) is available as a comprehensive toolkit, allowing local resolution estimations in proteins maps, 2D micrographs, tilt-series and tomograms. RESOLVE is shown to adequately represent data quality as local resolution estimates, which in particular describe quality-defining data characteristics such as lamella-edge distance, tomogram thickness, radiation damage and beam-induced motion.

Additionally, as a new framework for automated ultrastructural analysis, it is presented how 2D instance segmentation can be employed for 3D data. On that basis, it is shown how the SARS-CoV-2 replication organelle can be characterized in an automated manner, and thus the role of various viral and host proteins be studied.

Zusammenfassung

Die Kryo-Elektronentomographie erlaubt die Aufnahme biologischer Strukturen in 3D und ermöglicht die Untersuchung zellulärer Mechanismen in ihrem nativen zellulären Kontext im Nanometer-Bereich. Durch neue Methoden im Bereich der Mikroskopie und Bildverarbeitung können Daten schneller generiert werden, was dazu führt, dass mehr Daten für strukturelle Analysen bereitstehen. Die Daten manuell nach Qualität zu filtern ist jedoch zeitaufwändig und erfordert Erfahrung. Auch die Quantifizierung der Daten ist mühsame Arbeit und fehleranfällig, wenn manuell durchgeführt. Diese Arbeit zielt darauf ab, die Möglichkeiten zur Analyse von Tomogrammen zu verbessern. Dafür wird eine Software für die Analyse der Datenqualität entwickelt und ein Rahmen für die automatisierte Analyse und Quantifizierung von Tomogrammen untersucht.

Die hier präsentierte Methode zur Qualitätsbestimmung wird frei zu Verfügung gestellt als das umfassende Tookit RESOLVE (**RESO**lution estimation in **V**arying **E**nvironments). Die Methode basiert auf lokaler Auflösungsbestimmung und kann außer für 3D Tomogramme auch für 2D Messungen von einzelnen Bildern oder die Qualitätsbestimmung von Proteinstrukturen angewendet werden. Es wird gezeigt, dass die Auflösungsbestimmungen durch RESOLVE angemessene Qualtätsangaben macht, die insbesondere relevante Eigenschaften von Tomographie Daten wie den Abstand zum Lamellenrand, Tomogramm-Dicke, Schäden durch den Elektronenstrahl und die Bewegung in der Probe beschreiben.

Als Rahmen für die automatisierte Analyse von Tomogrammen wird gezeigt wie 2D-Instanzsegmentierung für 3D-Daten eingesetzt werden kann. Darauf aufbauend

wird gezeigt, wie die Replikationsorganellen von SARS-CoV-2 automatisiert charakterisiert werden können und somit die Rolle von verschiedenen viralen-und Wirtsproteinen untersucht werden kann.

Contents

List of Figures

List of Tables

List of Abbreviations

Cs	**C**osine **s**imilarity
cryo-EM	**cryo**genic-**e**lectron **m**icroscopy
cryo-TEM	**cryo**genic-**t**ransmission **e**lectron **m**icroscopy
CTF	**C**ontrast **t**ransfer **f**unction
DED	**D**irect **e**lectron **d**etector
DMV	**D**ouble **m**embrane **v**esicle
DFCP1	**D**ouble**F**YVE **c**ontaining **p**rotein **1**
ESCRT-III	**E**ndosomal **s**orting **c**omplex **r**equired for **t**ransport-**III**
ER	**E**ndoplasmic **r**eticulum
ERGIC	**E**ndoplasmic **r**eticulum–**G**olgi **i**ntermediate **c**ompartment
FEG	**F**ield **e**mission **g**un
FDR	**F**alse **d**iscovery **r**ate
FSC	**F**ourier **s**hell **c**orrelation
GUI	**G**raphical **u**ser **i**nterface
KD	**K**nock**d**own
KO	**K**nock**o**ut
LD	**L**ipid **d**roplet
NSP	**N**on **s**tructural **p**rotein
PI3P	**P**hosphatidylinositol **3**-**p**hosphate
RNA	**R**ibo**n**ucleic **a**cid
RO	**R**eplication **o**rganelle
RTC	**R**eplication-**t**ranscription **c**omplex
SARS-Cov-2	**S**evere **a**cute **r**espiratory **s**yndrome - **co**rona**v**irus - **2**

SNR	**s**ignal-to-**n**oise **r**atio
SSNR	**s**pectral **s**ignal-to-**n**oise **r**atio
Vipp1	**V**esicle **i**nducing **p**rotein in **p**lastids **1**
WBP	**W**eighted **b**ack **p**rojection
WT	**W**ild**t**ype

Chapter 1

Introduction

1.1 Cryo-electron microscopy

Cryo-electron microscopy (cryo-EM) enables the visualization of biological molecules at near-atomic resolution in their native, hydrated state. It significantly advanced the study of biological molecules, contributing to progress in drug discovery, vaccine development, and our understanding of cellular processes. Transmission electron microscopes operate by passing a beam of electrons through thin specimens. The interactions between the transmitted electrons and the sample contribute to image formation, and, due to the shorter wavelength of electrons compared to light, reveal structures at resolutions beyond the capabilities of light microscopy. Historically, the fragility of biological samples presented a challenge for electron microscopy, as radiation can quickly induce damage to the sample (Dubochet et al., 1988). While still remaining a limiting factor, this was largely overcome by rapidly freezing the specimen in vitreous ice, which preserves specimens in a near-native state (Adrian et al., 1984; Stark, Zemlin, and Boettcher, 1996). Thus, unlike for traditional electron microscopy, cryo-EM does not require specimens to be dehydrated or chemically fixed, processes that can distort biological structures (Dubochet et al., 1988). It also allows for protein structure determination without the need for crystallization, a requirement which historically limited X-ray crystallography.

The *resolution revolution*, driven by the development of direct electron detectors that dramatically improved electron capture and the refinement of computational

algorithms for protein structure determination, led to routinely achieved resolutions of 2-4 Å in single-particle cryo-EM with purified proteins (Kühlbrandt, 2014). Structures resolved at this resolution allow for atomic modeling, making single particle cryo-EM one of the most important protein structure determination techniques today. In cases where sufficient purification is not achievable or to study proteins in their native cellular context, subtomogram averaging can be employed, sometimes yielding similar resolutions (Tegunov et al., 2021; Burt et al., 2024).

1.1.1 Design of the transmission electron microscope

The optical electron path is kept under vacuum conditions, preventing electrons traveling through the microscope column from being scattered and decelerated by air molecules, which would degrade image quality and resolution (Carter and Williams, 2016). The electrons are emitted by an electron gun, generating the primary electron beam at energies typically ranging from 80 to 300 keV (Franken et al., 2020). The choice of acceleration voltage determines electron momentum and thus electron wavelength, limiting the resolution of the final image. Modern instruments typically employ field emission guns (FEGs) that provide the most coherent and brightest electron sources (Henderson, 2015; Egerton, 2016). The electron beam is controlled and focused through a series of electromagnetic lenses that function analogously to glass lenses in light microscopy, manipulating the electron beam's trajectory through magnetic fields. The condenser lens system, typically consisting of two or three condenser lenses, shapes and focuses the electron beam onto the specimen, with the current flowing through the electromagnetic coils determining the strength of the magnetic field and thus controlling beam convergence and magnification (Rose, 2008; Egerton, 2016; Carter and Williams, 2016). The sample stage is designed to maintain specimens at liquid nitrogen temperatures (approximately -180°C) to reduce radiation damage (Stark, Zemlin, and Boettcher, 1996) and to keep biological samples in their near-native hydrated state. After

electrons pass through the specimen, the objective lens re-focuses the scattered electrons and creates a real space image, that is further magnified and projected onto the detection system by a projection lens. Strategic placement of apertures throughout the system allow for selective filtering of the electron beam to enhance image contrast and reduce noise (Rose, 2008; Egerton, 2016; Carter and Williams, 2016). Ultimately, the transmitted electrons reach the detection system. While traditional detectors such as charge-coupled device (CCD) cameras or complementary metal-oxide-semiconductors (CMOS) requre the conversion from electrons into photons, direct electron detectors (DED) directly record the primary location of the electron hitting the detector. Therefore, direct electron detectors offer superior detective quantum efficiency (DQE) that translates into higher signal-to-noise ratio (SNR), as well as providing higher readout capabilities that enhance movie-mode data collection (McMullan et al., 2014; Carter and Williams, 2016).

1.1.2 Contrast transfer in the electron microscope

Electrons can either pass through the sample unhindered, or interact with the sample in various ways: Electrons can be elastically scattered, passing through the specimen and experiencing a change in direction and a phase shift. They can also be inelastically scattered, additionally losing energy. Elastically scattered electrons, together with the unscattered incident beam, create the phase contrast (Dubochet et al., 1988), that is the dominant source of contrast for biological samples. Inelastically scattered electrons contribute to amplitude contrast, which makes up only around 10 % and is thus comparatively small for biological samples (Carter and Williams, 2016). However, in a cryo-EM setup, they effectively introduce noise, as they damage the sample and are not focused on the image plane due to the loss of energy (Carter and Williams, 2016; Koning, Koster, and Sharp, 2018).

As the fragility of biological samples limits the electron dose, the SNR remains low. In order to enhance the contrast, in cryo-TEM, an underfocus is applied during

imaging. The defocus has the desirably effect of modulating the signal transfer, specifically increasing contrast for lower spatial frequencies. The transfer of contrast as a function of spatial frequency is described by the contrast transfer function (CTF) (Dubochet et al., 1988; Sigworth, 2016). Similar to a sine wave, the CTF oscillates between -1 and 1 (Figure 1.1), with an envelope function attenuating higher spatial frequencies due to microscope aberrations (Wade, 1992). The CTF is importantly modulated by the defocus: For low defocus, the transfer of low spatial frequencies is low, as well as the number of zero crossings. While negative values of the CTF can be dealt with by phase flipping, at zero crossings, the corresponding information is lost. Increasing the defocus importantly leads to an increase in low spatial frequency contrast, but also to an increased number of zero crossings (Sigworth, 2016; Carter and Williams, 2016; Rohou and Grigorieff, 2015). To deal with the frequency-dependent contrast transfer and recover the information, CTF-correction has to be applied. This requires estimation of the defocus, typically achieved by analyzing the power spectrum of the image (Figure 1.1).

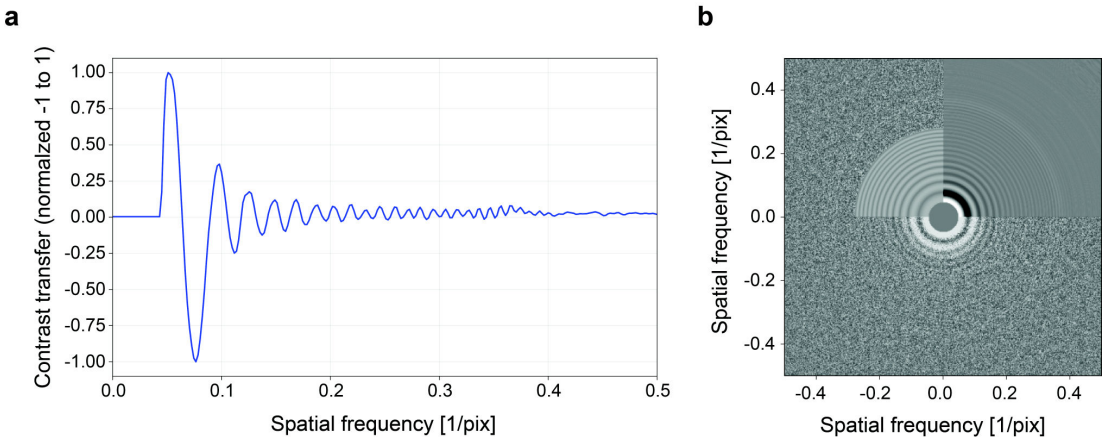

FIGURE 1.1: Typical CTF of a ribosome micrograph (EMPIAR-10524). Defocus of 1.6 μm. Measurements conducted with CTFFIND4 (Rohou and Grigorieff, 2015). a) radially averaged fit CTF. b) power spectrum (bottom half) with fit CTF (top half) in 2D.

1.1.3 Single particle cryo-EM

Despite impressive progress for instrumentation like direct electron detectors (Kühlbrandt, 2014), image contrast in cryo-EM is still limited. The most important limiting

factor is the fragility of biological samples, naturally limiting the electron dose. Additional difficulties may arise during sample preparation, like preferred protein orientation, ice thickness variation, sample heterogeneity, protein aggregation or denaturation. However, many shortcomings can be corrected for during image processing. The most important image processing workflow in structural biology is single particle cryo-EM, making protein structure determination on the atomic level possible. Single particle cryo-EM bypasses resolution limits induced by dose limitations by averaging many low exposure images and thereby increasing the SNR (Frank, 1975). Often, thousands of images are collected, containing hundreds of thousands of protein copies. The Fourier-slice theorem states that the Fourier transform of a projection image of an object equals a two-dimensional slice though the three-dimensional Fourier transform of that object. When all protein projection images are combined in the correct orientation in Fourier space, the three-dimensional Fourier transform of the protein can be back-projected into real space, yielding the three-dimensional structure of the protein (Nogales and Scheres, 2015). Although this technique led to impressive results early (Henderson et al., 1990), highlighting the potential of the technique, high resolution structures remained the exception. However, the last decade brought rapid algorithmic advances and the development of accessible software that have made single particle cryo-EM widely available. (Scheres, 2014; Grant and Grigorieff, 2015; Punjani et al., 2017; Tegunov and Cramer, 2019).

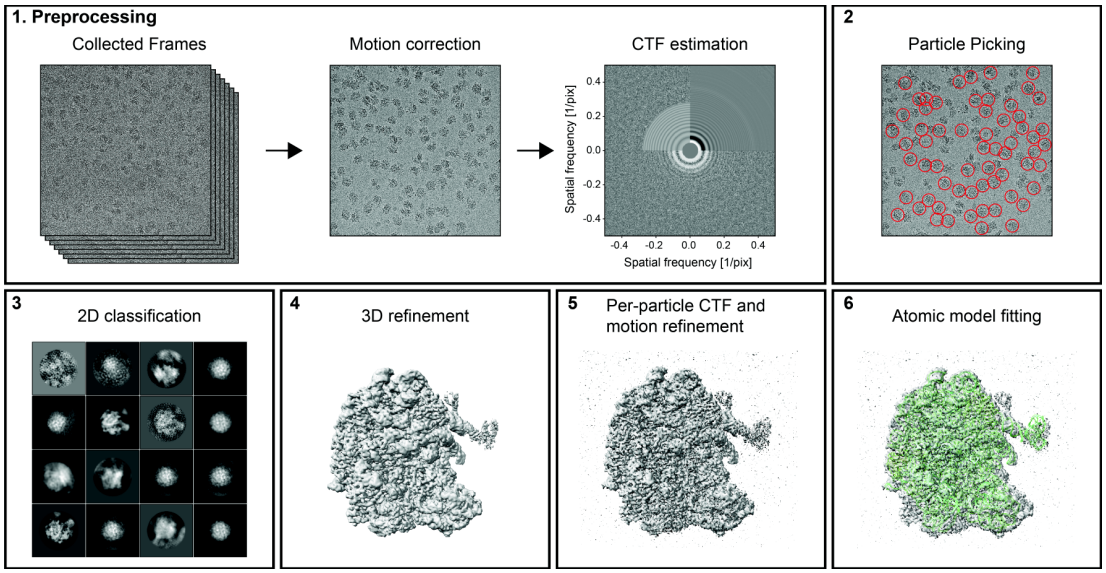

FIGURE 1.2: General single particle cryo-EM workflow (dataset taken from EMPIAR-10524). 1) Preprocessing includes motion correction and CTF estimation. 2) Particles are picked in an automated manner for the entire set of micrographs. Picked particles are highlighted with red circles. Note that automated particle picking in noisy micrographs remains challenging and is still subject to development. 3) Particles are classified and filtered (visualization via RELION5 (Burt et al., 2024). 4) The filtered particle set is aligned against a reference. 5) Per-particle corrections can be conducted once accurate information about particle structure, location and orientation are present, improving the resolution significantly. 6) If the workflow yielded a high resolution structure, an atomic model can be fit to the 3D map.

Single particle cryo-EM requires careful sample preparation and imaging (Weissenberger, Henderikx, and Peters, 2021; Xu and Dang, 2022). Once a set of micrographs is collected, the first step of image processing is usually motion correction (Figure 1.2). Image unblurring is required as the electron beam induces movement in the sample (Li et al., 2013; Scheres, 2014). Micrographs are collected in movie-mode, fractioning the electron dose to collect a set of frames instead of a single image, allowing to estimate motion between frames and to compensate for progressive radiation damage by weighting (Zheng et al., 2017). After alignment, the dose-weighted frames are combined, resulting in a single image with increased SNR. At this stage, usually CTF-correction is performed by careful defocus estimation (Rohou and Grigorieff, 2015). Once all micrographs are preprocessed in this way, particles need to be located, commonly referred to as *particle picking*. Particle picking is often based on either template-matching algorithms (Frank and

Wagenknecht, 1983) or, more recently, deep learning approaches (Wagner et al., 2019; Bepler et al., 2019). With the cropped projection images of these particles, 2D classification is performed: Images are classified by similarity, so that projection images from the same orientation can be averaged to increase the SNR (van Heel and Frank, 1981). At this stage, manual intervention is often necessary, filtering out noisy class averages or classes not containing protein projections (Singer and Sigworth, 2020). This stage is particularly important for heterogenous samples, as high resolution structure determination requires sample homogeneity. Then the refinement procedure can start: Each of the projection images is compared to projections from different orientations of a 3D reference structure. The initial reference structure can either be computed from the two-dimensional class averages, or existing structures with high similarity can be employed, albeit low-pass filtered in order to not bias refinements (Singer and Sigworth, 2020). Iteratively, the projection images are aligned to best fit the reference and, in a second step, used to create a new, better resolved reference. After refinement, for heterogeneous samples, classifications may be conducted in three dimensions, potentially increasing particle homogeneity. Once a good resolution has been obtained, the CTF, additional optical aberrations and motion can be further refined and corrected for on a particle level (Zivanov et al., 2018; Tegunov et al., 2021; Zivanov et al., 2022). The resolution of the final map can be calculated and once a high resolution map has been obtained, an existing atomic model can be fit or a new model can be built (Figure 1.2).

1.1.4 Cryo-electron tomography

Cryo-electron tomography (cryo-ET) is an imaging technique that does not require purified proteins and allows the visualization of vitrified cells, thus enabling imaging in the native cellular context. The workflow of *in situ* cryo-ET often requires prior localization of the regions of interest, achieved through correlative light and

electron microscopy (CLEM) techniques (Pierson, Yang, and Wright, 2024), and thinning of the imaged cell (Kudryashev, Castaño-Díez, and Stahlberg, 2012). For election microscopes operated at 200 or 300 kV, the mean free path (until a scattering event occurs) is around 300-400 nm (Schiøtz et al., 2024). To avoid multiple scattering events that degrade image resolution, the analyzed sample should ideally be thinner than 200 nm. To achieve this, the sample is thinned via focused ion beam (FIB) milling, generating so called *lamellae* suitable for imaging (Berger et al., 2023; Noble and Marco, 2024). While structure determination with *in situ* data via subtomogram averaging remains challenging, imaging proteins in their native context is often of special interest. Important contributions in the context of virology are the *in situ* structure of the SARS-CoV-2 spike protein (Turoňová et al., 2020b) or the pore of the replication organelle (Zimmermann et al., 2023). In some cases, with increasing maturity of processing software, high resolution *in situ* structures can be determined (Tegunov et al., 2021).

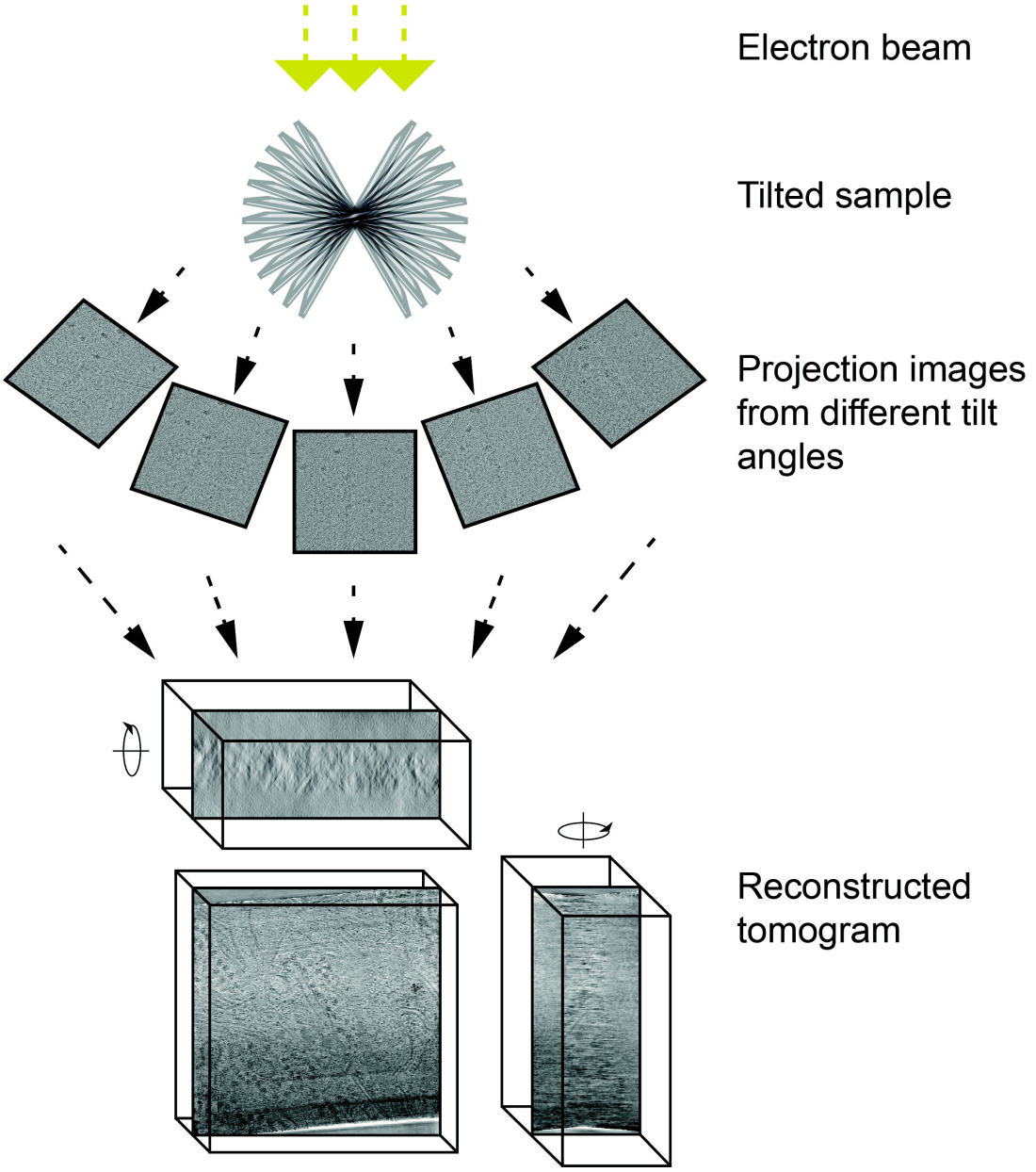

FIGURE 1.3: Data collection for cryo-ET. The stage is tilted in both directions up to ±60°. For different angular increments, images are recorded. Usually, 40-80 images are recorded, often with 2-3° increments. Yellow arrows indicate the electron beam. The reconstructed tomogram is portrayed as three 2D slices through the tomogram from the three different axes. For visualization purposes, images for the xz and yz plane are tilted by 90° around the indicated axis. Example data processed from EMPIAR-11306.

For cryo-ET, a set of images is collected from different orientations (Figure 1.3), with each images consisting of a set of frames to enable motion correction, resulting in a tilt-series. As the total electron dose is limited due to accumulating radiation damage to approximately $\sim 100\, e^{-}/\text{Å}^{2}$ (Koning, Koster, and Sharp, 2018), tilt

increment and tilt range are important tunable parameters, influencing the angular coverage and image resolution (Crowther, DeRosier, and Klug, 1970). One of the most commonly used techniques is the dose-symmetric acquisition scheme (Hagen, Wan, and Briggs, 2017; Turoňová et al., 2020a). The first image is recorded on the untilted stage, then recordings alternate between the increasingly positive and negative tilted stage, evenly distributing the electron dose in both tilt directions. To obtain full angular coverage, the stage needs to be tilted ±90°, which becomes impossible due to the increasing effective sample thickness and mechanical limitations of the stage. Therefore, the tilt range is limited to ±60-70°, always leaving parts of the views uncovered (Figure 1.3). This fact unavoidably leads to missing information along the z-axis (beam direction), assuming a wedge like shape in Fourier space, called the missing wedge (Lučič, Rigort, and Baumeister, 2013). After reconstructing a tomogram from the tilt-series, this loss of information manifests in characteristic artifacts: Membranes stretching across the xy-plane are invisible, and structures appear elongated along the z-axis. While this is not ideal, images reconstructed through the xy-plane appear mostly intact.

Tomogram reconstruction

The processing for tomography data is conducted similarly to the processing for single particle cryo-EM: It starts by correcting for beam induced motion and the CTF. However, due to the dose distribution over a large set of images, the SNR is much worse and the estimations can be more challenging (Koning, Koster, and Sharp, 2018). As the defocus for thick samples can vary, three-dimensional or local defocus estimation may be beneficial (Turoňová et al., 2017; Tegunov and Cramer, 2019). Different from single particle cryo-EM and essential for successful tomogram reconstruction is the alignment of the tilt-series (Koning, Koster, and Sharp, 2018; Wan and Briggs, 2016). If available, high contrast fiducial markers, added to the sample before freezing, can be located and tracked through the tilt series,

providing information about the spatial relations between the tilt images (Nickell et al., 2005). Alternatively, tilt images can be divided into patches, and the patch movements can be tracked via cross-correlating them across the tilt-series (Noble and Stagg, 2015). Then aligned tilt series is then used to reconstructed a 3D real-space tomogram. The most popular algorithm for tomogram reconstructed is weighted back-projection (WBP): The densities from the tilt images are projected into a real-space volume (Wan and Briggs, 2016). As the low frequencies (center of Fourier space) are sampled much denser within the tilt-series, frequency dependent weighting is applied for each tilt image before back-projection (Radermacher, 1988). Additionally, weighting on a tilt-level according to total accumulated dose can be conducted. The resulting 3D tomogram can be used for downstream analysis in two important ways: First, subtomogram averaging can be conducted for protein structure determination. Second, ultrastructural analysis can be conducted, often based on extensive tomogram labeling, called segmentation.

Subtomogram averaging

Subtomogram averaging (STA) is a structure determination workflow similar to single particle cryo-EM, only that it depends on three-dimensional tomograms instead of two-dimensional micrographs. Once tomograms are reconstructed, particle picking can be conducted. Similar as for single particle cryo-EM, for high resolution structure determination, a large number of particles is necessary. Thus, subtomogram averaging requires many thousand of particles, usually collected from a large set of tomograms. Due to the low SNR in tomography and thus reduced contrast in tomograms in addition to the crowded cellular environment for *in situ* cryo-ET, particle picking is often much more challenging and less streamlined than for single particle cryo-EM (Pyle and Zanetti, 2021). Once particle coordinates are found, small subvolumes around the particles are extracted, called subtomograms. This is different from single particle cryo-EM, where each particle is a two-dimensional image. As orientational and translational alignment procedures are complex even

in two dimensions, operating in three dimensions has the consequence of making subtomogram averaging much more computationally costly than single particle cryo-EM. Additionally, due to the lower SNR in tomography, subtomograms carry less information than the two-dimensional particle images in single particle cryo-EM, making alignments more difficult. For these reasons, it is often beneficial to acquire as many information about particle orientation as possible during particle picking. For example, particle poses are restrained if located on membranes, and geometrical picking approaches or membrane modeling can help to identify prior orientations (Navarro, Stahlberg, and Castaño-Díez, 2018; Castaño-Díez and Zanetti, 2019; Basanta et al., 2020; Martinez-Sanchez et al., 2020), making orientational alignments less computationally costly and more accurate. During subtomogram alignment to the reference, the missing wedge has to be masked, otherwise, it could potentially drive the alignment procedure (Bartesaghi et al., 2008). After the alignment procedure, similarly as for single particle cryo-EM, 3D classification can be performed, increasing the homogeneity of the particle set and allowing to reconstruct a better-resolved structure. Once a well resolved structure is generated, previously fixed parameters regarding the CTF, tilt-series alignment and sample motion can be refined and corrected for on a per-particle level (Tegunov et al., 2021; Zivanov et al., 2022), improving the final resolution.

Denoising

Before any downstream analysis of tomograms like segmentation or particle picking, it is often beneficial to increase the image contrast. Contrast improvements can for example be achieved by simply lowpass-filtering the tomogram, as high-frequency noise is suppressed (Figure 1.5). More elaborated approaches are based on Wiener-like filters, deconvoluting the tomogram to increase contrast (Tegunov and Cramer, 2019). Some of the most popular approaches are based on deep learning. Especially popular is the noise-2-noise approach (Lehtinen et al., 2018;

Buchholz et al., 2019; Bepler et al., 2020): Two independently reconstructed half-tomograms are created by splitting tilt-image frames into two subsets per tilt. Small subvolumes are then used to train the network, learning to predict subvolumes from one half-tomogram using subvolumes from the other half-tomogram as input. These approaches tend to perform well in tomography, especially for *in vitro* tomograms (Figure 1.5). Tomogram denoising can also be combined with modern deep learning approaches to regain missing wedge information (Liu et al., 2022; Wiedemann and Heckel, 2024).

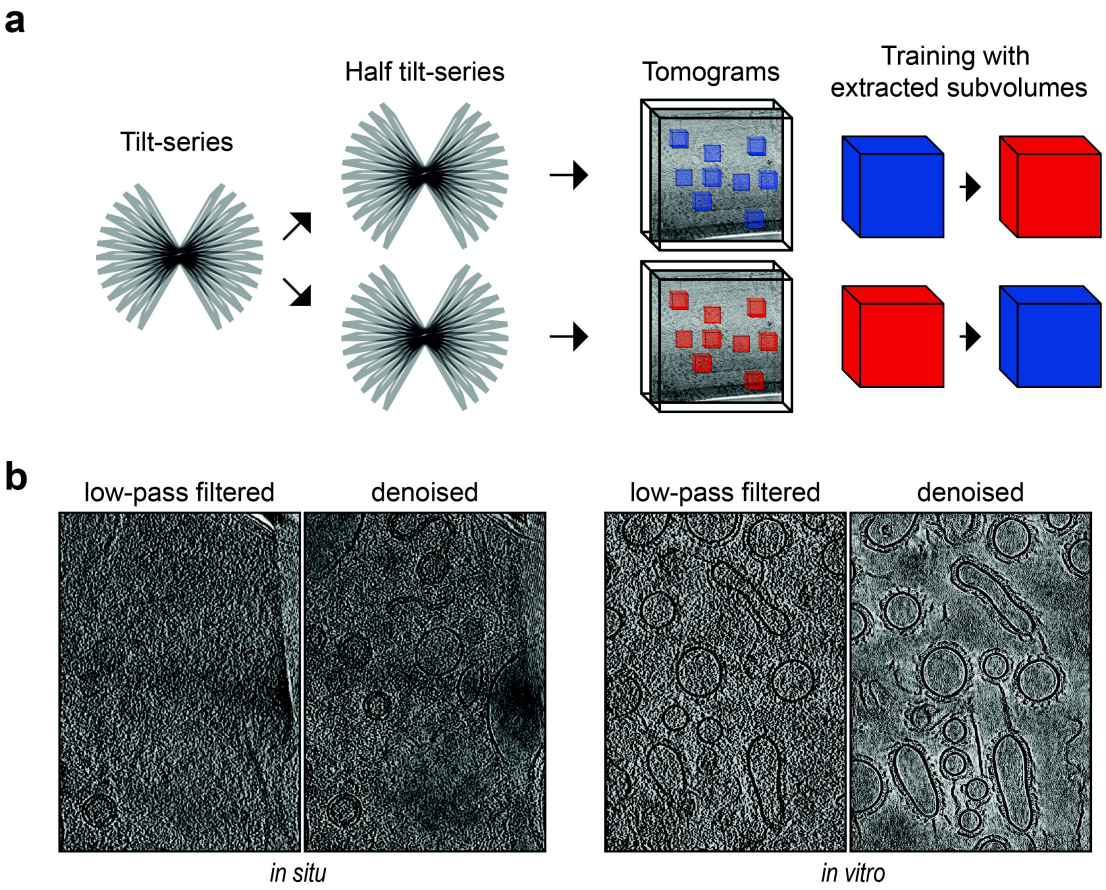

FIGURE 1.4: Denoising of tomograms. a) For denoising with the noise-2-noise method, the tilt-series is split across frames. From the two tilt-series, each tilt containing only half the image frames, two half-tomograms are reconstructed. Subvolume pairs are extracted, and the network learns to predict subvolumes from the second half-tomogram by being fed subvolumes from the first. b) Left: Low-pass filtered tomogram slice compared to a cryoCARE (Buchholz et al., 2019) denoised tomogram for an *in situ* tomogram (Tomogram provided by Kevin Boga). Right: Low-pass filtered tomogram slice compared to a cryoCARE (Buchholz et al., 2019) denoised tomogram for an *in vitro* tomogram (Tomogram provided by Dr. Benedikt Junglas).

Segmentation

Segmentation is the process of dividing images into smaller groups, based on shared characteristics: A group label needs to be assigned to each pixel of the image, or to each voxel in the case of three-dimensional data. In cryo-ET, segmentation often is the basis for further downstream analysis, as it allows for the quantification of data characteristics. Furthermore, particle picking approaches are often based on segmentation.

There are two important approaches used for cryo-ET data analysis: Semantic segmentation and instance segmentation. For semantic segmentation, a class label is assigned to each voxel of the tomogram, while for instance segmentation, a second label for different class instances is assigned as well. For particle picking, both approaches are regularly employed. However, particle picking is commonly conducted on two-dimensional tomogram slices, not on three-dimensional volumes. As tomograms are relatively large, usually from hundreds of megabytes up to more than one gigabyte, approaches that require the full tomogram feeding into the network at once are usually unfeasible. However, feeding small subvolumes into the network is usually not an issue for semantic segmentation, as long as the subvolumes are large enough to provide important spatial context. However, for instance segmentation, dividing the tomogram into subvolumes would split instances. Therefore, when different instances are required, often semantic segmentation is conducted first. In a second step, the segmentation is split into distinct instances, for example employing connected component analysis (Lamm et al., 2024). In many cases, this postprocessing approach is sufficient, as long as distinct entities are unconnected. A more elaborate approach is based on a dimensionless instance segmentation transformer (Kiewisz et al., 2025), applied on a point cloud based on the semantic segmentation. The transformer generates a probabilistic map of connections between graph points, which is then used as a basis for connected component analysis (Kiewisz et al., 2025).

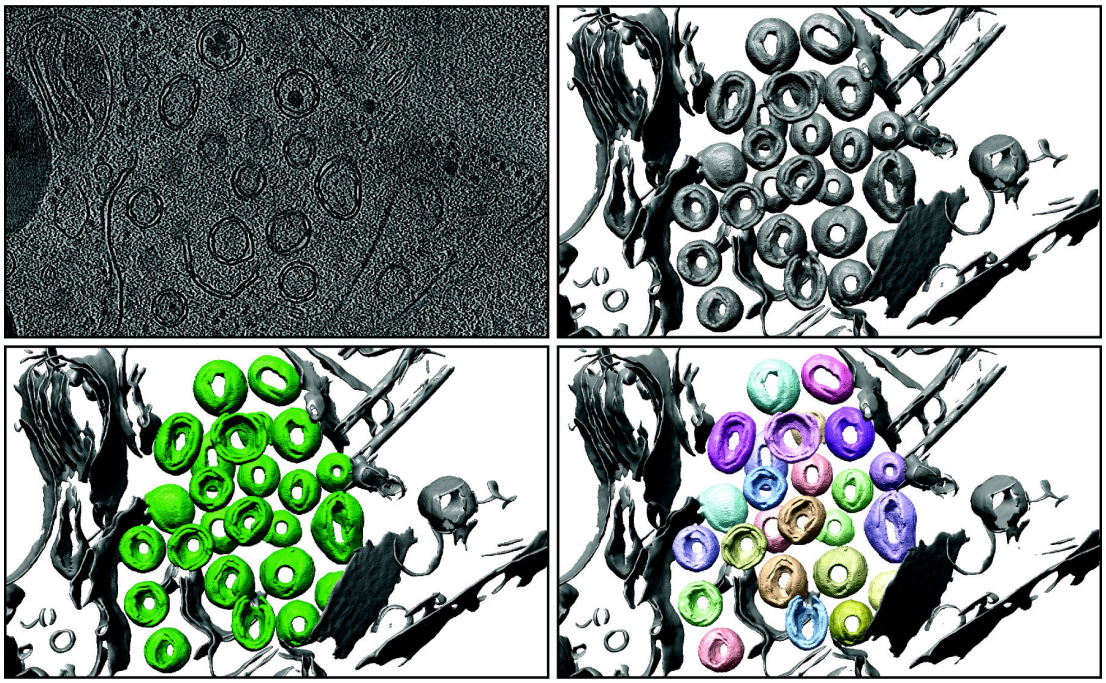

FIGURE 1.5: Segmentation at different stages. Tomogram provided by Kevin Boga. Top left: Z-slice of cryoCARE denoised tomogram (Buchholz et al., 2019). Top right: Segmented membranes via membrain-seg (Lamm et al., 2024). Bottom left: SARS-CoV-2 double-membrane vesicles labeled as one class (green), all other membranes as another class (grey). Bottom right: SARS-CoV-2 double-membrane vesicles labeled as distinct instances.

Once tomograms are labeled, quantitative analysis becomes possible. Software such as Dragonfly (Object Research Systems), Amira (Thermo Fisher Scientific) or the surface morphometrics toolkit (Barad et al., 2023) provide tools for quantitative analysis. However, different experiments often require unique analytical approaches. Calculating spatial parameters for uniquely shaped membrane organelles or spatial relationships between proteins and membranes of interest is difficult to generalize, making tailored analysis approaches necessary.

1.1.5 Resolution estimation

With single particle cryo-EM often quickly facilitating structure determination via streamlined routines (Punjani et al., 2017), and even subtomogram averaging becoming more accessible (Burt et al., 2024), map validation becomes an important

aspect. Map quality in cryo-EM is usually measured as resolution. Resolution, referring to the finest detail that can be reliably distinguished in a reconstructed map, is therefore an important indicator defining map suitability for biological interpretation. As resolution allows quick and easy comparisons of map quality, it also serves as performance metric, making structure determination routines comparable.

Conventionally, resolution estimation is conducted by splitting the data into two half-sets, only using half of the particles for reconstruction, allowing the creation of two half-maps. These half-maps are then transformed in Fourier space, and, for each shell k, the Fourier shell correlation (FSC) (Harauz and van Heel, 1986) is calculated as:

$$\mathrm{FSC}(k) = \frac{\langle F_1 \cdot F_2^* \rangle}{\sqrt{\langle |F_1|^2 \rangle \langle |F_2|^2 \rangle}} \tag{1.1}$$

With F_1 and F_2 being the Fourier transforms of the two independently reconstructed maps. A FSC curve typically starts with high correlation for low spatial frequencies, as the half-maps correlate almost perfectly. For higher frequencies, the correlation drops, as noise becomes more dominant. The highest-frequency shell still surpassing a predetermined threshold is then setting the resolution-determining frequency, with the resolution simply being equal to $\frac{1}{frequency}$ (Figure 1.6). However, the adequate method for threshold determination remains open to discussion. Historically, an intuitive threshold of 0.5 was established (Böttcher, Wynne, and Crowther, 1997), representing the resolution where correlation drops below 50%. Later, a FSC threshold of 0.143 was established, derived from the correlation expected between two random, independent noise distributions (Rosenthal and Henderson, 2003). The most commonly used threshold today is still 0.143, although it is argued that under certain circumstances, the more conservative threshold of 0.5 is more appropriate (van Heel and Schatz, 2005). Another line of argument takes the effect of sample size into account, arguing that no fixed threshold can be suitable for all maps. The proposition then was to use statistical, map-specific methods for threshold determination. One popular example is the σ-criterion (Saxton and Baumeister,

1982; Orlova et al., 1997; van Heel and Schatz, 2005), explicitly taking the sample size into account, or the usage of a permutated reference distribution for threshold determination, which implicitly also accounts for sample size (Beckers and Sachse, 2020).

Another often discussed feature of global resolution estimates concerns the effects of masking. It is known that masking introduces dependencies in Fourier space, leading to artificially inflated resolution estimates (Chen et al., 2013). In this context, it has been shown that correlation measurements can also be conducted in real space, as an alternative to the conventional FSC (Penczek, 2020). Conducting real-space measurements of shell-depended band-pass filtered maps has the advantage that masking is conducted after the Fourier-space operations, therefore not introducing dependencies. However, it has to be dealt with dependencies that arise due to band-pass filtering as a consequence. Another solution, employing permutation testing, uses an effective sample size for the creation of the reference distribution which is tested against (Beckers and Sachse, 2020). The creation of a reference distribution is achieved by conducting correlation measurements across randomly permutated shells, which leads to the dependencies introduced via masking being destroyed due to the random permutations. The proposed solution is to use an effective sample size, which is in a way an effort to simulate masking effects in order to keep the authenticity of the reference distribution intact: Reducing the information present in the current Fourier-shell by random sub-sampling has the effect of widening the resulting reference distribution, effectively simulating the mask-induced dependencies that lead to a wider distribution for the actual measurements as well.

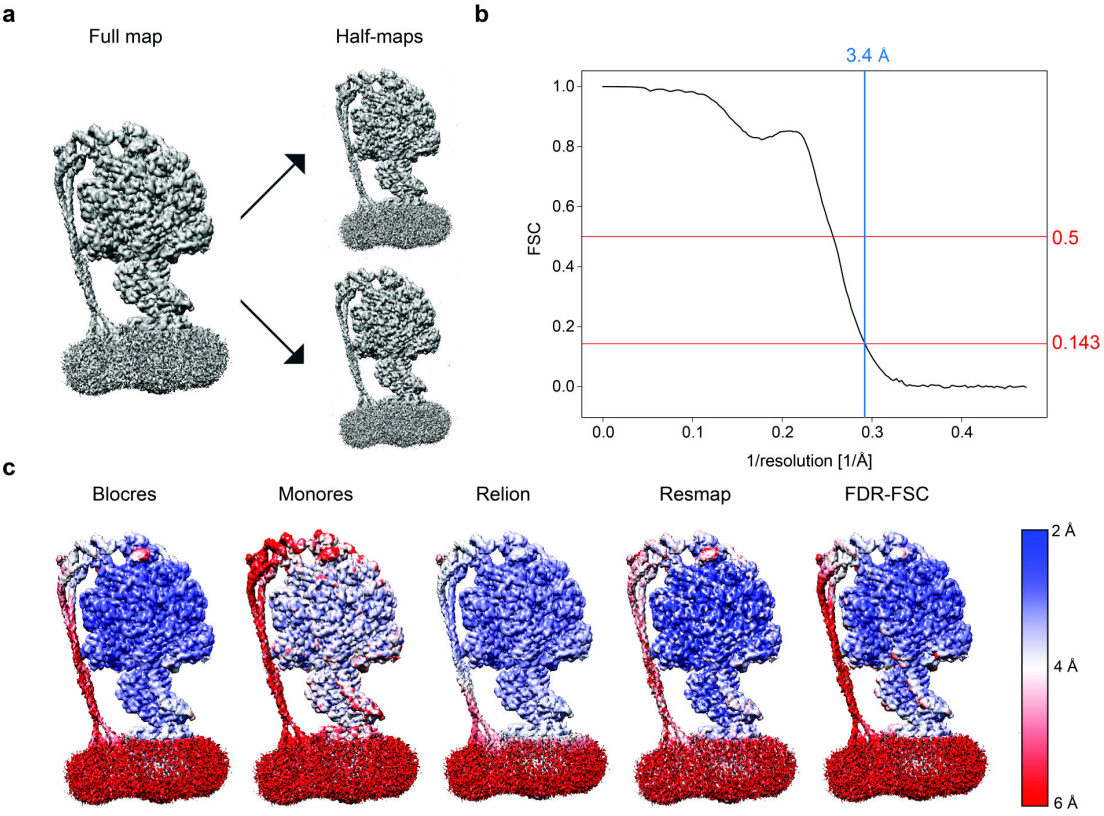

FIGURE 1.6: ATP synthase (EMD-9333) used as example for resolution estimates. a) Two half-maps can be generated from two random half-sets of the particles. b) Top: Formula for FSC (van Heel and Schatz, 2005). Bottom: Typical FSC curve, displaying high correlation for low frequencies, and dropping close to zero for high frequencies. FSC 0.143 and 0.5 are marked with red horizontal lines. The resolution value for a threshold of 0.143 is at 3.4 Å, marked with a blue vertical line. The measurement is conducted without masking, using FDR-FSC. c) Examples of popular tools for Resolution estimation. Usually, local resolution is visualized as projected onto the map. From left to right: Blocres (Cardone, Heymann, and Steven, 2013), Monores (Vilas et al., 2018), RELION (Zivanov et al., 2022), ResMap (Kucukelbir, Sigworth, and Tagare, 2014) and FDR-FSC (Beckers and Sachse, 2020).

Later, local resolution estimates became prominent as a complementary metric, as experimental cryo-EM maps very often do not display homogenous quality across the volume: Some parts of a protein map may be better resolved than others, visibly as blue regions (high resolution, around 2 Å) and red regions (lower resolutions, up to 6 Å) in Figure 1.6 c. Map heterogeneity can have many causes: Preferred particle orientation can lead to anisotropic map quality, and sample heterogeneity or particle flexibility may lead to large spatial quality differences within the map. In these cases, one global resolution value is not sufficient as a metric to define map

quality, and local resolution estimates are required: they allow for detailed quality analysis, provide guidance for model building and enable to identify structural features that can be interpreted with confidence.

Many different approaches for calculating local resolution exist (Figure 1.6, c): Often, local resolution estimations are conducted similar to global estimates by employing the FSC, only that repeated locally confined measurements with different thresholds are conducted (Blocres and FDR-FSC, Cardone, Heymann, and Steven, 2013; Beckers and Sachse, 2020). A second popular method is based on local feature evaluation as for Resmap and Monores (Kucukelbir, Sigworth, and Tagare, 2014; Vilas et al., 2018). One of the most important variables for local resolution estimates is the window size or radius, defining the range around each location within which local resolution estimates are conducted. This range is the defining factor for the available information content for correlation measurements, therefore limiting the measurement accuracy, or, if chosen too large, the measurement locality. Although resolution estimation for refined maps, as derived from structure determination algorithms such as single particle cryo-EM or subtomogram averaging, has been long established, quality estimation methods for images and volumes derived from single measurements, such as micrographs, tilt-series and tomograms, have barely been investigated. So far, mostly indirect indicators are used to judge quality, most commonly the CTF (Tegunov and Cramer, 2019; Elferich et al., 2024).

1.2 Biological membranes

Membranes are an essential component of living cells, fundamental to cell survival and functioning. They form the barrier between a cell and the environment and also define interior cell compartments. Nevertheless, membranes are more than passive barriers: They are fundamental for the communication between cell compartments, ions or molecules pass through them and membrane components adopt different conformations (Cantor and Gennis, 1989). The primary component of

cells are phospholipids. Most commonly, phospholipids arrange as lipid bilayers, where the hydrophilic head groups face outward on both sides of the membrane, while the hydrophobic fatty acid tails orient inward toward each other, creating a water-impermeable barrier (Alberts et al., 2003). Many enzymes with a variety of functions are attached to membranes. Some catalyze reactions with reactants on both sides of the membrane, some facilitate molecular transport or are involved in sequential reactions of several enzymes that are concentrated in the membrane (Cantor and Gennis, 1989).

1.2.1 The ESCRT-III protein complex and Vipp1

One such protein family that attaches peripherally to membranes is the endosomal sorting complex required for transport-III (ESCRT-III) family. The ESCRT-III protein complex is a key component of the evolutionarily conserved ESCRT machinery found in eukaryotic cells (Babst et al., 2002). This protein system plays an essential role in membrane dynamics, orchestrating the remodeling and scission of cellular membranes. ESCRT-III is as a central player in these processes. ESCRT-III function is mediated by the assembly of its subunits into dynamic higher-order polymers, that adopt distinct helical filament architectures to drive membrane deformation and scission events (Park et al., 2024).

Evolutionary and structural analyses show that bacterial proteins phage shock protein A (PspA) and vesicle-inducing protein in plastids 1 (Vipp1) are homologous to ESCRT-III (Liu et al., 2021). Vipp1 is a highly conserved protein in photosynthetic organisms (Zhang et al., 2012; Vothknecht et al., 2012). Functionally, it is known to be involved in thylakoid membrane biogenesis and remodeling, but also membrane stabilization and/or maintenance (Gutu, Chang, and O'Shea, 2018; Siebenaller, Junglas, and Schneider, 2019; Junglas and Schneider, 2018). Vipp1 binds to negatively charged lipds (Heidrich et al., 2016), and the binding is modulated by membrane stored curvature elastic stress. Stored curvature elastic stress is a

physical tension that develops within lipid bilayers when membrane components are forced into energetically unfavorable conformations, such as when lipids like phosphatidylethanolamine and diacylglycerols predominate within the membrane (McDonald et al., 2015). This stress creates hydrophobic cavities known as lipid-packing defects, that correlate with increased membrane curvature.

1.2.2 The membrane-enveloped virus SARS-CoV-2

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a betacoronavirus belonging to the family Coronaviridae. SARS-CoV-2 emerged as the causative agent of COVID-19 (Coronavirus Disease 2019) in 2019 as an enveloped, positive-sense RNA virus, with a genome of approximately 30,000 nucleotides (Wu et al., 2020; Kim et al., 2020). The SARS-CoV-2 genome encodes four structural proteins, sixteen non-structural proteins and multiple accessory proteins (Figure 1.7) (Mariano et al., 2020; Gorkhali et al., 2021; Kakavandi et al., 2023). The structural proteins that are essential for cell entry and virus assembly make up the spike (S) protein, nucleocapsid (N) protein, membrane (M) protein, and envelope (E) protein. The sixteen non-structural proteins (NSPs 1-16) are essential for viral replication, RNA transcription, and immune evasion.

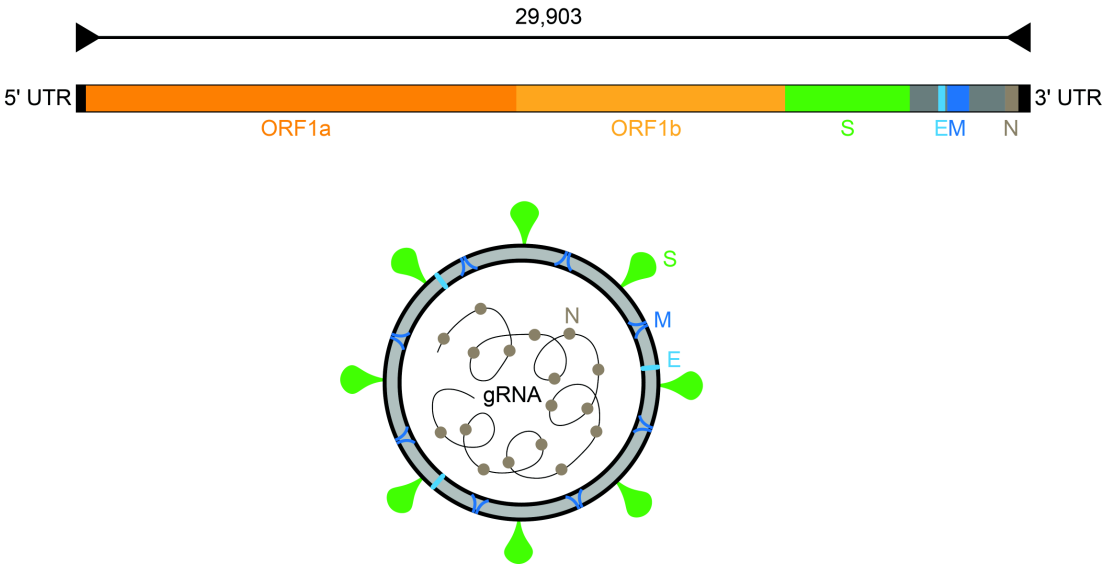

FIGURE 1.7: The SARS-CoV-2 genome and architecture. Top: Illustration of most important components of the SARS-CoV-2 genome. Bottom: Illustration of the architecture of the SARS-CoV-2 virion with the four structural proteins. The figure was conceptually inspired by Kim et al., 2020.

Infection cycle

SARS-CoV-2 initiates infection through the interaction between the S protein and the host cell receptor ACE2 (angiotensin-converting enzyme 2). The S protein consists of two subunits, S1 and S2. S1 contains the receptor-binding domain (RBD) which recognizes ACE2 (Hoffmann et al., 2020), while S2 mediates the fusion of viral and cellular membranes. Binding to ACE2 leads to conformational changes of S1, exposing the S2' cleavage site. The virus then enters the cell via one of two pathways: In the presence of transmembrane serine protease 2 (TMPRSS2), the S protein is cleaved at the exposed S2' site at the cell surface, allowing the fusion of the virion and the host cell membrane. Alternatively, when TMPRSS2 is not present, the virus-ACE2 complex is internalized via clathrin-mediated endocytosis. Within the resulting endolysosomes, S2' cleavage can be performed by cathepsins. Both pathways result in conformational changes of the S2 subunit, membrane fusion and ultimately to the release of viral RNA into the cytoplasm (Jackson et al., 2022).

Once in the cytoplasm, host ribosomes start translating the viral RNA. First, the

two large open reading frames (ORF) 1a and 1b (ORF1a and ORF1b) are translated into polyproteins pp1a and pp1ab, yielding the sixteen NSPs after proteolytic cleavages performed by viral proteases. NSP1 mediates the interruption of host mRNA translation, and other NSPs assemble into the replication-transcription complexes (RTCs), facilitating viral RNA synthesis (Malone et al., 2022). NSPs, together with host cell factors, are also responsible for cell membrane remodeling, initiating the biogenesis of replication organelles. This network of modulated endoplasmic reticulum (ER) membranes includes double-membrane vesicles (DMV), where viral RNA synthesis occurs shielded from immune responses (Malone et al., 2022). The synthesized RNA can be transported into the cytoplasm via a pore that spans the double membrane (Wolff et al., 2020), allowing more translation to occur or packaging into newly forming virions (V'kovski et al., 2021).

Virus assembly occurs at the ER–Golgi intermediate compartment (ERGIC), where the virus obtains its membrane envelope through budding processes that incorporate ERGIC membranes into the developing virions. (Sergio et al., 2024). M, E, and S are transmembrane proteins synthesized in the ER, while N is a protein synthesized in the cytosol (Katiyar et al., 2024). First, the M protein induces membrane curvature via multimerization, bending the ERGIC membrane into the ERGIC lumen and determining where assembly occurs. M also recruits the three other structural proteins to the budding site, N, E and S. The N protein binds to viral genomic RNA and the E protein is together with M and N necessary for proper viral particle formation (Boson et al., 2021; Katiyar et al., 2024). The last step is the virion egress trough exocytosis. Several egress pathways have been proposed, including the classical secretory route through the Golgi network, an alternative lysosomal exocytosis pathway, and a direct route via recycling endosomes (Sergio et al., 2024).

Replication organelle

SARS-CoV-2 replication heavily depends on the membrane-derived replication organelles (ROs). They provide adequate concentration of necessary macromolecules

for RNA-synthesis and prevent innate cytosolic immune detection (V'kovski et al., 2021). The primary membrane source for RO formation is the ER, that is heavily remodeled upon viral cell entry. The restructured membranes emerge as zippered membranes and DMVs, the primary sides for RNA synthesis (Figure 1.8). In the DMVs, molecular pore complexes clamp together the outer and inner membranes at several sites (Klein et al., 2020), acting as stabilizing spacers of the double membranes. The pore also allows RNA transportation across the double membranes. It is formed by NSP3 and NSP4, where NSP3 is localized in the outer membrane and NSP4 sits in the inner membrane (Huang et al., 2024). NSP3 and NSP4 are also the minimal viral components sufficient for the formation of native-like DMVs and sufficient to induce DMV formation. However, several studies have shown that NSP3-4 induced DMVs have a lower diameter of approximately 100 nm (Zimmermann et al., 2023), while the average diameter of DMVs in SARS-CoV-2 infected cells was found to be significantly larger with 300 nm (Klein et al., 2020).

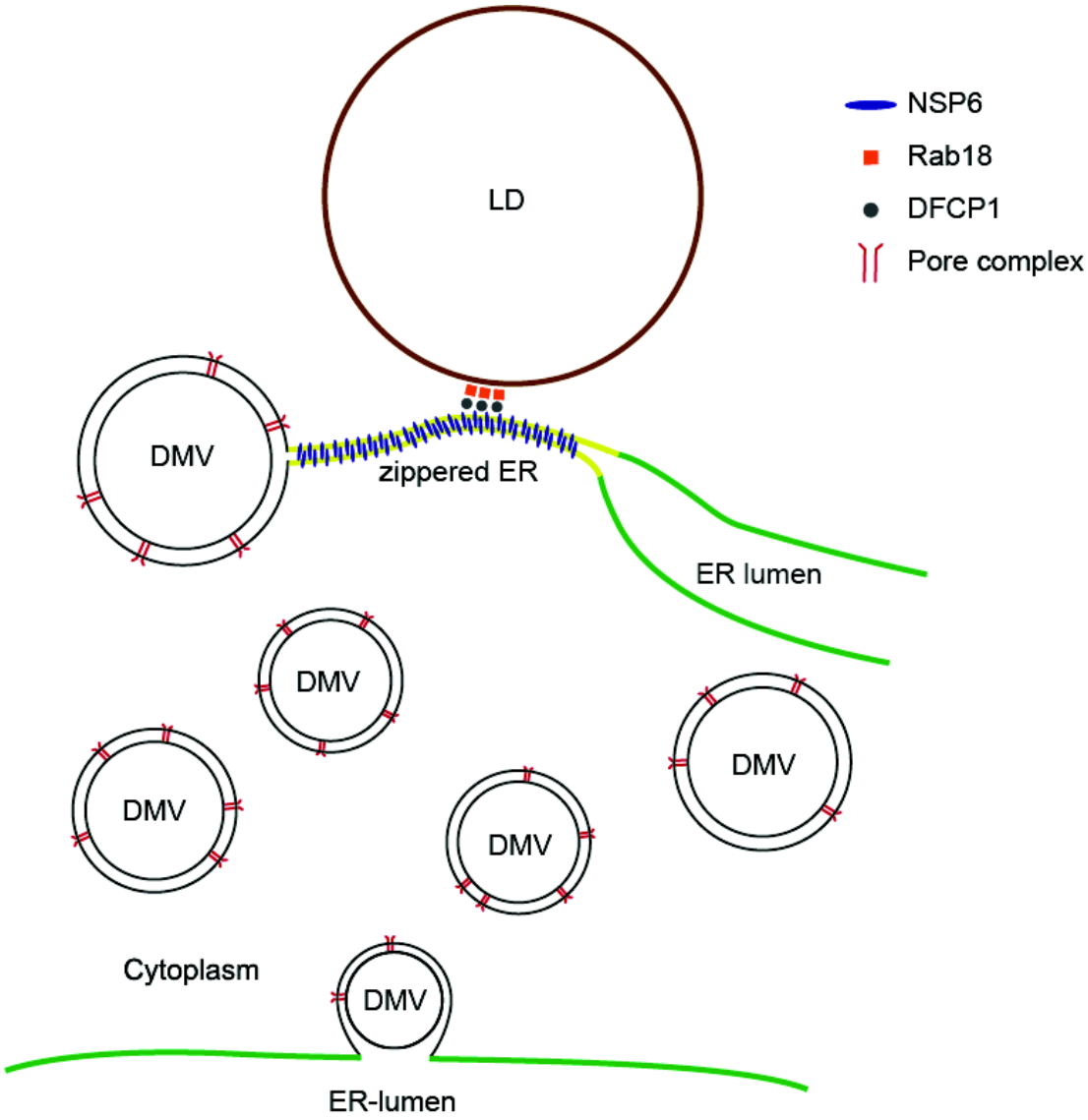

FIGURE 1.8: Cartoon depiction of a DMV cluster. DMVs slightly vary in diameter, each containing several pores (red), constituted by NSP3 and NSP4. Zippered ER (yellow) occasionally connects ER (green) with DMVs. DFCP1 interacts with Rab18, potentially recruiting LDs to the ER. The figure was conceptually inspired by Ricciardi et al., 2022.

Another viral protein contributes to RO formation in important ways: NSP6, the third transmembrane nonstructural protein. While NSP3/NSP4 are sufficient for DMV formation, NSP6 seems to be required for proper architectural organization of the RO. NSP6 forms homo-dimeric complexes, zippering ER membranes that connect DMVs and the ER. In addition to contributing to the spatial organization of DMV clusters, NSP6-mediated ER zippering establishes selective communication

tracks that regulate molecular traffic between the ER lumen and DMVs: While permitting free lipid flow, it excludes ER luminal proteins from accessing the DMVs (Ricciardi et al., 2022). Together with the host protein DFCP1 (double FYVE-containing protein 1), a component of the autophagy machinery, and Rab18, which together form a LD-tethering complex, NSP6 may also mediate LD recruitment to the ER zipper, providing an additional lipid source (Ricciardi et al., 2022). DFCP1 has been shown to play a crucial role in SARS-CoV-2 infection, and DFCP1 knockdown drastically reduces DMV biogenesis and SARS-CoV-2 replication (Twu et al., 2021).

DFCP1

DFCP1 is one of the host proteins required by SARS-CoV-2 to facilitate its replication. It is known to be part of the autophagy machinery, a pathway in eukaryotic cells for degrading their own obsolete proteins or organelles (Takeshige et al., 1992). Autophagy can be triggered by nutrient starvation, degrading cytoplasmic content in an unspecific manner, or selective targeting of damaged organelles and protein aggregates through specific receptors (selective autophagy) (Feng et al., 2014; Zaffagnini and Martens, 2016). Autophagosome are cargo engulfing, double-membrane vesicles that form during autophagy, ultimately fusing with lysosomes, leading to cargo degradation (Takeshige et al., 1992; Mizushima, Yoshimori, and Ohsumi, 2011). These autophagosomes show structural similarity to the viral DMVs, both featuring characteristic double-membrane architectures and sharing common membrane origins, predominantly from the ER (Axe et al., 2008; Roingeard et al., 2022). Thus, links between autophagosome formation mechanisms and DMV biogenesis have been suggested (Blanchard and Roingeard, 2015; Roingeard et al., 2022). DFCP1 contains two FYVE domains that specifically bind to phosphatidylinositol 3-phosphate (PI3P), a lipid enriched at specialized ER subdomains known as omegasomes. These omegasomes are omega-shaped ER structures that serve as autophagy initiation platforms where phagophores begin

their formation and expand to engulf cargo during the early stages of autophagy, ultimately developing into mature autophagosomes (Axe et al., 2008; Hayashi-Nishino et al., 2009). DFCP1 further exhibits ATPase activity, that is stimulated by membrane binding, and it has further been shown that DFCP1 dimerizes in an ATP-dependent manner (Nähse et al., 2023; Ismail, Naismith, and Kast, 2023). It is known that DFCP1 mediates ATPase-driven omegasome constriction during selective autophagy, and its knockdown inhibits selective autophagy (Nähse et al., 2023). DFCP1 also interacts with Rab18 for LD localization, mediating ER-LD contacts (Xu et al., 2018): It has been observed that DFCP1 overexpression increases LD size and enhances ER-LD contacts, whereas DFCP1 knockdown has the opposite effect, reducing ER-LD contacts (Li et al., 2019).

Chapter 2

Aims

This thesis aims to advance the field of cryo-electron tomography through the development and application of robust and automated analysis methods. Within that framework, the work encompasses two primary objectives: First, the establishment of automated segmentation and analysis workflows focused on membrane morphology characterization in the service of ultrastructural characterization. The developed methodologies are applied to address fundamental biological questions, specifically investigating the membrane remodeling functionality of the Vipp1 protein and elucidating critical aspects of the SARS-CoV-2 replication cycle. In the latter context, this work examines how host protein DFCP1 and viral non-structural proteins NSP3, NSP4, and NSP6 collectively influence the morphology of the SARS-CoV-2 replication organelle. Second, the creation of a robust resolution estimation tool to asses the quality of cryo-electron tomograms, enhancing data reliability and processing efficiency. This tools aims at facilitating more informed decision-making in data processing workflows and providing a framework for evaluating how imaging parameters and processing methodologies impact data quality. Through these integrated approaches, this work seeks to enhance both the technical capabilities and biological insights achievable through cryo-electron tomography, and to contribute to the understanding of membrane-associated cellular processes and viral replication mechanisms.

Chapter 3

Automated tomogram analysis: Membrane remodeling activity of Vipp1

The work detailed in this chapter is published, and some of the text sections, figures and tables in this chapter were taken from the publication (Junglas et al., 2025) as indicated by a star (*) in the section header, figure legend or table description. Section headers marked with double stars (**) indicate content that has been substantially adapted and expanded from the original publication. Figure legends or section headers with no star indicate figures or sections are not present in the publication. All tomograms presented in this chapter were collected by Dr. Benedikt Junglas.

3.1 Introduction**

For cryo-electron tomography data, multiple steps of computational processing are necessary to reconstruct three-dimensional tomograms from raw tilt-series. These include motion correction, contrast transfer function (CTF) estimation, tilt-series alignment and tomogram reconstruction. While these steps are largely streamlined and can be quickly conducted in an automated manner, the subsequent downstream analysis remains challenging. In the service of processing a large tomography dataset in an unbiased manner, we established a workflow automated tomogram

annotation and quantitative analysis. We further used the annotated tomogram (segmentation) to derive particle coordinates and orientations, making structure determination via subtomogram averaging possible. This workflow was applied to investigate the membrane remodeling activity of Vipp1.

Vipp1 is a member of the ESCRT-III superfamily (Liu et al., 2021). Vipp1 binds peripherally to the cytoplasmic membrane surface (Heidrich et al., 2016) and is known to be involved in thylakoid membrane biogenesis and remodeling (Gutu, Chang, and O'Shea, 2018; Siebenaller, Junglas, and Schneider, 2019; Junglas and Schneider, 2018). Vipp1 can form large homo-oligomeric complexes, such as rings of various diameters and rods (Gupta et al., 2021; Theis et al., 2019). Vipp1 rings consist of six to seven stacked layers that are formed from monomers arranged in radial symmetry (Figure 3.1, a). The ring diameters and rotational symmetries can differ (Gupta et al., 2021). The Vipp1 monomer was shown to contain six α-helices, and a seventh C-terminal α-helix has been predicted, but could not be resolved with cryo-EM (Vothknecht et al., 2012; Gupta et al., 2021). α0 is the N-terminal membrane-anchoring helix, and α5 forms the C-terminal end, located in the periphery (Figure 3.1, b). The monomers differ at the hinge regions between α3 and α4 and between α4 and α5 and in the length of α4 (Figure 3.1, c) (Liu et al., 2021). Vipp1 can assume different assembly states, with Vipp1 carpet type assemblies on membranes being of particular interest for this study (Junglas et al., 2020). As of the conditions for VIPP1 membrane binding, it has been shown to depend on the degree of membrane stored curvature elastic stress, which is linked to increased membrane curvature (McDonald et al., 2015).

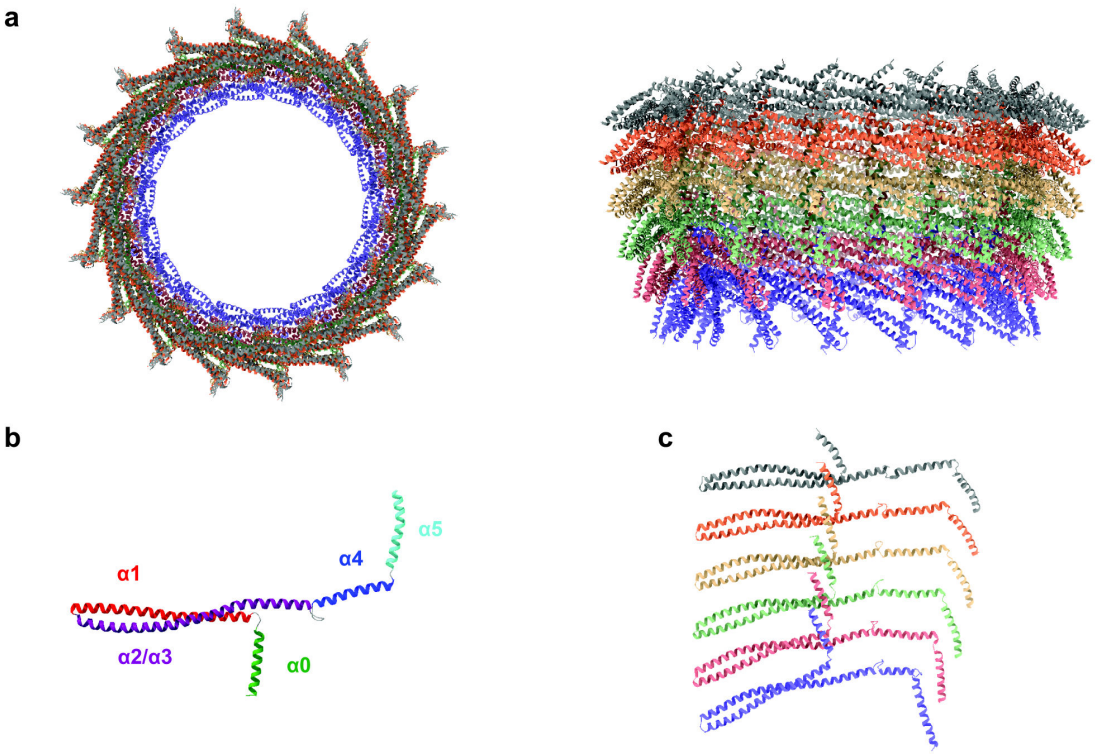

FIGURE 3.1: Vipp1 momomer and ring structure (Gupta et al., 2021, taken from pdb entry 7O3Z). a) Vipp1 stacked ring with C18 symmetry from different orientations. b) Vipp1 monomer with labeled alpha-helices. c) Five stacked Vipp1 monomers.

In this setting, a biochemically reconstituted system was used to investigate Vipp1 interactions with membranes: Vipp1 was purified under denaturing conditions and refolded in the presence of *Escherichia coli* polar lipid extract (lab work conducted by Dr. Benedikt Junglas). As a denaturing agent, urea was used (Bennion and Daggett, 2003) (More details about the sample can be found in the Methods section 3.4.1). In the presence of lipids, Vipp1 can assemble into various structures, and the potential mechanistic connections between them for Vipp1 to perform its functions have yet to be explored. The focus of this study were Vipp1 carpets (Junglas et al., 2020), an assembly type of Vipp1 into a stable structure on flexibly curved membranes. To that end, we investigated the relation between membrane curvature and Vipp1 presence in a quantitative manner. Additionally, for the structural analysis of the carpets, we conducted subtomogram averaging, yielding three structures.

As the membrane-attached Vipp1 assemblies were not suitable for 3D reconstruction from 2D images using single particle cryo-EM, due to the lack of consistent three-dimensional views, we elucidated the 3D structure of these assemblies by cryo-electron tomography, and a large tomogram dataset was recorded (tomograms recorded by Dr. Benedikt Junglas). Thus, in this work, I explore methods for the computational analysis of tomograms, while also investigating the before mentioned biological questions. The focus here is on the computational work that I have conducted, while I also try to outline the biological contributions this work yielded. If not mentioned explicitly otherwise, the described work was conducted by me.

3.2 Results

3.2.1 Analysis of membrane curvature in relation to Vipp1 presence**

In order to investigate changes in membrane curvature in relation to presence of membrane covering Vipp1 carpets, we used the surface morphometrics toolkit to estimate membrane curvedness (Barad et al., 2023). We investigated the relation between membrane curvature and presence of Vipp1 carpets in a quantitative manner with a custom Python script: Coordinates were sorted into fixed-size bins according to their curvedness and, for each bin, the ratio of membrane coordinates with a present proximal carpet segmentation was determined resulting in the corresponding occupancy value between 0 and 1 (0, no coverage; 1, full coverage). The resulting graph shows linearly increasing membrane occupancy with increasing membrane curvedness between curvedness values of 0.02–0.06 (Figure 3.2). For high-curvature regions (curvednesses larger than 0.06), the graph indicated a saturation of Vipp1 membrane binding. Our data thus shows that as Vipp1 occupancy increases, membrane curvature increases as well. It is thus likely that the

Vipp1 membrane reshaping activity induces higher membrane curvature.

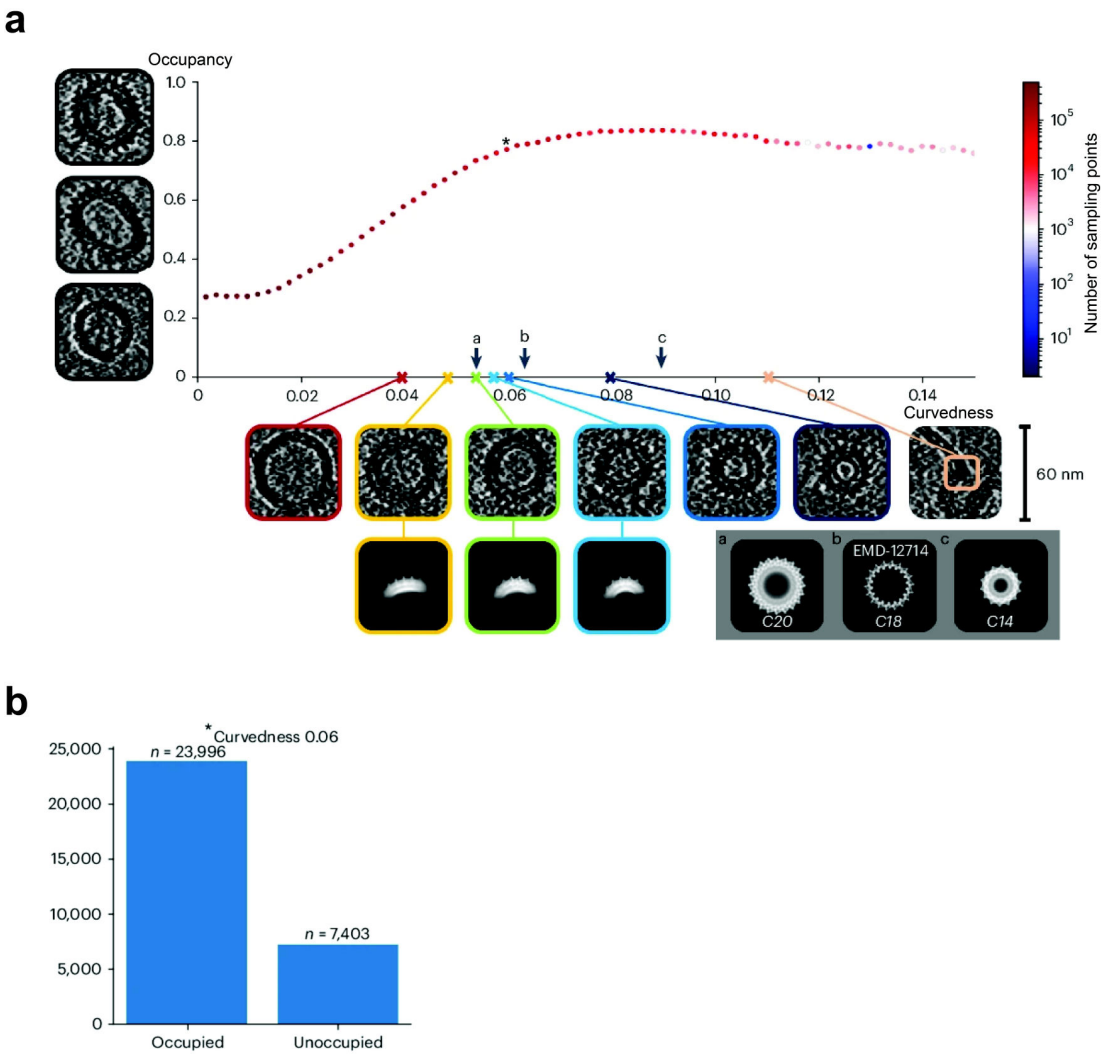

FIGURE 3.2: *Membrane curvedness with respect to occupancy, representing the Vipp1 membrane coverage. a) Left of y-axis: Example images of different occupancies with fully covered, partly covered and uncovered vesicles. Below s-axis: Example images of vesicles and other shapes with specified curvedness. Below that, the 2D averages of subtomogram averages corresponding to the indicated curvedness. Bottom right: The 2D average of three Vipp1 ring structures with different C-symmetry (and thus different curvature) as a comparison: *a* C14 ring containing membrane, *b* C18 ring without membrane and *c* C20 ring containing membrane. Taken from literature as a comparison (Junglas et al., 2025 for *a* and *c*, EMD accession 12714 for *b*. b) illustrative examination for all segmentation points with a curvedness of 0.06, resulting in the presented occupancy ratio.

Due to the continuous nature of the Vipp1 coating, the determined membrane coordinates could be used as particle locations for the Vipp1 carpet structure determination. The membrane normals estimated by the surface morphometrics toolkit

could additionally be used to facilitate this process, as these membrane normals allowed restriction of angular alignments during subtomogram averaging.

3.2.2 Determining the structure of Vipp1 carpets**

We used RELION 4 (Zivanov et al., 2022) to conduct subtomogram averaging of Vipp1 carpets. As initial particles, the membrane coordinates resulting from the above described membrane segmentation were used, and coordinates were selected according to the presence of Vipp1 carpet structures. We estimated the initial orientations of particles using their corresponding membrane normal with a custom Python script. To create an initial model from a subset of homogenous particle picks, we clustered the particles according to their corresponding membrane curvedness. After several rounds of 3D classifications and refinements, three subtomogram averages with different curvedness emerged, with the best having a resolution of around 18 Å (Figure 3.3, a). The carpet structures were remarkably similar to Vipp1 rings with respect to spike separation distance and monomer stacking distance, thus displayed the same characteristic features (that is, the typical α_1-α_4 contacts in the core, α_1-α_3 hairpin-to-α_5 contacts in the periphery and α_0-mediated membrane contacts, see Introduction 3.1. Finally, we fit an atomic Vipp1 ring model (PDB:7O3Z) into the best resolved structure. To facilitate rigid model fitting and get a better understanding of the Vipp1 carpet structure, we created a symmetrized version of our density map (details in the Methods section 3.4.7).

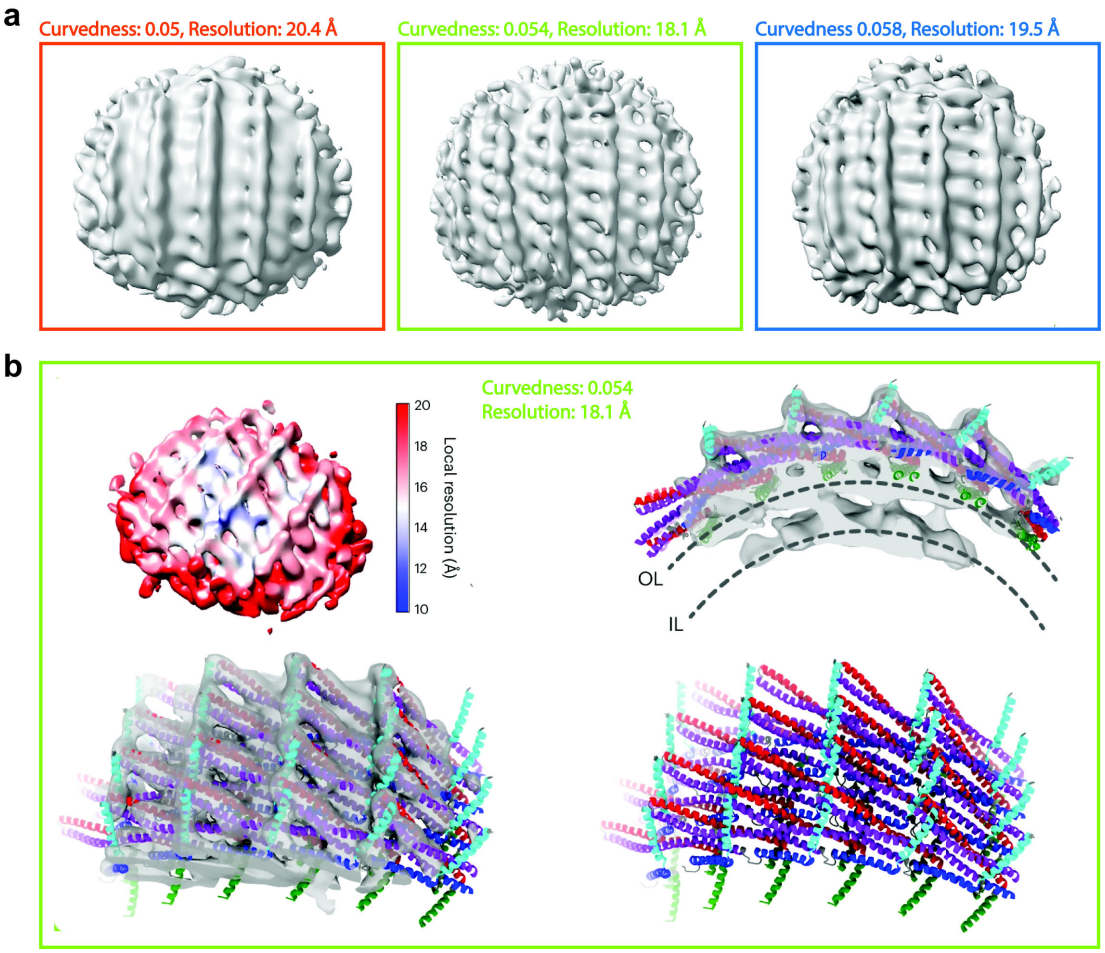

FIGURE 3.3: *Subtomogram averages of Vipp1 carpet structures. a) subtomogram averages for curvedness 0.05, 0.054 and 0.058. b) Local resolution map yielding an average resolution of 18 Å and density maps with fitted atomic model based on the C20 symmetrization (adapted from PDB 7O3Z). OL, outer leaflet; IL, inner leaflet.

3.3 Discussion

3.3.1 Establishing a workflow for membrane surface analysis

To the end of quantifying membrane properties in relation to protein presence in a large cryo-electron tomography dataset, we established a largely automated workflow, including membrane segmentation, particle localization and membrane curvature analysis.

For membrane segmentation and particle localization, we used Dragonfly (Object Research Systems, Heebner et al., 2022). Dragonfly offers convenient and effective

methods for the creation of labeled training data: It allows to iteratively label small 2D-frames within tomogram z-slices, retrain an already pretrained U-net, and predict new frames, which can then be corrected in order to extend the training set. As new methods develop, whenever available, generalized models are usually preferable, making manual labeling completely unnecessary and often even enhancing performance, especially if training sets are large and steadily updated as for the newly developed membrain-seg (Lamm et al., 2024). However, often the features of interest are specific and no generalized model available, and in this context, tools such as Dragonfly provide an efficient and convenient environment for model training. In this study, we were able to reliably locate Vipp1 attached to membranes, which results in a unique visual pattern, via retraining a pretrained U-net in Dragonfly. We were also able to successfully segment lipid bilayer membranes with Dragonfly, as no generalized model was available at the time the study was conducted.

We used the surface morphometrics toolkit (Barad et al., 2023) to determine curvedness as a measure for membrane bending. Such tools, providing ways for the quantification of 3D membrane surface properties, are integral to make the analysis of large tomography datasets possible. However, while there may be cases which are worth streamlining, like quantifying morphological properties of mitochondria (Barad et al., 2023), it is difficult to fully automate the quantification of tomography data, even in the case of membranes, where the combination of automated membrane segmentation via a generalized model and automated surface property quantification can be a powerful basis for further analysis. But especially in an *in situ* context, it then often requires extensive manual curation to extract the membrane structures of interest and discard the rest. As techniques for *in situ* cryo-electron tomography develop and data collection becomes faster, there is an increasing need for automating segmentation and quantification to the extend possible.

Due to the unique circumstances in this study, we conducted automated analysis

where possible, with network training for detecting the features of interest remaining as the biggest manual task. One further challenge was the interface between different software, and a major task for automation was integration. For the quantitative analysis in this study, there was a need for the integration of Warp, AreTomo, Amira, Dragonfly, surface morphopmetrics toolkit and the final data management and plotting conducted in Python. Integration either requires manual work or automation, which can be accomplished via pipeline creation, for example in Python or Bash. However, this is difficult to streamline, as tools constantly develop.

3.3.2 Approaches for determining the VIPP1 carpet structure

To facilitate the processing of this dataset, some unconventional processing approaches had to be explored. In order to conduct structure determination in cryo-EM, as we did via subtomogram averaging, particles need to be located within the tomograms. Conventionally, particles are located in cryo-EM via leveraging existing tools, using either template matching (Tegunov et al., 2021; Wan, Khavnekar, and Wagner, 2024) or neural networks (Wagner et al., 2019; Bepler et al., 2019). For membrane-bound particles, geometrical picking approaches can be used, allowing to assign approximate orientations by exploiting the orientational constraints imposed by membrane embedding (Castaño-Díez and Zanetti, 2019). However, geometrical approaches are usually either used for elliptical shapes or tubes. Other tools allow the detection of membrane-bound particles without any geometrical restraints (Martinez-Sanchez et al., 2020). However, in our specific case, Vipp1 assumes a variety of continuous assemblies on flexibly curved membranes. For this particular case, we used a novel approach: The output of two different tools was used, combining semantic segmentations of membranes and Vipp1 assemblies from Dragonfly (Object Research Systems, Heebner et al., 2022) and taking particle orientations from the surface morphometrics toolkit (Barad et al., 2023). We combined the output of both tools using custom Python scripts in order to determine

initial particle picks. Furthermore, we were then able to estimate the membrane curvedness determined by the surface morphometrics toolkit to create an initial particle set with increased homogeneity via clustering for curvedness, facilitating the creation of an initial model for subtomogram averaging.

Building an atomic model usually requires a well-resolved structure. However, in our case, we only had a low-resolution model, with local resolutions mostly ranging from 14 to 20 Å (Figure 3.3, b). There are many possible explanations for why we were not able to achieve a higher resolution. The main reason may be that the carpet structures within the dataset seem to be curved continuously, which makes it difficult to find a homogeneous particle set. Additionally, subtomograms are noisy by nature, and a large part of the structural density consisted of lipids. These factors made alignments of the Vipp1 carpet structure on top of the membrane difficult, as the membrane tended to drive the alignment procedure. However, for our purposes, a low resolution structure was sufficient: Due to its similarity to the existing Vipp1 ring structures, we suspected that the atomic structure is similar. In order to get a better understanding of the Vipp1 carpet structure and to facilitate model fitting, we sought to impose rotational symmetry on our map. In order to symmetrize the Vipp1 carpet into a full ring, a new approach was explored. The proposed method, based on cross-correlating our structure with different symmetrized versions of itself, yielded a realistic atomic model as a result, and we were able to fit five Vipp1 monomers from an existing ring structure into our density rigidly, achieving a good fit.

3.3.3 Vipp1 forms carpets which bind to highly curved membranes

While the focus of my work within the study was on the processing techniques, the study yielded interesting biological insights. First, we found strong evidence

through our quantitative analysis that Vipp1 preferably binds to highly curved membranes. This is in accordance with previous suggestions that Vipp1 binding depends on membrane stored curvature elastic stress (McDonald et al., 2015) which is associated with increased membrane curvature. The abundance of membrane-bound Vipp1 structures reached a plateau at a curvature corresponding to an approximate vesicle diameter of 30 nm (Figure 3.2). While Vipp1 arrangements were also found on lower curvature membranes (Figures 3.4, a; 3.2), we were only able to determine the structure of Vipp1 carpets for higher curvatures with a curvedness of 0.05-0.06 (Figure 3.3), indicating that Vipp1 assembles on low curvature membranes in loose coats first. Our subtomogram averages of the carpet structures suggest a monomer arrangement similar to that found in Vipp1 rings, except that the assemblies do not close into a ring. Instead, the lattice is relaxed to adapt to lower membrane curvatures and extended to cover larger areas (compare figure 3.6, b and figure 3.1, a).

While my contributions to the overall study were described here in detail, the biological implications of the presented results should be understood in the context of the full study (Junglas et al., 2025). Overall, this work contributed to the proposal of a model for the sequence for Vipp1 membrane interactions: Vipp1 monomers or small oligomers accumulate on low-curvature membranes in loose coats until a critical concentration is reached locally. At these membrane areas, Vipp1 monomers oligomerize in the plane of the membrane and the formed assembly patches start to induce curvature on the membrane, ultimately resembling the determined carpet structures. As Vipp1 preferentially binds to curved membranes (that is, membrane binding induces curvature), Vipp1 bulges start to grow on the membrane increasing the local curvature, stacking oligomers away from the membrane and thereby internalizing the membrane through helix $\alpha 0$ interactions. In this way, the initial bulges further emerge into the Vipp1 rings subsequently elongate into rods that finally bud off the membrane.

3.4 Materials and Methods

3.4.1 Sample preparation*

Note that the sample was prepared by Dr. Benedikt Junglas. As my contributions concern the computational part of the study, sample preparation is only recapitulated here briefly. Detailed descriptions are available in the published study (Junglas et al., 2025).

Vipp1 (*sll0617*) from *Synechocystis* sp. PCC 6803 was expressed in *E. coli* C41 cells using pET50(b)-derived plasmids with C-terminal His-tags. Proteins were purified under denaturing conditions (6 M urea) using Ni-NTA chromatography. After optional His-tag cleavage with 3C protease, proteins were concentrated and stored at −20 °C. For Vipp1 reconstitution with lipids, urea-unfolded Vipp1 was added to the *Escherichia coli* polar lipid extract film and incubated until the lipid film was resolved. Subsequently, the mixture was dialyzed overnight against 10 mM Tris-HCl pH 8.0 including three buffer exchanges to refold the protein.

3.4.2 Data collection and processing

Here, the detailed parameters used for imaging and during subtomogram averaging are listed. Note that images were recorded by Dr. Benedikt Junglas, while subtomogram averaging was conducted by me.

TABLE 3.1: *Parameters for subtomogram averaging. Data collection was conducted by Dr. Benedikt Junglas. Subtomogram Averaging and all associated procedures were conducted by me. Note that *sample 1*, *sample 2* and *sample 3* refer to the three obtained subtomogram averages. Resolution measurements were done in Relion4 (Zivanov et al., 2022), using FSC=0.143 as a threshold.

Parameter	**Sample 1**	**Sample 2**	**Sample 3**
Magnification	×53,000	×53,000	×53,000
Voltage (kV)	300	300	300
Total dose (e^- per Å^2)	152	152	152
Dose per tilt (e^- per Å^2)	3.7	3.7	3.7
Energy filter slit width (eV)	20	20	20
Defocus	3 μm	3 μm	3 μm
Physical pixel size (Å)	1.7	1.7	1.7
Detector	Gatan K3	Gatan K3	Gatan K3
Acquisition scheme	Dose-symmetric	Dose-symmetric	Dose-symmetric
Tilt range	−60° to +60°	−60° to +60°	−60° to +60°
Tomogram number	123	123	123
Initial subtomograms	526,864	526,864	526,864
Curvedness	0.05	0.054	0.058
Final subtomograms	3,548	7,349	6,544
Global resolution (Å)	20.4	18.1	19.5
Local resolution range (Å)	20–24	12–24	18–25
EMDB	–	EMD-18620	–

3.4.3 Tomogram preprocessing*

Tilt image frames were gain-corrected and motion-corrected using WARP (Tegunov and Cramer, 2019). The resulting tilt-series were aligned, 8× binned and reconstructed by the weighted backprojection method using AreTomo (Zheng et al., 2022). The reconstructed tomograms (Figure 3.4, a) were filtered with a recursive

exponential filter followed by a non-local means filter using Amira (Thermo Fisher Scientific) (Figure 3.4, b).

3.4.4 Segmentation**

For further analysis, it was necessary to label the membrane structures across the full dataset. Additionally, we labeled specifically those membranes which had Vipp1 assembled on top of them as a separate class: The filtered tomograms were segmented in Dragonfly (Object Research Systems, Heebner et al., 2022) by progressively training a U-Net with an increasing number of manually segmented tomogram frames. U-Net is a convolutional neural network architecture that has become the standard approach for segmentation tasks in biological and medical imaging (Ronneberger, Fischer, and Brox, 2015). To make the model robust across tomograms, frames (partial tomogram z-slices) from a total of eight tomograms were labeled. Feature detection focused on two major classes: First lipid bilayers, second Vipp1 located on membranes and third the ice background. The trained U-Net was used to predict these features within 123 tomograms. The resulting segmentations were coarsely corrected for errors, then they were exported as .tif files and assembled into .mrc files for further analysis (Figure 3.4, c).

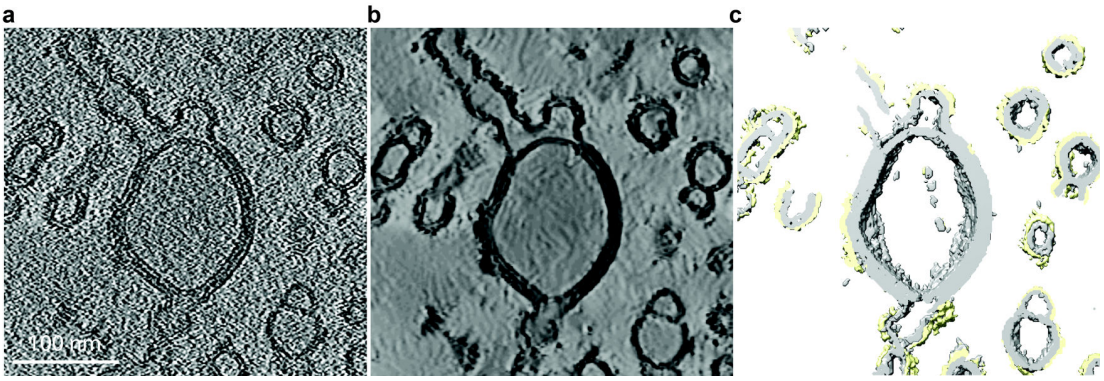

FIGURE 3.4: *Illustrative tomogram z-slices and segmentations. a) z-slice of a reconstructed tomogram, Vipp1 visible on membranes as spikes. b) Filtered tomogram slice. c) same area, now as 3D segmentation. Grey: Lipid bilayer membrane. Yellow: Vipp1.

3.4.5 Membrane curvedness estimation**

To investigate changes in membrane curvature in relation to presence of membrane covering Vipp1 carpets, the membrane curvature had to be estimated, using the surface morphometrics toolkit (Barad et al., 2023) (Figure 3.5). Conceptually, curvature defines how tightly a curve bends by measuring how quickly it deviates from a straight line. The unit of curvature is $\frac{1}{radius}$. For surfaces, infinitely many curves in any direction can be created. The principal directions are those directions where the curvature reaches its maximum and minimum value. With k_1, maximum curvature, and k_2, minimum curvature, the curvedness can be calculated as

$$c = \sqrt{\frac{k_1^2 + k_2^2}{2}} \tag{3.1}$$

quantifying the overall amount of bending at a surface point, regardless of the direction (Salfer et al., 2020). Thus, curvedness was used as a measure of curvature on 3D surfaces. The surface morphometrics toolkit generates a graph representation of membrane surfaces, generating a discrete set of nodes for which the local membrane characteristics are calculated. For each graph node, associated with a 3D coordinate, the corresponding curvedness value and membrane normal are estimated. To filter for adjacent Vipp1 carpet segmentations, the Vipp1 carpet segmentations were transformed into coordinates as well, using a custom Python script. The membrane coordinates with a distance up to 50 Å to Vipp1 carpet segmentations were judged as covered by Vipp1 carpet. To minimize errors, coordinates close to tomogram edges were excluded.

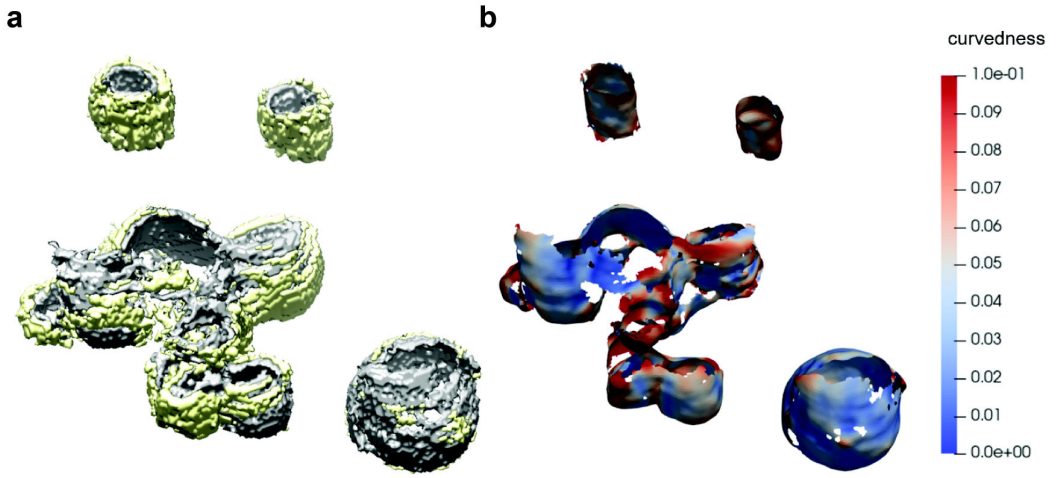

FIGURE 3.5: *Example segmentation and corresponding curvedness. a) Membranes as grey, Vipp1 as yellow. b) Curvedness mapped on membrane surface.

3.4.6 Subtomogram averaging**

For structure determination, RELION 4 (Zivanov et al., 2022) was used in order to conduct subtomogram averaging. CTF estimation was conducted with ctffind4 (Rohou and Grigorieff, 2015). As initial particles, the membrane coordinates resulting from the above described membrane segmentation were used. Coordinates were selected for the presence of Vipp1 carpet structures (within a radius of 100 Å) and a minimal inter-particle distance of 15 Å was enforced to avoid excessive overlaps. The orientation of particles was initially estimated using their corresponding membrane normal with a custom Python script. This normal could then be used to restrain angular searches for all subsequent angular alignment procedures: The angular alignments are defined through three Euler-angles, covering all possible rotational combinations. While the in-plane rotations (rotations around the axis perpendicular to the membrane surface) were left unrestricted, as we had no knowledge about the in-plane orientation of Vipp1, rotations around the other two axis were restrained. To create an initial model from a subset of homogenous particle picks, the coordinates were clustered according to their corresponding membrane curvedness (determined by the surface morphometrics toolkit (Barad et al., 2023) as described previously), using the Scikit-Learn implementation of k-means. After

particle averaging and several rounds of 3D classifications and refinements in 4× binning, a low-resolution model emerged. This model was then used as an initial reference and, from here on, all particles were included. Multiple rounds of 3D classifications and refinements for cleaning up the dataset followed, initially using 4× binning and later on 2× binning. As a further means to exclude poor particles, particles with diverging orientations compared to the majority of particles in close proximity were excluded. Finally, the resulting maps were used to improve predetermined parameters, using RELION's frame alignment and CTF refinement procedures followed by a final round of refinements.

3.4.7 Vipp1 carpet: Atomic model fitting**

In order to get a better understanding of the Vipp1 carpet structure and to facilitate model fitting, we sought to impose rotational symmetry on our map. For the highest-resolution map, the average curvedness of all points from which the final average was calculated (0.054) indicated a substantially higher diameter than the currently existing highest diameter Vipp1 ring model (C18, Protein Data Bank (PDB) 7O3Z). Therefore, PDB 7O3Z was used to initially set the distance to the symmetry axis and rotation. To find the correct symmetry parameters (C symmetry, distance and rotation in relation to the symmetry axis), the map was fitted in ChimeraX (Pettersen et al., 2021) along the x axis. The correlation of the original map with versions of itself with different symmetry parameters applied was then calculated and the correlation was maximized. With this brute-force approach, the best symmetry parameters based on the correlation scores were found. To be able to conduct this search effectively, custom ChimeraX scripts were used. First, translational searches along the x-axis in 1-Å increments were conducted for C18–C22 symmetries. The best correlation for the 0.054 curvedness reconstruction was identified for a C20 symmetry. After optimal distance and symmetry parameters were found, rotational searches were conducted (Figure 3.6, a). Finally, the symmetry,

shift and rotational parameters were applied to the reconstruction and PDB 7O3Z was rigid-body fitted into the resulting map (Figure 3.6, b).

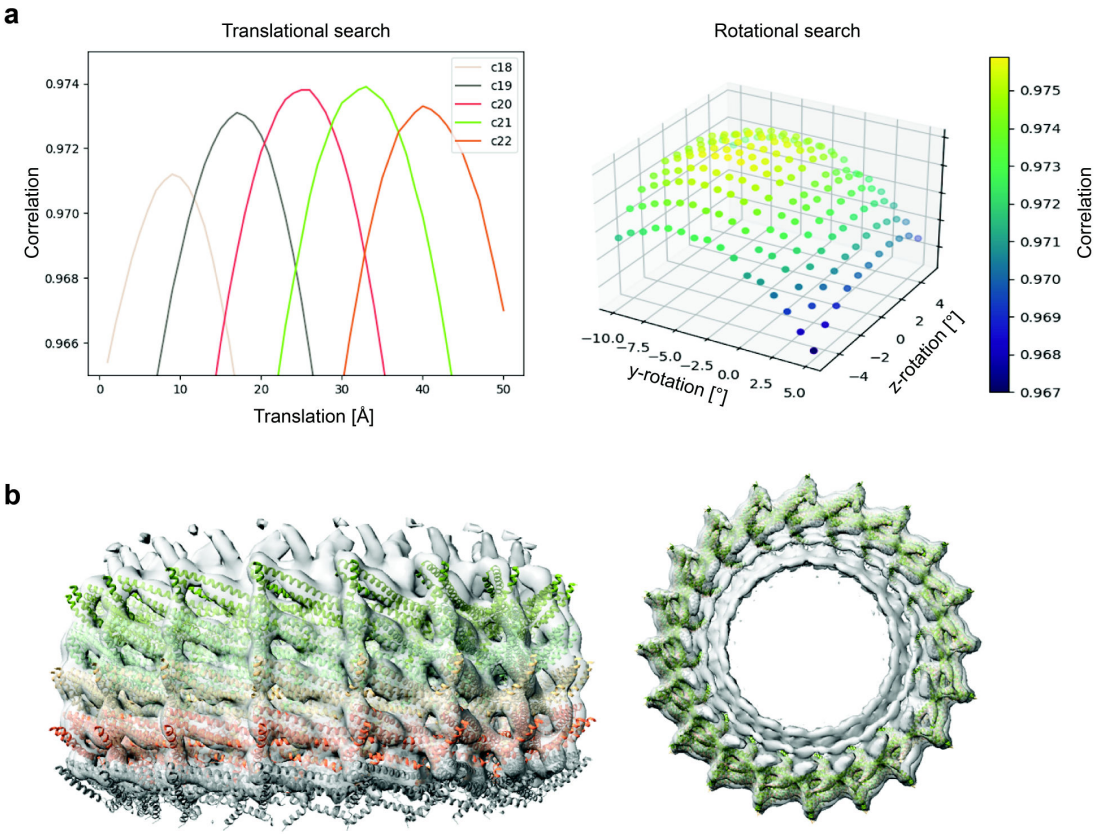

FIGURE 3.6: *Symmetrization and model fitting for the Vipp1 monomer. a) Left: correlation of translational steps with different cyclical symmetries C18, C19, C20, C21, C22 (C20 yielding the highest correlation). Right: translational and rotational search (here only plotted for y and z) over different cyclical symmetries (C18, C19, C20, C21, C22) of the 0.054 curvedness structure. The C20 symmetry gave the best correlation with the subtomogram average reconstruction. b) Symmetrized subtomogram average with rigid-body fitted and symmetrized model PDB:7O3Z.

Chapter 4

Automated tomogram analysis: The morphology of the SARS-CoV-2 replication organelle

This work focuses on developing a computational workflow for automated analysis of membrane morphology in SARS-CoV-2 replication organelles. Experimental setup and data collection were conducted by Kevin Boga.

4.1 Introduction

Double membrane vesicles (DMVs) constitute the SARS-CoV-2 replication organelle (RO), allowing viral RNA synthesis to occur protected from the hosts innate immune response. These DMVs contain molecular pores, which are comprised of two viral proteins: NSP3 and NSP4 (Huang et al., 2024). NSP3 and NSP4 expression is also sufficient for the formation of DMV-like, pore containing structures in transfected cells (Zimmermann et al., 2023). Another viral protein, NSP6, contributes to the formation of the RO: It facilitates the zippering of endoplasmic reticulum (ER). The zippered ER regulates molecular traffic between the ER lumen and DMVs: While permitting free lipid flow, it excludes ER luminal proteins from accessing the DMVs (Ricciardi et al., 2022). NSP6 may also play an additional role in RO formation, as

it potentially helps to ensure lipid supply by mediating lipid dropled (LD) recruitment to the ER zipper via interaction with the host protein DFCP1 (Ricciardi et al., 2022).

DFCP1 is known as an autophagy-related protein, where it localizes to PI3P-rich regions serving as platforms for autophagosome formation (Hayashi-Nishino et al., 2009). Links between viral RO biogenesis and the autophagy pathway have been suggested (Blanchard and Roingeard, 2015; Roingeard et al., 2022): However, large parts of the autophagy machinery seem to be dispensible for SARS-CoV-2 replication. Some specific autophagy components have been shown to play critical roles: Knockdown (KD) of either VPS34, Beclin1 or DFCP1 impair viral replication and DMV formation, with DFCP1 knockdown having the strongest effect (Twu et al., 2021). Thus, DFCP1 was the focus of this study. In the meantime, it has been demonstrated that DFCP1 mediates ATPase-driven omegasome constriction during selective autophagy, thus facilitating the proper formation of autophagosomes (Nähse et al., 2023). DFCP1 additionally mediates LD-ER contacts, and DFCP1 overexpression is known to increase LD size and enhance ER-LD contacts (Li et al., 2019). This work aims to investigate the role of DFCP1 in SARS-CoV-2 RO morphology, specifically focusing on sites of viral RNA synthesis, the DMVs (Klein et al., 2020).

Some interesting morphological differences have previously been reported to occur between DMVs in SARS-CoV-2 infection, DMVs in a NSP3-4 overexpression system and DMVs in a NSP3-4-6 overexpression system: In infected cells, DMVs are on average 200-400 nm in diameter (Klein et al., 2020), and for NSP3-4 and NSP3-4-6 overexpression, DMVs are much smaller. In the case of NSP3-4 overexpression, DMVs have been reported to have an average diameter of 80-100 nm (Ricciardi et al., 2022; Zimmermann et al., 2023; Yang et al., 2025); whereas in the case of NSP-3-4-6 overexpression, the diameter has been reported to be slightly lower (average diameter of 67.5 nm), while DMVs have been shown to more abundant (Ricciardi et al., 2022).

A large set of tomograms was collected to compare DMV morphologiy under different conditions, with the aim of quantifying the effect of DFCP1 on DMV morphology. Thus, differentially transfected VeroE6 cells were used. The viral proteins NSP3 and NSP4 were the focus, as they were sufficient to form a DMV-like structure with pores, and NSP6 was of special interest due to its known effect on DMV morphology and interactions with DFCP1 (Ricciardi et al., 2022). Thus, five different conditions were investigated: 1) NSP3-4 overexpression with endogenous DFCP1 expression 2) NSP3-4 overexpression with DFCP1 overexpression 3) NSP3-4-6 overexpression with endogenous DFCP1 expression 4) NSP3-4-6 overexpression with DFCP1 overexpression 5) NSP3-4-6 overexpression in a DFCP1 knockout (KO) cell line with overexpression of the mutant DFCP1 K193A, which is nucleotide-binding defective (Nähse et al., 2023).

The focus here is on the technical aspects of the downstream image analysis, starting from tomogram segmentation to the final quantification. Most importantly, methods for automating DMV shape analysis are explored. These can be summarized in two major steps: First, as a basis for further processing, DMV instances needed to be extracted. The extraction of isolated DMV instances was a prerequisite for quantitative analysis, and required the labeling of every relevant DMV. Labeling of DMVs as instances was also necessary for the automated localization of pores within the DMV double membranes. Second, the quantification of various DMV characteristics, such as circumference, minimal pore distance or membrane spacing had to be conducted.

4.2 Results

4.2.1 Quantifying the effects of DFCP1 overexpression

This section focuses on the analysis of a small tomography dataset in the service of investigating the effect of DFCP1 overexpression in wild type (WT) VeroE6 cells. As

the aim of this study is the investigation of how DFCP1 affects the morphology of the SARS-CoV-2 DMVs, we were first interested to investigate the effects of DFCP1 overexpression without viral proteins.

The occurrence of a large number of vesicles was striking, as well as the frequent occurrence of lipid droplets (LDs) (Figure 4.1, a). Thus, the subsequent analysis concerned the locations and sizes of these vesicles in order to investigate their spatial relation to LDs, and only tomograms containing LDs were used for the subsequent analysis (9 tomograms).

For the vesicle size, no obvious correlations to LD proximity could be found: The LD distance does not seem to have an influence on the LD size. However, it did become clear that vesicles accumulated in proximity to LDs, as their number decreased with increasing LD distance (Figure 4.1, b). Thus, this analysis facilitated new insights into the structural changes within the cell upon DFCP1 overexpression. Next, in order to examine the effect of DFCP1 on SARS-CoV-2 replication, we wanted to investigate how DFCP1 overexpression affects DMV morphology.

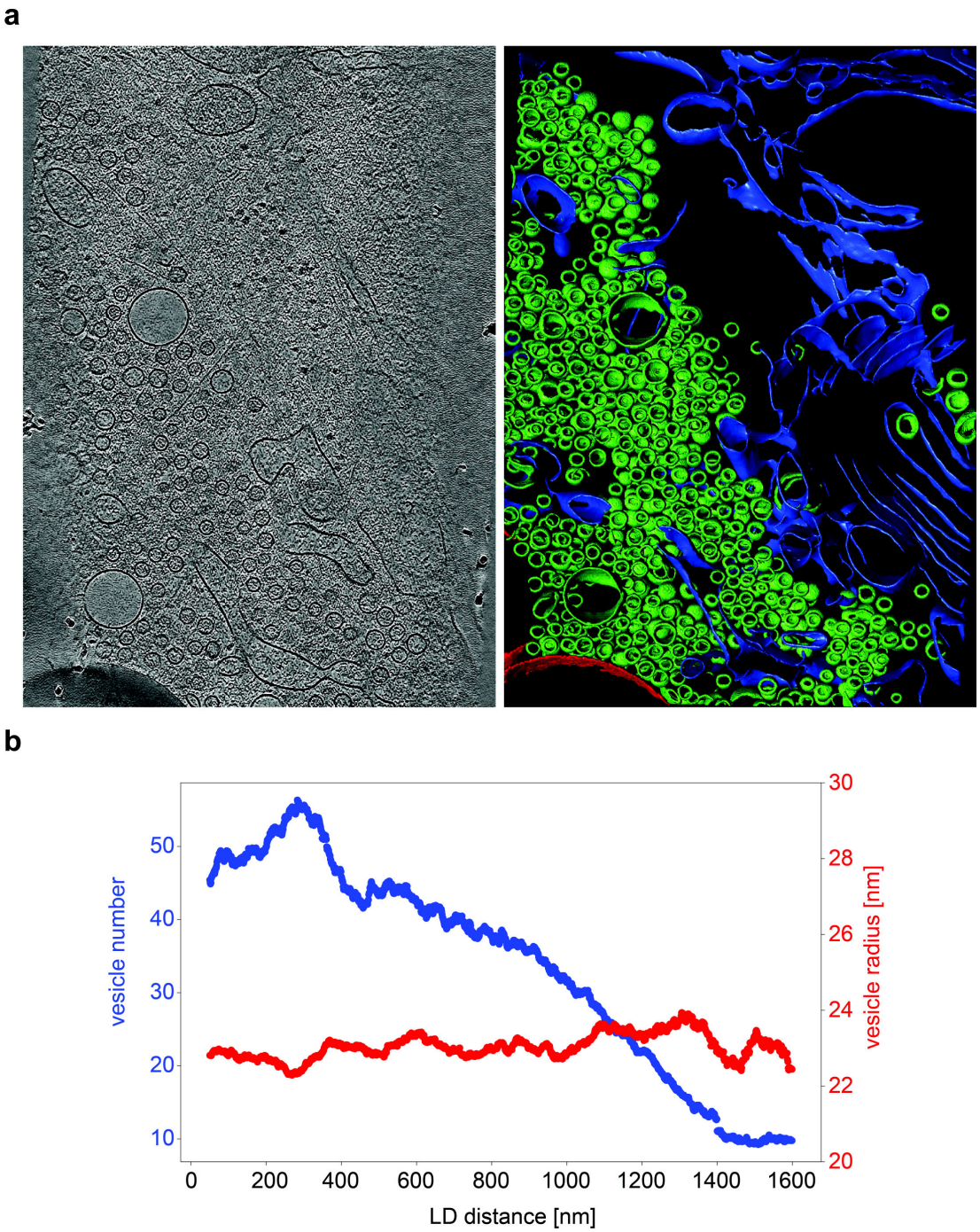

FIGURE 4.1: DFCP1 overexpression results in vesicle shedding. a) Left: z-slice through the denoised tomogram. Right: Segmented tomogram with distinctively colored membranes of lipid droplets (red), vesicles (green) and the remaining membranes (blue). b) Vesicle number (blue) and vesicle average size (red) plotted against lipid droplet (LD) distance. Vesicle number is averaged per tomogram. Overall, 9 tomograms, together containing 13 LDs, were analyzed.

4.2.2 Quantifying the effects of DFCP1 on DMV morphology

In this section, a tomogram analysis workflow is presented, aimed at fully automating the analysis of DMV characteristics. The detailed procedures are described in the Methods section. The workflow is based on semantically segmented membranes, that were easy to obtain via membrain-seg (Lamm et al., 2024), an existing general model for segmenting membranes. The initial membrane segmentation is therefore considered the last step of tomogram preprocessing, which includes streamlined and quick to execute processing steps such as motion correction and tomogram reconstruction. The full preprocessing was conducted using existing software, but automated with custom Python scripts, written in either Python or Bash, as outlined in the Methods section. For the downstream analysis (everything except tomogram preprocessing), dataset-specific methods were newly explored and implemented in Python to fully automate the quantification process.

The downstream analysis can be outlined with the following six steps: 1) The first step, essential for any subsequent analysis, was the DMV labeling as separate instances. 2) Then, the shape of the isolated DMVs could be analyzed with computational methods. The most important step here was the conversion of the DMV segmentation into a contour for each z-plane, which could then be used for circumference calculations and allowed to calculate a DMV graph representation. 3) In addition to the DMVs themselves, the DMV pores also needed to be located as separate instances. 4) The DMV pores could then be mapped onto the DMV graph, and pore-to-pore distances could be calculated. 5) With the DMV contour and the pore coordinates, the DMV intermembrane spacing and its relation to pore proximity could be investigated. 6) All investigated characteristics were compared, plotted and statistical tests were conducted.

This newly implemented workflow allowed for automating the full quantification process, and was applied to the five above mentioned conditions, overexpressing the viral proteins NSP3, NSP4 and NSP6 and the host protein DFCP1 in different

combinations. The aim was to investigate the importance and function of DFCP1 for SARS-CoV-2 DMV formation.

Instance segmentation for DMVs and pores

Labeling DMVs as separate instances is a prerequisite for further automated analysis. To extract the DMV instances, we performed 2D instance segmentation using a pretrained Mask2former model. We used semantically segmented membranes as input images. Since a well-functioning model for membrane segmentation already exists, our model could focus specifically on separating out the membrane complexes of interest (the DMVs). One potential limitation of applying 2D instance segmentation to 3D tomograms is the loss of spatial information along the z-axis. To provide the model with additional spatial context, we explored the following approach: Since the instance segmentation model expects RGB images as input and tomograms contain only grayscale values, we utilized the three color channels by populating each with a different tomogram z-plane. This strategy enabled the model to capture 3D information rather than receiving only 2D image data. Detailed descriptions of this methodology can be found in the Methods section (4.4.5). For extracting pore instances, we slightly adapted this approach: whereas we applied the model to semantically segmented membranes for DMV instance segmentation, for pores we used the denoised tomograms as input (detailed descriptions in the Methods section 4.4.7).

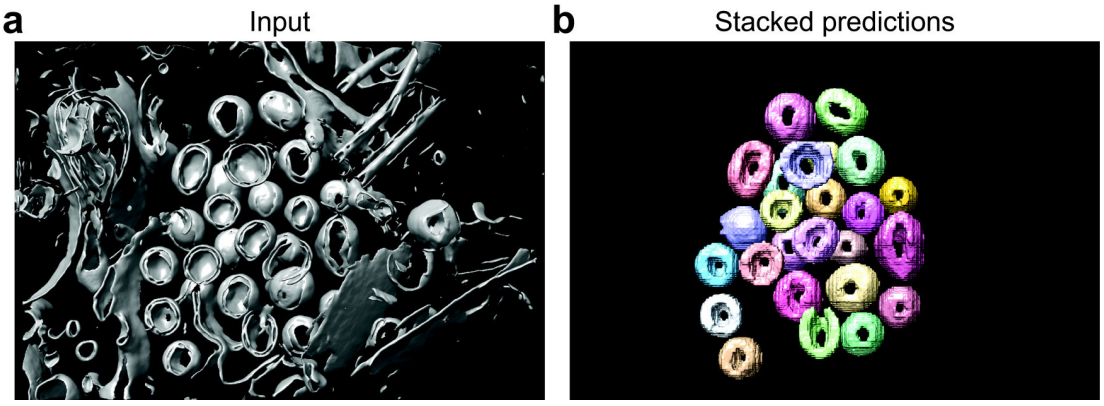

FIGURE 4.2: Illustration showing the segmentation of double membrane vesicles (DMVs) into distinct instances. a) The input for the instance segmentation network were the semantically segmented membranes. b) The predicted output contained distinct DMV instances, discarding any other membranes.

The established and retrained model for segmenting DMV and pore instances in 2D tomogram z-planes was then applied to the full dataset, comprising 41 tomograms from the five previously described conditions. To connect the resulting 2D DMV instances across the z-axis, we employed a distance-based approach that measured proximities between segmented instances on neighboring planes (detailed descriptions in the Methods section 4.4.5). As a result, we extracted individual DMVs as separate subvolumes, which served as the basis for DMV shape analysis (Figure 4.2). The pore instance segmentation enabled precise localization of pores on DMV surfaces, facilitating analysis of their spatial distribution. The DMV instance segmentation also enabled extraction of the associated semantically segmented DMV membranes, which was essential for membrane spacing analysis.

Analyzing DMV shape, pore distribution and membrane spacing

To estimate DMV circumference and facilitate downstream pore distribution and membrane spacing analyses, DMV contours were extracted on a per-plane basis. We used individual DMV instance subvolumes derived from the instance segmentation as input (4.3, a, right). The workflow included closing small holes and connecting open DMV membranes where necessary, while preserving the original DMV shape (detailed descriptions in the Methods section 4.4.6). Representative results are

shown in Figure 4.3, where DMV contours are displayed stacked across all planes in both 3D (b) and in 2D (c) views. This 3D contour representation was the primary output of the shape analysis, enabling straightforward circumference estimation and proving essential for DMV pore analysis: the DMV contour could be readily converted into a graph representation, facilitating pore localization and enabling pore distance calculations along the DMV surface.

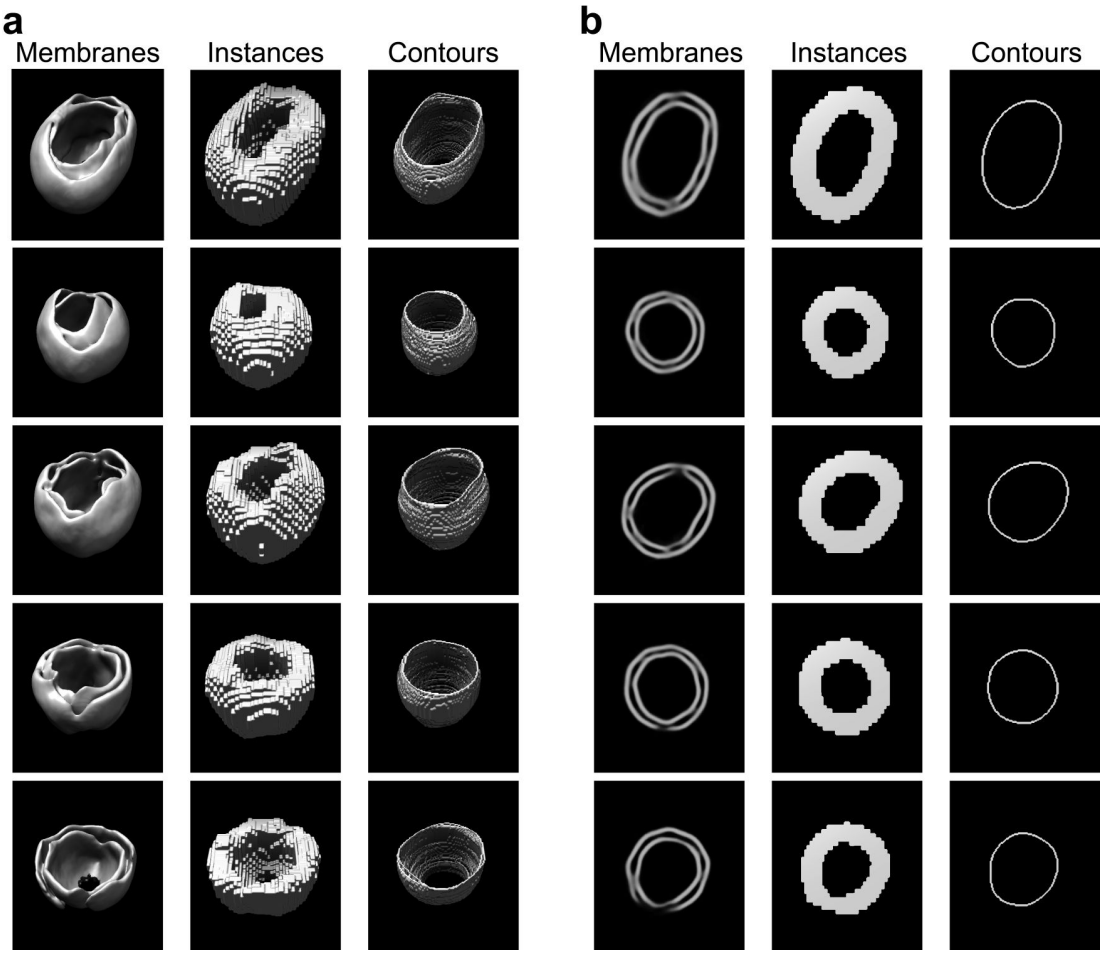

FIGURE 4.3: Visualization of intermediate steps of double membrane vesicle (DMV) shape analysis. a) Five example rows of extracted DMVs. Left: Subvolumes with original membrain-seg segmented membrane. Middle: DMV instance segmentation as detected by the Mask2Former model. Right: stacked DMV contour planes after completed 2D analysis. b) Five example rows of extracted DMVs, only portraying the highest circumference plane. Left: 2D plane with original membrain-seg segmented membrane. Middle: DMV 2D plane as detected by Mask2Former instance segmentation. Right: DMV contour after completed 2D analysis.

The basis for the pore analysis comprised the segmented pore instances and the 3D

DMV graph. We localized each segmented pore instance by averaging the corresponding voxel coordinates (Figure 4.4, a). This location was then mapped onto a node of the 3D DMV graph by proximity. We calculated the shortest paths between every pore pair, and assigned each pore the distance to its closest neighbor (detailed descriptions in the Methods section 4.4.8). The process from pore instance segmentation to pore localization and shortest path calculation is visualized in figure 4.4 b.

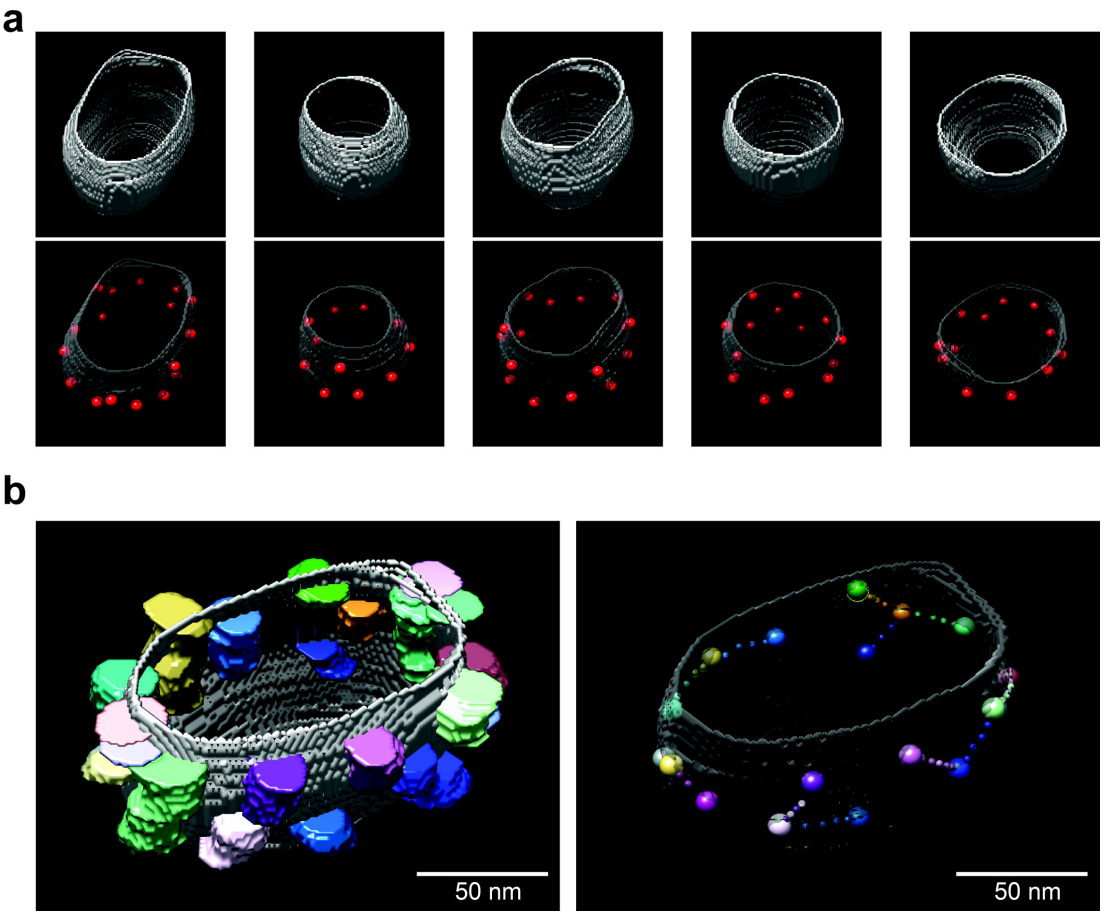

FIGURE 4.4: Double membrane vesicle (DMV) pore characterization. a) Five example DMVs as stacked contours (top). Pores localized on DMV surface are colored as red spheres (bottom). b) Left: Example DMV with segmented pore instances stacked along the z-axis. Right: The calculated locations of valid pores (large spheres), together with indicated paths to their closest neighbor (small spheres).

For the DMV double membrane spacing calculations, we used the 3D contour representation of each DMV instance and the associated semantic membrane segmentation, which provided clear separation between the two DMV membranes. We

conducted the analysis for every 2D z-slice of each DMV by extracting orthogonal line profiles centered on the DMV contour. A Gaussian filter was applied to each resulting density profile, and the two highest peaks were extracted to calculate the membrane spacing (detailed descriptions in the Methods section 4.4.9). While using a membrane segmentation as the basis for distance calculations may initially appear less precise than working directly with the denoised tomogram, this approach offers advantages in terms of consistency and accuracy. Membrain-seg, having been trained on a diverse dataset of membrane structures (Lamm et al., 2024), can be expected to consistently assign peak scores to the membrane center rather than arbitrarily favoring either the inner or outer lipid layers. This consistency ensures reliable and reproducible membrane spacing measurements. It should be noted though that this methodology differs from conventional approaches, that typically measure distances between outer lipid layers (Zimmermann et al., 2023). Thus, when comparing the conducted spacing measurements with previous studies, the membrane thickness has to considered. However, the conventional method has the drawback of incorporating membrane thickness into the intermembrane distance measurements. By measuring the distances between membrane centers, the segmentation-based approach provides an adequate representation of the actual separation between membrane pairs.

Quantification of DMV size, pore distribution and double membrane spacing

In order to investigate the relevance of the host protein DFCP1 for the formation of SARS-CoV-2 DMVs, we compared the morphological characteristics of DMVs for five different conditions of VeroE6 cells: 1) NSP3-4 overexpression with endogenous DFCP1 expression (in the following figures labeled as *NSP3-4*) 2) NSP3-4 overexpression with DFCP1 overexpression (in the following figures labeled as *NSP3-4 DFCP1*) 3) NSP3-4-6 overexpression with endogenous DFCP1 expression (in the following figures labeled as *NSP3-4-6*) 4) NSP3-4-6 overexpression with DFCP1 overexpression (in the following figures labeled as *NSP3-4-6 DFCP1*) 5)

NSP3-4-6 overexpression in a DFCP1 KO cell line with overexpression of the mutant DFCP1 K193A (in the following figures labeled as *NSP3-4-6 DFCP1-K193A*), which is nucleotide-binding defective, inhibiting the DFCP1 ATPase activity (Nähse et al., 2023). Overall, 41 tomograms were used, throughout which 708 DMVs and 7278 pore were analyzed in an automated manner with the above described methods. The DMV and pore count by condition is listed below (Table 4.1). Note that due to the necessity for more complete DMVs for pore analysis, as the analysis is invalid if multiple DMV planes are missing, more stringent filters were applied, resulting in less DMVs being available for the pore analysis than for the DMV circumference and DMV intermembrane spacing analysis.

Condition	**Tomograms**	**DMVs**	**DMVs for pore analysis**	**Pores**
NSP3-4	1	21	20	266
NSP3-4 DFCP1	17	445	391	5320
NSP3-4-6	1	21	14	108
NSP3-4-6 DFCP1	7	35	14	132
NSP3-4-6 DFCP1-K193A	15	186	139	1452

TABLE 4.1: Tomogram, DMV and pore counts for the five investigated conditions.

Visual investigation of the data revealed that NSP3-4 overexpression led to larger DMVs compared to NSP3-4-6 expression, while DFCP1 overexpression generally seemed to increase DMV size, often also leading to more irregularly shaped DMVs (Figure 4.5). Overexpression of the nucleotide-binding defective DFCP1-K193A mutant resulted in excessively large DMVs, and in some cases, multilamellar DMVs could be found (bottom right). These could be identified due to their double membrane structure and the embedded molecular pore structures.

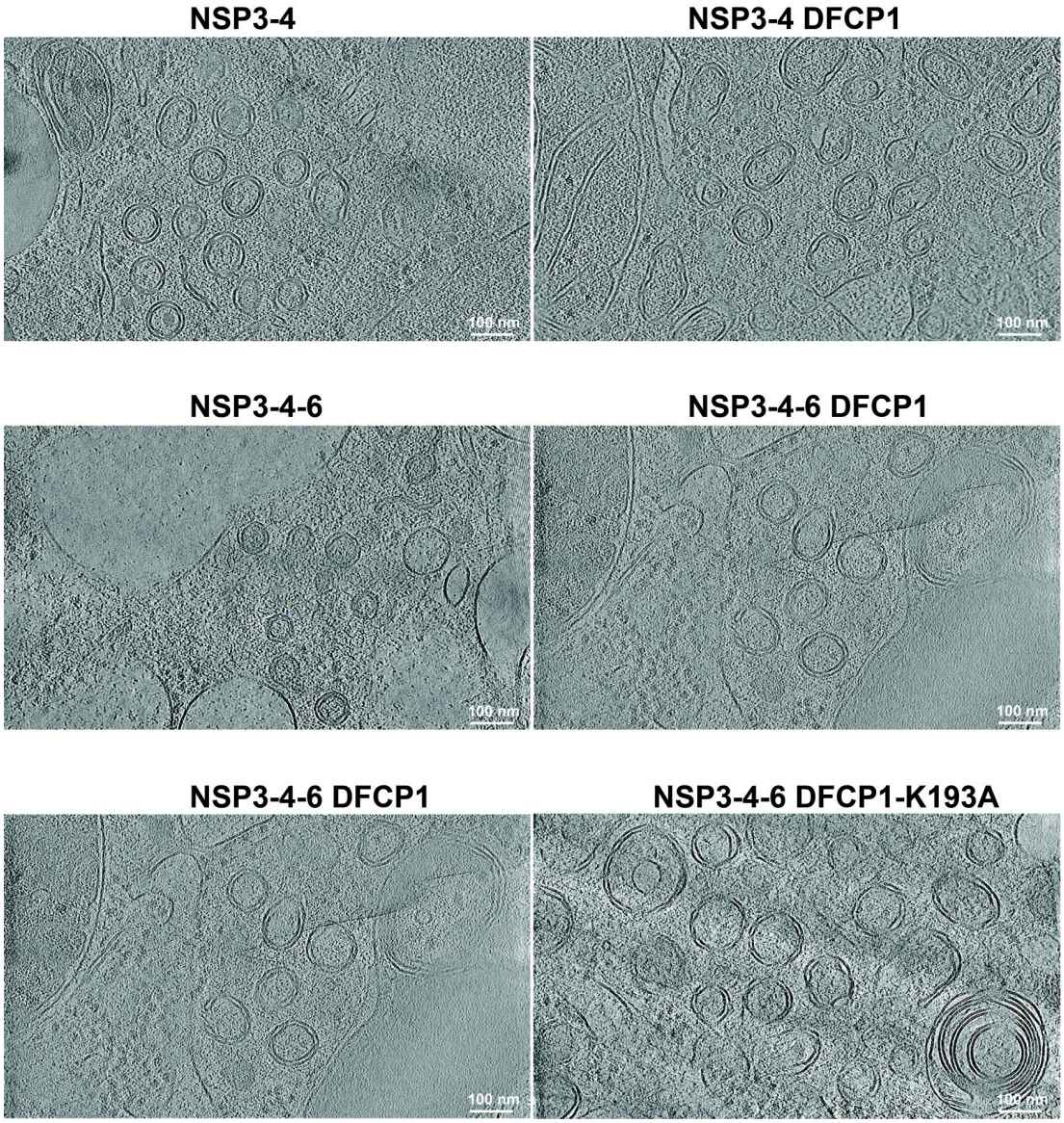

FIGURE 4.5: Illustrative tomogram z-slices focused on double membrane vesicle (DMV) clusters of the five investigated conditions, visualized as three comparisons. Top: NSP3-4 overexpression with endogenous DFCP1 expression (left) and with DFCP1 overexpression (right). NSP3-4-6 overexpression with endogenous DFCP expression (left) and with DFCP1 overexpression (right). Bottom: NSP3-4-6 overexpression with DFCP1 overexpression and NSP3-4-6 overexpression (left) with overexpression of the DFCP1 nucleotide-binding defective mutant (right). A multilamellar DMV - characterized by multiple stacked double membranes - is visible in the bottom right of the image.

The before described automated shape analysis was used to quantify these differences. Here, multilamellar DMVs were not included, as the analysis focused on distinctively identifiable DMVs. The analysis revealed that DFCP1 overexpression significantly impacts DMV circumference. When DFCP1 was overexpressed alongside

NSP3-4, mean DMV circumference expanded from 318 nm to 365 nm compared to NSP3-4 alone. The addition of NSP6 resulted in a mean circumference of 236 nm, which increased to 328 nm when DFCP1 was co-overexpressed. The largest DMVs were observed with NSP3-4-6 and DFCP1-K193A overexpression, reaching 380 nm in mean circumference. Beyond increasing overall size, DFCP1 overexpression also enhanced size diversity, as evidenced by increased standard deviation (Figure 4.6, a).

In contrast to the DFCP1 effects, NSP6 overexpression consistently reduced both DMV size and size diversity, regardless of whether DFCP1 was expressed endogenously or overexpressed. However, this moderating effect of NSP6 was lost when DFCP1 function was compromised. With the nucleotide-binding defective DFCP1-K193A mutant, NSP6 overexpression lost its regulating effect on DMV shape, resulting in large, highly variable DMV circumferences similar to those seen with functional DFCP1 overexpression.

Next, we investigated the pore distribution within the DMVs, as the pores may play a membrane-and shape-regulating role. Here, the primary measurement of interest was the closest pore-to-pore distance, which is a robust way to measure the pore number per DMV (lower pore-to-pore distance indicating more pores). DFCP1 overexpression made a slight but measurable difference (Figure 4.6, b): The mean distance was 27.7 nm for NSP3-4 overexpression alongside DFCP1 overexpression, instead of only 26.3 nm for endogenous DFCP1 expression. NSP6 also had a large impact: The mean closest pore-to-pore distance for NSP3-4-6 overexpression with endogenous DFCP1 expression was much lower at 22.7 nm, and 25.8 nm if DFCP1 was co-overexpressed. Thus, NSP6 overexpression had the effect of decreasing the pore-to-pore distance, while DFCP1 overexpression increased it. However, when the nucleotide-binding defective DFCP1-K193A mutant was overexpressed alongside NSP3-4-6, the pore-to-pore distance peaked at 30.2 nm, confirming the production of highly irregular DMVs for the case of DFCP1-K193A overexpression.

For the number of pores per DMV, we found no significant differences (Figure 4.6,

c) between endogenous DFCP1 expression and DFCP1 overexpression. It should however be noted that the closest pore distance for NSP3-4-6 was with 22.7 nm lower than for NSP3-4 expression (26.3 nm), although the pore number per DMV itself was lower as well (7.7 nm versus 13.3 nm). Although pore distribution differences could explain this finding, it more likely reflects data quality artifacts affecting DMV completeness along the z-axis, since better-imaged DMVs reveal more pores. Such an effect becomes especially relevant since the number of tomograms was low for some conditions, and data quality is expected to be relatively consistent within a tomogram. The pore-to-pore distance measurement is less dependent on data quality, as the pore distance measurements do not require complete DMVs for robustness. Thus, the pore-to-pore distance measurements should be regarded as the more accurate measurement for characterizing pore distribution.

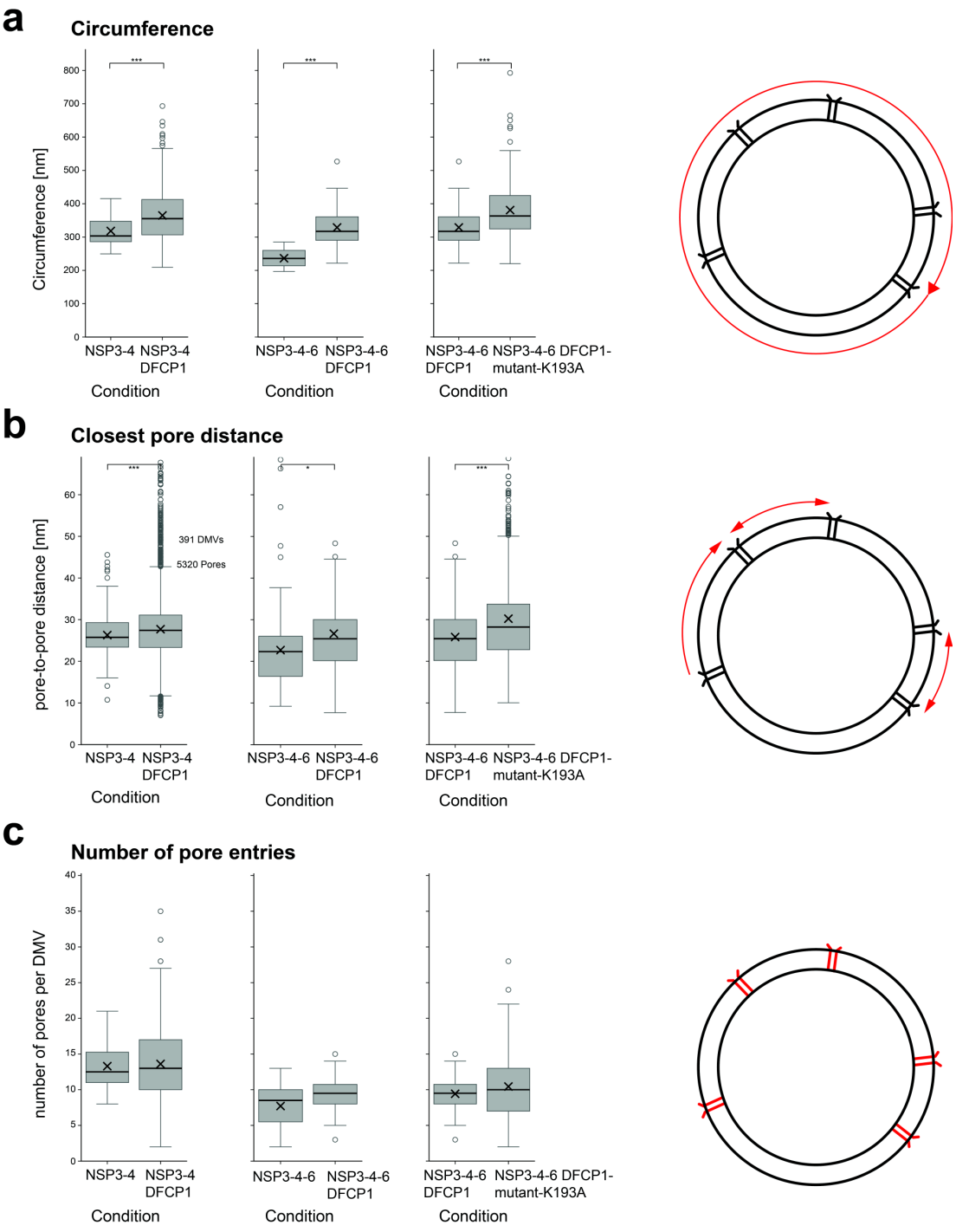

FIGURE 4.6: Comparisons of double membrane vesicle (DMV) circumference, DMV pore-to-pore distance and DMV pore count. (a-c) Left: NSP3-4 overexpression versus NSP3-4 with DFCP1 overexpression. Middle: NSP3-4-6 overexpression versus NSP3-4-6 with DFCP1 overexpression. Right: NSP3-4-6 with DFCP1 overexpression versus NSP3-4-6 with overexpression of DFCP1 nucleotide-binding defective mutant. On the far right, a sketch of the measured characteristic is portrayed. Within the boxplots, the mean value of the distribution is marked as x. Significance was tested with a two-sided Welch's t-test (Welch, 1938). Asterisks indicate the level of statistical significance for reported test statistics. One asterisk (*) denotes $p < 0.05$, two asterisks (**) denote $p < 0.01$ and three asterisks (***) denote $p < 0.001$. Results without asterisks are not statistically significant. a) Comparison of circumference. b) Closest pore distances within DMVs. Note that distant outlier values > 70 nm are not visualized, but still part of the significance test. c) Number of pore entries per DMV.

Finally, we investigated the distance between the DMV double membranes (Figure 4.7, a). First, we analyzed the membrane spacing independent of DMV affiliation (Figure 4.7, b). Throughout the conditions, significant differences could be observed when DFCP1 was overexpressed. However, due to the very large dataset (as spacing investigations were conducted for every second DMV-contour-belongig voxel), it is difficult to judge whether the measured significance was of biological relevance. The only condition where a clear difference was observable by eye was when the nucleotide-binding defective DFCP1-K193A mutant was overexpressed alongside NSP3-4-6. With endogenous DFCP1, the average membrane spacing was 11 nm, and for co-overexpression of DFCP1-K193A, it reached 13.8 nm. Another difference was the slightly higher standard deviation for DFCP1 overexpression compared to endogenous DFCP1 expression, and the lower standard deviation if NSP6 was included, except when NSP3-4-6 was overexpressed with the DFCP1-K193A mutant, where the standard deviation peaked.

Measuring membrane spacing without differentiating between DMV instances resulted in large DMVs contributing disproportionately to the measurements (due to having more sampling points). Additionally, the spacing values were distributed across a wide range, making it more difficult to detect meaningful differences. Thus, as a second approach, we averaged the spacing measurements per DMV, and then compared the averaged spacing values across conditions. Consequently, the value range was much smaller, making it easier to detect differences (Figure 4.7, c). These measurements revealed consistently increased membrane spacing when DFCP1 was overexpressed, with differences that were both visually apparent and statistically significant. Comparing NSP3-4 overexpression with endogenous DFCP1 expression and NSP3-4 with DFCP1 co-overexpression, the mean intermembrane spacing per DMV increased from 11.8 nm to 12.6 nm. Including NSP6 in the analysis, the membrane spacing increased from 10.8 nm to 11.4 nm, again comparing endogenous DFCP1 expression to DFCP1 co-overexpression. For the DFCP1-K193A mutant, the previously observed increase in membrane spacing

was confirmed.

To investigate whether DMV pores regulate intermembrane spacing uniformity, we examined the relationship between intermembrane spacing and pore distance directly. (Figure 4.7, d). Membrane spacing clearly increased in all investigated cases with increasing pore distance. This effect was less pronounced when NSP6 was overexpressed in addition to NSP3-4, and the standard deviation was much lower, indicating more regular intermembrane spacing. For NSP3-4 with DFCP1 co-overexpression, a consistent increase in membrane spacing could be observed when compared to endogenous DFCP1 expression, independent of pore distance. Also, the standard deviation was clearly increased, indicating a higher variability in intermembrane spacing. However, with additional NSP6 overexpression, the intermembrane distance between endogenous DFCP1 expression and DFCP1 overexpression was very similar. While this observation seems to be in conflict with the previously described difference between the two conditions, is has to kept in mind that here, only membranes in proximity to pores were tracked (up to 30 nm distance). Thus, taken together, these results indicate that in the case of NSP3-4-6 overexpression, membranes farther away from pores portray higher intermembrane spacing when DFCP1 is overexpressed compared to endogenous DFCP1 expression. Finally, for overexpression of NSP3-4-6 alongside the DFCP1-K193A mutant, the intermembrane spacing was consistently highest, as was the standard deviation.

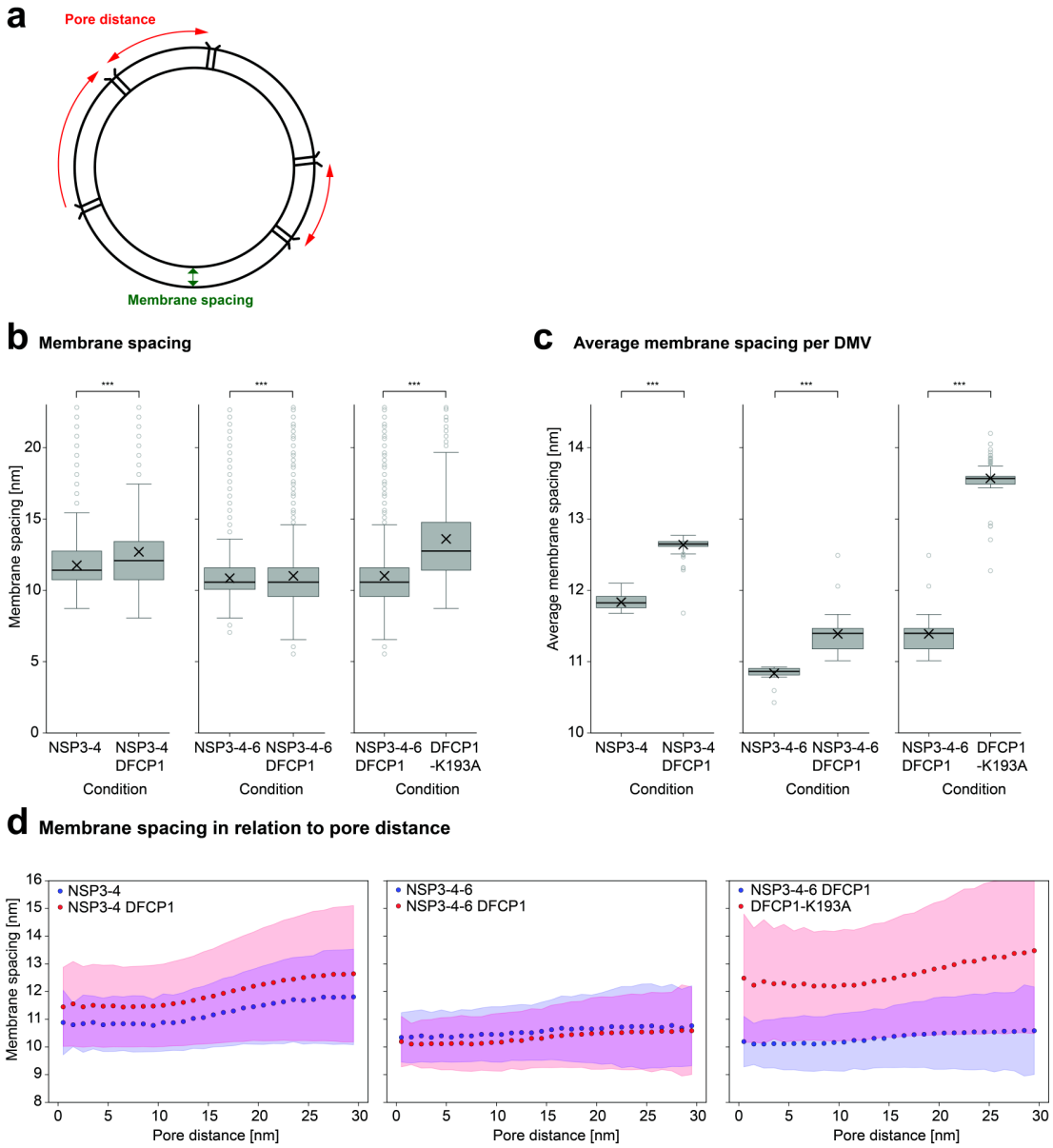

FIGURE 4.7: Comparisons of of the double membrane vesicle (DMV) double membrane spacing and its relation to pore distance. a) The characteristics compared visualized: Membrane spacing (green) concerns the distance between DMV double membranes, and pore distance the closest pore-to-pore distance as measured along the DMV surface (red). (b-c) Comparison of membrane spacing. Within the boxplots, the mean value of the distribution is marked as x. Significance was tested with a two-sided Welch's t-test (Welch, 1938). Asterisks indicate the level of statistical significance for reported test statistics. One asterisk (*) denotes $p < 0.05$, two asterisks (**) denote $p < 0.01$ and three asterisks (***) denote $p < 0.001$. Results without asterisks are not statistically significant. b) Membrane spacing per condition independent of DMV affiliation. c) Average membrane spacing per DMV. d) Membrane spacing as it relates to pore distance. (b-c) Left: NSP3-4 overexpression versus NSP3-4 with DFCP1 overexpression. Middle: NSP3-4-6 overexpression versus NSP3-4-6 with DFCP1 overexpression. Right: NSP3-4-6 with DFCP1 overexpression versus NSP3-4-6 with overexpression of the nucleotide-binding defective DFCP1-K193A mutant.

4.3 Discussion

4.3.1 Establishing a workflow for 3D instance segmentation and DMV morphology analysis

In the context of DMV morphology analysis, measurements such as DMV diameter or DMV pore distance calculations were conventionally conducted manually or only partially automated (Zimmermann et al., 2023). Some measurements however are difficult to conduct manually, for example circumference calculations of complex shapes. Another example are intermembrane spacing calculations, where manual measurements are usually less reliable as the number of locations one can inspect is limited, and the chosen locations are therefore subject to bias. Thus, manual labor is a limiting factor, as conducting these measurements for large datasets is time consuming and often subject to bias. For these reasons, in this work we focus on the development of a workflow for fully automated DMV morphology analysis. One essential step for the quantification of DMV characteristics is the automated DMV extraction as separate instances, and similarly for pores. There is currently no tool available that performs these specific tasks: Whereas TARDIS (Kiewisz et al., 2025) provides a general instance segmentation model, it is incapable of separating out partially connected DMVs and does not identify DMVs as distinct from any other membrane complex. Membrain-seg (Lamm et al., 2024) offers a similar approach via connected component analysis. Until now, no platform for training instance segmentation models for complex mambrane formations in three-dimensional data exists in the cryo-EM community. Therefore, we explored an approach for 3D instance segmentation, based on training a 2D instance segmentation model and connecting the instances across three dimensions. For the presented cases, this approach was capable of adequately segmenting DMV and pore instances, allowing for further automated analysis.

For automating the quantification of DMV morphology, we developed a workflow

for the following DMV characteristics: DMV circumference, DMV curvature, minimal pore distance and intermembrane spacing. Such quantifications were previously conducted with a variety of different tools and scripts. DMV and pore localization and counting, extracting locations for membrane spacing calculations as well as DMV diameter calculations were often conducted manually (Klein et al., 2020; Zimmermann et al., 2023). In this work, we only needed to conduct DMV extraction and pore localization manually to the extend that a network could be trained. We implemented the full workflow, including segmenting DMV and pore instances and quantifying the described characteristics in Python as a fully automated pipeline, only requiring the denoised tomograms and its semantic membrane segmentation as inputs. With this workflow, the efficient processing of much larger datasets is possible, as the work-intensive manual tasks are automated. It also may have the additional advantage of being more objective: As the exact same workflow is applied independent of dataset, the risks for subjective biases are decreased. However, full automation is not easy to achieve. Cryo-ET data tends to be noisy, and although the initial semantic membrane segmentations by membrain-seg (Lamm et al., 2024) are generally of high quality, segmentations were sometimes of insufficient quality, especially for thick lamellae and noisier tomograms. Additionally, sometimes ice contamination or various image artifacts may be identified as membranes, and sometimes DMV segmentations are incomplete because the DMV may still be in the process of forming. Thus, for various reasons there are many cases of only partial or faulty DMV segmentations included in the data. For these reasons, smoothing of segmentations and filtering out low quality segmentations were essential parts of the automated workflow. However, considering all these cases leading to incomplete segmentations was difficult. While manual measurements naturally accommodate these variations, implementing automated analysis necessitated establishing standardized criteria to filter these cases consistently. Ultimately, accounting for every single anomaly is impractical, and overly stringent

filtering risks excluding valid data. When sufficient data is available, a few remaining erroneous instances can typically be ignored without significantly affecting the overall results. However, to verify data validity, in this work, we briefly investigated the DMV contours for validity before subjecting them to further analysis.

4.3.2 The role of DFCP1 in DMV formation

We first investigated DFCP1 overexpression in wild-type VeroE6 cells. Our observations highlight the previously described function of DFCP1 as a membrane remodeling protein (Axe et al., 2008, Nähse et al., 2023), as DFCP1 overexpression led to the shedding of many small vesicles. This vesicle shedding occurred predominantly near lipid droplets (Figure 4.1), potentially reflecting DFCP1-mediated lipid droplet recruitment to the ER, as previously described (Li et al., 2019). In that context, lipid droplets may serve as an additional lipid source for vesicle formation or compensate for the increased lipid consumption. Whether the ER or lipid droplets serve as the primary lipid source for these vesicles remains an open question.

Furthermore, the outcome of our quantification demonstrates that DFCP1 activity is a crucial factor for determining DMV size and shape (Figure 4.6), confirming its critical role during DMV biogenesis (Twu et al., 2021). DFCP1 co-overexpression with the viral non-structural proteins consistently led to enlarged DMVs and higher pore-to-pore distances. This increased distance between pores may be a direct consequence of the increased DMV size. DFCP1 overexpression further had the effect of increasing the double membrane spacing, leading to membrane bulges. These were overwhelmingly located between pores, as we observed an increase in double membrane spacing with increased pore distance (Figure 4.7). This observation indicates that the molecular pores function as clamps, holding the double membranes together and defining the distance of the intermembrane spacing, thus leading to more regularly shaped DMVs. Although the underlying mechanism leading to enlarged DMVs remains unclear, we can rule out the previously described pathway of

lipid droplet recruitment through DFCP1-NSP6 interactions as a major contributor. In this pathway, lipid droplets are recruited to NSP6-zippered ER membranes via DFCP1 and Rab18 (Ricciardi et al., 2022). However, as this pathway requires NSP6, and we observe even larger DMVs upon DFCP1 co-overexpression with viral proteins NSP3-4 only, this pathway does not seem to contribute to DMV enlargement. Lipid droplets could independently serve as a lipid source for NSP6-mediated RO organization, consistent with the reported NSP6-induced increase in DMV numbers in the aforementioned study.

We further evaluated the effects of the viral protein NSP6, co-overexpressing it together with either NSP3-4, or with NSP3-4 and DFCP1. It has previously been shown that NSP3-4-6 overexpression leads to smaller DMVs with more regular shapes compared to NSP3-4 overexpression (Ricciardi et al., 2022), and our data shows the same trends: The observed DMVs were significantly smaller when NSP6 was overexpressed in addition to NSP3-4 (Figure 4.6). This shape regulating effect of NSP6 was also present in the case of DFCP1 overexpression: Although DMVs were still increased in size, the increase was much lower than for NSP3-4 with DFCP1 overexpression. The standard deviation was also lower, indicating increased size regularity. Additionally, the membrane spacing stayed relatively similar (Figure 4.7). This shape regulating influence of NSP6 may stem both from its role in the organization of DMV clusters and its ER-zippering properties (Ricciardi et al., 2022). A direct contributor to more regular DMV shapes — potentially a consequence of enhanced control over DMV biogenesis through zippered ER and smaller DMVs — appears to be the NSP3/NSP4 pore complex. NSP6 overexpression consistently reduced the mean closest pore-to-pore distance (Figure 4.6), suggesting a more uniform pore spacing that likely promotes regular DMV morphology. Overall, the NSP6-induced RO organization and the associated decrease in DMV size likely leads to an increased number of molecular pores per unit of DMV membrane surface area, and the pores function to clamp the DMV double membranes together and define the intermembrane spacing, leading to more regularly shaped DMVs.

NSP6 may thus regulate DMV morphology indirectly, consistent with previous reports showing no co-localization between NSP6 and DMVs (Ricciardi et al., 2022). Nevertheless, additional studies are required to establish whether NSP6 regulates DMV shape through direct or indirect mechanisms.

In the case of abolishing the DFCP1 ATPase activity, the three investigated parameters (DMV size, the pore-to-pore distance, and the intermembrane spacing) reached their highest values (Figures 4.6; 4.7), thus resulting in excessively enlarged DMVs with irregular shapes. Additionally, we observed the occurrence of multilamellar DMVs, where multiple double membranes are stacked. These observations indicate that ATP-dependent DFCP1 dimerization plays a major role in proper DMV formation, possibly similar to how it mediates omegasome constriction in selective autophagy (Nähse et al., 2023). These constriction processes facilitated by the DFCP1 ATPase activity may similarly contribute to explaining the above described vesicle shedding for DFCP1 overexpression in wild-type cells (Figure 4.1). The DFCP1 ATPase activity seems therefore critical for the consistent formation of proper, regularly sized and shaped DMVs. A subset of more regularly shaped DMVs was also found (Figure 4.5), possibly as a result of residual DFCP1 ATPase activity or an additional, independent constriction mechanism. Notably, DFCP1 was observed to affect autophagosome biogenesis in selective autophagy, but not in bulk autophagy (Axe et al., 2008, Nähse et al., 2023). These findings together implicate additional mechanisms in autophagosome and DMV constriction beyond DFCP1, although DFCP1 seems to be critical for DMV formation in particular (Twu et al., 2021). Members of the ESCRT family are potential candidates and have been shown to be involved in DMV biogenesis (Kifer et al., 2025). However, the exact mechanisms and conditions triggering specific pathways during DMV biogenesis require further investigation.

Finally, it has to be noted that the presented data for NSP3-4 overexpression and NSP3-4-6 overexpression was limited (only 21 DMVs from a single tomogram were available for both analyses), and the results need to be corroborated. Additionally,

the described effects were observed in a minimal system, expressing only a small subset of viral proteins, and not necessarily resemble the characteristics of a viral infection. However, several previous studies demonstrated the utility of this minimal system, and it is considered well suited to study SARS-CoV-2 DMVs (Zimmermann et al., 2023, Huang et al., 2024).

In summary, both the host protein DFCP1 and viral protein NSP6 influence DMV morphology throughout the biogenesis process (Figure 4.8). Two complementary models that describe DMV formation were previously proposed (Oudshoorn et al., 2017): the *membrane wrapping* model, where ER-membranes first zipper, then curve and close into DMVs; and the *budding* model, involving sequential inward budding into the ER lumen followed by outward budding. Another form of ER-membrane zippering induced by NSP6 may maintain DMV-ER connectivity (Ricciardi et al., 2022), providing a potential mechanism for the NSP6 shape-regulatory function. Additionally, DFCP1 and Rab18 are thought to facilitate lipid droplet recruitment to the NSP6-zippered ER, and may thus secure the lipid supply for RO organization and DMV formation. Beyond its role in lipid provision, the DFCP1 ATPase activity likely drives membrane constriction during DMV formation. Loss of the DFCP1 ATPase activity results in aberrant structures, including malformed DMVs and multilamellar DMVs. Our findings thus identify DFCP1 as a critical host factor that regulates DMV architecture during SARS-CoV-2 infection, controlling both DMV size and morphology, thereby positioning DFCP1 as a potential target for antiviral intervention.

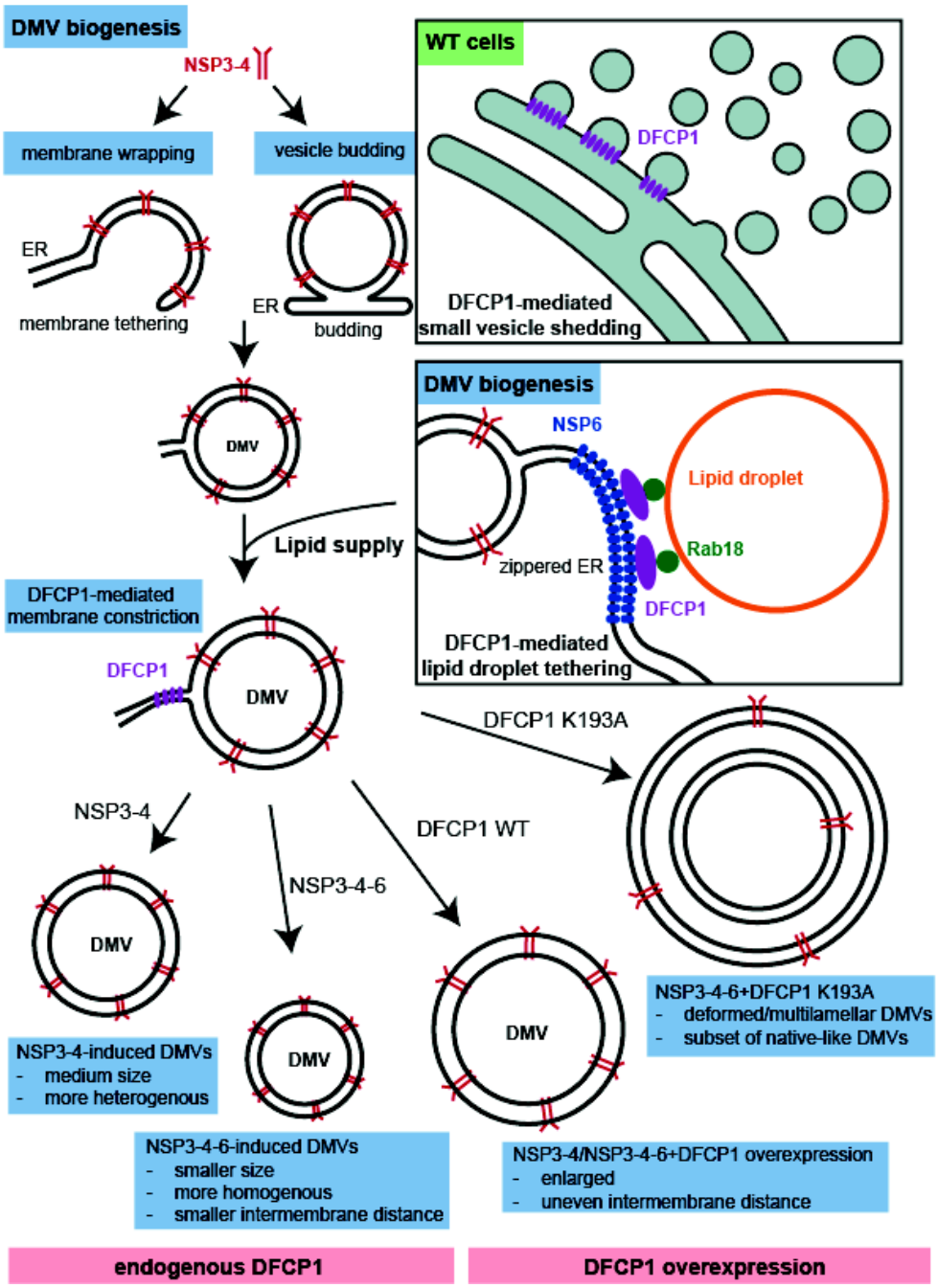

FIGURE 4.8: Summary of the role of DFCP1 in double membrane vesicle (DMV) biogenesis. Figure provided by Kevin Boga. In wild-type (WT) cells, DFCP1 overexpression induces shedding of small vesicles. NSP3-4 may induce DMVs by either membrane wrapping or vesicle budding. DFCP1 together with Rab18 tethers lipid droplets (orange) to the ER. Membrane constriction is likely mediated by DFCP1 (violet) and lack of DFCP1 ATPase activity led to the formation of deformed DMVs and multilamellar DMVs. NSP3-4 overexpression with additional NSP6 overexpression led to smaller DMVs with homogeneous shape, while DFCP1 overexpression generally led to an increase in DMV size.

4.4 Materials and Methods

Sample preparation and data collection was conducted by Kevin Boga. As this is not the focus of this thesis, the associated procedures are only briefly summarized for a more complete understanding.

4.4.1 Sample preparation

VeroE6 cells were grown under standard conditions (37°C, 5% CO2) in supplemented DMEM GLutamax medium (Gibco, Thermo Fisher Scientific, Waltham, MA, USA). The cells were transiently transfected with various plasmid constructs using commercial transfection reagents like Fugene HD (Promega, No. E2311) or Mirus TransIT-LT1 (Mirus Bio, No. MIR 2304). Experiments were typically conducted 20-24 hours after transfection to allow for protein expression. Cells were harvested, counted, and seeded onto micropatterned grids in glass-bottom dishes. After a short period of incubation for cell attachment, the grids were quickly washed and then vitrified (rapidly frozen) in liquid ethane using an automated plunge-freezer under controlled temperature and humidity conditions. The frozen grids were stored in liquid nitrogen and mounted in specialized holders for subsequent imaging. To prepare samples thin enough for transmission electron microscopy, they were processed using focused ion beam (FIB) milling to create ultra-thin sections called lamellae. Frozen grids were loaded into an Aquilos FIB system equipped with a fluorescence microscope to identify and confirm transfected cells. Cells were semi-automatically milled using stepwise reduction of ion beam currents to create progressively thinner lamellae, starting from rough milling and ending with fine polishing to achieve final thicknesses of 100-200 nm. The process included protective platinum coating and stress-relief features to maintain lamella integrity while preserving the fluorescent features of interest.

4.4.2 Data collection

Samples were imaged using either a Krios G4 or Talos Arctica transmission electron microscope, both equipped with energy filters and direct electron detectors. Tilt-series were collected using a dose-symmetric scheme (Hagen, Wan, and Briggs, 2017). The starting tilt-angle was 10° to compensate for the lamella pretilt. Details are listed below.

TABLE 4.2: Summary of data collection parameters. Data collection was conducted by Kevin Boga.

Parameter	**Krios G4**	**Talos Arctica**
Magnification	×26,000	×24,000
Voltage (kV)	300	200
Total dose (e^- per Å^2)	∼130	∼140
Dose per tilt (e^- per Å^2)	∼2.5	∼2.5
Energy filter slit width (eV)	20	30
Defocus	5 μm	5 μm
Physical pixel size (Å)	1.678	1.7435
Detector	Gatan K3	Gatan K3
Acquisition scheme	Dose-symmetric	Dose-symmetric
Tilt range	−40° to +60°	−40° to +60°

4.4.3 Tomogram preprocessing

MotionCorr3 was used for correcting beam-induced motion, IMOD (Kremer, Mastronarde, and McIntosh, 1996) was used for stacking up the averaged tilts, and AreTomo2 (Zheng et al., 2022) for tilt-series alignment and tomogram reconstruction, using a pixel size of either 10.46 Å or 13.5 Å. The reconstructed tomograms were denoised with cryoCARE (Buchholz et al., 2019), enhancing contrast and providing a solid basis for tomogram segmentation. These tasks were conducted with

a custom Python script, allowing to fully automate the process. The script for automated processing was written by me and applied to the data by Kevin Boga. The processing included separating the tomogram frames to create two half-tomograms, suited for the noise-2-noise denoising approach of cryoCARE. Semantic segmentations were conducted with a custom batch script written in bash, using membrain-seg (Lamm et al., 2024) to label the membrane-belonging voxels. Membrain-seg by default outputs the segmented membranes as a pre-thresholded binary file, yielding a conventional binary membrane segmentation (ones for membranes, zeros for background). As an alternative, it can produce a score file. This score file is very similar to a binary segmentation, only that it contains continuous scores representing the probability for each voxel of belonging to a membrane, instead of just ones and zeros as for the binary sgmentation. This score file was used for further analysis.

4.4.4 Quantification of vesicles distribution and size

To investigate the spatial distribution of vesicles and their size in relation to lipid droplet (LD) Locations, we focused this analysis on the subset of 9 LD-containing tomograms. As the dataset size was limited, Chimera (Pettersen et al., 2004) was used to manually pick 1) the center coordinate and 2) one coordinate in the vesicle membrane for radius calculations. In the same way, LD center coordinates and radii were collected.

We implemented the following procedures in custom Python scripts. For each tomogram, vesicles (by their center coordinates) were sorted according to LD distance: the number and average radius of vesicles within 100 nm shells was estimated, and shells with increasing LD center distances were calculated, with the two borders of each shell (S1, S2) being located at the following distances:

$$S1 = LD_r + LD_d - \frac{sd}{2} \tag{4.1}$$

$$S2 = LD_r + LD_d + \frac{sd}{2} \tag{4.2}$$

where LD_r is the LD radius, LD_d the distance from the LD center and sd the shell diameter. For the shell diameter, 100 nm was used, and the LD distances increased with a step size of 1, starting by 50 nm (half the shell diameter). Each data point in the resulting graph then represented either the number of vesicles within such a 100 nm shell (Figure 4.1, b, blue) or their average radius (Figure 4.1, b, red), yielding a continuous graph due to a new shell being sampled at each nanometer distance. We conducted the analysis in a 3D coordinate system.

4.4.5 Instance segmentation of DMVs

We conducted DMV instance labeling via 2D instance segmentation, employing a pretrained Mask2Former model and its training interface (Pavel Iakubovskii, April, 2025). Mask2Former is a neural network architecture for 2D instance segmentation tasks, employing a transformer-based approach with attention restricted towards localized features via masking (Cheng et al., 2022; Molina et al., 2025). Transformer-based architectures were extremely successful in recent years for a variety of deep learning tasks, most popularly for large language models (Vaswani et al., 2017; Singh, Kumar, and Mehra, 2023). In the context of image segmentation, the transformer-based approach utilizes attention mechanisms to capture dependencies between different parts of an image, allowing the model to understand contextual relationships across spatial locations. The employed model was pretrained on the widely adopted COCO dataset (Lin et al., 2014), a comprehensive dataset of color images specifically designed for instance segmentation tasks. This dataset provides high-quality per-instance segmentation annotations across diverse object categories, establishing a robust foundation for subsequent fine-tuning.

Thus, the model expected a colored image as input, meaning each pixel had to contain three values, one for each of the color channels. The three color channels represent red, green and blue (RGB), and each pixel value was expected to be within the conventional image intensity range between 0 and 255. Notably, the output (instance labels) of the model were also encoded as images: Here, the red channel had to contain a class identifier. For this application, a single class identifier for DMVs (ones) together with a background identifier (zeros) was sufficient. The green channel had to contain the instance identifiers, meaning each DMV instance was labeled with a separate integer value. The blue channel was left empty.

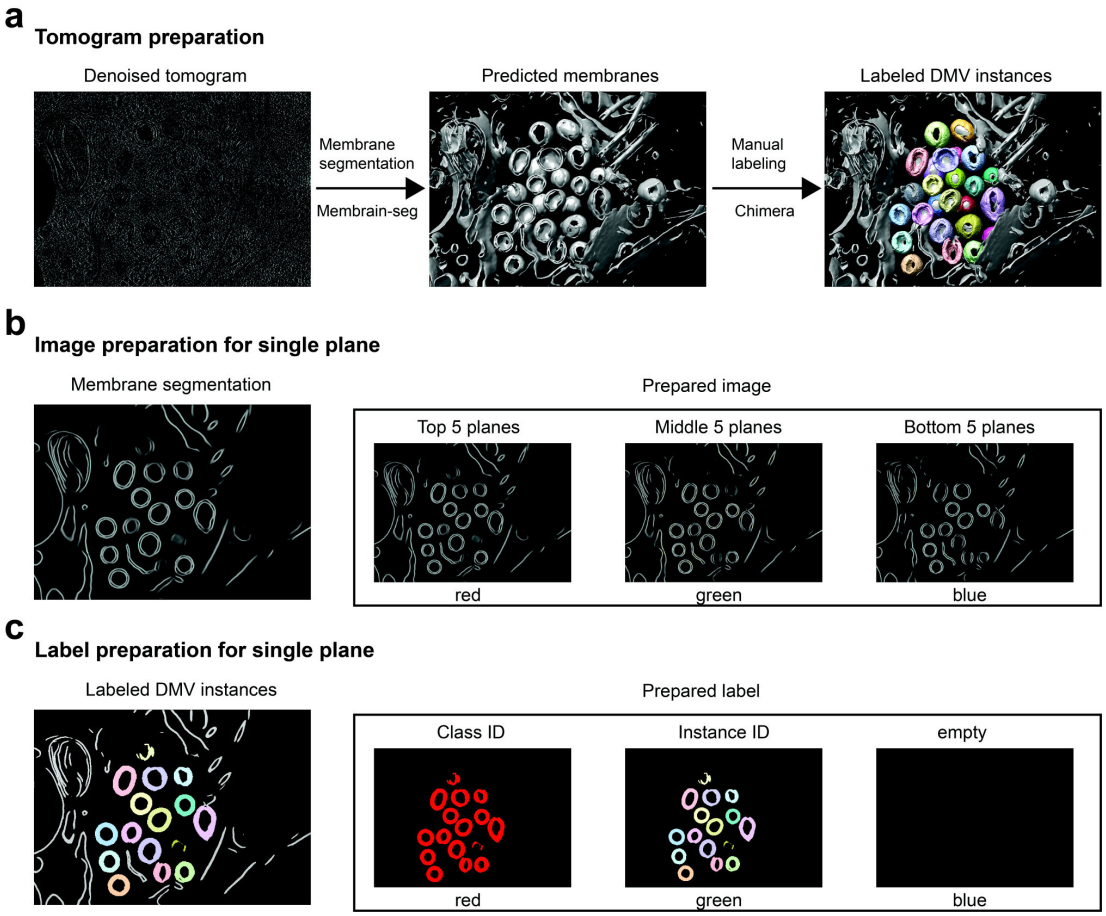

FIGURE 4.9: Data preparation for training a Mask2Former model for double membrane vesicle (DMV) instance prediction. a) The denoised tomogram was segmented with Membrain-seg (Lamm et al., 2024). The predicted membranes were then used as a basis for manual DMV instance labeling. b) The labels for a single slice were prepared by using a DMV class identifier in the red channel, separate identifiers (integers from '0' to 'n DMVs') for separate DMV instances in the green channel. The blue channel was left empty. c) The input image for a single slice was prepared by cropping its intensity values between -50 and 15, followed by normalization between 0 and 255. For each color channels, 5 neighboring planes were averaged.

As a well-functioning model for membrane segmentation already exists (Lamm et al., 2024), we used the membrain-seg semantic segmentations as a basis (Figure 4.9, a). Relying on a well-established general semantic segmentation for membrane identification allowed our model to focus on separating out the membrane complexes of interest (the DMVs).

For the creation of a training set (input images and output labels), we created the input images as follows, using customly written Python scripts: We split the 3D

tomograms into 2D z-slices and then transformed them into proper images as required for the model (Figure 4.9, b). After investigation of the membrain-score files, we concluded that the relevant range for score values was between -50 and 15. To focus the instance segmentation on the relevant values, this value range was scaled to the conventional image intensity scale of 0-255. Values below -50 were set to 0 and values above 15 were set to 255. One potential issue with 2D instance segmentation when applied to 3D tomograms is that the model misses spatial information across the third dimension. To give the model some more spatial information, we explored the following approach: As tomograms only contain grey values, and thus only one color channel would be needed, the required three color channels were made use of by filling them with different tomogram z-plane averages. A tomogram is computationally just represented as a 3D grey-value array, and the 2D extracted images used for instance segmentation only conatin a single tomogram z-plane. Thus, the idea was to 1) average several z-planes to create 2D images with 3D information and 2) use the three color channels for different averaged z-planes for more 3D information. How this was done is best illustrated with an example: Assuming a z-coordinate of 10 (the 10th 2D z-plane of a 3D tomogram), the red color channel of the image contained the average of planes 3-7, the green channel 8-12, and the blue channel 13-17. That way, the model received information across 15 z-planes. As neighboring tomogram planes portrayed high similarity, every fifth plane was used for training only in order to prevent leakage between training and validation set.

Additional to the input images, model training also required the image labels. Thus, we labeled a large number of DMVs within a set of diverse tomograms as described in the following section, using custom Python scripts. To create this labeled training data, we used semantic segmentations by membrain-seg of five tomograms across the different conditions as a basis. The above described membrain-seg score files were used, setting the score threshold to -5 (instead of the pre-set threshold of 0

by membrain-seg) to create a binary membrane segmentation. As lower values indicated a lower probability of a voxel belonging to a membrane, this conservative threshold led to the DMVs being captured to the fullest extent possible. For producing cleanly labeled DMVs, any other membranes were carefully deleted manually with the volume eraser in Chimera (Pettersen et al., 2004). The result was a clean, binary semantic segmentation of separated DMVs. Then, simple connected component analysis was used to save the DMVs as separate instances. Instead of a binary 3D membrane segmentation (containing ones for membranes and zeros for background), the tomograms now contained a distinct integer number for each DMV (Figure 4.9, a). We prepared the separated DMVs as labels for training in the following manner (Figure 4.9, c): The 3D tomograms were separated into 2D z-slices. As described previously, the target labels were expected to be represented as proper images with three color channels scaled between 0 and 255, with red containing the class identifiers and green the instance identifiers (and blue staying empty). Thus, for each z-plane, an empty image (containing only zeros in the three color channels) with the shape of a tomogram z-plane was created. The red channel (class label) was set to one for the DMV-belonging pixels. The green channel was set to the requisite integer identifier for each DMV-belonging pixel, and the blue channel was left empty. Every z-slices across the five training tomograms was prepared in this way.

After predicting the z-slices for each single tomogram, the segmented DMV instances were only present in 2D. However, for the analysis, 3D DMV instances were required. Thus, the labeled DMVs in 2D z-planes needed to be connected along the z-axis. To achieve this, we implemented an approach based on DMV distance measurements across z-planes in custom Python scripts: The average closest distance between any segmented instances on neighboring planes was checked. For calculating the distance between two 2D segmented instance pairs, for each pixel of instance 'a' the distance to the closest pixel of instance 'b' was calculated (Figure 4.10). Using the average of these closest distances, a cutoff of 3 pixels distance was

implemented to decide whether these instances separated across z-planes belonged to the same DMV instance. As this required significant computational effort, instead of using every DMV-belonging pixel for distance calculations, every 10th pixel was used instead to calculate the average distance between two instances from neighboring planes. Finally, the DMVs could be extracted as separate subvolumes for further analysis.

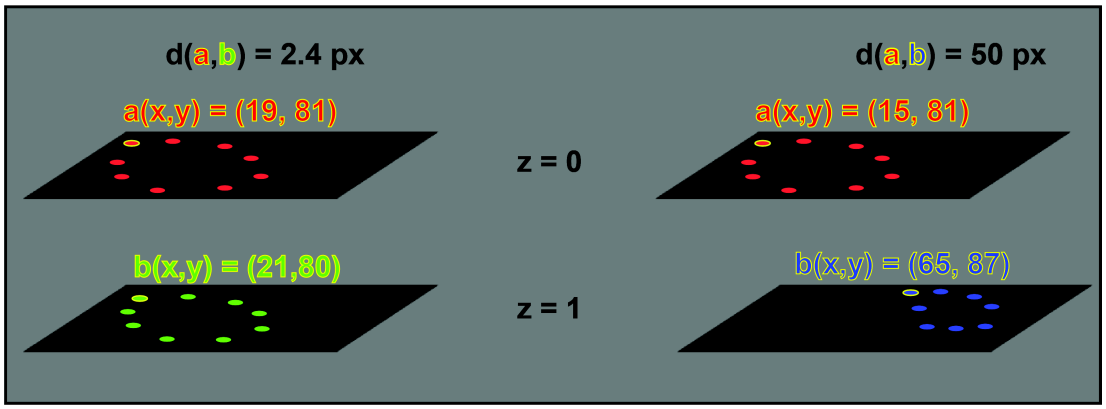

FIGURE 4.10: Connecting labeled instances across neighboring 2D planes. Left: Two neighboring z-planes (distance 1 pixel) carry neighboring slices of the same double membrane vesicle (DMV). The distance between the marked point of the red 2D DMV instance 'a' and the closest marked point of the green DMV instance 'b' is calculated. As their x-y coordinates are very similar, their distance is close to 1. This is repeated for every point (in practice, every 10th pixel). The average of these calculated distances is the instance distance. Right: The same procedure, this time 'b' (blue) does not belong to the same DMV instance as 'a' (red). Although 'b' is the closest point to 'a' from the neighboring plane, their x-y coordinates are far apart, resulting in a large distance between a and b, ultimately contributing to a large average distance between the DMV instances.

4.4.6 DMV shape analysis and contour creation

For analyzing the DMV shape, we used singular DMV instance subvolumes, as derived from the instance segmentation, as a basis. Here, the difference between the semantic segmentations derived from an existing tool, membrain-seg, and the instance segmentation derived from our retrained model is crucial (Compare *membranes* and *instances* in Figure 4.3, b and c). Visually, for the semantic segmentation, the double membranes were separated, and for the instance segmentation, there

is no separation of the double membranes. The use of the semantic segmentation is 1) as a basis for training the instance segmentation model more effectively (as opposed to using the denoised tomogram itself) 2) for later membrane spacing analysis, which requires separation of the double membranes. The value of the instance segmentation lies in its ability to identify DMVs as separate instances. The instance segmentation yields an instance label and a class label for each DMV belonging voxel, as opposed to the membrane belonging voxels just containing a class label for the semantic segmentation. Additionally, as the instance segmentation model was trained to identify DMV-like shapes, DMV-connected or DMV-adjacent membranes were cut off. As the instance segmentation and the semantic segmentation occupy the same coordinate system, the instance segmentation could be used to extract the semantically segmented DMVs as separate as well (or even the associated subvolume from the denoised tomograms), so that the more detailed semantic segmentation could be used for later intermembrane spacing analysis.

We conducted the described operations in Python, unless stated otherwise. The DMV instance segmentation was used to extract the DMV subvolume from three different kinds of tomograms, relevant for later analysis: First, we extracted each DMV subvolume from the corresponding denoised tomogram. While not relevant for the DMV shape analysis itself, we used the denoised subvolumes as the input for pore localization later. Second, we extracted each DMV subvolume from the corresponding semantic membrane segmentation (derived from membrain-seg (Lamm et al., 2024)), which was later used to calculate the DMV intermembrane spacing. Third, we extracted each DMV subvolume as the segmented instance itself for the shape analysis (as opposed to the semantic segmentation).

The automated DMV shape analysis consisted of a series of steps we implemented in a custom Python script, mostly focused on filtering out DMVs too incompletely segmented for further analysis, filling holes in the DMV segmentation, and deriving the DMV contour representation. First, any DMVs touching tomogram edges were removed from analysis. The remaining DMVs were analyzed as binarized 2D z-slices

(Figure 4.11, step 1). Then, a few steps were conducted to create a one-pixel-wide DMV representation. First, to smooth the segmentation, a gaussian filtered with a sigma of 1.5 was used on the binary segmentation (membranes labeled as ones, background as zeros). Then, any pixels with intensity values < 0.7 were removed from the segmentation. Connected component analysis was performed, and any small components with less than 100 pixels were removed as noise (Figure 4.11, step 2). To smooth object boundaries, binary opening (morphological erosion, then dilation) was conducted, followed by binary closing (dilation, then erosion). Dilation expands an object by adding pixels at its boundaries, and erosion shrinks an object, eroding its boundaries (Soille, 1999). Finally, the segmented DMV double membrane was thinned down to a one pixel wide representation using skeletonization (Lam, Lee, and Suen, 1992), iteratively removing pixels from the segmentation border, using the Zhang algorithm (Zhang and Suen, 1984). The DMV was now represented as a one-pixel-wide line, preserving the topological structure of the DMV (Figure 4.11, step 3).

In order to create a realistic and complete representation of the 2D DMV plane, it was useful to work with a graph instead of a simple segmentation. Thus, we created a DMV graph representation: Each skeleton pixel became a node, and edges were connecting each node with its horizontal, vertical, and diagonal neighbors. The edge weights here represented the Euclidean distances between nodes. Each connected component in the graph (separate component of the DMV segmentation), represented as a connected sub-graph, was simplified by calculating the minimum spanning tree (Kruskal, 1956) with Borůvkas algorithm (Nešetřil, Milková, and Nešetřilová, 2001). The minimum spanning tree is a subset of graph edges, selected so that the overall edge distance is minimized and the nodes stay connected: Unnecessary edges were removed from the graph, resulting in a sparser graph, although containing many sticking out branches. To remove these extra branches, the graph diameter path was calculated (Cormen et al., 2009) in three steps: 1) starting with a random graph node, the most distant node 'a' was calculated with

Dijkstra's algorithm, which finds the shortest path from a starting node to every other node in a weighted graph (Dijkstra, 1959). 2) From this node 'a', the most distant node 'b' was calculated, again by using Dijkstra's algorithm. 3) The shortest path between 'a' and 'b' was the diameter path. Each node not being on that path was deleted (Figure 4.11, step 4). To identify the graph endpoints for each component, any nodes with only one edge were identified. These endpoints were then connected among each other, but, to preserve a realistic vesicle-like shape, only if the connecting line did not intersect any existing edges (as there should not be any intersecting membranes in DMVs). These connections were added as new edges to the graph (Figure 4.11, step 5). By using the Christofides approximation of the traveling salesman problem (Christofides, 2022), we found the shortest Hamiltonian cycle (Cormen et al., 2009) through all graph nodes. The traveling salesman problem asks for the shortest possible route that visits each graph node at least once. Its solution allows us here to find the shortest cycle within the graph that visits every node exactly once (shortest Hamiltonian cycle), returning to the starting node at the end, and was thus used to mark the unused edges for removal. To ensure that edges were added in a way that the original DMV shape was preserved, the combined edge weight added during this procedure was only allowed to make up 10 % of the combined edges of the existing graph. If this condition was not fulfilled, the current DMV z-slice was judged as too incompletely segmented to be processed further.

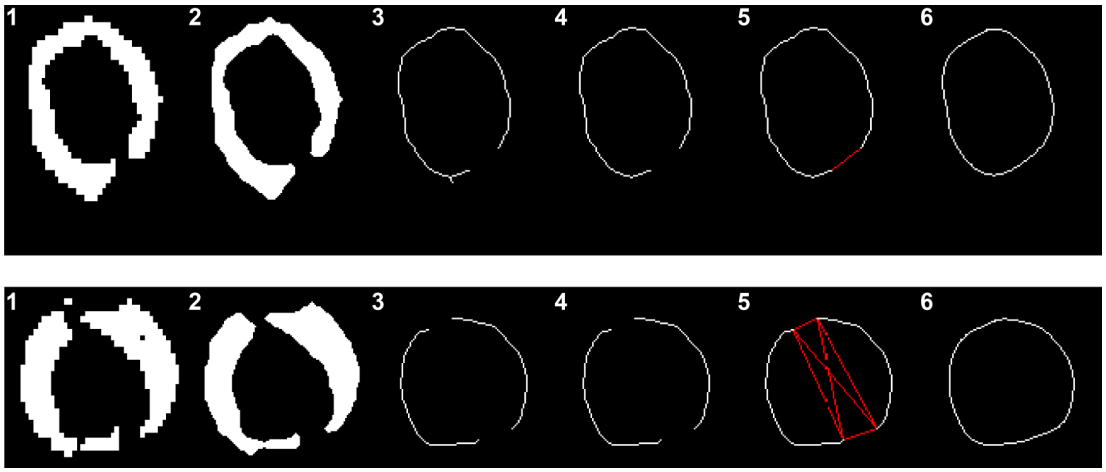

FIGURE 4.11: Creating a 2D double membrane vesicle (DMV) contour from an instance segmentation. The two rows are examples of segmented DMV instances. Contour creation is illustrated in six steps: 1) Original 2D DMV instance segmentation. 2) Smoothed segmentation. 3) Skeletonized segmentation converted into a graph. 4) DMV graph after removing nodes that were not on the diameter path. 5) DMV graph with added edges between unconnected nodes (edges in red). 6) Using the traveling salesman algorithm to select the correct graph edges and smoothing of the resulting DMV contour.

Finally, we converted the graph, now being a completed cycle and truthfully representing the 2D DMV shape (Figure 4.11, step 6), back to pixel coordinates: Any pixel along any of the graph edges was added to a list. The resulting pixel sequence was smoothed with a strong gaussian filter with a sigma of 5. By summing the Euclidean distances between consecutive pixels along the path, the circumference was calculated.

To ensure the validity of our result, two further calculations were conducted. First, the curvature of each point was calculated as the normalized cross product of velocity and acceleration vectors:

$$\kappa = \frac{x'y'' - y'x''}{(x'^2 + y'^2)^{3/2}} \tag{4.3}$$

To avoid irregular shapes leading to inconsistent measurements, for example sometimes happening in the case of DMV-connected membranes, only contours were included where the maximum curvature was < 1. Second, as an additional safety

measure to filter out unrealistic contours, the cirularity of each contour was calculated as:

$$\kappa = \frac{4\pi A}{C^2} \tag{4.4}$$

With 'A' being the contour area, and 'C' being the contour circumference. The area was calculated with the shoelace formula (Lee and Lim, 2017):

$$A = \frac{1}{2}\left|\sum_{i=0}^{n-1}(x_i y_{i+1} - x_{i+1} y_i)\right| \tag{4.5}$$

The circularity is a measurement designed to judge how much a shape resembles a circle. It can takes values between 0 and 1, with 1 being a perfect circle. Circularity was used as a measure for contour quality: valid DMV contours with a circularity $<$ 0.5 were not found in the data. Therefore, 0.5 was used as a cutoff, and the DMV contours with a lower circularity were discarded. Finally, to each DMV, the largest circumference across the investigated planes was assigned. To verify for further analysis, each DMV had to have a valid contour across 20 z-planes ($\sim$20 nm), with a maximum of 3 consecutive invalid contours.

Even with the presented filters, due to the noise present in cryo-ET data or erroneous semantic segmentation, in a few cases, faulty or incompletely segmented DMV instances were not filtered out. Although not impacting the final analysis in a significant manner, the DMV contours were briefly investigated for validity in ChimeraX (Pettersen et al., 2021) and faulty segmentations were excluded from the analysis.

4.4.7 Instance segmentation of molecular pores

We conducted the pore instance segmentation in a similar way as the DMV instance segmentation. Again, 2D instance segmentation was used, employing the before used pretrained Mask2Former model. We again implemented the described

operations in Python, unless stated otherwise. For creating the input images (as the model required proper images with three color channels, with pixel values scaled between 0 and 255), again 15 planes were distributed across the three image color channels, averaging 5 planes for each. Different from the DMV extraction, for the pores, the denoised tomograms were used instead of the membrain-seg score files, with normalized intensities between 0 and 255 (Figure 4.12, b). Extracted DMV subvolumes (with help of the before conducted DMV instance segmentations) were used as a basis. For creating the training labels, across 9 tomograms, the pores in 178 DMV planes were manually labeled (labeling done by Kevin Boga with Dragonfly (Object Research Systems), Figure 4.12, a). Additionally, 102 DMV planes without pores were added to the training set. The labeled image as required by the model was created in the same way as for the DMV instance segmentation: The red channel (class label) was set to one for any pore-belonging pixels. The green channel was set to the requisite integer identifier for each pore belonging pixel, and the blue channel was left empty.

We then used the established and retrained model for segmenting pore instances in 2D DMV z-planes for the full dataset (containing 41 tomograms of the previously described five conditions). The pores were stacked up across the z-axis in the exact same way as the DMVs were stacked up (Figure 4.10), only that a cutoff of 2 pixels distance instead of 3 was implemented to decide whether these instances separated across planes belonged to the same pore instance. We made this choice due to the more compact pore shape compared to DMVs, making it more likely that two pores separated across the z-axis were recognized as one.

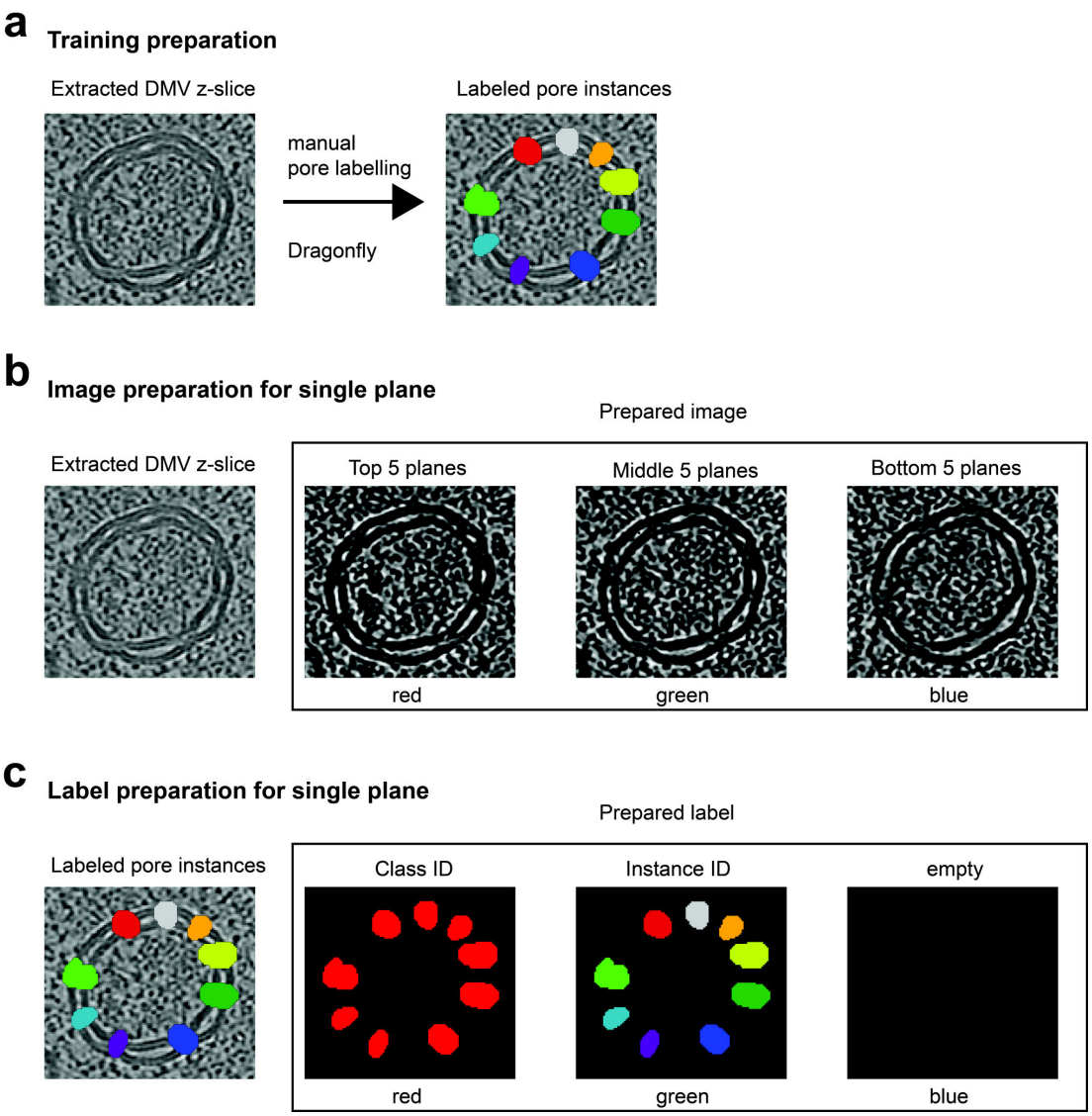

FIGURE 4.12: Data preparation for training a Mask2Former model for predicting pore instances within double membrane vesicles (DMVs). From the denoised tomograms, extracted DMV-containing subvolumes were prepared. a) For each extracted DMV z-slice, the pores were manually labeled in Dragonfly (Object Research Systems) (pore labeling conducted by Kevin Boga). b) The input image for a single slice was prepared by normalizing between 0 and 255. For each color channel, 5 neighboring planes were averaged. c) The labels for a single slice were prepared by using a pore class identifier in the red channel, separate identifiers (integers from '0' to 'n pores') for separate pore instances in the green channel. The blue channel was left empty.

4.4.8 DMV pore analysis

We implemented the analysis for the spatial distribution of pores in custom Python scripts. We localized the pores within 3D segmented DMVs by identifying individual pore instances via instance segmentation as described above, mapping them onto

the DMV surface, and computing shortest surface distances between pores. This required a precise and complete skeletonized DMV segmentation. To achieve this, we stacked the previously computed 2D z-slice contours for each DMV to create a skeletonized 3D DMV segmentation (Figure 4.4 a, top). A new 3D DMV graph was then derived from this segmentation by creating edges between each segmented voxel and its neighboring voxels within a distance threshold of 3.5 voxels (slightly larger than the maximum allowed number of consecutive missing DMV contours). This distance was small enough to avoid cutting corners in highly curved DMVs while still yielding a well-connected graph, leading to accurate distance estimations. For each pore instance, its precise location was calculated by averaging the voxel coordinates of the associated segmentation (Figure 4.4 a, bottom). Only pores detected across at least five z-planes were considered. Each pore location was mapped onto the DMV graph by proximity, using a cutoff of 5 voxels (∼5 nm) to the closest node to exclude invalid pore segmentations. Dijkstra's algorithm (Dijkstra, 1959) was used to calculate the shortest paths between pores mapped onto the DMV graph (Figure 4.4 b), which then served as pore distance measurements along the DMV surface.

4.4.9 DMV intermembrane spacing analysis

DMV instances were located within the semantic segmentation of the corresponding tomograms via the before described instance segmentation approach, and DMV subvolumes could be extracted, containing the DMV membrane scores (each voxel intensity is a score, which describes its probability of belonging to a membrane). These subvolumes were used as a basis for investigating the membrane spacing between the DMV double membranes. Additionally, the before created DMV membrane contour, derived from the skeletonized and cleaned DMV instance segmentation, was used. The subsequent analysis was conducted for every 2D z-slice of each DMV (Figure 4.13, a). For every second point contained in the contour, an

orthogonal line profile was extracted. This was done by placing an orthogonal line across the double membrane, using both neighboring contour points as references, reaching 30 pixels in each direction (Figure 4.13, b). The orthogonal line was sampled with a frequency of 0.5 pixels, and for each sampled position a radius of 2 pixels was used to calculate the average score of any included pixels within the membrain-seg score file. Finally, the resulting list containing the average scores was smoothed with a gaussian filter, using a sigma of 4. Some filters needed to be implemented to detect invalid segmentations: For the resulting density profiles (1D lists), the highest three peaks were detected by simple comparison with neighboring values. These peaks were then used for filtering, and to ensure measurements between valid double membranes, several conditions needed to be fulfilled: First, it was checked whether the second highest peak had at least 70 % of the height of the first. Second, if there was a third peak, it needed to be less than half the height of the second peak. Third, the lowest point between the two highest peaks needed to be less than half of the height of the second peak. The distance between the two highest peaks was the intermembrane spacing (Figure 4.13, c)

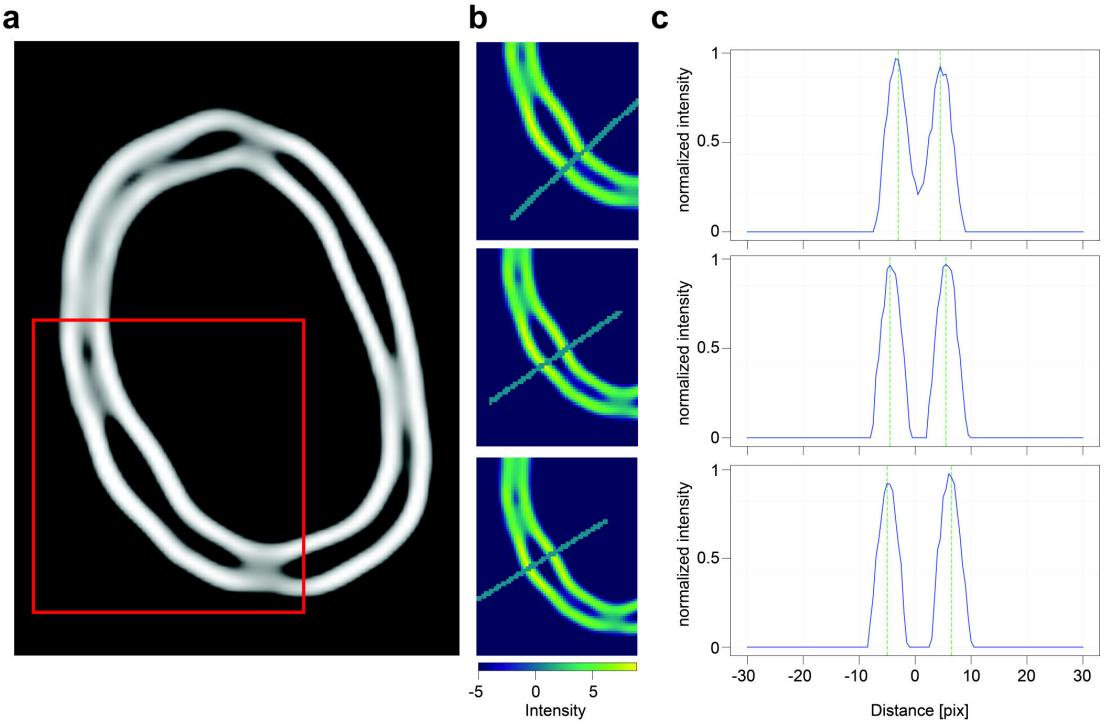

FIGURE 4.13: Illustrative visualization of membrane spacing calculation for three different sampling points. a) The membrane segmentation of the inspected double membrane vesicle (DMV) as 2D z-slice. The investigated region is highlighted in red. b) Three examples of spacing analysis. A line is placed orthogonal to the DMV contour, reaching 30 pixels in each direction. c) The intensity profiles with cutoffs as green dotted lines.

Additionally, to investigate the dependence of DMV intermembrane spacing on molecular pore distance, for each point of the DMV graph that had an assigned intermembrane spacing, the distance to the closest pore was calculated, using a cutoff of 40 pixels. This was done by calculating the Euclidean distance using the 3D position of the current contour point and the previously calculated pore positions.

Chapter 5

Image quality assessment in variable conditions via resolution estimation

The work detailed in this chapter was partially present as an unpublished manuscript (written by me and Prof. Dr. Carsten Sachse), and some of the text sections and figures in this chapter were taken from the manuscript as indicated by a star (*) in the section header or figure legend. Section headers marked with double stars (**) indicate content that has been substantially adapted and expanded from the manuscript. Figure legends or section headers with no star indicate figures or sections are not present in the manuscript.

5.1 Introduction**

Resolution estimation is an essential metric in cryo-EM, allowing for thorough data quality estimation. So far, resolution estimation has largely been used for protein structure maps, determined either via single-particle cryo-EM or subtomogram averaging. Resolution estimation has many advantages: It is indicative of the suitability for atomic model building and biological interpretation (Jamali et al., 2024). Local resolution estimates also allow for local low-pass filtering (Beckers and Sachse, 2020), making local resolution cutoff determination possible and thus reducing noise. Resolution estimates further allow for direct map comparisons, and therefore

facilitate benchmarking of different computational steps during structure determination (Tegunov et al., 2021) or even of data collection and imaging parameters (Tuijtel et al., 2024). However, the fact that resolution estimation is commonly used for refined maps yielded via either single-particle cryo-EM or subtomogram averaging only, is also a limiting factor: Comparing processing steps or data acquisition parameters usually rests on running the complete structure determination workflow (Berger et al., 2023; Tuijtel et al., 2024), which is work-intensive and can make direct and accurate comparisons more difficult.

Conventionally, resolution estimation is conducted by splitting the data into half-maps and calculating the Fourier shell correlation (FSC) between them (Chapter 1, section 1.1.5). Resolution estimates can be conducted on a global or on a local level. While global estimates are convenient, yielding a single easy-to-compare number, in many cases, local resolution estimates are required (Chapter 1, section 1.1.5). Local estimates can be, similar to global resolution estimates, conducted via Fourier shell correlation, however, on a local level instead. One essential parameter hereby is the local window size or radius. This choice is often left to the user who can choose in a case-dependent manner. Window size has previously also been tuned in a signal dependent manner, based on the assumption that smaller windows are sufficient for higher-quality maps, as the signal-to-noise ratio is increased (Cardone, Heymann, and Steven, 2013). Another approach is to tune the window size in a resolution dependent manner on a per-shell level. Window size remains an important factor which should be chosen carefully: A large window contains enough information for reliable measurements, but measurement locality suffers. A small window on the other hand runs in danger of containing insufficient information, and finding an appropriate balance is key for adequate resolution measurements (Beckers and Sachse, 2020).

One interesting recent example of validating image acquisition parameters via structure determination is the characterization of lamella quality. For *in situ* cryo-electron tomography (cryo-ET), the vitrified cells often need to be thinned before imaging

via focused ion beam (FIB) milling, yielding a so called lamella. This thinning is often necessary because the electron beam cannot adequately pass thick samples (Kudryashev, Castaño-Díez, and Stahlberg, 2012), as electrons scattered multiple times degrade image resolution (Schiøtz et al., 2024). However, FIB-milling also induces damage to the sample, especially at the lamella edges. This damage and its localization has been characterized via structure determination, with the help of global resolution estimates (Berger et al., 2023; Tuijtel et al., 2024).

In this work, we present a new algorithm for local resolution estimation, termed RESOLVE (**RESO**lution estimation in **V**arying **E**nvironments). We used extensive simulations to empirically derive resolution-dependent window radii on a per-shell level, optimizing the balance between measurement accuracy and locality. While yielding good results for averaged maps as derived from single-particle cryo-EM or subtomogram averaging, RESOLVE can also be applied to micrographs, tilt-series and tomograms, thus measuring resolution in two or three dimensions. We show that RESOLVE's resolution measurements yield meaningful results, allowing to judge particle quality on a early level by investigating the respective micrographs or tomograms. RESOLVE allows the characterization of radiation damage induced quality deterioration during tilt-series collection, and is able to capture sample-thickness-dependent quality differences. Finally, RESOLVE allows the characterization of FIB-damage without the application of subtomogram averaging.

5.2 Results

5.2.1 The procedure of resolution estimation**

In order to develop a robust resolution determination tool suitable for multiple data types in cryo-EM, we started to work with typical half-maps of single-particle cryo-EM maps, e.g. of the ribosome (EMD-0194). In the procedure, we iterate through every resolution shell (Figure 5.1): (1) we apply a bandpass filter to the

half-maps corresponding to the resolution shell and (2) we compute running correlation values for every map voxel, employing Cosine Similarity (CS) for real-space measurements:

$$Cs = \frac{\mathbf{A} \cdot \mathbf{B}}{|\mathbf{A}| \cdot |\mathbf{B}|} \tag{5.1}$$

Where A is comprised of all factors within the local window of the first half-map, and B of the associated factors in the second half-map. After looking up the predetermined resolution-dependent optimal window radius, this radius is employed to every map measurement location in order to estimate the correlation for a locally masked area. By including a permutated half-map correlation, (3) we obtain statistically relevant *p*-values for every measurement location.

Once steps 1-3 are iterated over for every resolution shell, we identify the highest resolution shell in analogy to the previously applied multiplied testing correction (Beckers and Sachse, 2020): for every measurement location, we analyze obtained *p*-values for all resolution shells and use multiple testing correction to derive *q*-values from our *p*-values by false-discovery rate (FDR) control (Benjamini and Yekutieli, 2001). Finally, we assign a resolution value to every voxel of the cryo-EM map when the *q*-value of the FSC estimate drops below a defined threshold for the first time, e.g. 1 % FDR.

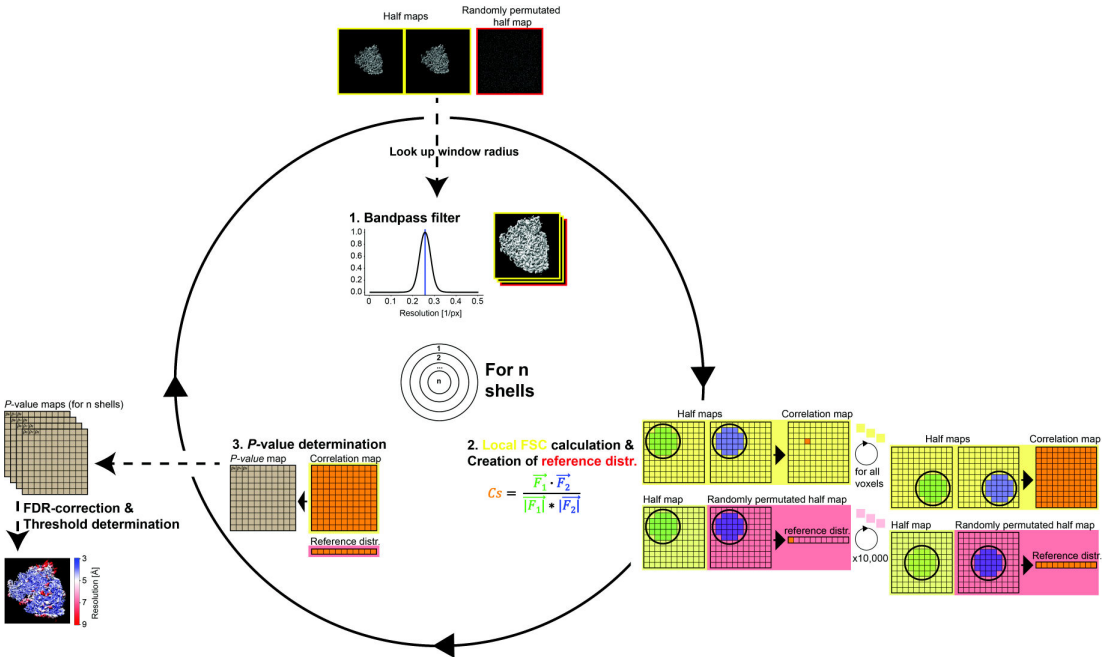

FIGURE 5.1: *Schematic presentation of the RESOLVE local resolution estimation workflow. Steps 1 - 3 are conducted for every resolution shell of the map. (1) Half maps and their permutated counterpart are bandpass-filtered according to the current resolution shell. Resolution-dependent window radii are obtained for every resolution shell. (2) Running correlation calculations are conducted and a reference distribution created. (3) Map of p-values (q-values after multiple testing correction) are determined. Finally, after iterating through every shell, FDR-correction is applied and for every voxel of the map, a resolution threshold is determined for when the highest shell drops below $q < 0.01$ (or 1 %).

In order to implement the adaptable window radius efficiently, we turned to previously proposed reverse-order approach (Penczek, 2020) that is an equivalent way of obtaining FSC values using band-pass filtered real-space volumes. Instead of excising local volumes and then conducting the Fourier correlation measurements, we generated a series of band-pass filtered half-maps keeping only the information band of the shell of interest (steps 1 and 2).

For the implementation of the band-pass filter operation in Fourier space (step 1), we used a hyperbolic tangent to avoid artifacts through spectral leakage. The filter was applied to both half-maps, introducing dependencies that are known to affect the correlation measurements (Beckers and Sachse, 2020; Penczek, 2020). Here, we dealt with these dependencies in the following manner (step 2): we created a correlation reference distribution for each resolution shell by permutating one of

the original input half-maps, applying the same bandpass filters as for the original half-maps, and conducting correlation measurements between the permutated half-map and the original half-map for 10,000 random locations. As this reference correlation distribution now contains the same dependencies as the originally bandpass-filtered half-maps, we were able to derive the correct *p*-value map by determining the position of each measurement value within the sorted reference distribution. In this context, data from refined maps, derived from single-particle cryo-EM and subtomogram averaging, is treated different from data derived from single recordings, namely micrographs, tilt-series and tomograms. As these maps are derived from a single recording, they portray measurement specific noise characteristics, which are necessary to retain in the reference distribution. Keeping the noise characteristics intact is especially relevant for tomograms: The reconstruction in Fourier space and the missing wedge introduce additional artificial dependencies represented as continuous changes in the power spectrum. To include these effects, phase-permutation in Fourier space was employed instead of real-space permutation, keeping the power spectrum of the permutated map intact.

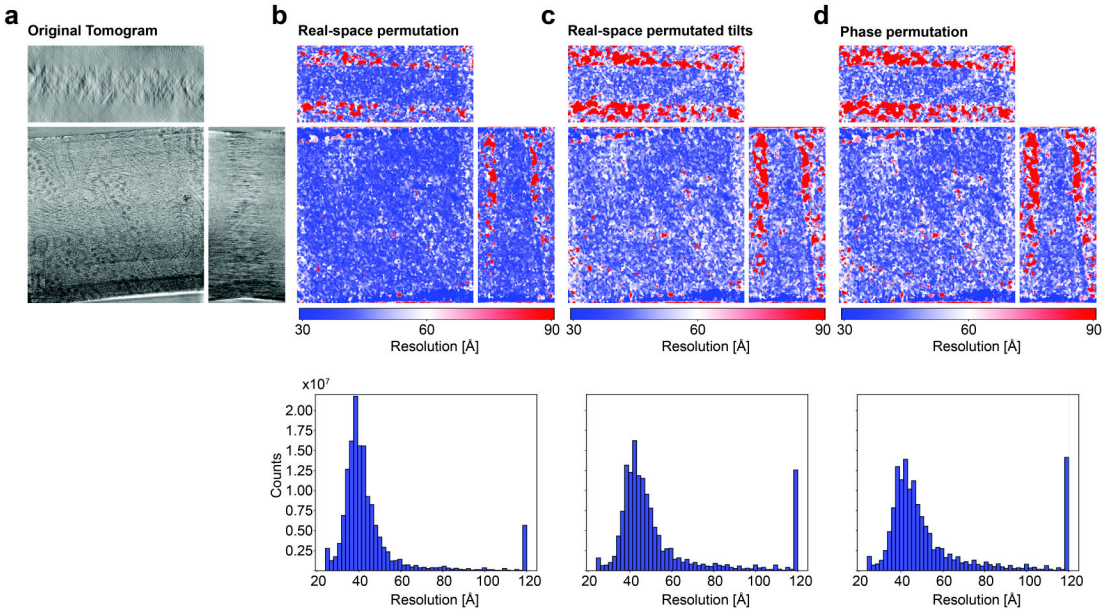

FIGURE 5.2: *Effect of tomogram reconstruction on resolution estimates. Top row: center slices of the three axes. Bottom row: Histograms of resolutions of the complete tomograms. a) Original, denoised tomogram. b) Resolution measurement using real-space permutation for reference distribution creation. c) Resolution measurement conducted by using a tomogram as reference distribution created by real-space permutation of the single tilt images before reconstruction, ensuring any potentially biases introduced via tomogram reconstruction are included in the reference distribution. d) Resolution measurement using phase-permutation for reference distribution creation.

To show that this procedure is able to compensate for reconstruction and missing wedge artifacts, we created a test case by permutating every tilt of one half-map individually, reconstructing the tomogram, and using it unaltered as permutated half-tomogram for reference distribution creation. In this case, any artifacts introduced via tomogram reconstruction, including the missing wedge, would be expected to be present in the reference distribution, leading to a realistic null distribution. Indeed, comparing the resulting resolution measurements with the case of reference distribution creation by real-space permutation (compare Figure 5.2 b and c), we measured more conservative resolutions (worse or lower resolutions refer to higher values, blue represents high resolution and red low resolution in the figure), indicating that artificial dependencies introduced by tomogram reconstruction inflate correlation measurements. It is reasonable to assume that phase-permutation instead of real-space permutation is an adequate and practical shortcut, as it keeps

the power spectrum with is systematic biases intact, and it indeed yields very similar results (5.2, d).

Finally, after running through every shell, multiple testing correction was applied to the *p*-value map to derive the *q*-values for the final per-voxel threshold determination.

5.2.2 Window radii determination**

A critical parameter for the reliability of the FSC or, more generally, local correlation calculations, is the choice of the real-space window size or radius (Beckers and Sachse, 2020; Cardone, Heymann, and Steven, 2013). Therefore, we developed a strategy to adapt the window radius depending on the spatial frequency. Our basic aim was to generate a lookup table for the window radii in order keep the window as small as possible for obtaining resolution values of high locality while still yielding reliable resolution measurements. To approximate the optimal radius, measurements were conducted on 1000 simulated half-map pairs with varying spectral signal-to-noise ratios (SSNR) and the smallest window radius still capable of reliably estimating the simulated SSNR was identified. For the simulations, we used the atomic model of human parechovirus shell (PDB-ID 4udf), that exhibited a high surface-to-volume ratio and differently distributed signal across the map. For each different shell tested, we created a new set of simulated maps with differing SSNR, in order for the SSNR of each map only corresponding to the map-specific shell of interest. We sampled the frequency domain for 10 shells, covering the frequency spectrum evenly, with the lowest at a normalized frequency of 0.1 and the highest at 0.47, close to the Nyquist limit. For each of these 10 shells, 100 different half-map pairs of varying SSNRs were simulated, ranging from roughly 0.9 to 4.4 by scaling the mean of the white noise power spectrum relative to the mean of the shell-specific power spectrum of each simulated map. Given that SSNR is related to the Fourier shell correlation at a given resolution shell:

$$\mathrm{SSNR} = \frac{\mathrm{FSC}}{1-\mathrm{FSC}} \tag{5.2}$$

the simulated SSNRs corresponded to practically relevant FSC values between 0.1 and 0.8. For each shell-specific simulated map with known SSNR, we performed the RESOLVE algorithm for local correlation estimation, only focusing on the map-specific shell, using a series of differing window radii ranging from 2 to 30 pixels for 3D. We defined correct signal detection as 95% of signal regions derived from the noise-free maps passing the $p<0.01$ threshold. This way, the smallest window still detecting the known SSNR correctly ($p<0.01$ when testing correlation against the reference distribution) was identified. The SSNR/FSC specific sampling points were linearly interpolated to create an extensive lookup table (Figure 5.3, a) from which we obtained the smoothed curve for the relevant FSC value of 0.143. To smooth the derived window radii relationship with spatial frequency, we used a Savitzky-Golay filter (Savitzky and Golay, 1964). In this way, using simulations of known SSNRs, we generated a lookup table with the smallest appropriate window size to still reliably estimate resolution at the corresponding spatial frequency.

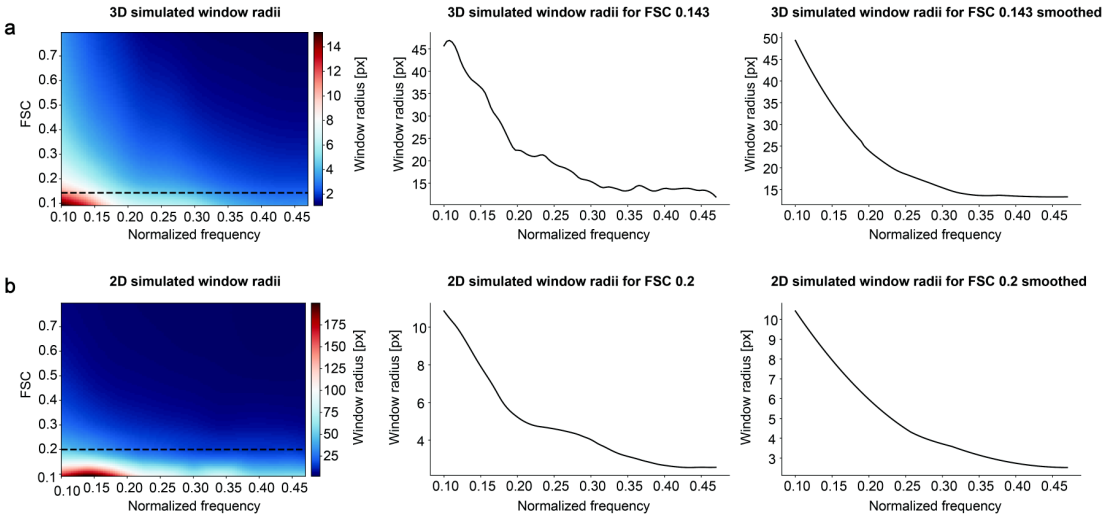

FIGURE 5.3: *Window radii determined from simulations for 2D and 3D. a) Window radii for 3D measurements. Left: Table with the determined frequency dependent window radii for different spectral noise levels, as represented by the FSC. FSC 0.143 is highlighted with a black dotted line. Middle: Extracted original curve for FSC 0.143. Right: Smoothed curve via a Savitzky-Golay filter for FSC 0.143. b) Window radii for 2D measurements. Left: Table with the determined frequency dependent window radii for different spectral noise levels, as represented by the FSC. FSC 0.2 is highlighted with a black dotted line. Middle: Extracted original curve for FSC 0.2. Right: Smoothed curve via a Savitzky-Golay filter for FSC 0.2.

For micrograph measurements, due to the missing third dimension, larger window radii are required to include a sufficient number of pixels for reliable measurements. In particular with lower frequencies, it goes at the expense of the locality of the measurements. Consequently, the window radii necessary for reliable correlation calculations are naturally much higher in 2D, and for the simulation, we tested radii between 2 and 100 pixels. Therefore, for the window radii assessment in tilt-series and micrographs, the parameters during the simulation were adjusted to be more lenient, i.e. choosing a SSNR according to a FSC of 0.2 (instead of 0.143 for 3D) and accepting a window radius when 90% of measurements passed (rather than the 95% for 3D). This design was chosen in order to limit the window size for the 2D measurements (Figure 5.3, b).

5.2.3 Application and benchmarking of RESOLVE with single-particle cryo-EM maps**

To benchmark our proposed approach for local resolution estimation, RESOLVE's resolution measurements of the ribosome (EMD-0194) and Sars-CoV-2 spike protein (EMD-34658) maps were compared with commonly used programs: blocres (Cardone, Heymann, and Steven, 2013), Monores (Vilas et al., 2018), RELION (Zivanov et al., 2022), ResMap (Kucukelbir, Sigworth, and Tagare, 2014) and SPOC/FDR-FSC (Beckers and Sachse, 2020). In comparison, when the local resolutions were mapped on the isosurface, the different programs yield overall similar results in the relative variation of resolution while absolute resolution values differed (Figure 5.4). However, differences could be more clearly visualized by inspecting local resolution values obtained from a single map slice cut through the volume.

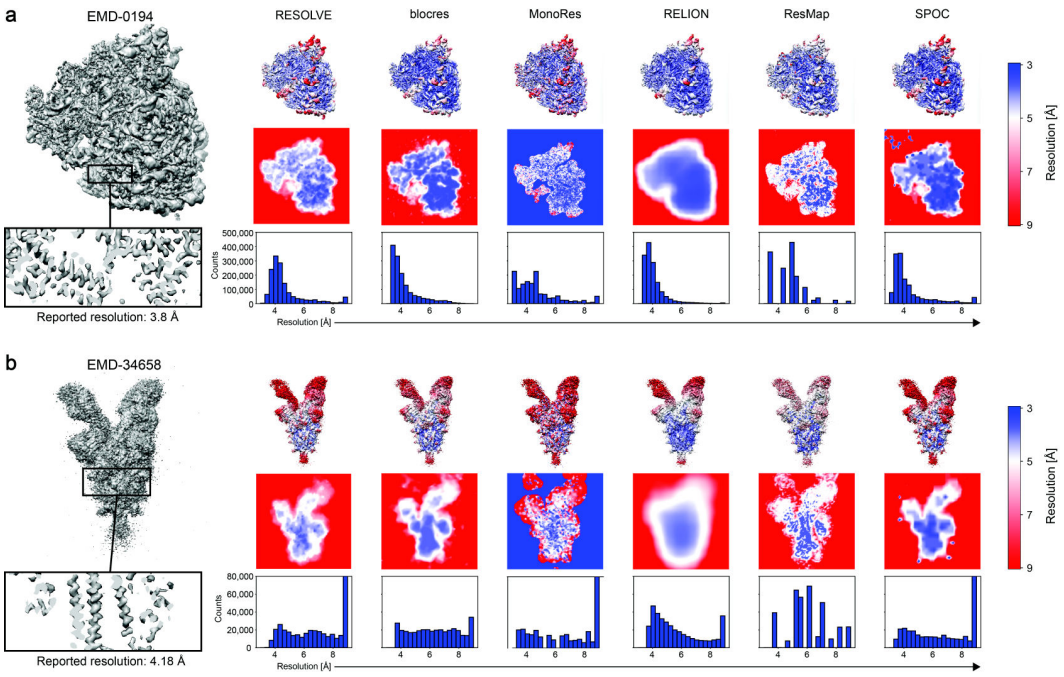

FIGURE 5.4: *Benchmarking RESOLVE's local resolution estimates for 3D cryo-EM maps. For a) and b), the following information is shown in the panels: Top left: cryo-EM map with inset slice through α-helical density. Top right: Results of local resolution estimation using RESOLVE, blocres, MonoRes, RELION, ResMap and SPOC. Top row: isosurface with mapped local resolution values (colorbar right). Center row: slice through volume center with color density display according to local resolution. Bottom row: Histogram of local resolution values for each program above. a) Rabbit 80S ribosome (EMD-0194). b) SARS-CoV-2 Omicron BA.1 spike trimer (EMD-34658).

High variations within short distances represent resolution assignments with high locality (e.g. Monores and ResMap), while using a large window leads to a blurred appearance (e.g. RELION). Clearly, excessively large resolution differences between neighboring pixels are the result of unreliable measurements, and in such circumstances, high resolution outliers can be found in unexpected locations. The central parameter that determines the balance between high locality and information content is the window size, which can be tuned for most programs. Therefore, RESOLVE opted for adaptable window radii. The effect of adaptable window sizes could be best observed when comparing RESOLVE and blocres: At first glance, they yield similar results, especially for low resolution regions. For higher resolutions, however, RESOLVE revealed more locality. The adaptable window size used in RESOLVE can make use of smaller windows while still yielding reliable results in high resolution regions, and for lower resolution regions, it can refer to larger window sizes. In contrast, blocres remains conservative for high-resolution shells, using the same window size required for reliable measurements of the lowest resolution shell, sacrificing locality and therefore available high-resolution information. RESOLVE's optimized adaptable window sizes for different resolution shells strike the accommodate balance between high-resolution information alongside robust low-resolution resolution estimation.

5.2.4 Application of RESOLVE to (2D) micrograph movies**

After establishing the RESOLVE work-flow for three-dimensional (3D) cryo-EM maps, we applied RESOLVE to micrograph movies to determine local resolution estimates and derive a global micrograph quality criterion. Here, we used the 2D equivalent of the Fourier ring correlation in real space, obtained from half-sets made of even and odd frames of each micrograph (Figure 5.5, a). We also adjusted the generation of the lookup table for the optimal window radius to account for the smaller number of pixels involved in the before described manner. The remaining

workflow steps are identical to the way described above for 3D maps. First, we tested the method on a previously recorded TMV micrograph 'TMV_004_Sep18' (EMPIAR-10305). The local resolution mapped on the motion-corrected micrographs showed resolutions up to 6 Å in areas of TMV as opposed to empty areas of vitreous ice between 20 and 30 Å. When comparing the motion-corrected sum of the aligned frames versus the simple sum of the unaligned frames, we detected a significant deterioration in resolution in areas of TMV and vitreous ice (Figure 5.5, b). However, it has to be kept in mind that a single high-resolution measurement cannot not be fully representative of the information content due to the commonly applied underfocus in cryo-EM image acquisition. To investigate the effect of the CTF, we plotted the frequency-dependent changes of the median *q*-values obtained from the entire micrograph against the estimated CTF by CTFFIND4 (Rohou and Grigorieff 2015), showing a clear inverse correlation: The *q*-values are lowest in the regions of high contrast transfer and highest in the regions of the zero intercepts. To mitigate this effect for the case of micrographs measurements (as well as tilt-series or tomograms below), we adjusted the final threshold determination step to be more lenient, using a *q*-value cutoff of 0.05 instead of 0.01 as previously described. In this context, RESOLVE also allows the determination of a single global resolution value based on the local estimates of the micrograph. We term that measurement *median resolution*, as it is derived from a list of shell-associated median *p*-values (for details, see methods section 5.4.3).

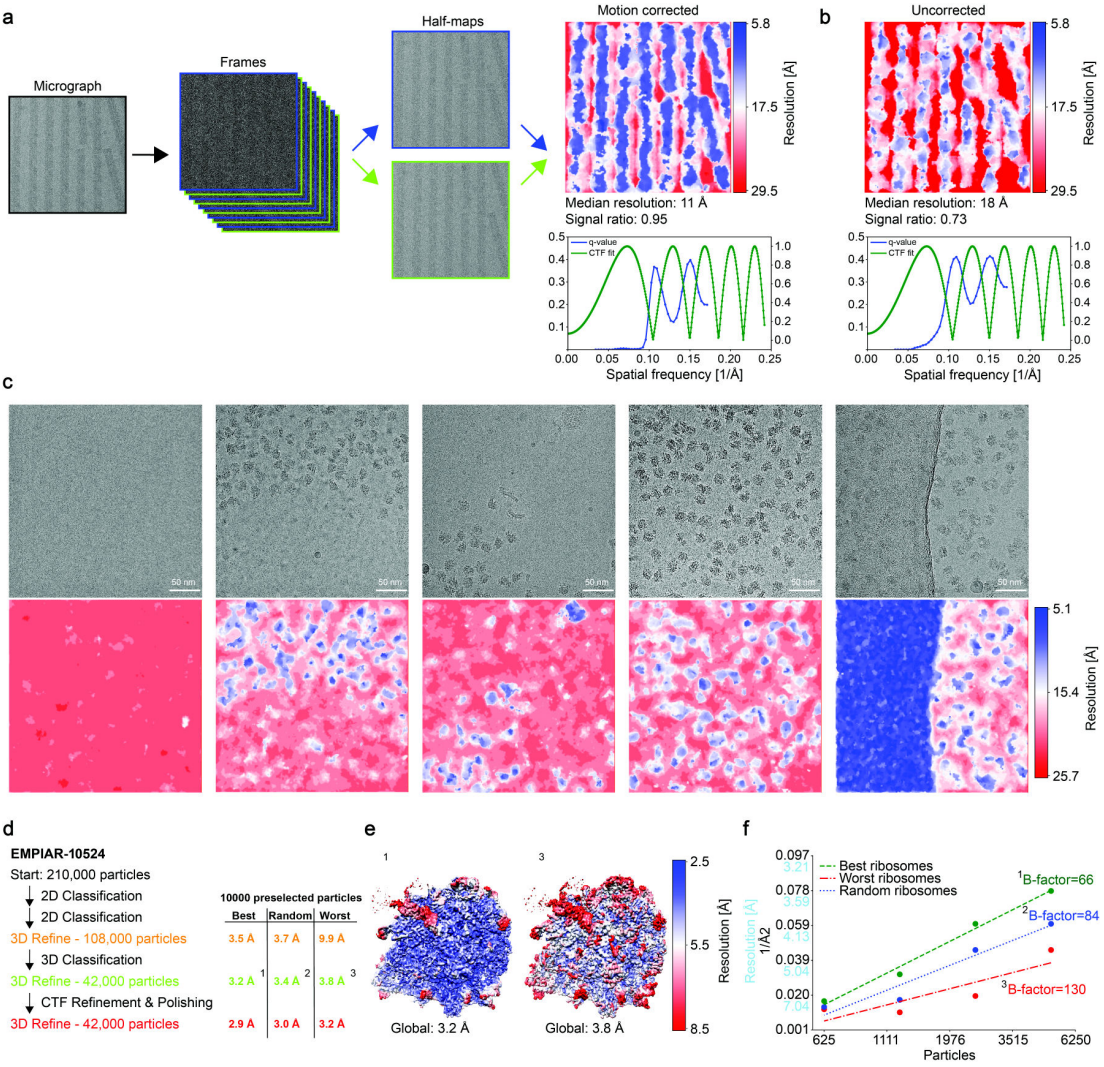

FIGURE 5.5: *Application of RESOLVE to 2D electron micrographs. a) Micrograph movies are split into odd and even for half-set creation. Local resolution map of a motion corrected micrograph from tobacco mosaic virus (TMV) (EMPIAR-10305). Below the micrograph: plot of the median q-values from our measurements compared with a CTF-fit of CTFFIND4 for the same micrograph. b) In contrast, resolution estimation of a simple sum of aligned TMV micrograph frames (not motion-corrected). Below the micrograph: plot of the median q-values from our measurements compared with a CTF-fit of CTFFIND4. c) A set of micrographs (top) with their resolution estimates (bottom) of E. coli 50S ribosomes (EMPIAR-10524). The columns show from left to right: an empty micrograph, a density gradient, a mostly empty micrograph with a strong signal from an artifact in the top right corner, a regular micrograph and a micrograph partially containing carbon film. d) Left: Basic procedure of single-particle workflow (EMPIAR-10524). Right: Global resolution estimates for 10,000 best, random and worst preselected particles by RESOLVE on a micrograph level. e) Local resolution as estimated by RESOLVE mapped on 3D refinement results of best and worst particles. f) B-factor comparison for 'best', 'random' and 'worst' particles as determined by RESOLVE.

When RESOLVE estimated the local resolution of micrographs from E. coli 50S ribosomes (EMPIAR-10524), it yielded higher resolution estimates for areas with particles present and lower in case they were absent (Figure 5.5, c). However, two additional cases led to high resolution as well: first, in the case of devitrified ice and second when carbon film was present. While operating with median q-values, those effects can be minimized depending on the extent present in the micrographs. For the purpose of the following tests, all micrographs including carbon film were excluded.

After showing the principal utility of RESOLVE for estimating local resolution on micrographs, we applied RESOLVE as a quality criterion for imaged particles on a micrograph level, testing whether better resolved single particles as judged by RESOLVE give rise to improved 3D maps. For that purpose, we processed a ribosome dataset (EMPIAR-10524) and the best and worst 10,000 particles as determined by RESOLVE on a micrograph level were selected. These two subsets 'best' and 'worst' subsets, in addition to a third 'random' subset with randomly chosen 10,000 particles were processed. In order assess the effect of image processing, we conducted this procedure at three different points of the processing workflow (Figure 5.5, d): First, after two rounds of 2D classification, second, after one further round of 3D classification and third, this set of particles was subjected to a round of CTF-refinements and particle polishing following the RELION documentation. For those three types of refinements, the highest resolution was consistently obtained when using the 'best' particle set when compared with the 'random' or 'worst', where the lowest resolution was obtained. For instance, after 3D classification with a total of 42,000 particles, RELION's resolution estimates (FSC 0.143) of the cryo-EM map were 3.2, 3.4 and 3.8 Å for the 'best', 'random' and 'worst' particle sets, respectively. Visual comparison of the 'best' vs. 'worst' maps confirms the resolution estimate (Figure 5.5, e). The fitted Rosenthal-Henderson B-factors at 66, 84 and 130 Å^2 for 'best', 'random' and 'worst' particle sets, respectively, further supports the notion that preselection based on RESOLVE's local resolution can potentially be used to

improve the 3D map quality (Figure 5.5, f). After CTF refinement and particle polishing, the difference became smaller yielding 2.9, 3.0 and 3.2 Å, respectively, as particle polishing improves further the signal from individual particles. Together, RESOLVE is capable of estimating the local resolution on micrograph movies. The obtained global and local quality criteria can be used to select for better subsets of micrographs and particles, resulting in improved cryo-EM reconstructions.

5.2.5 Application of RESOLVE to tilt-series and tomograms**

After successfully applying the local resolution estimation tool RESOLVE to 2D micrograph movies, we investigated RESOLVE's utility for cryo-electron tomography. First, we sought to evaluate RESOLVE for tilt-series quality measurements and later on for entire tomographic reconstructions. For the tilt-series acquired on a lamella from -61 to 40°, we applied RESOLVE to each tilt micrograph based on the half-set frames, estimating local resolution in analogy to 2D micrographs (Figure 5.6, a).

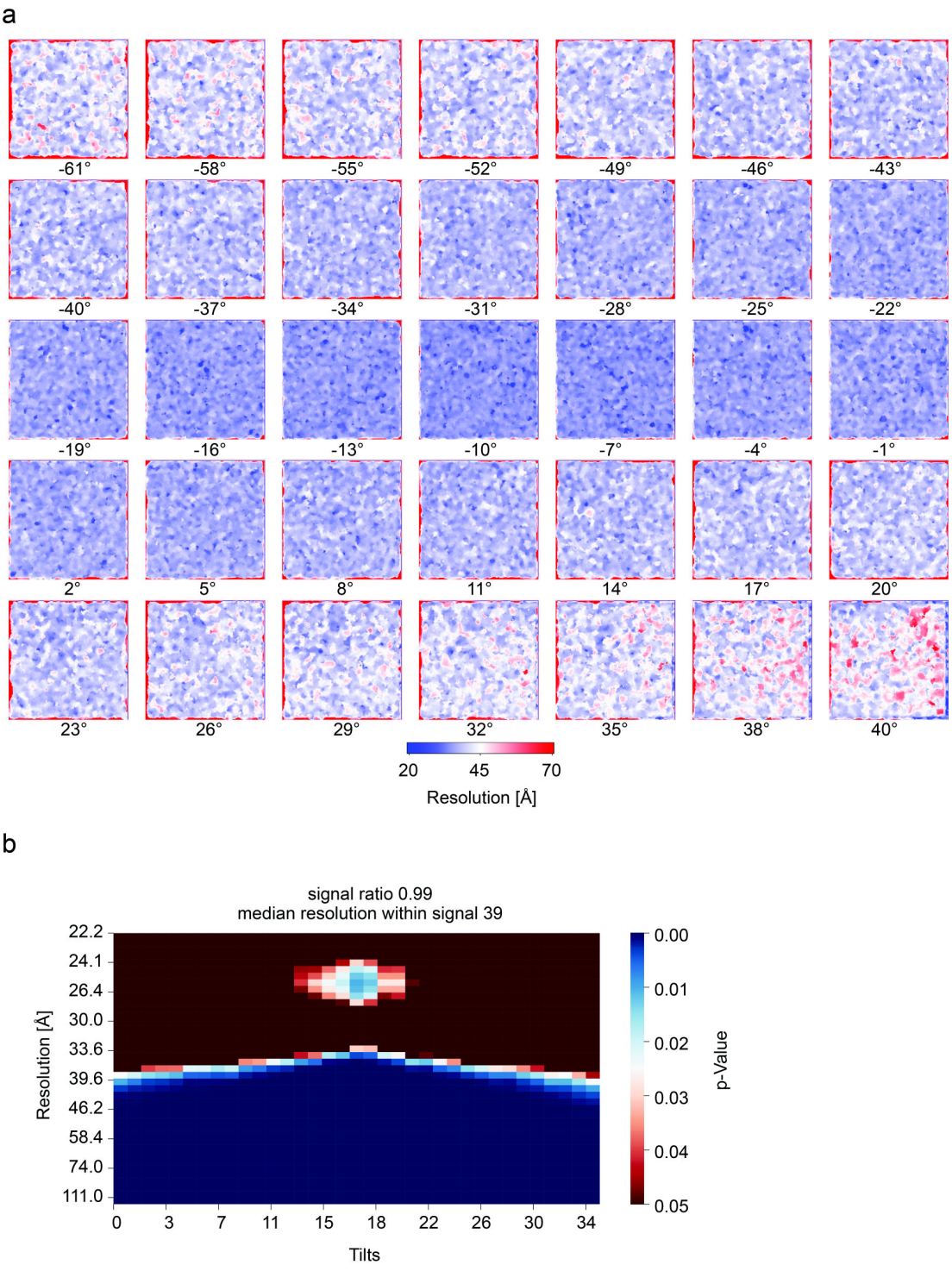

FIGURE 5.6: *Local resolution estimates of tilt-series. a) Resolution maps (color bar below from 20 to 70 Å) for full tilt series of 35 tilted micrographs from -61 to 40°. b) Stacked median *p*-values for each spatial frequency, resulting in a heatmap, the *p*-value map.

For a more compact visualization, we stacked the median *p*-values for each spatial frequency vertically for each tilt, resulting in a heat map, from hereon called *p*-value

map (Figure 5.6, b). The map has a pyramid-like shape, where the central tilts show highest resolution. This finding can be rationalized because of the fact that data was acquired dose symmetrically, i.e. the planar tilts were recorded first due to radiation damage. Therefore, likely due to radiation damage and sample thickness, tilts that were recorded later showed consistently lower resolution estimates, giving rise to the pyramid shape of the heat map. As expected, the above-described effects of the CTF could be observed as well, as p-values inversely correlate with the defocus and drop almost to 0.01 for higher frequencies a second time. RESOLVE's local resolution estimates capture the deterioration of resolution for tilted micrograph acquisitions due to buildup of radiation damage at higher tilt angles.

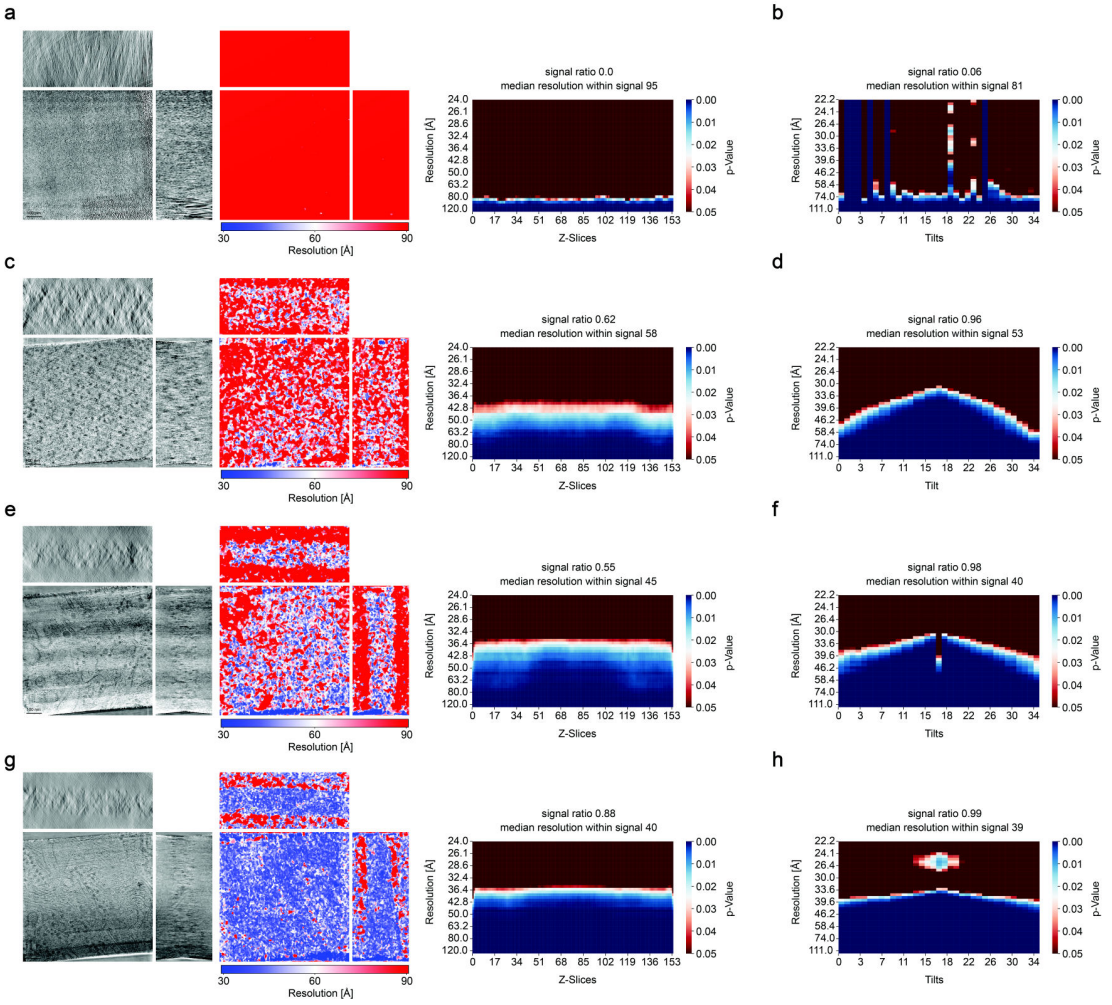

FIGURE 5.7: *Local resolution estimation in tomogramas. Four illustrative tomograms with their central slices through the three axis visualized with and without resolution next to the stacked median *p*-values for each spatial frequency per z-slice (a, c, e, g). On the right, the same stacked median *p*-values for each spatial frequency per tilt of the associated tilt-series (b, d, f, h).

Next, we applied RESOLVE for 3D local resolution estimations of reconstructed 3D electron tomograms. In some cases in the investigated dataset, when the tomogram quality was poor due to the lack of clear features during tracking, the quality can be easily identified by visual inspection of the tomogram, and was also well-represented by our resolution estimates, which were consistently at the low-resolution end (<90 Å) (Figure 5.7, a). However, in many cases, it requires expert knowledge to judge tomogram quality by eye. Here, RESOLVE may provide valuable support for tomogram quality assessment (Figure 5.7, c, e, g).

The same visualization of portraying the median *p*-values for each spatial frequency

that was used for tilt-series, could be applied to tomograms, only that here, z-slices were used instead of tilts. Interestingly, outlier results for tilts, as for single artificially well-resolved tilts or tilts that show exceptionally bad resolution (Figure 5.7, b, f), were not visible in the corresponding *p*-value map for tomograms. This behavior is to be expected, as location within single tilts, which are projection images, cannot be correlated with real space-locations in the associated tomograms, they rather seem to contribute equally to the outcome of the tomogram resolution measurement. Within the tilt-series *p*-value maps, high variations regarding the steepness of the resolution falloff for earlier and later tilts can be observed. For tomograms *p*-value maps on the other hand, there only seems to be a slight falloff for the lowest and highest z-slices. Notably, CTF-associated resolution fluctuations were also visible in tomograms, although less strikingly than for tilt-series (compare Figure 5.8 a and b).

To validate our resolution estimates, we turned to another benchmarking approach involving the local quality assessment of particles located in subtomograms, analogous to the presented cases on single-particle cryo-EM.

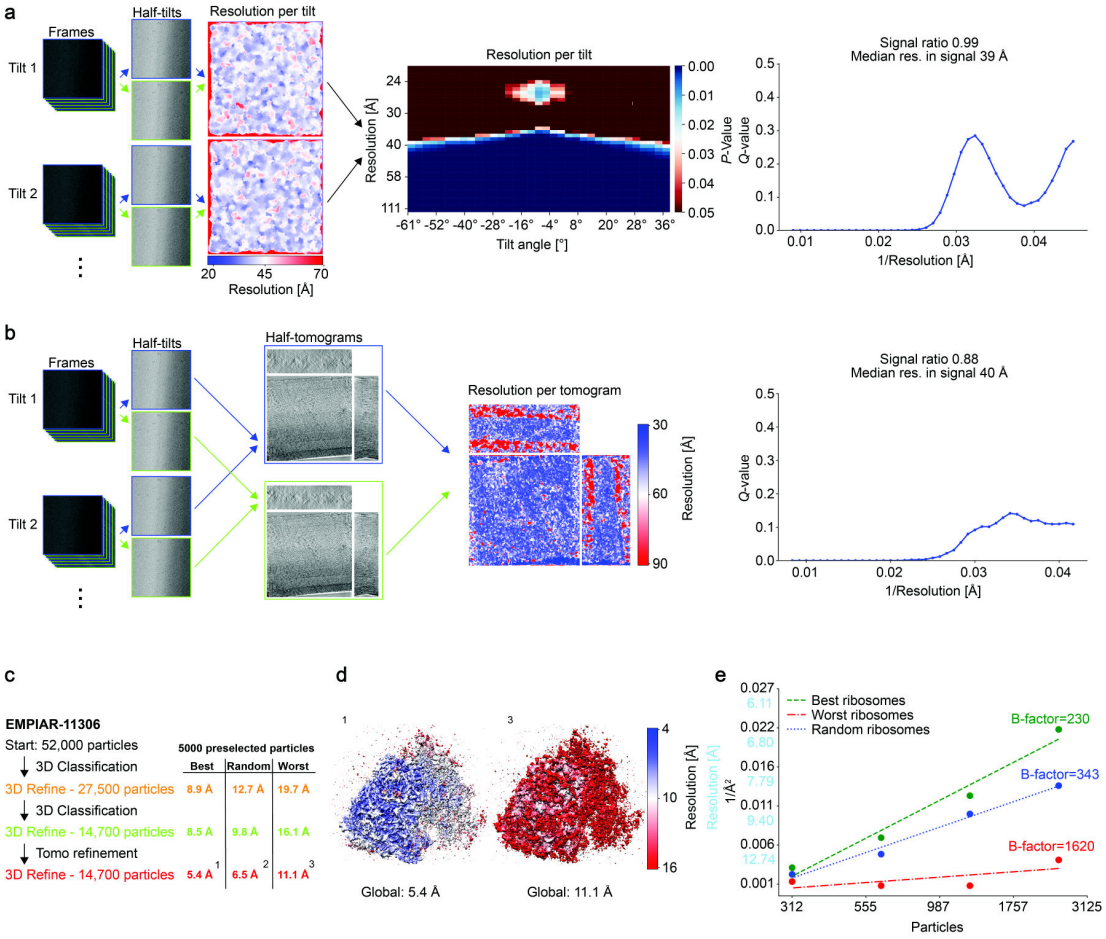

FIGURE 5.8: *Application of RESOLVE to cryo-electron tilt-series, tomograms and subtomograms. a) Procedure of dividing a tomographic tilt series into two half-tilt-series. Middle: resolution measurements displayed as heat map of vertically stacked tilts with median *p*-value per resolution. Right: median *q*-values per resolution. b) Procedure for dividing a tomogram into half-tomograms. Middle: local resolution measurements displayed in tomograms. Right: median *q*-values per resolution. c) Left: Basic procedure of subtomogram averaging (EMPIAR-11306). Right: RELION global resolution estimates for 5000 'best', 'random' and 'worst' preselected ribosome subtomogram based on RESOLVE's local resolution assessment. d) Local resolution as estimated by RESOLVE mapped on 3D refinement results of 'best' and 'worst' particles. e) B-factor comparison for 'best', 'random' and 'worst' particles as determined by RESOLVE.

Using ribosomes from the dataset EMPIAR-11306, we initially picked a total of 52,000 particles, from which a cleaner subset of 14,700 particles was generated by filtering for heterogeneity by two rounds of 3D classification (Figure 5.8, c). Next, we used RESOLVE to assess local tomogram quality, again creating three subtomogram subsets of 5000 particles, i.e. according to the 'best', 'random' and 'worst' local resolution criteria. Subtomogram averaging was performed before running the

RELION tomo refinement cycle, yielding 8.5, 9.8 and 16.1 Å resolution (FSC 0.143 criterion RELION), and after running the RELION tomo refinement cycle, yielding 5.4, 6.5, 11.1 Å for 'best', 'random' and 'worst' particles, respectively. Surprisingly, we detected large differences in resolution and in map quality between best and worst particles at all three stages of refinement (Figure 5.8, d). The Rosenthal-Henderson plot and the fitted B-factors of 230, 343 and 1620 $Å^2$ (best, random, worst, respectively) confirmed the previous observations (Figure 5.8, e) that the preselection based on RESOLVE's tomogram local resolution estimates presents an alternative way to assess particle quality at an early stage.

As RESOLVE was able to demonstrate robust local tomogram quality assessment useful for subtomogram averaging, we wanted to investigate other matters related to resolution in cryo-tomography. In order to examine the effects of damage inferred by the high-energy beam of a focused ion beam (FIB) milling instrument (Berger et al., 2023; Lucas and Grigorieff, 2023), EMPIAR-11306 was further investigated. First, the local ribosome resolutions as determined by RESOLVE on a tomomgram level were plotted into a graph, to see whether an obvious relation between lamella edge distance (from top or bottom surface) and measured ribosome quality could be determined. To increase the number of sampling points within the lamella, the ribosome particle set from subtomogram averaging after the first 3D classification was used, including approximately 27,400 particles. In figure 5.9 some example tomograms containing a large number of ribosomes are portrayed. Here, the quality drop close to the lamella edges (top or bottom surfaces) can be clearly observed, especially for the thin lamellae (Figure 5.9, a). This observation is in accordance with previous studies (Berger et al., 2023; Lucas and Grigorieff, 2023), showing that FIB-damage occurs in the regions close to the lamella edge, degrading resolution. However, for thicker lamellae, the results are not as clear, and it seems that resolution decreases at the lamella edges, but also in the center to some extend (Figure 5.9, b).

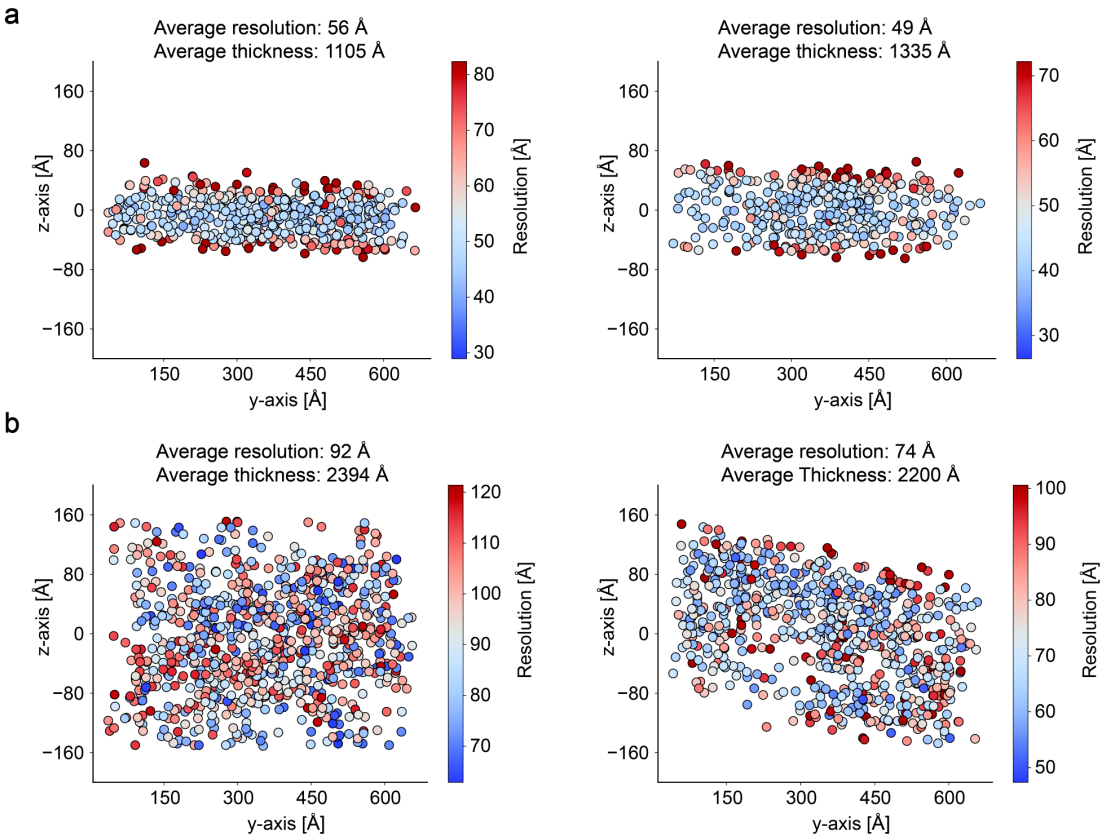

FIGURE 5.9: Ribosome resolution estimates for illustrative lamellae. Only z-axis and y-axis are portrayed. a) Two examples of thin lamellae. b) Two examples of thick lamellae.

For a more thorough investigation of the quality dependence on lamella edge distance, we conducted quantifications across the full dataset. For that, the lamellae were modeled so that the distance of each voxel within its lamella to the edges (top or bottom surfaces) could be measured (Figure 5.10). Investigating the measurements for single tomograms (without focusing the measurement on ribosomes, but using all local resolution measurements inside the lamella and sorting them after lamella edge distance), there again seems to be a clear drop in resolution with increasing lamella edge distance. Interestingly, for the investigated thick tomogram, the highest resolution was not in the lamella center: Quality seemed to decrease again close to the center (Figure 5.10, a).

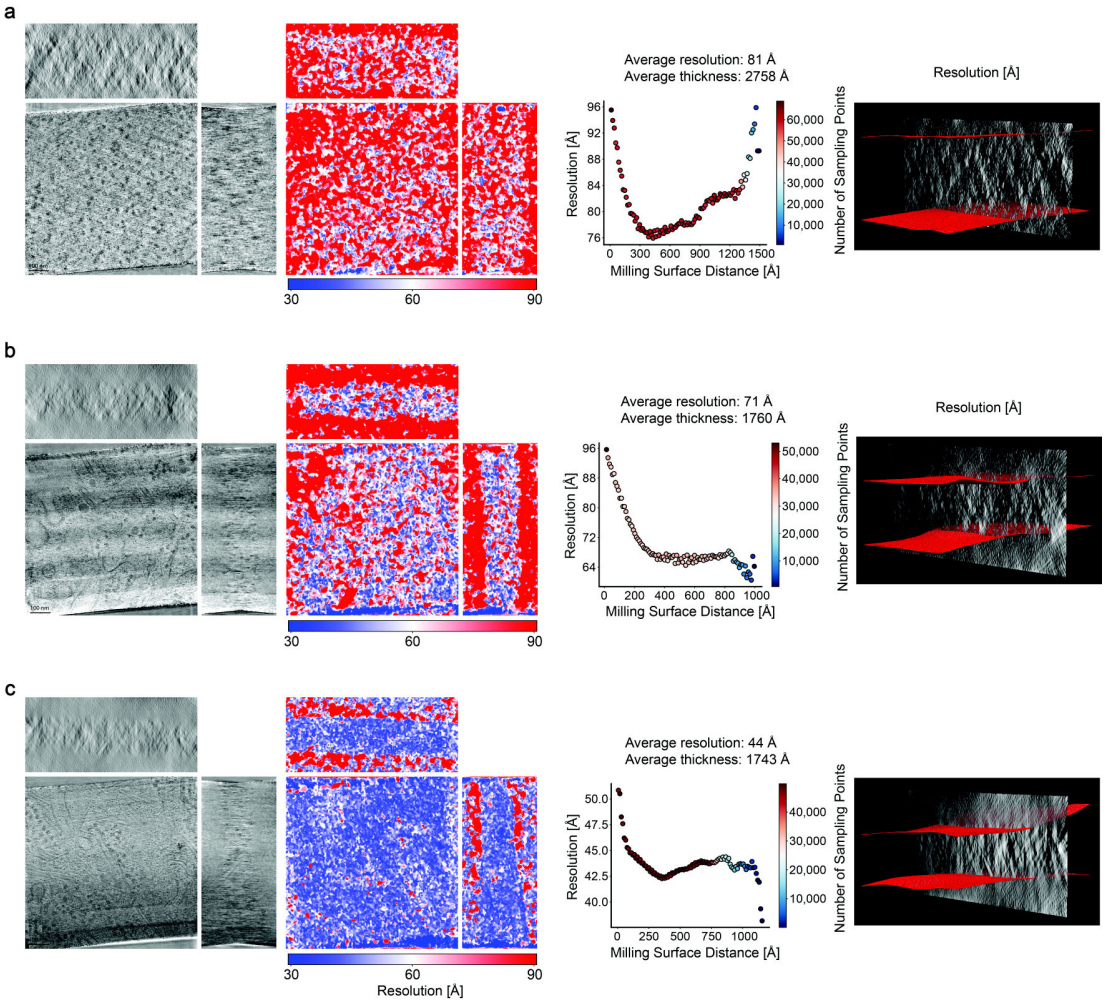

FIGURE 5.10: *Resolution estimates of three illustrative tomograms (a-c). Left: Central slices through the three axis visualized with and without resolution. Middle: Resolution vs. closest milling surface distance. Right: Lamella model though central xz-slice. a) Tomogram of a thick lamella. b) tomogram of a thin lamella, lower quality. c) Tomogram of a thin lamella, lower quality.

For an in-depth analysis of these findings, every tomogram where the lamella could be modeled was included in the analysis (total of 106 tomograms). First, we plotted the average resolution for voxels inside FIB-milled lamellae as determined by RESOLVE per corresponding lamella thickness (Figure 5.11, a). While the relation is not obvious at first glance, it does seem that lamella thickness limits the achievable resolution. The reason a clearer relationship cannot be observed may be due to thinner lamellae having a higher relative amount of FIB-damaged content. This reasoning is in accordance with the shape of the high-resolution limit as it relates to lamella thickness, which is especially pronounced when outliers are filtered out

(Figure 5.11, a, bottom), and the observation that the highest resolved tomograms are found for lamellae around 160 nm thickness, not for thinner lamellae. Additionally, diverse tomogram content contributes to resolution differences. Therefore, when only focusing on the position of ribosomes, i.e. within 10 nm radius of ribosome coordinates, the relationship became more pronounced (Figure 5.11, b).

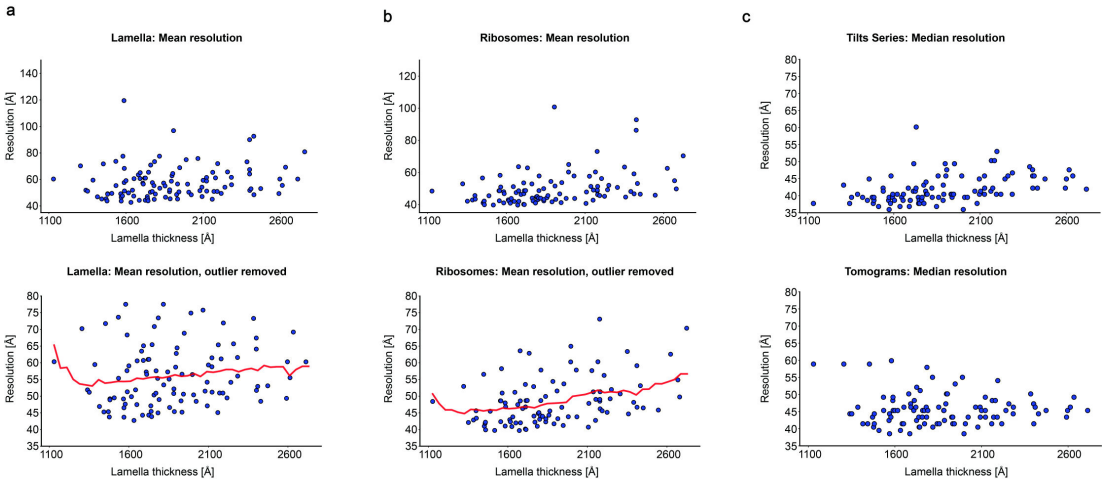

FIGURE 5.11: *RESOLVE measurements for global median resolution for cryo-electron tilt-series and tomograms, compared to average resolution measurements focused on the lamella. a) Scatter plot of average resolution per tomogram within the lamella vs. lamella thickness. Top: full dataset (106 tomograms). Bottom: Outliers > 80 Å removed, moving average as red line (window size 400 Å). b) Scatter plot of average resolution of ribosomes (averaged per tomogram, using the particle set determined from subtomogram averaging with 14,700 ribosome locations) within the lamella vs. lamella thickness. Top: full dataset (102 tomograms). Bottom: Outliers > 80 Å removed, moving median as red line (window size 400 Å). c) Top: Scatter plot of median global resolution (as described in methods) in tilt-series (n=106). Not focused on lamella. Outliers > 80 Å removed. Bottom: Scatter plot of median global resolution in tomograms (as described in Methods) (n=106). Not focused on lamella. Outliers > 80 Å removed.

These results are specifically focused on the lamellae, which is difficult to achieve in an automated setting. When comparing these results with RESOLVE's globally derived median resolution estimates, which are not specifically focused on the lamellae and therefore expected to be less accurate, we found the shape of the resulting distribution to be similar to the averaged lamellae estimates (compare Figure 5.11, a and b, bottom), indicating the measurement viability: The best resolved tomograms again contain lamellae of around 160 nm thickness, with resolution dropping

for thinner and for thicker tomograms. Interestingly, when measuring the globally derived median resolution for tilt-series instead of for reconstructed tomograms, the resolution drop for thinner lamellae is less pronounced (Figure 5.11, c, top).

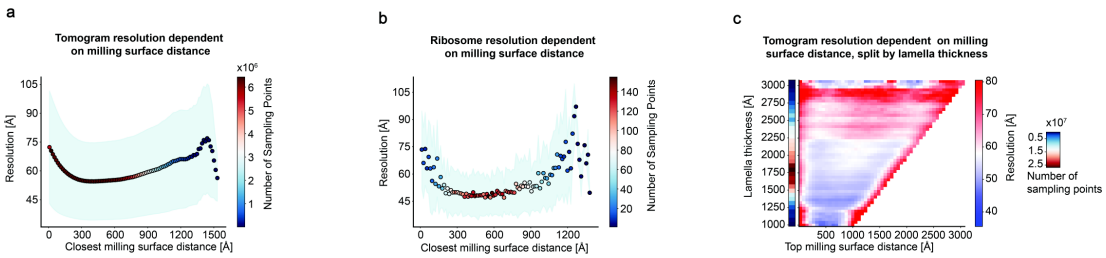

FIGURE 5.12: *Quantification of resolution dependence on lamella edge distance. a) Average resolution vs. distance to closest milling surface for 106 tomograms. Blue background coloring portrays standard deviation. b) Average resolution vs. distance to closest milling surface focused on ribosomes. Blue background coloring portrays standard deviation. c) Lamella thickness vs. distance to top milling surface, resolution and particle numbers are each color coded separately. Full set of 106 tomograms.

Next, we plotted the average resolution derived from the full set of tomograms (associated with voxels located within the lamellae from 106 tomograms where lamellae edges could be identified) against the voxel distance to the closest lamella surface (top or bottom) (Figure 5.12, a). When the distance from the closest surface increased, an improvement in resolution could be observed: From 70 Å until it reached a resolution plateau around 50 Å between 35-60 nm distance. Beyond 60 nm distance, the resolution became worse again at a slower rate. One possible explanation for that is the relatively higher contribution of thicker lamellae, the higher the lamella surface distance, due to the maximum surface distance being limited by lamellae thickness. When focusing only on voxels around ribosome positions, the result was very similar (Figure 5.12, b). In addition to analyzing the effect of lamella surface distance on resolution, local lamella thickness was included in the analysis. Here, one needs to be cautious as for a certain lamellae thickness range, the number of sampling points was inconclusively low, as indicated in Figure 5.12 c. With that in mind, only lamellae $>$ 150 nm and $<$ 225 nm should be considered. When not split up by local lamella thickness, the slight drop in resolution beyond 60

nm distance from lamella surface could be explained by the relatively higher contribution of thicker tomograms. However, even if split up by thickness, the quality decrease is visible: For lamellae of above 175 nm thickness, a clear quality decrease in the lamella center could be observed again.

Together, RESOLVE is capable of capturing resolution deterioration induced by the radiation damage of milling, which can be best observed quantitatively for the full dataset or for single tomograms.

5.3 Discussion**

Local resolution estimation has been mostly applied as a common tool for measuring quality of reconstructed cryo-EM maps (Beckers and Sachse, 2020; Cardone, Heymann, and Steven, 2013; Kucukelbir, Sigworth, and Tagare, 2014; Vilas et al., 2018). Direct quality estimates for single measurements, such as micrographs, tilt-series or tomograms, have so far barely been conducted. Existing quality metrics are often of an indirect nature, for example, CTF estimation is often used to judge data quality (Tegunov and Cramer, 2019; Elferich et al., 2024). For tomography, there are two noteworthy exceptions: One recent approach estimates tomogram quality on a per-slice level, employing a neural network trained with expert-labeled data (Tan et al., 2025), therefore relying on human judgement and excluding the possibility of local quality analysis. Another older approach estimates local tomogram quality (Vilas et al., 2020), however, without adequate benchmarking and application cases, and without considering inflated correlations due to tomogram reconstruction artifacts and the missing wedge.

We here present a new algorithm termed RESOLVE. RESOLVE provides a robust method for local resolution estimation in various environments: We apply and benchmark RESOLVE for local resolution estimation beyond refined maps, applying it to micrographs, tilt-series and tomograms (Figures 5.4, 5.5, 5.8). We then show RESOLVE to be beneficial for automated quality assessment of micrographs and

tomograms, pre-selecting particles on micrographs, pre-selecting subtomograms within tomograms and revealing radiation damage effects through FIB milling. The RESOLVE algorithm is based on optimizing the tradeoff between measurement accuracy and measurement locality per resolution shell by tuning the window size (Figure 5.3).

The optimal window radius for a local and reliable measurement is expected to differ based on spectral signal-to-noise ratio (SSNR) of the involved signal and the spatial frequency due to the number of available samples in the respective resolution shell. In contrast to previous considerations (Cardone, Heymann, and Steven, 2013), RESOLVE sets the window radii based on spatial frequency, rather than SSNR. Choosing window radii dependent on spatial frequency instead of SSNR has the advantage of using similar measurement conditions and statistics for each measured frequency, as a reliable measurement requires a minimum number of waves present as well as sufficient sampling of the latter. This strategy also avoids the difficulties in comparing resolution measurements derived from different window sizes among maps with differing SSNRs.

To make a conclusive case for the viability of RESOLVE's resolution measurements for micrographs and tomograms, we conducted single-particle cryo-EM and subtomogram averaging to confirm the correctness of the particle quality ranking via RESOLVE's local resolution estimates on a micrograph and tomogram level. RESOLVE also provide a median global resolution estimate derived from the local resolution measurements. Such quality assessments could be useful for early data cleaning, be it in preparation for structure determination or for automated tomogram analysis, facilitating faster and more accurate downstream processing. As tomography workflows become more and more automated, RESOLVE thus provides a new automated approach of judging data quality at an early stage.

We showed that RESOLVE can facilitate micrograph, tilt-series and tomogram quality characterization in the service of validating image processing or data acquisition parameters. For micrographs, RESOLVE's measurements clearly capture the effects

of motion correction, thus representing the quality-decreasing effects of beam induced motion. For measurements of 2D tilt-series, RESOLVE adequately captures quality differences between earlier and later recorded tilts. This finding indicates that the measurements account for the effects of sample thickness (Tuijtel et al., 2024) and radiation damage on data quality (Koning, Koster, and Sharp, 2018). For tomograms, RESOLVE can be used for the accurate characterization of quality dependence on distance from the lamella edge. Such local quality studies for *in situ* tomograms have previously been conducted via subtomogram averaging (Berger et al., 2023; Tuijtel et al., 2024) or 2D template matching (Lucas and Grigorieff, 2023) and have therefore been laborious and limited to regions containing the protein of interest. Furthermore, relying on proteins decreases, dependent on protein size and distribution, measurement locality. In this study, we used RESOLVE to quantitatively characterize tomogram quality with lamella edge distance and lamella thickness. The results indicate tomogram quality becomes increasingly worse with less than 35 nm distance from the lamella edge and, interestingly, the measurements also document a decrease of quality in the lamella center, especially for lamellae thicker than 175 nm. This observation has not been reported in previous studies and requires validation. However, it is also not necessarily in conflict with previous studies: The observation of the quality drop in the center of thick lamellae may not be stark enough to be visible without splitting up the data according to locally measured lamella thickness. This requires a certain sampling frequency and measurement locality, which is more difficult to achieve when focusing on randomly distributed ribosomes and not taking the rest of the volume into account.

5.3.1 Limitations**

Resolution estimation as used in cryo-EM effectively characterizes the signal-to-noise ratio (SNR), which is not necessarily the same as data quality. For refined maps, SNR usually accurately represents data quality, as many data chunks from

different sources are used to increase the SNR in the service of determining one common structure. In that context, any aberration or artifact, even itself with high SNR, is automatically understood as noise if deviating from the underlying common structure. However, if only one data source is present, the situation is different: High SNR regions will automatically be assigned high resolution to and are impossible to differentiate from the signal of interest. Some examples are carbon film (Figures 5.5, 5.7), devitrified ice (Figures 5.5, 5.7), reconstruction artifacts at tomogram edges (Figure 5.7), or sometimes regions in tomograms far outside the lamella (Figure 5.7). We give an additional global metric, the signal ratio, which can be used complementarily and can help to identify erroneous data. However, the impact of these cases is limited for several reasons: First, in most cases, these regions are regionally limited, and there is no or very low impact on the median global resolution. Second, if one is interested in detailed manual investigation of local resolution distribution, these artifacts are usually obvious and quickly identified by eye. Third, for larger scale studies, such as for the presented local quality dependence on lamella edge distance, quantification over many tomograms make such effects negligible. However, it should be noted that in the context of single-particle cryo-EM or subtomogram averaging, RESOLVE's resolution estimates do not necessarily capture structural homogeneity for the above-mentioned reasons, which is one important factor constituting map quality in the context of structure determination. Thus, RESOLVE should be understood as a complementary tool for particle selection rather than an alternative to conventional classification algorithms.

One potential limitation of tomogram resolution measurements based on half-maps derived from split frames is their relationship to image consistency across frames. Frame-split half-maps provide very useful measurement and, as reported, appear to capture much of what we care about regarding the contributing factors to image quality. However, due to the procedure of how aligned tilt-series are split - typically, tilt-series alignment is conducted on the full tilt-series, and then the same alignment parameters are applied to both half-tilt-series prior to reconstructing

the half-tomograms - tilt-series alignment quality is not captured. One alternative would be to reconstruct half-tomograms by dividing the tilt-series by tilts, which, however, comes with its own challenges due to the lower angular sampling, and is not expected to perform as well in every other respect.

5.4 Methods

5.4.1 Bandpass filtering*

As shape of the employed bandpass filters, we used a hyperbolic tangent. As bandwidth, 10 % of Nyquist centered at the current spatial frequency of interest was used, combined with a two-sided symmetric falloff width of 1.5: This design was chosen in order to employ a consistent bandpass filter independent of map size, in contrast to choosing a fixed number of pixels in Fourier space. Considerations were to choose the band width wide enough for reliable results, while still being small enough to be precise in Fourier space. Moreover, the input half-maps were padded to the closest dimension for efficient fast Fourier transformation (FFT) calculations, and small maps received zero padding up to a minimum of 160 pixels in each dimension in 3D and 300 pixels in 2D.

5.4.2 Threshold determination and reference distribution creation*

For *p*-value determination, we require a reference distribution created via correlation calculations between one of the half-maps and its permutated counterpart. For refined maps (derived from single-particle cryo-EM or subtomogram averaging), we create this permutated half-map by random permutation in real space, preserving only the intensity distribution, as we do not expect systematic artifacts to be present due to the averaging of many different maps. However, for maps derived

from single recordings, such as for micrographs, tilt-series and tomograms, systematic artifacts may arise from both acquisition and processing. One particularly relevant case is the missing wedge in tomograms, which creates anisotropic resolution and therefore causes correlated noise patterns across half-maps, artificially inflating correlations. For accurate statistics, it is necessary that the permutated half-map with which we derive our reference distribution also contains these artifacts. To capture these changes, we use phase-permutation instead of real-space permutation, preserving the power spectrum while disrupting spatial correlations. Consequently, both the original and permuted maps share the same spectral properties and systematic biases, ensuring that our correlation measurements are tested against a realistic null distribution.

To keep the statistical properties of the permutated half-map truthful to the experimental input half-maps, it also needs to be subjected to the same band-pass filter as the input half-maps. Then, 10,000 correlation measurements between the bandpass-filtered permutated half-map and one of the original half-maps are conducted to generate the reference distribution, using the same window radius as for the actual half-maps, but at random positions. One potential limitation with this approach for reference distribution creation arises for situations with large windows and small maps, where the number of unique non-overlapping measurement locations is small, so especially in the case of micrographs: repeatedly using the same or overlapping windows for conducting correlation measurements narrows the reference distribution. One simple solution we implemented for such cases is the creation of multiple permutated maps. Additionally, it is noteworthy that Fourier-space dependencies arise independent of real-space location. Therefore, the real-space window locations between both maps do not need to be identical. This random pairing dramatically increases the number of possible combinations. Subsequently, we determine the p-value for each local measurement between the experimental half-maps based on the obtained reference distribution. We iterate through every shell of interest for the procedure. In the interest of computational

efficiency, for input 3D maps, we use every second voxel as a measurement location, in order to still have overlapping windows for even the highest resolution shells. For 2D maps, every fifth pixel is sufficient, as the windows used are much larger. We also offer a fast mode intended for large maps or batch processing of large datasets. Here, only every third voxel is used for 3D, and, both for 2D and 3D measurements, the number of investigated shells is halved.

After iterating though every resolution shell, we end up with a new map, consisting of a list of shell-derived *p*-values at each location. As established before (Beckers and Sachse, 2020), for each location, we correct for multiple testing converting *p*-value into *q*-values (Benjamini and Yekutieli, 2001). Finally, a threshold of q=0.01 is used for the identification of significant signal. We use the conventional way of threshold determination: we iterate through shell-derived *q*-values list, starting with the lowest-frequency shell, in order to identify the highest resolution shell containing significant signal, stopping when the first *q*-value loses significance. For tomograms, tilt-series and micrographs, we instead use a cutoff of q=0.05, and, additionally, a requirement of two subsequent shells to be above the threshold is added. The cutoff is then chosen as the first of these two shells. This modulation is useful as for noisier maps, the risk of prematurely assigning low resolution is much higher. Additionally, the risk of a loss of significance following the first CTF zero-transition, leading to lower *p*-values, is reduced.

5.4.3 Deriving global resolution estimates: Median resolution estimation*

We established a way to derive a global median resolution estimate we termed *median resolution*,from RESOLVE's local measurements for micrographs, tilt-series and tomograms. We base this approach on a procedure of frequency-dependent *p*-value determination: For each frequency shell, we determine the median *p*-value (from the full set of *p*-values from all measurement locations within the map) and from

this simple 1D list, we derive the q-values via multiple testing correction. This list then allows us to determine a resolution the same way as for single pixels or voxels in the above described practice for micrographs, tilt-series and tomograms, also employing a cutoff value of q=0.05. For better comparability across datasets, by default RESOLVE focuses on signal regions only (regions were our measurements did pass the lowest frequency shell). For more stringent comparisons, we also allow the user to supply an input mask, defining the signal regions for the median resolution estimate, and additionally an option to use the mean instead of the median. As tilt-series present a set of processed 2D images, we adjust the procedure slightly: Instead of taking just the median p-value per frequency, we first take the median p-value for each 2D tilt image, and then the mean p-value across all tilts to derive a single p-value per frequency, resulting in a simple 1D q-value list after multiple testing correction, which we use for cutoff determination as described previously. For *in situ* tomograms, quality variability can be expected dependent on the z-plane due to positioning of the lamella. For that reason, similar as for tilt-series, we take the median p-value for each z-slice, and then the mean p-value across all z-slices to derive a single p-value per frequency. This procedure is, as for tilt-series, followed by multiple testing correction and cutoff determination. Generally, for tomograms, tilt-series and micrographs, RESOLVE automatically produces a frequency dependent median q-value plot from which the median resolution is derived. Additionally, for tilt-series and tomograms, RESOLVE also produces a p-value plot, displaying tilt- or z-slice-dependent p-value behavior. Together with the optional user-defined input mask, allowing to focus these measurements on regions of interest, this setup allows for maximal flexibility and sensitivity in detecting small quality changes.

5.4.4 Implementation and graphical user interface

RESOLVE is implemented in Python and offers a command-line mode or a simple graphical user interface (GUI) for running. RESOLVE offers four distinct processing modes: "Refined-Maps", "Micrographs", "Tilt-Series" and "Tomograms" (Figure 5.13), each coming with its own, adapted implementation, as presented before. For larger datasets, command-line mode is preferred: In command-line mode, it is possible to run RESOLVE on a large batch of input data at once, without need for manual intervention. When running RESOLVE for a single pair of input half-maps, the graphical user interface may be preferred, as it may be simpler to operate for many users (Figure 5.13).

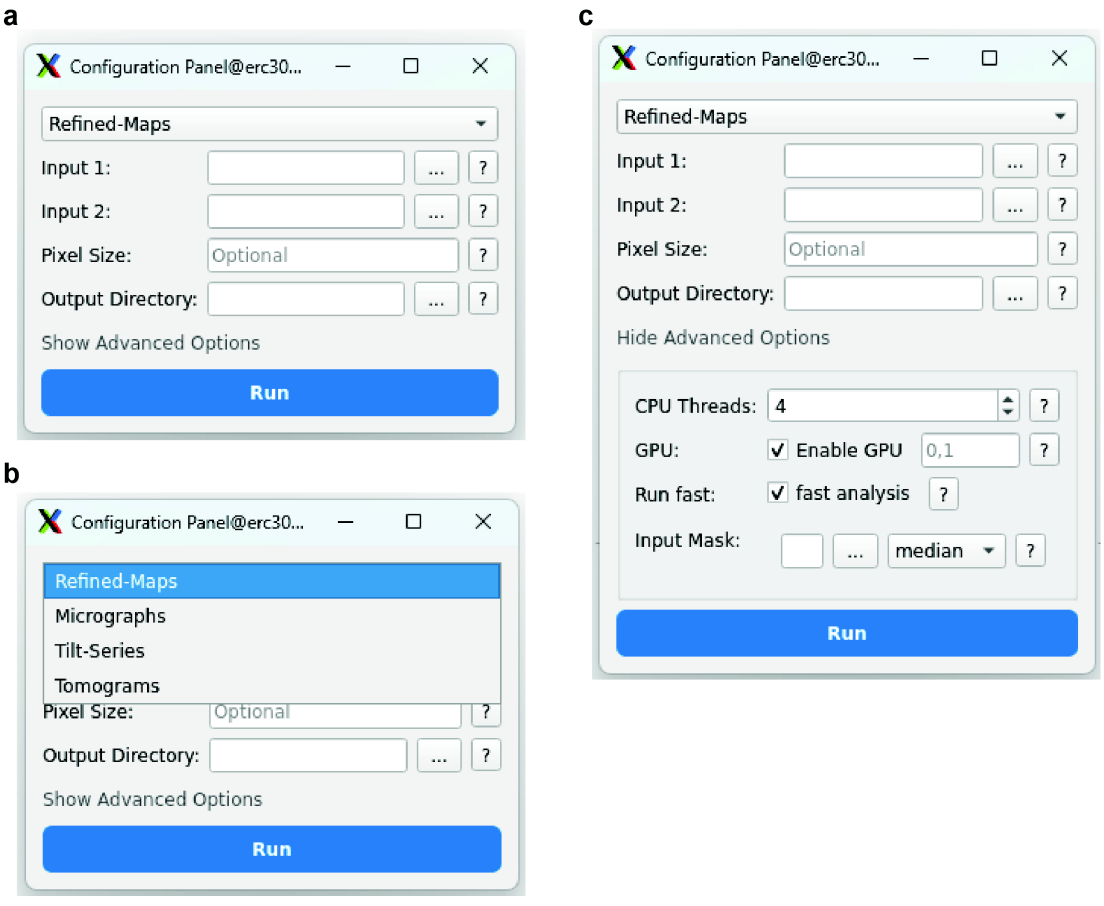

FIGURE 5.13: Presentation of the RESOLVE GUI. a) Graphical user interface without advanced options, as it pops up after starting. b) Four options of processing, can be selected via the drop-down menu. c) Advanced options made visible, mostly referring to computational preferences such as CPU threads, selected GPUs, or fast mode. Additionally, it is possible to input a mask within which global median resolution is estimated. For specific cases, one may also choose the mean instead.

RESOLVE can run on the CPU and, additionally, the computationally demanding steps may be executed on the GPU. Running RESOLVE with a GPU is strongly recommended for 3D input, especially for larger tomograms. Two processing modes are available, a fast mode and an default mode, differing in Fourier-space and real-space sampling. For most cases, fast mode should be sufficient, and especially for large tomography datasets, it is recommended. While computational time strongly depends on the hardware, on the machines tested, fast mode took less than 2 minutes per tomogram of 560 megabytes size.

For 2D micrographs, GPU usage is not recommended, as data transfer from GPU to CPU is too slow to be computationally beneficial. On the machines tested, per micrograph, 5-20 seconds (fast versus default mode) were needed for RESOLVE to run.

5.4.5 Single-particle image processing*

For the TMV micrographs (EMPIAR-10305), frames were motion corrected and split into half-frames with MotionCor3 (Zheng et al., 2017) using a pixel size of 2.925 Å for resolution measurements. For CTF estimation, CTFFIND4 was used (Rohou and Grigorieff, 2015) and for subsequent single-particle workflows with EMPIAR-10524, RELION-5 (Burt et al., 2024) was used. Micrograph half-sets were reconstructed from raw frames with MotionCor3 (Zheng et al., 2017). After motion correction, half-micrographs were binned with IMOD (Kremer, Mastronarde, and McIntosh, 1996) by a factor of 2 to a pixel size of 2.57 Å. RESOLVE was applied to the binned, non CTF-corrected micrograph half-sets. For refinements with preselected particles by RESOLVE's micrograph resolution estimates, particles on micrographs with carbon film were excluded. For estimating particle quality on a micrograph level, a custom Python script was used to associate star file coordinates and micrograph resolution. Particle quality was estimated by averaging all

resolution estimates within an approximate 10 nm diameter box around the particle location. Micrographs were visualized with IMOD (Kremer, Mastronarde, and McIntosh, 1996).

5.4.6 Tomography image processing*

For the tomography dataset EMPIAR-11306, we employed the tomography pipeline of RELION-5 (Burt et al., 2024), including half-tilt series and half-tomogram reconstructions. Half-tilts were stacked to produce a half-tilt-series and binned using IMOD (Kremer, Mastronarde, and McIntosh, 1996). Resolution measurements were conducted for the tilt-series with a final pixel size of 11.1 Å after binning. Half-tomograms were reconstructed with a pixel size of 12 Å. We used CTFFIND4 for CTF estimation (Rohou and Grigorieff, 2015) and AreTomo (Zheng et al., 2022) for tilt-series alignment. For estimating particle quality on a tomogram level, a custom Python script was used to associate star file coordinates and tomogram resolution. Resolution estimates on tomograms were conducted with the RELION-5 reconstructed tomograms. Particle quality was estimated by averaging all resolution estimates within an approximate 20 nm diameter box around the particle location. Lamellae were modelled for 106 tomograms. For better lamella visibility, tomograms were denoised with cryoCARE (Buchholz et al., 2019). All tomograms were excluded where lamella edges were difficult to identify. For the remaining 106 tomograms, top and bottom lamella edges (top and bottom surface) were marked manually for multiple y-slices throughout the tomograms, so that the surfaces could be interpolated (5.10). Distances to surfaces were determined for every second voxel with a custom Python script, using the closest surface coordinate as reference. Local lamella thickness was estimated as the distance between closest top surface coordinate and closest bottom surface coordinate. All tomogram visualizations were created with either IMOD or chimera (Pettersen et al., 2004).

Chapter 6

Discussion

This work focuses on advancing the analysis of cryo-ET data in two methodological respects: First, we presented a framework for fully automated DMV morphology analysis, including instance segmentation and complete quantitative analysis. Second, we developed a new method for data quality assessment, allowing local resolution estimates of micrographs, tilt-series and tomograms.

Cryo-ET requires extensive image processing, and part of the required techniques are well established. The more streamlined and automated initial processing steps (Zheng et al., 2022; Burt et al., 2024), including correcting for beam-induced motion, alignment of the tilt-series, tomogram reconstruction and tomogram denoising go together with methodical developments that allow for increased data-collection speed (Bouvette et al., 2021; Eisenstein, Fukuda, and Danev, 2024). One potential analysis bottleneck for more quickly collected and larger tomography datasets in the future may be the downstream analysis, specifically concerning ultrastructural analysis and quantification. Large parts of the downstream analysis of tomograms still have to be conducted manually today.

Usually, the first step of downstream analysis requires extensive data labeling, and here, important advancements have been made: Often, ultrastructural analysis concerns membrane structures such as vesicles, mitochondria or autophagosomes and recently, some general tools for membrane labeling emerged (Lamm et al., 2024;

(Kiewisz et al., 2025). These tools allow for much faster and improved general segmentations and thus improve downstream analysis. For more specific cases, established tools such as Amira (Thermo Fisher Scientific) or Dragonfly (Object Research Systems) provide platforms for adapting existing models for effective semantic segmentation. However, semantic segmentation is often not sufficient. In many cases, membrane structures have to be identified as distinct instances for further automated quantification, as it is the case for shape analysis of vesicles, mitochondria or autophagosomes. To some extent, instance segmentation can be generalized or convenient platforms for retraining pretrained neural networks can be provided in a case-specific manner. For labeling particles as distinct instances or objects (particle picking), this is well established in the cryo-EM community. However, for identifying complex membrane structures in 3D, this remains challenging. The commonly used technique for semantic segmentation, based on dividing a large tomogram into smaller subvolumes, is difficult to apply to large membrane structures, as it removes important context and splits instances. Recently established techniques are too general and cannot be adapted in a case-specific manner (Kiewisz et al., 2025; Lamm et al., 2024): A platform for training proper instance segmentation models in three dimensions is still lacking. In this work, we suggest another approach for 3D instance segmentation for large membrane structures: Based on existing and easy to come by semantic membrane segmentations and a retrained 2D instance segmentation model, 3D instance segmentations can be derived. We show that these segmentations can be used for adequate quantitative analysis and how such quantifications can be conducted. Such an instance segmentation model could in principle be trained for any membrane structure, and providing a user-friendly platform similar to those already available for semantic segmentation would greatly benefit the field.

One major issue inhibiting platforms for generalized quantification is that the characteristics of interest are usually dataset-dependent and thus the analysis steps are difficult to generalize. In recent works, automated quantification capabilities were

explored: The surface morphometrics toolkit offers some options for quantifying mitochondrial characteristics (Barad et al., 2023) for segmented membranes, but not without the manual effort of separating mitochondria instances as well as the inner and outer mitochondrial membranes. TARDIS (Kiewisz et al., 2025) offers Python wrapper for the analysis of segmented filaments and vesicles. One recent tool even covers full vesicle analysis for 2D data, from segmentation to quantification of separated instances (Schönnenbeck, Junglas, and Sachse, 2025). However, for the case of the SARS-CoV-2 DMVs, no such automated quantification tools exist. The presented approach, while not currently available as a packaged tool, offers a viable framework for tool development.

The other focus of this thesis was on establishing a procedure for resolution estimation for tomograms, in order to enable automated tomogram analysis. The presented approach can however be generalized and works for tomograms, as well as for micographs, tilt-series and refined maps. Conventionally, resolution estimation has been conducted for refined maps, as derived from structure determination algorithms such as single-particle cryo-EM and subtomogram averaging. The quality of micrographs or tomograms has usually either been assessed by eye, or in an indirect manner, for example via CTF estimation (Tegunov and Cramer, 2019; Elferich et al., 2024). Here, we show that local resolution estimates via the presented approach yield adequate quality estimates in each given environment, and that these estimates describe, among other contributing factors, FIB-induced lamella damage, radiation damage, sample thickness and the effects of motion within the sample. Consequently, these local quality estimates can be used to judge particle quality in micrographs or tomograms, allowing to focus structure determination workflows on a high-quality particle set from the start. Our results demonstrate that the developed procedure accurately assesses tomogram quality without bias from the missing wedge. We provide the full framework for local resolution estimation in either refined maps, micrographs, tilt-series or tomograms as a computationally optimized and accessible tool, called RESOLVE.

Chapter 7

Outlook

Data generation for *in situ* cryo-electron tomography (cryo-ET) is complex and time-consuming, as it requires elaborate sample preparation including specimen thinning, followed by lengthy recording procedures where the specimen is tilted and imaged at multiple angles. Recent advancements in high-throughput data generation have significantly improved cryo-ET workflows through innovations in both sample preparation and data acquisition. The Thermo Fisher Scientific Arctis cryo-plasma FIB enables automated, high-throughput lamella production. For data acquisition, the parallel cryo-electron tomography method (PACE-tomo) combines geometrical sample modeling with optical image shift to record tens of tilt-series in parallel, achieving throughputs exceeding one tilt-series per 5 minutes (Eisenstein et al., 2023). Recently developed machine learning approaches can further automate lamella detection, target selection, and parallel tilt-series acquisition (Eisenstein, Fukuda, and Danev, 2024). As cryo-ET techniques mature and the field rapidly moves towards automation, it is thus essential to facilitate faster and more objective downstream analysis of tomography data. Today, tomogram analysis requires extensive manual work, with typically only individual steps being automated. Deficiencies in large-scale data analysis capabilities potentially constrain both study impact and reproducibility, preventing investigators from fully exploiting the rich information contained in *in situ* cryo-ET data.

While some important advancements have been made on the software side - including general segmentation models with some basic capability for instance separation

(Lamm et al., 2024; Kiewisz et al., 2025) or membrane surface characterization (Barad et al., 2023) - the field is still lacking general and accessible methods, particularly for instance identification of cellular structures and streamlined quantification approaches. An important next step would be the provision of a dedicated software platform for training instance segmentation models on 3D data in a case-specific manner. The availability of a pretrained general model would allow rapid retraining to adapt the model to specific use cases, making quick and accessible instance segmentation of any cellular structure possible. Combined with specialized software for streamlined and automated quantification of individual membrane structures, this would provide a comprehensive analytical pipeline. Many features such as membrane thickness, membrane curvedness, and circumference can be calculated in a generalized manner, and automatically assigning these features to specific membrane structures would be immensely helpful. Furthermore, automated differentiation between instances of distinct membrane structures would vastly enhance the study of spatial relationships among them, including the characterization of contact sites. This spatial information could open up new possibilities for contextual protein localization, serving both ultrastructural analysis and structure determination workflows. In the context of the presented work on SARS-CoV-2 replication, such streamlined automated analysis would enable rapid quantification of drug effects on DMV formation, providing mechanistic insights that complement viral replication studies.

In order to facilitate comprehensive and streamlined automated analysis, the presented extension of resolution estimation to micrographs, tilt-series and tomograms could serve as automated data quality control. In the future, further aspects of the proposed method for quality estimation could be explored, advancing the utility of such quality estimates in several respects. It would be desirable to enhance or complement the presented approach in ways that include special cases that are currently not addressed, like ice contamination or quality of tilt-series alignment. Future approaches could focus on computational methods for tilt-series splitting

based on full volumes that do not require the user to conduct the half-map splitting. Half-map splitting, although offered by many of the popular tools (Tegunov et al., 2021; Burt et al., 2024), requires extra work. Some basic approaches already exist for measuring resolution on a global level for single-particle cryo-EM maps (Sousa and Grigorieff, 2007; Verbeke et al., 2024). However, while half-map generation is usually straightforward and easy to achieve for single-particle cryo-EM, new half-map generation methods for micrograph or tomogram resolution measurements could make a real qualitative difference — for example, by potentially capturing the quality of tilt-series alignment as well.

Bibliography

Adrian, M., J. Dubochet, J. Lepault, and A. W. McDowall (1984). “Cryo-electron microscopy of viruses”. In: *Nature* 308.5954, pp. 32–36. ISSN: 0028-0836. DOI: 10.1038/308032a0.

Alberts, B., A. Johnson, J. Lewis, Raff, M., Roberts, and P. K. and Walter (2003). “Molecular biology of the cell. 4th edn”. In: *Annals of Botany* 91.3, p. 401. ISSN: 03057364. DOI: 10.1093/aob/mcg023.

Axe, Elizabeth L., Simon A. Walker, Maria Manifava, Priya Chandra, H. Llewelyn Roderick, Anja Habermann, Gareth Griffiths, and Nicholas T. Ktistakis (2008). “Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum”. In: *The Journal of cell biology* 182.4, pp. 685–701. DOI: 10.1083/jcb.200803137.

Babst, Markus, David J. Katzmann, Eden J. Estepa-Sabal, Timo Meerloo, and Scott D. Emr (2002). “Escrt-III: an endosome-associated heterooligomeric protein complex required for mvb sorting”. In: *Developmental cell* 3.2, pp. 271–282. ISSN: 1534-5807. DOI: 10.1016/s1534-5807(02)00220-4.

Barad, Benjamin A., Michaela Medina, Daniel Fuentes, R. Luke Wiseman, and Danielle A. Grotjahn (2023). “Quantifying organellar ultrastructure in cryo-electron tomography using a surface morphometrics pipeline”. In: *The Journal of cell biology* 222.4. DOI: 10.1083/jcb.202204093.

Bartesaghi, A., P. Sprechmann, J. Liu, G. Randall, G. Sapiro, and S. Subramaniam (2008). "Classification and 3D averaging with missing wedge correction in biological electron tomography". In: *Journal of structural biology* 162.3, pp. 436–450. DOI: 10.1016/j.jsb.2008.02.008.

Basanta, Benjamin, Saikat Chowdhury, Gabriel C. Lander, and Danielle A. Grotjahn (2020). "A guided approach for subtomogram averaging of challenging macromolecular assemblies". In: *Journal of structural biology: X* 4, p. 100041. DOI: 10.1016/j.yjsbx.2020.100041.

Beckers, Maximilian and Carsten Sachse (2020). "Permutation testing of Fourier shell correlation for resolution estimation of cryo-EM maps". In: *Journal of structural biology* 212.1, p. 107579. DOI: 10.1016/j.jsb.2020.107579.

Benjamini, Yoav and Daniel Yekutieli (2001). "The control of the false discovery rate in multiple testing under dependency". In: *The Annals of Statistics* 29.4. ISSN: 0090-5364. DOI: 10.1214/aos/1013699998.

Bennion, Brian J. and Valerie Daggett (2003). "The molecular basis for the chemical denaturation of proteins by urea". In: *Proceedings of the National Academy of Sciences of the United States of America* 100.9, pp. 5142–5147. DOI: 10.1073/pnas.0930122100.

Bepler, Tristan, Kotaro Kelley, Alex J. Noble, and Bonnie Berger (2020). "Topaz-Denoise: general deep denoising models for cryoEM and cryoET". In: *Nature communications* 11.1, p. 5208. DOI: 10.1038/s41467-020-18952-1.

Bepler, Tristan, Andrew Morin, Micah Rapp, Julia Brasch, Lawrence Shapiro, Alex J. Noble, and Bonnie Berger (2019). "Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs". In: *Nature methods* 16.11, pp. 1153–1160. DOI: 10.1038/s41592-019-0575-8.

Berger, Casper, Navya Premaraj, Raimond B. G. Ravelli, Kèvin Knoops, Carmen López-Iglesias, and Peter J. Peters (2023). "Cryo-electron tomography on focused ion beam lamellae transforms structural cell biology". In: *Nature methods* 20.4, pp. 499–511. DOI: 10.1038/s41592-023-01783-5.

Blanchard, Emmanuelle and Philippe Roingeard (2015). "Virus-induced double-membrane vesicles". In: *Cellular microbiology* 17.1, pp. 45–50. DOI: 10.1111/cmi.12372.

Boson, Bertrand, Vincent Legros, Bingjie Zhou, Eglantine Siret, Cyrille Mathieu, François-Loïc Cosset, Dimitri Lavillette, and Solène Denolly (2021). "The SARS-CoV-2 envelope and membrane proteins modulate maturation and retention of the spike protein, allowing assembly of virus-like particles". In: *The Journal of biological chemistry* 296, p. 100111. DOI: 10.1074/jbc.RA120.016175.

Böttcher, B., S. A. Wynne, and R. A. Crowther (1997). "Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy". In: *Nature* 386.6620, pp. 88–91. ISSN: 0028-0836. DOI: 10.1038/386088a0.

Bouvette, Jonathan, Hsuan-Fu Liu, Xiaochen Du, Ye Zhou, Andrew P. Sikkema, Juliana Da Fonseca Rezende E Mello, Bradley P. Klemm, Rick Huang, Roel M. Schaaper, Mario J. Borgnia, and Alberto Bartesaghi (2021). "Beam image-shift accelerated data acquisition for near-atomic resolution single-particle cryo-electron tomography". In: *Nature communications* 12.1, p. 1957. DOI: 10.1038/s41467-021-22251-8.

Buchholz, Tim-Oliver, Alexander Krull, Réza Shahidi, Gaia Pigino, Gáspár Jékely, and Florian Jug (2019). "Content-aware image restoration for electron microscopy". In: *Methods in cell biology* 152, pp. 277–289. ISSN: 0091-679X. DOI: 10.1016/bs.mcb.2019.05.001.

Burt, Alister, Bogdan Toader, Rangana Warshamanage, Andriko von Kügelgen, Euan Pyle, Jasenko Zivanov, Dari Kimanius, Tanmay A. M. Bharat, and Sjors H. W. Scheres (2024). "An image processing pipeline for electron cryo-tomography in RELION-5". In: *FEBS open bio* 14.11, pp. 1788–1804. DOI: 10.1002/2211-5463.13873.

Cantor, Charles R. and Robert B. Gennis (1989). *Biomembranes*. New York, NY: Springer New York. ISBN: 978-1-4757-2067-9. DOI: 10.1007/978-1-4757-2065-5.

Cardone, Giovanni, J. Bernard Heymann, and Alasdair C. Steven (2013). "One number does not fit all: mapping local variations in resolution in cryo-EM reconstructions". In: *Journal of structural biology* 184.2, pp. 226–236. DOI: 10.1016/j.jsb.2013.08.002.

Carter, C. Barry and David B. Williams (2016). *Transmission Electron Microscopy*. Cham: Springer International Publishing. ISBN: 978-3-319-26649-7. DOI: 10.1007/978-3-319-26651-0.

Castaño-Díez, Daniel and Giulia Zanetti (2019). "In situ structure determination by subtomogram averaging". In: *Current opinion in structural biology* 58, pp. 68–75. DOI: 10.1016/j.sbi.2019.05.011.

Chen, Shaoxia, Greg McMullan, Abdul R. Faruqi, Garib N. Murshudov, Judith M. Short, Sjors H. W. Scheres, and Richard Henderson (2013). "High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy". In: *Ultramicroscopy* 135, pp. 24–35. ISSN: 03043991. DOI: 10.1016/j.ultramic.2013.06.004.

Cheng, Bowen, Ishan Misra, Alexander G. Schwing, Alexander Kirillov, and Rohit Girdhar (2022). "Masked-attention Mask Transformer for Universal Image Segmentation". In: *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, pp. 1280–1289. ISBN: 978-1-6654-6946-3. DOI: 10.1109/CVPR52688.2022.00135.

Christofides, Nicos (2022). "Worst-Case Analysis of a New Heuristic for the Travelling Salesman Problem". In: *Operations Research Forum* 3.1. DOI: 10.1007/s43069-021-00101-z.

Cormen, Thomas, Charles Leiserson, Ronald Rivest, and Clifford Stein (2009). *Introduction to Algorithms*.

Crowther, R. A., D. J. DeRosier, and A. Klug (1970). "The reconstruction of a three-dimensional structure from projections and its application to electron microscopy".

In: *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 317.1530, pp. 319–340. ISSN: 0080-4630. DOI: 10.1098/rspa.1970.0119.

Dijkstra, E. W. (1959). "A note on two problems in connexion with graphs". In: *Numerische Mathematik* 1.1, pp. 269–271. ISSN: 0029-599X. DOI: 10.1007/BF01386390.

Dubochet, J., M. Adrian, J. J. Chang, J. C. Homo, J. Lepault, A. W. McDowall, and P. Schultz (1988). "Cryo-electron microscopy of vitrified specimens". In: *Quarterly reviews of biophysics* 21.2, pp. 129–228. ISSN: 0033-5835. DOI: 10.1017/s0033583500004297.

Egerton, R. F. (2016). *Physical Principles of Electron Microscopy*. Cham: Springer International Publishing. ISBN: 978-3-319-39876-1. DOI: 10.1007/978-3-319-39877-8.

Eisenstein, Fabian, Yoshiyuki Fukuda, and Radostin Danev (2024). "Smart parallel automated cryo-electron tomography". In: *Nature methods* 21.9, pp. 1612–1615. DOI: 10.1038/s41592-024-02373-9.

Eisenstein, Fabian, Haruaki Yanagisawa, Hiroka Kashihara, Masahide Kikkawa, Sachiko Tsukita, and Radostin Danev (2023). "Parallel cryo electron tomography on in situ lamellae". In: *Nature methods* 20.1, pp. 131–138. DOI: 10.1038/s41592-022-01690-1.

Elferich, Johannes, Lingli Kong, Ximena Zottig, and Nikolaus Grigorieff (2024). "CTFFIND5 provides improved insight into quality, tilt, and thickness of TEM samples". In: *eLife* 13. DOI: 10.7554/eLife.97227.

Feng, Yuchen, Ding He, Zhiyuan Yao, and Daniel J. Klionsky (2014). "The machinery of macroautophagy". In: *Cell research* 24.1, pp. 24–41. DOI: 10.1038/cr.2013.168.

Frank, J. (1975). "Averaging of low exposure electron micrographs of non-periodic objects". In: *Ultramicroscopy* 1.2, pp. 159–162. ISSN: 03043991. DOI: 10.1016/S0304-3991(75)80020-9.

Frank, Joachim and Terence Wagenknecht (1983). “Automatic selection of molecular images from electron micrographs”. In: *Ultramicroscopy* 12.3, pp. 169–175. ISSN: 03043991. DOI: 10.1016/0304-3991(83)90256-5.

Franken, Linda E., Kay Grünewald, Egbert J. Boekema, and Marc C. A. Stuart (2020). “A Technical Introduction to Transmission Electron Microscopy for Soft-Matter: Imaging, Possibilities, Choices, and Technical Developments”. In: *Small (Weinheim an der Bergstrasse, Germany)* 16.14, e1906198. DOI: 10.1002/smll.201906198.

Gorkhali, Ritesh, Prashanna Koirala, Sadikshya Rijal, Ashmita Mainali, Adesh Baral, and Hitesh Kumar Bhattarai (2021). “Structure and Function of Major SARS-CoV-2 and SARS-CoV Proteins”. In: *Bioinformatics and biology insights* 15, p. 11779322211025876. ISSN: 1177-9322. DOI: 10.1177/11779322211025876.

Grant, Timothy and Nikolaus Grigorieff (2015). “Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6”. In: *eLife* 4, e06980. DOI: 10.7554/eLife.06980.

Gupta, Tilak Kumar, Sven Klumpe, Karin Gries, Steffen Heinz, Wojciech Wietrzynski, Norikazu Ohnishi, Justus Niemeyer, Benjamin Spaniol, Miroslava Schaffer, Anna Rast, Matthias Ostermeier, Mike Strauss, Jürgen M. Plitzko, Wolfgang Baumeister, Till Rudack, Wataru Sakamoto, Jörg Nickelsen, Jan M. Schuller, Michael Schroda, and Benjamin D. Engel (2021). “Structural basis for VIPP1 oligomerization and maintenance of thylakoid membrane integrity”. In: *Cell* 184.14, 3643–3659.e23. DOI: 10.1016/j.cell.2021.05.011.

Gutu, Andrian, Frederick Chang, and Erin K. O’Shea (2018). “Dynamical localization of a thylakoid membrane binding protein is required for acquisition of photosynthetic competency”. In: *Molecular microbiology* 108.1, pp. 16–31. DOI: 10.1111/mmi.13912.

Hagen, Wim J. H., William Wan, and John A. G. Briggs (2017). "Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging". In: *Journal of structural biology* 197.2, pp. 191–198. DOI: 10.1016/j.jsb.2016.06.007.

Harauz, George and Marin van Heel (1986). "Exact filters for general geometry three dimensional reconstruction". In: *Optik*.

Hayashi-Nishino, Mitsuko, Naonobu Fujita, Takeshi Noda, Akihito Yamaguchi, Tamotsu Yoshimori, and Akitsugu Yamamoto (2009). "A subdomain of the endoplasmic reticulum forms a cradle for autophagosome formation". In: *Nature cell biology* 11.12, pp. 1433–1437. DOI: 10.1038/ncb1991.

Heebner, Jessica E., Carson Purnell, Ryan K. Hylton, Mike Marsh, Michael A. Grillo, and Matthew T. Swulius (2022). "Deep Learning-Based Segmentation of Cryo-Electron Tomograms". In: *Journal of visualized experiments : JoVE* 189. DOI: 10.3791/64435.

Heidrich, Jennifer, Verena Wulf, Raoul Hennig, Michael Saur, Jürgen Markl, Carsten Sönnichsen, and Dirk Schneider (2016). "Organization into Higher Ordered Ring Structures Counteracts Membrane Binding of IM30, a Protein Associated with Inner Membranes in Chloroplasts and Cyanobacteria". In: *The Journal of biological chemistry* 291.29, pp. 14954–14962. DOI: 10.1074/jbc.M116.722686.

Henderson, R., J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckmann, and K. H. Downing (1990). "Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy". In: *Journal of molecular biology* 213.4, pp. 899–929. ISSN: 0022-2836. DOI: 10.1016/S0022-2836(05)80271-2.

Henderson, Richard (2015). "Overview and future of single particle electron cryomicroscopy". In: *Archives of biochemistry and biophysics* 581, pp. 19–24. DOI: 10.1016/j.abb.2015.02.036.

Hoffmann, Markus, Hannah Kleine-Weber, Simon Schroeder, Nadine Krüger, Tanja Herrler, Sandra Erichsen, Tobias S. Schiergens, Georg Herrler, Nai-Huei Wu,

Andreas Nitsche, Marcel A. Müller, Christian Drosten, and Stefan Pöhlmann (2020). "SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor". In: *Cell* 181.2, 271–280.e8. DOI: 10.1016/j.cell.2020.02.052.

Huang, Yixin, Tongyun Wang, Lijie Zhong, Wenxin Zhang, Yu Zhang, Xiulian Yu, Shuofeng Yuan, and Tao Ni (2024). "Molecular architecture of coronavirus double-membrane vesicle pore complex". In: *Nature* 633.8028, pp. 224–231. ISSN: 0028-0836. DOI: 10.1038/s41586-024-07817-y.

Ismail, V. A., T. Naismith, and D. J. Kast (2023). "The NTPase activity of the double FYVE domain-containing protein 1 regulates lipid droplet metabolism". In: *The Journal of biological chemistry* 299.2, p. 102830. DOI: 10.1016/j.jbc.2022.102830.

Jackson, Cody B., Michael Farzan, Bing Chen, and Hyeryun Choe (2022). "Mechanisms of SARS-CoV-2 entry into cells". In: *Nature reviews. Molecular cell biology* 23.1, pp. 3–20. DOI: 10.1038/s41580-021-00418-x.

Jamali, Kiarash, Lukas Käll, Rui Zhang, Alan Brown, Dari Kimanius, and Sjors H. W. Scheres (2024). "Automated model building and protein identification in cryo-EM maps". In: *Nature* 628.8007, pp. 450–457. ISSN: 0028-0836. DOI: 10.1038/s41586-024-07215-4.

Junglas, Benedikt, David Kartte, Mirka Kutzner, Nadja Hellmann, Ilona Ritter, Dirk Schneider, and Carsten Sachse (2025). "Structural basis for Vipp1 membrane binding: from loose coats and carpets to ring and rod assemblies". In: *Nature structural & molecular biology* 32.3, pp. 555–570. DOI: 10.1038/s41594-024-01399-z.

Junglas, Benedikt, Roberto Orru, Amelie Axt, Carmen Siebenaller, Wieland Steinchen, Jennifer Heidrich, Ute A. Hellmich, Nadja Hellmann, Eva Wolf, Stefan A. L. Weber, and Dirk Schneider (2020). "IM30 IDPs form a membrane-protective carpet upon super-complex disassembly". In: *Communications biology* 3.1, p. 595. DOI: 10.1038/s42003-020-01314-4.

Junglas, Benedikt and Dirk Schneider (2018). "What is Vipp1 good for?" In: *Molecular microbiology* 108.1, pp. 1–5. DOI: 10.1111/mmi.13924.

Kakavandi, Sareh, Iman Zare, Maryam VaezJalali, Masoud Dadashi, Maryam Azarian, Abdullatif Akbari, Marzieh Ramezani Farani, Hamidreza Zalpoor, and Bahareh Hajikhani (2023). "Structural and non-structural proteins in SARS-CoV-2: potential aspects to COVID-19 treatment or prevention of progression of related diseases". In: *Cell communication and signaling : CCS* 21.1, p. 110. DOI: 10.1186/s12964-023-01104-5.

Katiyar, Harshita, Ariana Arduini, Yichen Li, and Chen Liang (2024). "SARS-CoV-2 Assembly: Gaining Infectivity and Beyond". In: *Viruses* 16.11. DOI: 10.3390/v16111648.

Kiewisz, Robert, Gunar Fabig, Will Conway, Jake Johnston, Victor A. Kostyuchenko, Aaron Tan, Cyril Bařinka, Oliver Clarke, Magdalena Magaj, Hossein Yazdkhasti, Francesca Vallese, Shee-Mei Lok, Stefanie Redemann, Thomas Müller-Reichert, and Tristan Bepler (2025). "Accurate and fast segmentation of filaments and membranes in micrographs and tomograms with TARDIS". In: *bioRxiv : the preprint server for biology*. DOI: 10.1101/2024.12.19.629196.

Kifer, Allison, Franciso Pina, Nicholas Codallos, Anita Hermann, Lauren Ziegler, and Maho Niwa (2025). "Orchestration of SARS-CoV-2 Nsp4 and host cell ESCRT proteins induces morphological changes of the endoplasmic reticulum". In: *Molecular biology of the cell* 36.4, ar40. DOI: 10.1091/mbc.E24-12-0542.

Kim, Dongwan, Joo-Yeon Lee, Jeong-Sun Yang, Jun Won Kim, V. Narry Kim, and Hyeshik Chang (2020). "The Architecture of SARS-CoV-2 Transcriptome". In: *Cell* 181.4, 914–921.e10. DOI: 10.1016/j.cell.2020.04.011.

Klein, Steffen, Mirko Cortese, Sophie L. Winter, Moritz Wachsmuth-Melm, Christopher J. Neufeldt, Berati Cerikan, Megan L. Stanifer, Steeve Boulant, Ralf Bartenschlager, and Petr Chlanda (2020). "SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography". In: *Nature communications* 11.1, p. 5885. DOI: 10.1038/s41467-020-19619-7.

Koning, Roman I., Abraham J. Koster, and Thomas H. Sharp (2018). “Advances in cryo-electron tomography for biology and medicine”. In: *Annals of anatomy = Anatomischer Anzeiger : official organ of the Anatomische Gesellschaft* 217, pp. 82–96. DOI: 10.1016/j.aanat.2018.02.004.

Kremer, J. R., D. N. Mastronarde, and J. R. McIntosh (1996). “Computer visualization of three-dimensional image data using IMOD”. In: *Journal of structural biology* 116.1, pp. 71–76. DOI: 10.1006/jsbi.1996.0013.

Kruskal, Joseph B. (1956). “On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem”. In: *Proceedings of the American Mathematical Society* 7.1, p. 48. ISSN: 00029939. DOI: 10.2307/2033241.

Kucukelbir, Alp, Fred J. Sigworth, and Hemant D. Tagare (2014). “Quantifying the local resolution of cryo-EM density maps”. In: *Nature methods* 11.1, pp. 63–65. DOI: 10.1038/nmeth.2727.

Kudryashev, Mikhail, Daniel Castaño-Díez, and Henning Stahlberg (2012). “Limiting factors in single particle cryo electron tomography”. In: *Computational and structural biotechnology journal* 1, e201207002. ISSN: 2001-0370. DOI: 10.5936/csbj.201207002.

Kühlbrandt, Werner (2014). “Biochemistry. The resolution revolution”. In: *Science (New York, N.Y.)* 343.6178, pp. 1443–1444. DOI: 10.1126/science.1251652.

Lam, L., S.-W. Lee, and C. Y. Suen (1992). “Thinning methodologies-a comprehensive survey”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14.9, pp. 869–885. ISSN: 01628828. DOI: 10.1109/34.161346.

Lamm, Lorenz, Simon Zufferey, Hanyi Zhang, Ricardo D. Righetto, Florent Waltz, Wojciech Wietrzynski, Kevin A. Yamauchi, Alister Burt, Ye Liu, Antonio Martinez-Sanchez, Sebastian Ziegler, Fabian Isensee, Julia A. Schnabel, Benjamin D. Engel, and Tingying Peng (2024). *MemBrain v2: an end-to-end tool for the analysis of membranes in cryo-electron tomography*. DOI: 10.1101/2024.01.05.574336.

Lee, Younhee and Woong Lim (2017). “Shoelace Formula: Connecting the Area of a Polygon and the Vector Cross Product”. In: *The Mathematics Teacher* 110.8, pp. 631–636. ISSN: 0025-5769. DOI: 10.5951/mathteacher.110.8.0631.

Lehtinen, Jaakko, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila (2018). *Noise2Noise: Learning Image Restoration without Clean Data*. DOI: 10.48550/arXiv.1803.04189.

Li, Dongfang, Yan G. Zhao, Di Li, Hongyu Zhao, Jie Huang, Guangyan Miao, Du Feng, Pingsheng Liu, Dong Li, and Hong Zhang (2019). “The ER-Localized Protein DFCP1 Modulates ER-Lipid Droplet Contact Formation”. In: *Cell reports* 27.2, 343–358.e5. DOI: 10.1016/j.celrep.2019.03.025.

Li, Xueming, Paul Mooney, Shawn Zheng, Christopher R. Booth, Michael B. Braunfeld, Sander Gubbens, David A. Agard, and Yifan Cheng (2013). “Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM”. In: *Nature methods* 10.6, pp. 584–590. DOI: 10.1038/nmeth.2472.

Lin, Tsung-Yi, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick (2014). “Microsoft COCO: Common Objects in Context”. In: *Computer Vision – ECCV 2014*. Ed. by David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars. Vol. 8693. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 740–755. ISBN: 978-3-319-10601-4. DOI: 10.1007/978-3-319-10602-1{\textunderscore}48.

Liu, Jiwei, Matteo Tassinari, Diorge P. Souza, Souvik Naskar, Jeffrey K. Noel, Olga Bohuszewicz, Martin Buck, Tom A. Williams, Buzz Baum, and Harry H. Low (2021). “Bacterial Vipp1 and PspA are members of the ancient ESCRT-III membrane-remodeling superfamily”. In: *Cell* 184.14, 3660–3673.e18. DOI: 10.1016/j.cell.2021.05.041.

Liu, Yun-Tao, Heng Zhang, Hui Wang, Chang-Lu Tao, Guo-Qiang Bi, and Z. Hong Zhou (2022). "Isotropic reconstruction for electron tomography with deep learning". In: *Nature communications* 13.1, p. 6482. DOI: 10.1038/s41467-022-33957-8.

Lucas, Bronwyn A. and Nikolaus Grigorieff (2023). "Quantification of gallium cryo-FIB milling damage in biological lamellae". In: *Proceedings of the National Academy of Sciences of the United States of America* 120.23, e2301852120. DOI: 10.1073/pnas.2301852120.

Lučič, Vladan, Alexander Rigort, and Wolfgang Baumeister (2013). "Cryo-electron tomography: the challenge of doing structural biology in situ". In: *The Journal of cell biology* 202.3, pp. 407–419. DOI: 10.1083/jcb.201304193.

Malone, Brandon, Nadya Urakova, Eric J. Snijder, and Elizabeth A. Campbell (2022). "Structures and functions of coronavirus replication-transcription complexes and their relevance for SARS-CoV-2 drug design". In: *Nature reviews. Molecular cell biology* 23.1, pp. 21–39. DOI: 10.1038/s41580-021-00432-z.

Mariano, Giuseppina, Rebecca J. Farthing, Shamar L. M. Lale-Farjat, and Julien R. C. Bergeron (2020). "Structural Characterization of SARS-CoV-2: Where We Are, and Where We Need to Be". In: *Frontiers in molecular biosciences* 7, p. 605236. ISSN: 2296-889X. DOI: 10.3389/fmolb.2020.605236.

Martinez-Sanchez, Antonio, Zdravko Kochovski, Ulrike Laugks, Johannes zum Meyer Alten Borgloh, Saikat Chakraborty, Stefan Pfeffer, Wolfgang Baumeister, and Vladan Lučić (2020). "Template-free detection and classification of membrane-bound complexes in cryo-electron tomograms". In: *Nature methods* 17.2, pp. 209–216. DOI: 10.1038/s41592-019-0675-5.

McDonald, Christopher, Goran Jovanovic, Oscar Ces, and Martin Buck (2015). "Membrane Stored Curvature Elastic Stress Modulates Recruitment of Maintenance Proteins PspA and Vipp1". In: *mBio* 6.5, e01188–15. DOI: 10.1128/mBio.01188-15.

McMullan, G., A. R. Faruqi, D. Clare, and R. Henderson (2014). "Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy". In: *Ultramicroscopy* 147, pp. 156–163. ISSN: 03043991. DOI: 10.1016/j.ultramic.2014.08.002.

Mizushima, Noboru, Tamotsu Yoshimori, and Yoshinori Ohsumi (2011). "The role of Atg proteins in autophagosome formation". In: *Annual review of cell and developmental biology* 27, pp. 107–132. DOI: 10.1146/annurev-cellbio-092910-154005.

Molina, José M., Juan P. Llerena, Luis Usero, and Miguel A. Patricio (2025). "Advances in instance segmentation: Technologies, metrics and applications in computer vision". In: *Neurocomputing* 625, p. 129584. ISSN: 09252312. DOI: 10.1016/j.neucom.2025.129584.

Nähse, Viola, Camilla Raiborg, Kia Wee Tan, Sissel Mørk, Maria Lyngaas Torgersen, Eva Maria Wenzel, Mireia Nager, Veijo T. Salo, Terje Johansen, Elina Ikonen, Kay Oliver Schink, and Harald Stenmark (2023). "ATPase activity of DFCP1 controls selective autophagy". In: *Nature communications* 14.1, p. 4051. DOI: 10.1038/s41467-023-39641-9.

Navarro, Paula P., Henning Stahlberg, and Daniel Castaño-Díez (2018). "Protocols for Subtomogram Averaging of Membrane Proteins in the Dynamo Software Package". In: *Frontiers in molecular biosciences* 5, p. 82. ISSN: 2296-889X. DOI: 10.3389/fmolb.2018.00082.

Nešetřil, Jaroslav, Eva Milková, and Helena Nešetřilová (2001). "Otakar Borůvka on minimum spanning tree problem Translation of both the 1926 papers, comments, history". In: *Discrete Mathematics* 233.1-3, pp. 3–36. ISSN: 0012365X. DOI: 10.1016/S0012-365X(00)00224-7.

Nickell, Stephan, Friedrich Förster, Alexandros Linaroudis, William Del Net, Florian Beck, Reiner Hegerl, Wolfgang Baumeister, and Jürgen M. Plitzko (2005). "TOM software toolbox: acquisition and analysis for electron tomography". In: *Journal of structural biology* 149.3, pp. 227–234. DOI: 10.1016/j.jsb.2004.10.006.

Noble, Alex J. and Alex de Marco (2024). “Cryo-focused ion beam for in situ structural biology: State of the art, challenges, and perspectives”. In: *Current opinion in structural biology* 87, p. 102864. DOI: 10.1016/j.sbi.2024.102864.

Noble, Alex J. and Scott M. Stagg (2015). “Automated batch fiducial-less tilt-series alignment in Appion using Protomo”. In: *Journal of structural biology* 192.2, pp. 270–278. DOI: 10.1016/j.jsb.2015.10.003.

Nogales, Eva and Sjors H. W. Scheres (2015). “Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity”. In: *Molecular cell* 58.4, pp. 677–689. DOI: 10.1016/j.molcel.2015.02.019.

Orlova, E. V., P. Dube, J. R. Harris, E. Beckman, F. Zemlin, J. Markl, and M. van Heel (1997). “Structure of keyhole limpet hemocyanin type 1 (KLH1) at 15 A resolution by electron cryomicroscopy and angular reconstitution”. In: *Journal of molecular biology* 271.3, pp. 417–437. ISSN: 0022-2836. DOI: 10.1006/jmbi.1997.1182.

Oudshoorn, Diede, Kevin Rijs, Ronald W. A. L. Limpens, Kevin Groen, Abraham J. Koster, Eric J. Snijder, Marjolein Kikkert, and Montserrat Bárcena (2017). “Expression and Cleavage of Middle East Respiratory Syndrome Coronavirus nsp3-4 Polyprotein Induce the Formation of Double-Membrane Vesicles That Mimic Those Associated with Coronaviral RNA Replication”. In: *mBio* 8.6. DOI: 10.1128/mbio.01658-17.

Park, Jisoo, Jongyoon Kim, Hyungsun Park, Taewan Kim, and Seongju Lee (2024). “ESCRT-III: a versatile membrane remodeling machinery and its implications in cellular processes and diseases”. In: *Animal cells and systems* 28.1, pp. 367–380. ISSN: 1976-8354. DOI: 10.1080/19768354.2024.2380294.

Pavel Iakubovskii (April, 2025). *finetune-instance-segmentation-ade20k-mini-mask2former*.

Penczek, Pawel A. (2020). “Reliable cryo-EM resolution estimation with modified Fourier shell correlation”. In: *IUCrJ* 7.Pt 6, pp. 995–1008. ISSN: 2052-2525. DOI: 10.1107/S2052252520011574.

Pettersen, Eric F., Thomas D. Goddard, Conrad C. Huang, Gregory S. Couch, Daniel M. Greenblatt, Elaine C. Meng, and Thomas E. Ferrin (2004). "UCSF Chimera–a visualization system for exploratory research and analysis". In: *Journal of computational chemistry* 25.13, pp. 1605–1612. ISSN: 0192-8651. DOI: 10.1002/jcc.20084.

Pettersen, Eric F., Thomas D. Goddard, Conrad C. Huang, Elaine C. Meng, Gregory S. Couch, Tristan I. Croll, John H. Morris, and Thomas E. Ferrin (2021). "UCSF ChimeraX: Structure visualization for researchers, educators, and developers". In: *Protein science : a publication of the Protein Society* 30.1, pp. 70–82. DOI: 10.1002/pro.3943.

Pierson, Joshua A., Jie E. Yang, and Elizabeth R. Wright (2024). "Recent advances in correlative cryo-light and electron microscopy". In: *Current opinion in structural biology* 89, p. 102934. DOI: 10.1016/j.sbi.2024.102934.

Punjani, Ali, John L. Rubinstein, David J. Fleet, and Marcus A. Brubaker (2017). "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination". In: *Nature methods* 14.3, pp. 290–296. DOI: 10.1038/nmeth.4169.

Pyle, Euan and Giulia Zanetti (2021). "Current data processing strategies for cryo-electron tomography and subtomogram averaging". In: *The Biochemical journal* 478.10, pp. 1827–1845. DOI: 10.1042/BCJ20200715.

Radermacher, M. (1988). "Three-dimensional reconstruction of single particles from random and nonrandom tilt series". In: *Journal of electron microscopy technique* 9.4, pp. 359–394. ISSN: 0741-0581. DOI: 10.1002/jemt.1060090405.

Ricciardi, Simona, Andrea Maria Guarino, Laura Giaquinto, Elena V. Polishchuk, Michele Santoro, Giuseppe Di Tullio, Cathal Wilson, Francesco Panariello, Vinicius C. Soares, Suelen S. G. Dias, Julia C. Santos, Thiago M. L. Souza, Giovanna Fusco, Maurizio Viscardi, Sergio Brandi, Patrícia T. Bozza, Roman S. Polishchuk, Rossella Venditti, and Maria Antonietta de Matteis (2022). "The role of NSP6 in the biogenesis of the SARS-CoV-2 replication organelle". In: *Nature* 606.7915, pp. 761–768. ISSN: 0028-0836. DOI: 10.1038/s41586-022-04835-6.

Rohou, Alexis and Nikolaus Grigorieff (2015). "CTFFIND4: Fast and accurate defocus estimation from electron micrographs". In: *Journal of structural biology* 192.2, pp. 216–221. DOI: 10.1016/j.jsb.2015.08.008.

Roingeard, Philippe, Sébastien Eymieux, Julien Burlaud-Gaillard, Christophe Hourioux, Romuald Patient, and Emmanuelle Blanchard (2022). "The double-membrane vesicle (DMV): a virus-induced organelle dedicated to the replication of SARS-CoV-2 and other positive-sense single-stranded RNA viruses". In: *Cellular and molecular life sciences : CMLS* 79.8, p. 425. DOI: 10.1007/s00018-022-04469-x.

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi. Vol. 9351. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 234–241. ISBN: 978-3-319-24573-7. DOI: 10.1007/978-3-319-24574-4{\textunderscore}28.

Rose, H. H. (2008). "Optics of high-performance electron microscopes". In: *Science and Technology of Advanced Materials* 9.1, p. 014107. DOI: 10.1088/0031-8949/9/1/014107.

Rosenthal, Peter B. and Richard Henderson (2003). "Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy". In: *Journal of molecular biology* 333.4, pp. 721–745. ISSN: 0022-2836. DOI: 10.1016/j.jmb.2003.07.013.

Salfer, Maria, Javier F. Collado, Wolfgang Baumeister, Rubén Fernández-Busnadiego, and Antonio Martínez-Sánchez (2020). "Reliable estimation of membrane curvature for cryo-electron tomography". In: *PLoS computational biology* 16.8, e1007962. DOI: 10.1371/journal.pcbi.1007962.

Savitzky, Abraham. and M. J. E. Golay (1964). “Smoothing and Differentiation of Data by Simplified Least Squares Procedures”. In: *Analytical Chemistry* 36.8, pp. 1627–1639. ISSN: 0003-2700. DOI: 10.1021/ac60214a047.

Saxton, W. O. and W. Baumeister (1982). “The correlation averaging of a regularly arranged bacterial cell envelope protein”. In: *Journal of microscopy* 127.Pt 2, pp. 127–138. ISSN: 0022-2720. DOI: 10.1111/j.1365-2818.1982.tb00405.x.

Scheres, Sjors Hw (2014). “Beam-induced motion correction for sub-megadalton cryo-EM particles”. In: *eLife* 3, e03665. DOI: 10.7554/eLife.03665.

Schiøtz, Oda Helene, Sven Klumpe, Juergen M. Plitzko, and Christoph J. O. Kaiser (2024). “Cryo-electron tomography: en route to the molecular anatomy of organisms and tissues”. In: *Biochemical Society transactions* 52.6, pp. 2415–2425. DOI: 10.1042/BST20240173.

Schönnenbeck, Philipp, Benedikt Junglas, and Carsten Sachse (2025). “CryoVIA: An image analysis toolkit for the quantification of membrane structures from cryo-EM micrographs”. In: *Structure (London, England : 1993)* 33.4, 808–819.e4. DOI: 10.1016/j.str.2025.01.013.

Sergio, Maria Concetta, Simona Ricciardi, Andrea M. Guarino, Laura Giaquinto, and Maria Antonietta de Matteis (2024). “Membrane remodeling and trafficking piloted by SARS-CoV-2”. In: *Trends in cell biology* 34.9, pp. 785–800. DOI: 10.1016/j.tcb.2023.12.006.

Siebenaller, Carmen, Benedikt Junglas, and Dirk Schneider (2019). “Functional Implications of Multiple IM30 Oligomeric States”. In: *Frontiers in plant science* 10, p. 1500. ISSN: 1664-462X. DOI: 10.3389/fpls.2019.01500.

Sigworth, Fred J. (2016). “Principles of cryo-EM single-particle image processing”. In: *Microscopy (Oxford, England)* 65.1, pp. 57–67. DOI: 10.1093/jmicro/dfv370.

Singer, Amit and Fred J. Sigworth (2020). “Computational Methods for Single-Particle Electron Cryomicroscopy”. In: *Annual review of biomedical data science* 3, pp. 163–190. DOI: 10.1146/annurev-biodatasci-021020-093826.

Singh, Shashi Kant, Shubham Kumar, and Pawan Singh Mehra (2023). “Chat GPT & Google Bard AI: A Review”. In: *2023 International Conference on IoT, Communication and Automation Technology (ICICAT)*. IEEE, pp. 1–6. ISBN: 979-8-3503-0282-0. DOI: 10.1109/ICICAT57735.2023.10263706.

Soille, Pierre (1999). *Morphological image analysis: Principles and applications ; with 12 tables*. Berlin et al.: Springer. ISBN: 3-540-65671-5.

Sousa, Duncan and Nikolaus Grigorieff (2007). “Ab initio resolution measurement for single particle structures”. In: *Journal of structural biology* 157.1, pp. 201–210. DOI: 10.1016/j.jsb.2006.08.003.

Stark, Holger, Friedrich Zemlin, and Christoph Boettcher (1996). “Electron radiation damage to protein crystals of bacteriorhodopsin at different temperatures”. In: *Ultramicroscopy* 63.2, pp. 75–79. ISSN: 03043991. DOI: 10.1016/0304-3991(96)00045-9.

Takeshige, K., M. Baba, S. Tsuboi, T. Noda, and Y. Ohsumi (1992). “Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction”. In: *The Journal of cell biology* 119.2, pp. 301–311. DOI: 10.1083/jcb.119.2.301.

Tan, Xuqian, Ethan Boniuk, Anisha Abraham, Xueting Zhou, Zhili Yu, Steven J. Ludtke, and Zhao Wang (2025). “TomoScore: A Neural Network Approach for Quality Assessment of Cellular cryo-ET”. In: *Research square*. DOI: 10.21203/rs.3.rs-5405930/v1.

Tegunov, Dimitry and Patrick Cramer (2019). “Real-time cryo-electron microscopy data preprocessing with Warp”. In: *Nature methods* 16.11, pp. 1146–1152. DOI: 10.1038/s41592-019-0580-y.

Tegunov, Dimitry, Liang Xue, Christian Dienemann, Patrick Cramer, and Julia Mahamid (2021). "Multi-particle cryo-EM refinement with M visualizes ribosome-antibiotic complex at 3.5 Å in cells". In: *Nature methods* 18.2, pp. 186–193. DOI: 10.1038/s41592-020-01054-7.

Theis, Jasmine, Tilak Kumar Gupta, Johannes Klingler, William Wan, Sahradha Albert, Sandro Keller, Benjamin D. Engel, and Michael Schroda (2019). "VIPP1 rods engulf membranes containing phosphatidylinositol phosphates". In: *Scientific reports* 9.1, p. 8725. DOI: 10.1038/s41598-019-44259-3.

Tuijtel, Maarten W., Sergio Cruz-León, Jan Philipp Kreysing, Sonja Welsch, Gerhard Hummer, Martin Beck, and Beata Turoňová (2024). "Thinner is not always better: Optimizing cryo-lamellae for subtomogram averaging". In: *Science advances* 10.17, eadk6285. DOI: 10.1126/sciadv.adk6285.

Turoňová, Beata, Wim J. H. Hagen, Martin Obr, Shyamal Mosalaganti, J. Wouter Beugelink, Christian E. Zimmerli, Hans-Georg Kräusslich, and Martin Beck (2020a). "Benchmarking tomographic acquisition schemes for high-resolution structural biology". In: *Nature communications* 11.1, p. 876. DOI: 10.1038/s41467-020-14535-2.

Turoňová, Beata, Florian K. M. Schur, William Wan, and John A. G. Briggs (2017). "Efficient 3D-CTF correction for cryo-electron tomography using NovaCTF improves subtomogram averaging resolution to 3.4Å". In: *Journal of structural biology* 199.3, pp. 187–195. DOI: 10.1016/j.jsb.2017.07.007.

Turoňová, Beata, Mateusz Sikora, Christoph Schürmann, Wim J. H. Hagen, Sonja Welsch, Florian E. C. Blanc, Sören von Bülow, Michael Gecht, Katrin Bagola, Cindy Hörner, Ger van Zandbergen, Jonathan Landry, Nayara Trevisan Doimo de Azevedo, Shyamal Mosalaganti, Andre Schwarz, Roberto Covino, Michael D. Mühlebach, Gerhard Hummer, Jacomine Krijnse Locker, and Martin Beck (2020b). "In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges". In: *Science (New York, N.Y.)* 370.6513, pp. 203–208. DOI: 10.1126/science.abd5223.

Twu, Woan-Ing, Ji-Young Lee, Heeyoung Kim, Vibhu Prasad, Berati Cerikan, Uta Haselmann, Keisuke Tabata, and Ralf Bartenschlager (2021). "Contribution of autophagy machinery factors to HCV and SARS-CoV-2 replication organelle formation". In: *Cell reports* 37.8, p. 110049. DOI: 10.1016/j.celrep.2021.110049.

van Heel, Marin and Joachim Frank (1981). "Use of multivariates statistics in analysing the images of biological macromolecules". In: *Ultramicroscopy* 6.1, pp. 187–194. ISSN: 03043991. DOI: 10.1016/S0304-3991(81)80197-0.

van Heel, Marin and Michael Schatz (2005). "Fourier shell correlation threshold criteria". In: *Journal of structural biology* 151.3, pp. 250–262. DOI: 10.1016/j.jsb.2005.05.009.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin (2017). "Attention is all you need". In: *Advances in Neural Information Processing Systems 30*.

Verbeke, Eric J., Marc Aurèle Gilles, Tamir Bendory, and Amit Singer (2024). "Self Fourier shell correlation: properties and application to cryo-ET". In: *Communications biology* 7.1, p. 101. DOI: 10.1038/s42003-023-05724-y.

Vilas, J. L., J. Oton, C. Messaoudi, R. Melero, P. Conesa, E. Ramirez-Aportela, J. Mota, M. Martinez, A. Jimenez, R. Marabini, J. M. Carazo, J. Vargas, and C. O. S. Sorzano (2020). "Measurement of local resolution in electron tomography". In: *Journal of structural biology: X* 4, p. 100016. DOI: 10.1016/j.yjsbx.2019.100016.

Vilas, Jose Luis, Josué Gómez-Blanco, Pablo Conesa, Roberto Melero, José La Miguel de Rosa-Trevín, Joaquin Otón, Jesús Cuenca, Roberto Marabini, José María Carazo, Javier Vargas, and Carlos Oscar S. Sorzano (2018). "MonoRes: Automatic and Accurate Estimation of Local Resolution for Electron Microscopy Maps". In: *Structure (London, England : 1993)* 26.2, 337–344.e4. DOI: 10.1016/j.str.2017.12.018.

V'kovski, Philip, Annika Kratzel, Silvio Steiner, Hanspeter Stalder, and Volker Thiel (2021). "Coronavirus biology and replication: implications for SARS-CoV-2". In: *Nature reviews. Microbiology* 19.3, pp. 155–170. DOI: 10.1038/s41579-020-00468-6.

Vothknecht, Ute C., Stephanie Otters, Raoul Hennig, and Dirk Schneider (2012). "Vipp1: a very important protein in plastids?!" In: *Journal of experimental botany* 63.4, pp. 1699–1712. DOI: 10.1093/jxb/err357.

Wade, R. H. (1992). "A brief look at imaging and contrast transfer". In: *Ultramicroscopy* 46.1-4, pp. 145–156. ISSN: 03043991. DOI: 10.1016/0304-3991(92)90011-8.

Wagner, Thorsten, Felipe Merino, Markus Stabrin, Toshio Moriya, Claudia Antoni, Amir Apelbaum, Philine Hagel, Oleg Sitsel, Tobias Raisch, Daniel Prumbaum, Dennis Quentin, Daniel Roderer, Sebastian Tacke, Birte Siebolds, Evelyn Schubert, Tanvir R. Shaikh, Pascal Lill, Christos Gatsogiannis, and Stefan Raunser (2019). "SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM". In: *Communications biology* 2, p. 218. DOI: 10.1038/s42003-019-0437-z.

Wan, W. and J. A. G. Briggs (2016). "Cryo-Electron Tomography and Subtomogram Averaging". In: *Methods in enzymology* 579, pp. 329–367. DOI: 10.1016/bs.mie.2016.04.014.

Wan, William, Sagar Khavnekar, and Jonathan Wagner (2024). "STOPGAP: an open-source package for template matching, subtomogram alignment and classification". In: *Acta crystallographica. Section D, Structural biology* 80.Pt 5, pp. 336–349. DOI: 10.1107/S205979832400295X.

Weissenberger, Giulia, Rene J. M. Henderikx, and Peter J. Peters (2021). "Understanding the invisible hands of sample preparation for cryo-EM". In: *Nature methods* 18.5, pp. 463–471. DOI: 10.1038/s41592-021-01130-6.

Welch, B. L. (1938). “The Significance of the Difference Between Two Means when the Population Variances are Unequal”. In: *Biometrika* 29.3/4, p. 350. ISSN: 00063444. DOI: 10.2307/2332010.

Wiedemann, Simon and Reinhard Heckel (2024). “A deep learning method for simultaneous denoising and missing wedge reconstruction in cryogenic electron tomography”. In: *Nature communications* 15.1, p. 8255. DOI: 10.1038/s41467-024-51438-y.

Wolff, Georg, Ronald W. A. L. Limpens, Jessika C. Zevenhoven-Dobbe, Ulrike Laugks, Shawn Zheng, Anja W. M. de Jong, Roman I. Koning, David A. Agard, Kay Grünewald, Abraham J. Koster, Eric J. Snijder, and Montserrat Bárcena (2020). “A molecular pore spans the double membrane of the coronavirus replication organelle”. In: *Science (New York, N.Y.)* 369.6509, pp. 1395–1398. DOI: 10.1126/science.abd3629.

Wu, Fan, Su Zhao, Bin Yu, Yan-Mei Chen, Wen Wang, Zhi-Gang Song, Yi Hu, Zhao-Wu Tao, Jun-Hua Tian, Yuan-Yuan Pei, Ming-Li Yuan, Yu-Ling Zhang, Fa-Hui Dai, Yi Liu, Qi-Min Wang, Jiao-Jiao Zheng, Lin Xu, Edward C. Holmes, and Yong-Zhen Zhang (2020). “Author Correction: A new coronavirus associated with human respiratory disease in China”. In: *Nature* 580.7803, E7. ISSN: 0028-0836. DOI: 10.1038/s41586-020-2202-3.

Xu, Dijin, Yuqi Li, Lizhen Wu, Ying Li, Dongyu Zhao, Jinhai Yu, Tuozhi Huang, Charles Ferguson, Robert G. Parton, Hongyuan Yang, and Peng Li (2018). “Rab18 promotes lipid droplet (LD) growth by tethering the ER to LDs through SNARE and NRZ interactions”. In: *The Journal of cell biology* 217.3, pp. 975–995. DOI: 10.1083/jcb.201704184.

Xu, Yixin and Shangyu Dang (2022). “Recent Technical Advances in Sample Preparation for Single-Particle Cryo-EM”. In: *Frontiers in molecular biosciences* 9, p. 892459. ISSN: 2296-889X. DOI: 10.3389/fmolb.2022.892459.

Yang, Jinping, Buyun Tian, Pei Wang, Rongrong Chen, Ke Xiao, Xubing Long, Xinyue Zheng, Yun Zhu, Fei Sun, Yongxia Shi, Yaming Jiu, Wei Ji, Yanhong Xue, Tao Xu,

and Zonghong Li (2025). “SARS-CoV-2 NSP3/4 control formation of replication organelle and recruitment of RNA polymerase NSP12”. In: *The Journal of cell biology* 224.3. DOI: 10.1083/jcb.202306101.

Zaffagnini, Gabriele and Sascha Martens (2016). “Mechanisms of Selective Autophagy”. In: *Journal of molecular biology* 428.9 Pt A, pp. 1714–1724. ISSN: 0022-2836. DOI: 10.1016/j.jmb.2016.02.004.

Zhang, Lingang, Yusuke Kato, Stephanie Otters, Ute C. Vothknecht, and Wataru Sakamoto (2012). “Essential role of VIPP1 in chloroplast envelope maintenance in Arabidopsis”. In: *The Plant cell* 24.9, pp. 3695–3707. DOI: 10.1105/tpc.112.103606.

Zhang, T. Y. and C. Y. Suen (1984). “A fast parallel algorithm for thinning digital patterns”. In: *Communications of the ACM* 27.3, pp. 236–239. ISSN: 0001-0782. DOI: 10.1145/357994.358023.

Zheng, Shawn, Georg Wolff, Garrett Greenan, Zhen Chen, Frank G. A. Faas, Montserrat Bárcena, Abraham J. Koster, Yifan Cheng, and David A. Agard (2022). “AreTomo: An integrated software package for automated marker-free, motion-corrected cryo-electron tomographic alignment and reconstruction”. In: *Journal of structural biology: X* 6, p. 100068. DOI: 10.1016/j.yjsbx.2022.100068.

Zheng, Shawn Q., Eugene Palovcak, Jean-Paul Armache, Kliment A. Verba, Yifan Cheng, and David A. Agard (2017). “MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy”. In: *Nature methods* 14.4, pp. 331–332. DOI: 10.1038/nmeth.4193.

Zimmermann, Liv, Xiaohan Zhao, Jana Makroczyova, Moritz Wachsmuth-Melm, Vibhu Prasad, Zach Hensel, Ralf Bartenschlager, and Petr Chlanda (2023). “SARS-CoV-2 nsp3 and nsp4 are minimal constituents of a pore spanning replication organelle”. In: *Nature communications* 14.1, p. 7894. DOI: 10.1038/s41467-023-43666-5.

Zivanov, Jasenko, Takanori Nakane, Björn O. Forsberg, Dari Kimanius, Wim Jh Hagen, Erik Lindahl, and Sjors Hw Scheres (2018). “New tools for automated

high-resolution cryo-EM structure determination in RELION-3". In: *eLife* 7. DOI: 10.7554/eLife.42166.

Zivanov, Jasenko, Joaquín Otón, Zunlong Ke, Andriko von Kügelgen, Euan Pyle, Kun Qu, Dustin Morado, Daniel Castaño-Díez, Giulia Zanetti, Tanmay A. M. Bharat, John A. G. Briggs, and Sjors H. W. Scheres (2022). "A Bayesian approach to single-particle electron cryo-tomography in RELION-4.0". In: *eLife* 11. DOI: 10.7554/eLife.83724.