Elucidating the structural determinants of the poxvirus core using multi-modal cryo-EM

by

Julia Datler

December, 2024

*A thesis submitted to the
Graduate School
of the
Institute of Science and Technology Austria
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy*

Committee in charge:
Prof. Dr. Edouard Hannezo, Chair
Ass.-Prof. Dr. Carrie Bernecky
Prof. Dr. Andreas Bergthaler
Prof. Dr. Florian Schur

**Institute of
Science and
Technology
Austria**

The thesis of Julia Datler, titled "*Elucidating the structural determinants of the poxvirus core using multi-modal cryo-EM*", is approved by:

Supervisor: Florian Schur, ISTA, Klosterneuburg, Austria

Signature: _____

Committee Member: Carrie Bernecky, ISTA, Klosterneuburg, Austria

Signature: _____

Committee Member: Andreas Bergthaler, Medical University of Vienna, Austria

Signature: _____

Defense Chair: Edouard Hannezo, ISTA, Klosterneuburg, Austria

Signature: _____

Signed page is on file

I hereby declare that this thesis is my own work and that it does not contain other people's work without this being so stated; this thesis does not contain my previous work without this being stated, and the bibliography contains all the literature that I used in writing the dissertation.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee, and that this thesis has not been submitted for a higher degree to any other university or institution.

I certify that any republication of materials presented in this thesis has been approved by the relevant publishers and co-authors.

Signature: _____

Julia Datler

December, 2024

Signed page is on file

Abstract

Poxviruses are large pleomorphic double-stranded DNA viruses that include well-known members such as variola virus, the causative agent of smallpox, Mpox virus, as well as Vaccinia virus (VACV), which serves as a vaccination strain for formerly mentioned viruses. VACV is a valuable model for studying large pleomorphic DNA viruses in general and poxviruses specifically, as many features, such as core morphology and structural proteins, are well conserved within this family.

Despite decades of research, our understanding of the structural components and proteins that comprise the poxvirus core in mature virions remains limited. Although major core proteins were identified via indirect experimental evidence, the core's complexity, with its large size, structure and number of involved proteins, has hindered efforts to achieve high-resolution insights and to define the roles of the individual proteins. The specific protein composition of the core's individual layers, including the palisade layer and the inner core wall, has remained unclear.

In this study, we have merged multiple approaches, including single particle cryo-electron microscopy of purified virus cores, cryo-electron tomography and subtomogram averaging of mature virions and molecular modeling to elucidate the structural determinants of the VACV core. Due to the lack of experimentally derived structures, either *in situ* or reconstituted *in vitro*, we used Alphafold to predict models of the putative major core protein candidates, A10, 23k, A3, A4, and L4.

Our results show that the VACV core is composed of several layers with varying local symmetries, forming more intricate interactions than observed previously. This allowed us to identify several molecular building blocks forming the viral core lattice. In particular, we identified trimers of protein A10 as a major core structure that forms the palisade layer of the viral core. Additionally, we revealed that six petals of a flower-shaped core pore within the core wall are composed of A10 trimers. Furthermore, we obtained a cryo-EM density for the inner core wall that could potentially accommodate an A3 dimer. Integrating descriptions of protein interactions from previous studies enabled us to provide a detailed structural model of the poxvirus core wall, and our findings indicate that the interactions within A10 trimers are likely consistent across orthopox- and parapoxviruses.

This combined application of cryo-SPA and cryo-ET can help overcome obstacles in studying complex virus structures in the future, including their key assembly proteins, interactions, and the formation into a core lattice. Our work provides important fundamental new insights into poxvirus core architecture, also considering the recent re-emergence of poxviruses.

Acknowledgments

I want to thank Flo for giving me the opportunity to perform my PhD thesis in his lab and providing guidance, feedback and vital discussions on the project. Thank you for always being so optimistic, it has helped me tremendously to stay motivated throughout these years. I must admit that the 'funny' jokes also helped their fair share to keep the mood up. I thank you for all of your support, especially in the last 2 years, for helping me achieve both my scientific goals and being there for my family.

Additionally, I would like to thank my PhD thesis committee members, Carrie Bernecky and Andreas Bergthaler, for their support, scientific discussions, and annual feedback during my progress reviews.

I also want to thank all current and former lab members of the Schur group Betty, Dario, Florian, Andi, Georgi, Martin, Manjunath, Rahma, Mathias, Jesse, Darya, Lukas, Francisco, Benedikt and Kain for their continuous support, scientific discussions, fabulous lunch breaks, Billa walks and cake celebrations. I will truly miss these in the future. A big thank you to Jesse, without you our achievement of publishing the manuscript in NSMB would not have been possible. I especially want to thank the experienced "oldies but goldies" Betty and Dario for their friendship and support. You always knew when it was time to lift my mood with cute Emmi pics, pep talks and 'funny' jokes (that made me and Betty laugh and Dario cringe :D)

Further, I would like to thank Beata and Fleur, my fellow PhD buddies for their friendship, coffee breaks and city escapes. Additionally, I want to thank all of my close friends for their support and encouragement. A special thank you goes to Heike who often took care of Anton while I was writing this thesis.

I cannot express how thankful I am to my family, especially my parents Silvia and Heinz for always believing in me and encouraging me to reach my goals. Thank you for supporting me when I came up with this idea that I wanted to become a scientist at an early age. This accomplishment would not have been possible without you.

Finally, I would like to thank my husband Hannes and my son Anton, for always being there for me, supporting and encouraging me no matter what. I could not imagine a life without you and I am so happy to come home to you every day.

About the Author

Julia Datler completed her BSc in Molecular Biology at the Technical University of Graz in Austria. She then earned a MSc in Molecular Medicine from the University of Vienna and another MSc in Molecular Microbiology from the Technical University of Graz. During this period, Julia developed expertise in molecular biology and virology, working on a project investigating the N-myristoylation of picornaviruses, leading to a publication in the Proceedings of the National Academy of Sciences (PNAS). Before joining the Institute of Science and Technology Austria in September 2018, she spent a year at the University of Cambridge and the Francis Crick Institute, where she researched metabolomics. After joining Florian Schur's group at ISTA, Julia gained expertise in cryo-electron microscopy and computational processing. Her primary research interest is the structural virology of poxviruses. In this project, she combined her previous virology experience with her newly acquired cryo-EM skills to analyze the structural core morphology of poxviruses. Julia's work *"Multi-modal cryo-EM reveals trimers of protein A10 to form the palisade layer in poxvirus cores."* was published in the journal Nature Molecular & Structural Biology in 2024. During her PhD studies, she also presented her research findings at the European Virology Congress held in Gdansk, Poland, in 2023. Julia additionally contributed to two additional papers for one of them she used the vaccinia virus (VACV) tail model to examine the regulator of actin branching.

List of Publications

I am a shared first author on a paper published in the Journal Nature Structural & Molecular Biology, which is the central part of my PhD thesis.

Datler, J.*, Hansen, J.M.*, Thader, A., Schlögl, A., Bauer, L.W., Hodirnau, V.-V., and Schur, F.K.M. (2024). *"Multi-modal cryo-EM reveals trimers of protein A10 to form the palisade layer in poxvirus cores."* Nature Structural & Molecular Biology, 1–10. https://doi.org/10.1038/s41594-023-01201-6.
*Equal contribution

I am a co-author of the following paper published in Science Advances, which is described in Appendix 1:

Fäßler, F., Javoor, M.G., **Datler, J.,** Döring, H., Hofer, F.W., Dimchev, G., Hodirnau, V.-V., Faix, J., Rottner, K., and Schur, F.K.M. (2023*). "ArpC5 isoforms regulate Arp2/3 complex–dependent protrusion through differential Ena/VASP positioning."* Science Advances 9, eadd6495. https://doi.org/10.1126/sciadv.add6495.

List of Contributors

Jesse Hansen (Schur Group)
Jesse is shared first author of the publication *"Multi-modal cryo-EM reveals trimers of protein A10 to form the palisade layer in poxvirus cores."* that we published in Nature Structural & Molecular Biology in 2024. Jesse contributed to the Cryo-SPA data acquisition and single particle analysis as well as the structure fitting, structure modeling and interaction analysis described in chapter 4.6, chapter 4.7, chapter 4.8 and chapter 4.10.

Lukas Bauer (Schur Group)
Lukas helped with Alphafold predictions in Colabfold 2.3 of the described Entomopoxviruses (Figure 27).

Victor-Valentin Hodirnau (Electron Microscopy Facility)
Valentin provided technical support and helped with data acquisition at the Titan Krios G3i.

Alois Schlögl (ISTA Scientific Computing facility)
Alois provided help and support with scripts and software used in this thesis.

Armel Nicolas (ISTA proteomics facility)
Armel performed the mass spectrometry experiments of the soluble fraction of the VACV cores described in chapter 3.9.

Florian Fäßler (Schur Group, now Fäßler Group IGBMC Strasbourg)
Florian is the first author of the paper *"ArpC5 isoforms regulate Arp2/3 complex–dependent protrusion through differential Ena/VASP positioning."* published in Science Advances 2023. My contribution to the paper is outlined in Appendix I.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AF2	AlphaFold 2
AI	Artificial Intelligence
AmEPV	Amsacta moorei Entomopoxvirus
Å	Angstrom
BSL	Biosafety Level
CASP	Critical Assessment of Protein Structure Prediction
CCC	Cross-Correlation Coefficient
CEV	Cell-Associated Enveloped Virions
Cm	Centimeter
CO_2	Carbon Dioxide
CNN	Convolutional Neural Network
cryo-EM	Cryo-Electron Microscopy
cryo-ET	Cryo-Electron Tomography
cryo-SPA	Cryo-Single Particle Analysis
CTF	Contrast Transfer Function
CPE	Cytopathic Effect
°C	Degrees Celsius
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
EEV	Extracellular Enveloped Virions
EM	Electron Microscopy
EMDB	Electron Microscopy Data Bank
FBS	Fetal Bovine Serum
FSC	Fourier Shell Correlation
GUI	Graphical User Interface
HCl	Hydrochloric Acid
IEV	Intracellular Enveloped Virions
IV	Immature Virions
kDa	Kilodalton
KO	knock out

kV	Kilovolt
M	Molar
Mg	Milligram
mL	Milliliter
mm	Millimeter
mM	Millimolar
ML	Machine Learning
MOI	Multiplicity of Infection
MPXV	Monkeypox Virus
MS	Mass Spectrometry
MSA	Multiple Sequence Alignment
MSEPV	Melanoplus sanguinipes Entomopoxvirus
MVA	Modified Vaccinia Virus Ankara
MV	Mature Virions
NaCl	Sodium Chloride
Nm	Nanometer
NMR	Nuclear Magnetic Resonance
NP40	Nonidet P-40
pH	Potential of Hydrogen
PBS	Phosphate-Buffered Saline
PDB	Protein Data Bank
PFA	Paraformaldehyde
pLDDT	Predicted Local Distance Difference Test
RNA	Ribonucleic Acid
RT-EM	Room-Temperature Electron Microscopy
SD	Standard Deviation
SEM	Scanning Electron Microscopy
SIRT	Simultaneous Iterative Reconstruction Technique
SNR	Signal-to-Noise Ratio
SPA	Single Particle Analysis
STA	Sub-Tomogram Averaging
TCEP	Tris(2-carboxyethyl)phosphine
TEM	Transmission Electron Microscopy
Tris-HCl	Tris(hydroxymethyl)aminomethane-Hydrochloride

VACV	Vaccinia Virus
VACV WR	Vaccinia Virus Western Reserve
WASP	Wiskott-Aldrich Syndrome Protein
WBP	Weighted Back Projection
WHO	World Health Organization
WIP	WASP-Interacting Protein
wt	wildtype
XL-MS	Cross-Linking Mass Spectrometry
µg	Microgram
µL	Microliter
µm	Micrometer
µM	Micromolar
2D	Two-Dimensional
3D	Three-Dimensional

1 Introduction

1.1 *Poxviridae*

1.1.1 Orthopoxviruses – State of the art

Orthopoxviruses, members of the family *Poxviridae*, are enveloped pleomorphic double stranded deoxyribonucleic acid (dsDNA) viruses that can infect various hosts from vertebrates to arthropods (Condit et al. 2006). The most prominent member of the genus orthopoxviruses is variola virus, the causative agent of smallpox, which was, until its eradication in the 1970's, responsible for millions of deaths worldwide (Strassburg 1982). In 1796, Edward Jenner demonstrated for the first time that people inoculated with lesions from the cowpox virus were immune to smallpox and invented the concept of vaccination (Edward Jenner 1802). Eradication of smallpox was achieved using another representative of the orthopoxvirus genus as a live virus vaccine, vaccinia virus (VACV). In 1980 the World Health Organization (WHO) requested all laboratories to destroy variola virus samples to prevent new outbreaks and potential misuse and today only two remaining variola virus samples exist in US and Russian government-controlled laboratories (Damon et al. 2014).

VACV remains a valuable model for studying large pleomorphic DNA viruses in general, and poxviruses in particular, as many features such as core morphology and structural proteins are well conserved within this family. Additionally, VACV also has a high importance as a potential recombinant vaccine vector for various diseases such as influenza or herpes (Smith et al. 1984) and is also used as a protein expression system in biotechnological applications (Belongia and Naleway 2003, Hebben et al. 2007). This is due to its large genome of about 195 kilo base pairs (kbp), which allows packing of large exogenous DNA segments. Moreover, VACV also serves as an ideal model for studying assembly mechanisms and the infectious life cycle of other poxviruses, including variola virus and monkeypox virus.

The most frequently referenced strain in scientific publications is the VACV Western Reserve (WR) strain, known for its protein homology to variola virus, high virulence, and the potential for genetic modifications. For biomedical applications, such as in recombinant vaccination research, the attenuated modified vaccinia Ankara virus (MVA) strain is most commonly used. This strain was developed after 500 passages in chicken embryo fibroblasts and is still used as a smallpox vaccination. MVA is highly attenuated, featuring multiple mutations, and is unable to replicate in human cells for reasons that remain unclear (Blanchard et al. 1998).

Despite extensive research to understand viral assembly mechanisms and key aspects of the viral life cycle, which could improve the use of VACV as a vaccine vector or in biotechnology, several fundamental questions remain unanswered. The recent re-emergence of the monkeypox virus and the global outbreak of Mpox shows that fundamental research in *Poxviridae* is still essential to prevent and treat further outbreaks.

1.1.2   The poxvirus life cycle

Poxviruses have a large host tropism and the transmission occurs either as species-to-species spread or as zoonosis, depending on the virus strain. VACV in particular, spreads from cattle to human or through person-to-person contact and can infect a broad range of host cells (Silva et al. 2010). Upon binding to the cell membrane, VACV can enter the cell via different mechanisms, including direct fusion with the plasma membrane (Carter et al. 2005), a low pH endosomal entry path (Townsley et al. 2006) or actin-dependent micropinocytosis (Mercer and Helenius 2008). After access to the cell cytoplasm, the lateral bodies are disassembled and the intact viral cores travel to a perinuclear site via microtubules, establishing the viral factory. Early genes are transcribed, capped and polyadenylated within the core and the messenger RNA (mRNA) is released into the cytoplasm for translation at cellular polysomes. Those early proteins include proteins necessary for uncoating, transcription factors facilitating the intermediate gene expression and all proteins necessary for viral genome replication. (Mallardo et al. 2001, 2002, Katsafanas and Moss 2007, Greseth and Traktman 2022)

After DNA replication, late gene expression starts, including structural proteins, resulting in the assembly of immature virus particles (IVs). Notably, VACV (as all poxviruses), has a complicated assembly and maturation cycle that results in an immature and various mature virus forms.

Immature virus particle (IVs) are the first virus particles built in the viral factories and are non-infectious. Membrane crescents with integral viral membrane proteins assemble onto D13 protein scaffolds showing a honeycomblike pattern, resulting in spherical particles of ~300 nm diameter (Hyun et al. 2022). The crescent formation is controlled by viral protein kinase F10 which phosphorylates some cellular and viral proteins (Maruri-Avidal et al. 2011) and DNA is encapsidated inside the virus particle (Condit et al. 2006).

Subsequently, IVs are transformed into **Intracellular mature virus particles (MVs)**, the first and largest group of infectious virions produced within the host cell. Upon MV formation, the virion decreases in size compared to IVs and adopts a typical brick-like form with a dumbbell-shaped viral core. Infectious MVs can either emerge after cell lysis or are transported along microtubules to a wrapping site near a nuclear region (McKelvey et al. 2002). At these nuclear proximal wrapping sites MVs are enveloped with an additional double-membrane bilayer, derived from the Golgi apparatus or endosomal compartments (Chung et al. 1998, Sanderson et al. 2000).

These **intracellular enveloped virions (IEV)** are then transported to the cell periphery on microtubules with a kinesin-1-dependent mechanism. After crossing the actin cortex (Arakawa et al. 2007), membrane fusion **releases extracellular enveloped virus (EEV)** from the cell to facilitate long-range spread in the host organism.

Some IEVs remain associated with the plasma membrane as **cell associated enveloped viruses (CEV)** to support superspreading in an actin-dependent manner (Cudmore et al. 1995).
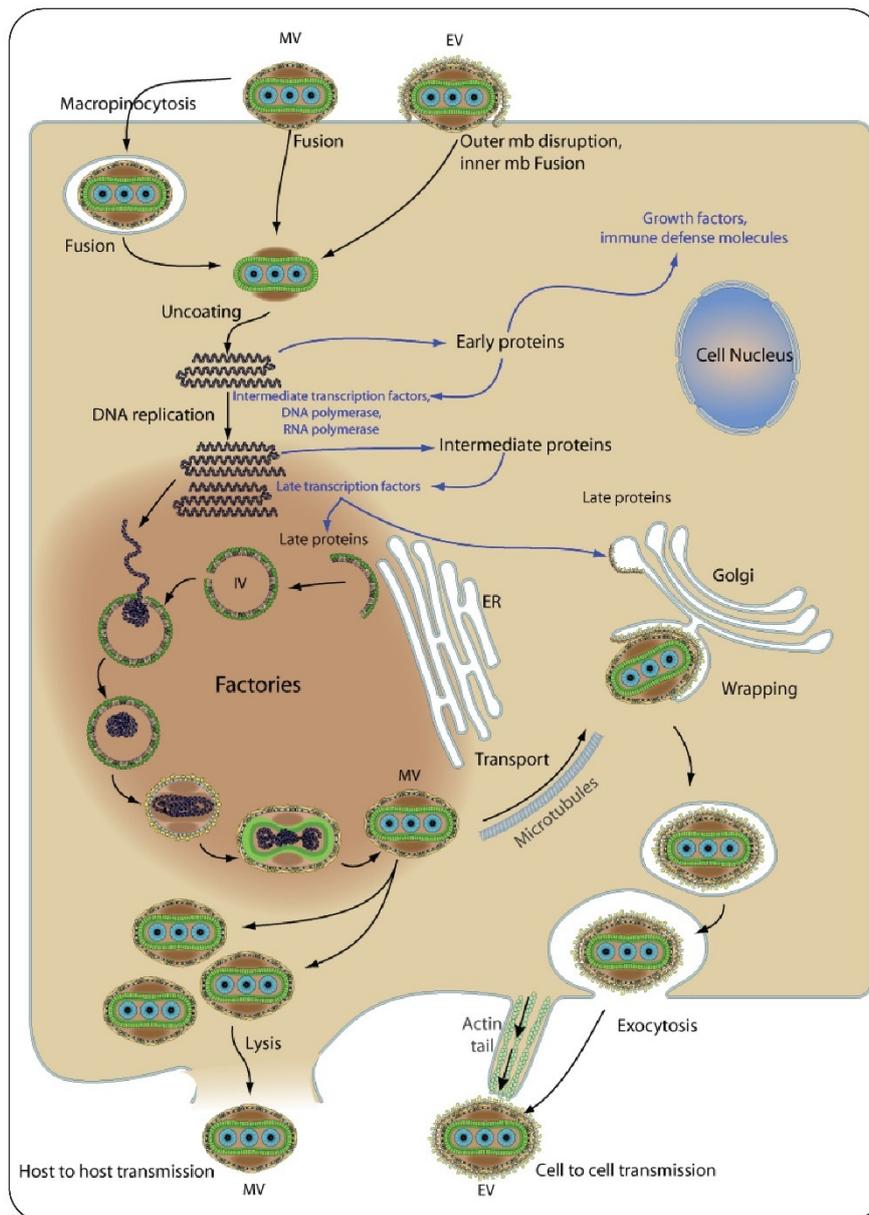
Figure 1: Life cycle of VACV

Graphical Illustration of a VACV infected cell and and the different virions build throughout the life cycle from entry to egress of the virus. The life cycle of VACV involves the entry of virus particles into the cell, either through macropinocytosis or fusion. Once inside the cell, the viral DNA replication begins, followed by the transcription of early and intermediate proteins. This leads to the formation of spherical immature virions (IVs) containing encapsidated DNA. As the viral core condenses, the IVs mature into intracellular mature virions (MVs), which exhibit a brick-like form with a dumbbell-shaped core. The MVs can exit the cell through lysis or be transported to the Golgi apparatus via microtubules, where they acquire an additional membrane layer to become enveloped virions (EVs). Upon exocytosis, the extracellular enveloped virions (EEVs) can either remain cell-associated as cell-associated enveloped virions (CEVs) and move along the cells in an actin-dependent manner, facilitating superspreading.

1.1.3   Poxviruses and the Actin cytoskeleton

As stated in the previous chapter, CEVs contribute to the local spread of infection in an actin-dependent manner (Cudmore et al. 1995). Early electron microscopy studies of VACV infected cells showed microvilli that displayed a virus particle on their tip (Stokes 1976). Actin tail formation in VACV infected cells occurs inside the cell underneath the plasma membrane, triggered by signaling of VACV on top of it The actin tails can extend up to 20μm and project toward uninfected cells, similar to other pathogens such as *Listeria* or *Shigella* (Cudmore et al. 1995, Welch and Way 2013). The VACV structural protein A36 is associated with the outer envelope of IEVs, and after virus release, it accumulates beneath the EEVs. A36 signals back into the cell and becomes phosphorylated by Src and Abl family kinases (Frischknecht et al. 1999, van Eijl et al. 2000, Newsome et al. 2006).
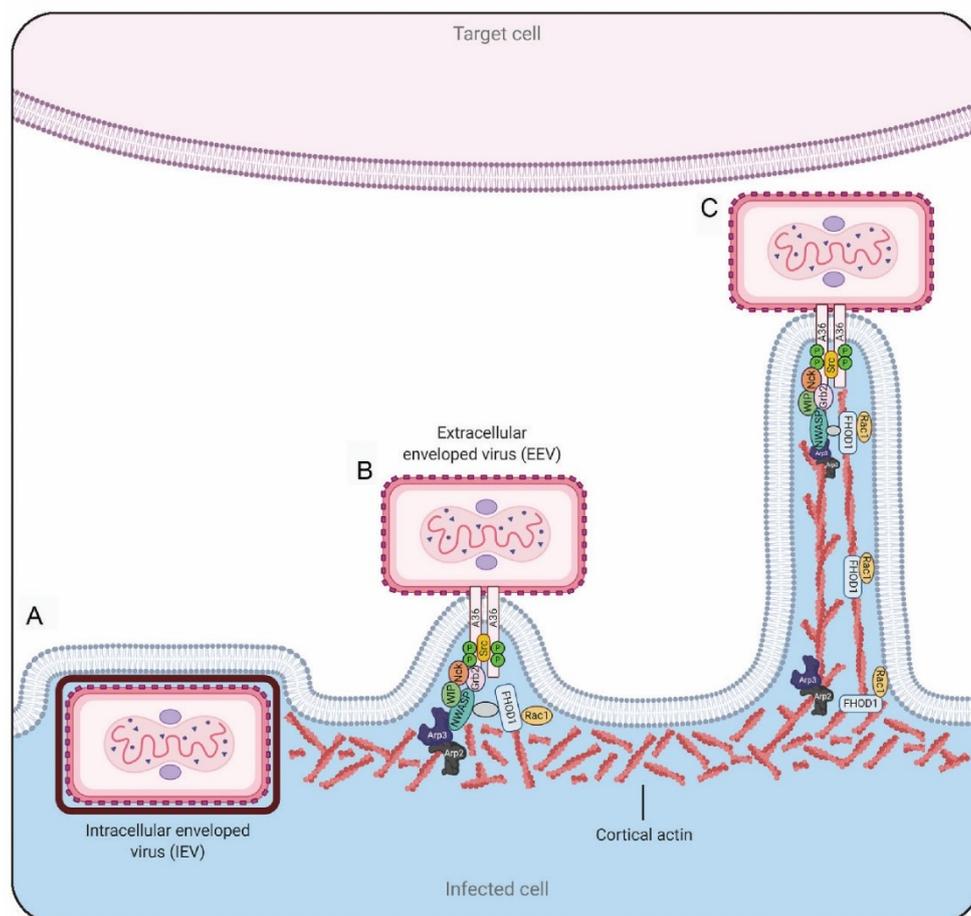
Figure 2: VACV cell-associated enveloped virions form actin tails for protrusion

(A) Intracellular enveloped virions (B) Extracellular enveloped virions signal back into the infected cell with their structural protein A36. A36 proteins are phosphorylated by kinases Src and Abl1 and recruit adapter proteins Grb2, Nck, and effector proteins WASP and WIP. (C) WASP stimulates the nucleation of Arp2/3 complex and Rac1 GTPase activates FHOD1 formin that enables the elongation of the actin tail

Image source:(Cifuentes-Munoz et al. 2020), License: 5835311283914

4

The phosphorylated A36 protein recruits adaptor proteins Grb2, Nck and downstream effectors WASP (Wiscott-Aldrich-Syndrome protein) and WIP (WASP interacting protein) (Frischknecht et al. 1999, Scaplehorn et al. 2002, Donnelly et al. 2013). WASP stimulates the actin nucleation of the Arp2/3 complex (Taylor et al. 2011). Additionally the small GTPase Rac1 activates the FHOD1 formin, which also enables the initiation and elongation of the actin tails (Alvarez and Agaisse 2013, Cifuentes-Munoz et al. 2020) (Figure 2).

The Arp2/3-dependent actin tail propels surface virions away from the host cell towards uninfected cells. To avoid infection of already infected cells, viral protein A56 is anchored in the infected cell's surface and protects from complement attacks, facilitating efficient viral superinfection (Dehaven et al. 2011). Previous studies have highlighted that actin based motility of VACV is a robust model to study actin filament nucleation and polymerization (Cudmore et al. 1995, Abella et al. 2016).

1.2   Structural unknowns of Orthopoxvirus assembly

As noted above, during its life cycle, VACV adopts different virus forms, with varying structure and morphology. They are pleomorphic, asymmetric and of substantial size. Naturally, this poses challenges for structural biology methods like cryo-electron microscopy (cryo-EM) single particle analyses (cryo-SPA) and X-ray crystallography, both suited to solve structures of icosahedral or helical virus particles. In comparison, cryo-EM tomography (cryo-ET) enables visualizing the overall structure of a single virus particle of such a large size in 3D, aiding the understanding of the virions ultrastructural organization. (Obr and Schur 2019).

The first attempts to study VACV using negative staining electron microscopy (EM) provided basic information about virus surface features (Peters and Mueller 1963, Westwood et al. 1964, Wilton et al. 1995, Griffiths et al. 2001). Studies to visualize assemblies within cells performed with thin plastic sectioning EM revealed the abundance of differently structured VACV particles and their subcellular location (Rodríguez et al. 1998). Notably, the varying assemblies of immature and mature forms made it challenging to precisely describe their structural composition, i.e., the exact contribution of specific structural proteins to each virus morphology has remained unclear. For a long time the limited available experimental methods resulted in artifacts caused by damaging sample preparation steps and prevented a more detailed description of virus particles and their components. More native insights could only later be revealed with the introduction of cryo-EM and in particular cryo-ET methods, which allow to study specimens in a near native state via a process called vitrification, i.e. the cryo-fixation of samples without any ice crystal formation (Dubochet et al. 1994, Cyrklaff et al. 2005).

Structural information on IVs

Immature virions (IV) with a diameter of 300 nm are derived from membrane crescents within the viral factory. The IV is suggested to comprise two membrane bilayers with an overall thickness of 10 to 15 nm (Rodríguez et al. 1998, Condit et al. 2006). The inner smooth layer consists of a trilamellar structure and the outer layer is covered with spikes that are regularly spaced (Risco et al. 2002). Upon assembly of the IVs, scaffolding protein D13 exists as individual trimers in the cytoplasm and gets recruited

to the viral membrane through an A17 interaction (Szajner et al. 2005, Bisht et al. 2009). A honeycomb lattice of pseudohexagonal D13 trimers is built and induces the curvature of the spherical-shaped IVs (Hyun et al. 2011, 2022). The origin of the membrane crescent has been controversial for a long time. Origination of membranes from cellular organelles (Risco et al. 2002) has been favored as hypothesis over *de novo* membrane formation theory (Dales and Mosbach 1968), as studies suggested that crescents are derived from preexisting cisternae and collapse through the action of virus proteins (Sodeik et al. 1993). Recent thin section EM and ET studies of VACV deletion mutants of viral membrane assembly complex proteins that are necessary to form IVs, showed connections between the viral membrane and the endoplasmatic reticulum, proving the origin of the viral membranes (Weisberg et al. 2017). DNA gets encapsidated during IV formation when the viral membrane crescent and D13 form the spherical particle. In EM micrographs local electron densities in the shape of spherical granular inclusions or paracrystalline arrays were detected inside IVs and identified as the DNA containing compartment (Dales 1963, Ericsson et al. 1995). During maturation the IV undergoes a substantial structural change where the spherical IV particle gets condensed into a brick-shaped MV particle.

Structural information on MVs

The MV is the structurally most studied virion of the VACV life cycle. MVs are usually encountered outside the viral factories (Sodeik et al. 1993, Ericsson et al. 1997) and are brick-shaped with approximate dimensions of 360 x 250 x 220 nm (Cyrklaff et al. 2005, Hernandez-Gonzalez et al. 2023). The transition from IVs to MVs is rapid (hence aggravating experiments to understand the exact transition mechanism) and includes proteolytic cleavage of the structural core proteins through viral protease I7 and the formation of disulfide bonds that facilitate the reorganization to MVs (Senkevich et al. 2002, Ansarah-Sobrinho and Moss 2004). The virus particle loses around 50% of its volume during maturation, and the viral core condenses to a dumbbell-shaped form flanked by lateral bodies (Hernandez-Gonzalez et al. 2023) (Figure 3).

The D13 hexameric lattice of IVs is lost throughout this process, and the viral core is condensed into a dumbbell-shaped form containing the viral genome (Condit et al. 2006, Hernandez-Gonzalez et al. 2023). Despite this knowledge, the morphogenesis of MVs is still incompletely understood and different models have been proposed. One of these models is based on cryo-ET studies from 2005 (Cyrklaff et al. 2005). It suggests that the IVs are surrounded by a double lipid bilayer and during maturation one lipid bilayer collapses around the encapsidated DNA, building the viral core, resulting in an outer layer also surrounded by a single membrane bilayer. Recent studies confirmed that one membrane layer of the IV forms the outer membrane of the MVs, but it is still unknown how the viral core wall is formed (Hernandez-Gonzalez et al. 2023). Another model is based on very early negative stain EM studies and proposes a single membrane bilayer and no membrane around the core. This model also suggests that crescents arise *de novo*, which later studies disproved (Dales and Mosbach 1968, Wilton et al. 1995, Weisberg et al. 2017). Overall, no clear consensus on the morphogenesis of MV has been reached, highlighting the importance of addressing this question further. Controlled degradation and immuno-localization with negative stain EM revealed that the viral core is flanked by lateral bodies of amorphous structure composed of proteins of unknown function (Easterbrook 1966, Cyrklaff et al. 2005, Condit et al. 2006). The MV is surrounded by a 5-6 nm lipidic membrane with

non periodic extensions of 3 to 6 nm on the outside, that could be part of the receptor binding components and a smooth inner layer (Cyrklaff et al. 2005). More recent studies described the outer layer of 6-7nm to be more continuous (Hernandez-Gonzalez et al. 2023). The differences could be due to different sample preparations and visualization techniques (Figure 3).

1.2.1   The poxvirus core

The VACV core is a key part of all infectious poxvirus forms and plays a central role in the viral lifecycle by securely transporting the viral genome and necessary accessory proteins to newly infected cells. The dumbbell-shaped viral core displays an inner cavity that is suggested to be filled with electron-dense coils of DNA-protein complexes (Condit et al. 2006). Visualization of viral core substructures by *in situ* atomic force microscopy indicated that the association of linear viral DNA with segmented proteins form a 16 nm diameter filament with a helical surface topography (Malkin et al. 2003). These described filaments are enclosed in 30 - 40 nm tubules condensed in the core. Previous cryo-EM studies could not confirm the formation of enclosed tubules, they show that the DNA is condensed and accumulated to the inner core wall (Cyrklaff et al. 2005, Hernandez-Gonzalez et al. 2023).

The viral core is surrounded by two layers with an overall thickness of 18-19nm: a discontinuous outer palisade layer described to consist of what appeared to be T-shaped spikes and an inner core wall. These spikes of the palisade layer were reported to have dimensions of 5 nm width and 8-12.5 nm height and to be arranged in hexagonal crystallites and the inner core wall to have a thickness of 3-4 nm (Dubochet et al. 1994, Cyrklaff et al. 2005, Hernandez-Gonzalez et al. 2023). The discontinuous outer layer is proposed to be built by side interaction of the spike proteins as they are anchored in the inner core wall and attached to lateral bodies is possible.(Cyrklaff et al. 2005)

Further insights into core wall composition suggested that pore-like structures exist in the viral core. The pores are embedded in the viral core wall with extended protrusions on the exterior side (Figure 3). These findings were provided by early cryo-ET studies of whole VACV (Cyrklaff et al. 2005), followed by negative stain experiments on stripped VACV cores (Moussatche and Condit 2015). However, none of these studies could provide a conclusive answer to the potential role and composition of the viral core pores. It was suggested that these pores could be responsible for extrusion of viral mRNA during viral transcription (Cyrklaff et al. 2005). Alternatively, pores could be involved in import and export of molecules to increase or decrease volume, when viral cores increase in size after being internalized into the cell or when IVs condense during their transition to MVs (Heuser 2005, Chichón et al. 2009). During the contraction step it would be possible that molecules get excluded and the MVs are impermeable. However, the pore structure and its protein composition are still unknown; therefore, no clear assumption about its function can be made.(Peters and Mueller 1963, Heuser 2005, Chichón et al. 2009) This highlights that despite significant efforts in examining the VACV structure, several fundamental questions such as the function and the protein composition of the pores as well as overall MV core wall composition are still open.

Figure 3: Schematic structure of VACV MVs and the composition of the core wall

Schematic depiction of an intact mature virion with annotated palisade layer, inner core wall, pore-like structure, and lateral bodies in comparison to the isolated viral core. (left). Suggested protein composition of the VACV core wall with A4 on the exterior, A10 in the middle, and A3 on the interior side. L4 is displayed as DNA binding protein inside the core. Protease I7 is responsible for the proteolytic cleavage of proteins A10, A3 and L4. (right) The question mark and yellow background highlight the uncertainties in the precise composition of the core wall, indicating areas that require further investigation.

Source: adapted from Viralzone, Swiss Institute of Bioinformatics (Hulo et al. 2011) (https://viralzone.expasy.org/4399 and Datler et al. 2024.

License: CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/)

1.2.2 Putative poxvirus core proteins

Despite not knowing exactly which proteins form the core layers, putative candidates for major core proteins have been described in literature using indirect experimental evidence. Protein names varied in different studies and the genes for these putative core proteins are referred to as A4L, A10L, A3L and L4R. All except A4L encode for precursor proteins p4a (A10L), p4b (A3L), and VP8 (L4R). These proteins are cleaved by protease I7 during maturation into A10 (p4a residues 1-614), 23k (p4a residues 698-891), A3 (p4b residues 62-641), and L4 (VP8 residues 33-291) (Ansarah-Sobrinho and Moss 2004). The protein A4 (also known as p39) does not undergo proteolytic cleavage. For simplicity and to adjust to the currently accepted nomenclature, I will use the protein names A10, A3, A4, 23K and L4 from now on. (Figure 4)

In previous studies using spatial proteomics, researchers identified the proteins present in the viral core and the lateral bodies. Western blot analysis of SDS-treated MVs revealed A3, A10, and L4 as major core proteins (Ichihashi et al. 1984, 1994). Gel-free liquid chromatography and tandem mass spectroscopy confirmed these findings and showed that A10, A3, and A4 are the most abundant viral proteins, with a ratio of 1:1:10 (Chung et al. 2006). Mass spectrometry of viral cores purified with NP-40 and DTT indicated the presence of A10 and A3 in the viral core fraction and A4 in the combined fraction of core and lateral bodies, suggesting that A4 is located at the outer layer of the core wall (Jensen et al. 1996, Bidgood et al. 2022).

These results align with immunogold labeling studies of purified viral cores, which detected A10 and A4 at the surface of the core wall and suggested that A4 forms a palisade-like structure (Cudmore et al. 1996, Roos et al. 1996, Pedersen et al. 2000, Moussatche and Condit 2015). Protein A3 has been suggested to be located at the inner part of the core wall, as immunolabeling was only feasible after the cores were broken due to hypertonic shock (Moussatche and Condit 2015). Immunolabeling of cryosections and broken viral cores indicate that L4 is located at the inner core wall and is a major DNA binding protein, as previously described in *in vitro* studies (Bayliss and Smith 1997, Pedersen et al. 2000, Moussatche and Condit 2015). A cross-linking mass spectrometry (XL-MS) paper described interactions between A10 and A4, 23k and A4, and A3 with 23k (Mirzakhanyan and Gershon 2019).

These findings suggested a protein core wall model as shown in Figure 4, where A4 builds the outer layer of the viral core wall, forming the palisade layer. A10 is situated in the middle and A3 is located on the inside of the viral core wall. L4 was suggested to be found on the interior side of the viral core, as a DNA binding protein.

Despite decades of research, our understanding of the structural determinants and proteins that form the virus core in MVs remains limited. While these studies have suggested potential major core proteins, the structural composition of the MV core wall is still unknown. Furthermore, there are no experimentally derived structures of the putative core protein candidates, neither *in situ* or *in vitro*, A10, 23k, A3, A4 and L4. This underscores the need for further research in this area, as we strive to unravel the complexities of the virus core in MVs.

Putative major structural core proteins & their precursor proteins

a

b

Protein (UniProt name)	Residues	MW (kDa)	UniProt ID
A10 (4a)	1 - 614	62	P16715
A3 (4b)	62 - 643	60	P06440
23K (cleavage product of A10 precursor)	698 - 891	23	P16715
A4 (p39)	1-281	31	P29191
L4 (VP8)	33 -251	25	P03295
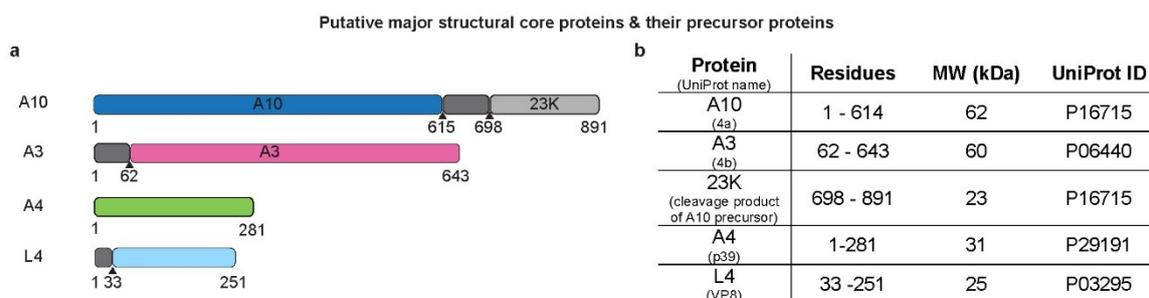
Figure 4: Putative structural core proteins

(a) Schematic of the protein domain architecture and cleavage sites of A10, A3, A4 and L4. (b) Protein names and Uniprot ID of proteins A10 (4a), A3 (4b), 23K (cleavage product of A10 precursor), A4 (p39) and L4 (VP8) and their molecular weight and residues. Adapted from Datler et al., 2024.

1.3   Structural characterization of the poxvirus core

Poxviruses are very large and complex, featuring a multilayered architecture that presents significant challenges for structural analysis. As discussed in previous chapters, early studies using various EM techniques revealed that VACV MVs have a distinctive barrel-shaped structure characterized by an internal viral core surrounded by an outer membrane layer. During the maturation process, the DNA containing viral core, condenses into a dumbbell-shaped form and is flanked by protein-rich lateral bodies (Cyrklaff et al. 2005, Moussatche and Condit 2015). Subsequent cryo-EM studies confirmed these observations and identified a regular palisade layer surrounding the viral core, with pore-like structures traversing the core wall (Cyrklaff et al. 2005, Hernandez-Gonzalez et al. 2023). More recent in *situ* cryo-ET studies further detailed this palisade layer, showing that it is composed of a pseudohexagonal lattice of trimers consistently present in all infectious virions (Hernandez-Gonzalez et al. 2023). Despite these advances, structural studies have so far been unable to determine the precise protein composition of the viral core due to limited resolution and the absence of available protein structures for accurate fitting within the observed densities. This chapter focuses on the various methods used to elucidate the structures of pleomorphic viruses like VACV and their technical evolution over time, highlighting recent advancements, including state-of-the-art techniques, computational approaches, and artificial intelligence (AI) based technologies.

1.3.1   Electron microscopy studies of the poxvirus core

EM is a powerful imaging technique employing an electron beam as an illumination source to acquire high-resolution images of biological and non-biological samples. Two main types of EM are transmission electron microscopy (TEM) and scanning electron microscopy (SEM). SEM scans the surface of a specimen, providing detailed images of the specimen surface, while TEM transmits electrons through thin samples and produces high-resolution 2D projection images. TEM is used to display the internal structures of thin-sectioned samples, isolated viruses and single particles. Overall, it is an essential tool for studying cells, tissues, viruses and macromolecular complexes at a nanometer scale. Preparing samples for conventional TEM and SEM can alter the structure of biological specimens as they must be dehydrated and stained. Different

techniques, such as freeze substitution, thin-sectioning, negative staining, and immunogold labeling, can enhance contrast and reveal structural details. The EM sample preparation can affect the specimen's integrity, leading to a loss of its native state. One effective solution to this problem is to utilize TEM under cryogenic conditions. In this method, the samples are vitrified in thin vitreous ice, free of crystals, which helps preserve their native state. Cryo-EM is a cutting-edge technique used to reveal the structures of protein complexes in a near-native, hydrated state. Thin specimens, such as viruses, cells, proteins, or molecular complexes, are rapidly plunge-frozen on cryo-EM grids in liquid ethane at approximately -180° C, effectively preventing ice crystal formation and preserving the native state of the sample without chemical modifications.

Single particle cryo-electron microscopy

Single particle cryo-EM (cryo-SPA) is a method for determining high-resolution structures of proteins and molecular complexes. In contrast to cryo-ET, where a tilt series is acquired to produce a 3D structure, cryo-SPA collects numerous 2D images of proteins, or other features of interest, that are differently oriented within the sample. Then, these 2D images, displaying variant orientations of the particles, are computationally processed and aligned to reconstruct a detailed 3D structure. The technique is especially effective for isolated or purified complexes, making it ideal for detailed structural studies of ribosomes, membrane proteins, enzymes, and large macromolecular assemblies. Cryo-SPA can achieve near-atomic resolution, revealing intricate details such as side-chain positioning, conformational states, and binding sites. (Figure 5c)

Cryo-Electron Tomography

Cryo-ET tomograms are obtained by acquiring a series of 2D micrographs at various tilt angles, typically ranging from -60° to +60° with increments of 2 to 3 degrees. After acquisition, the micrographs are aligned and computationally reconstructed into a 3D tomogram, revealing the structural features of the sample at the resolution of several nanometers. Regions of interest within these tomograms can be identified manually or through automated template- or shape-recognition algorithms. Once identified, subtomograms are extracted and aligned relative to a predefined reference. Through iterative alignment and refinement, a higher-resolution 3D structure of the target feature can be resolved. (Figure 5a, b)

In most recent cryo-ET studies on VACV, only low resolutions could be achieved due to methodological limitations. Therefore, no detailed conclusions could be made about the composition and function of the viral core (Cyrklaff et al. 2005, Moussatche and Condit 2015). In recent years, cryo-ET underwent major improvements, allowing the description of capsid assemblies in pleomorphic viruses like retroviruses (Schur et al. 2016), influenza virus (Peukes et al. 2020) and filovirus (Wan et al. 2017, Watanabe et al. 2024) at sub nanometer or even higher resolution. One main advancement was the implementation of new direct electron detection cameras to achieve a higher signal-to-noise ratio (SNR).
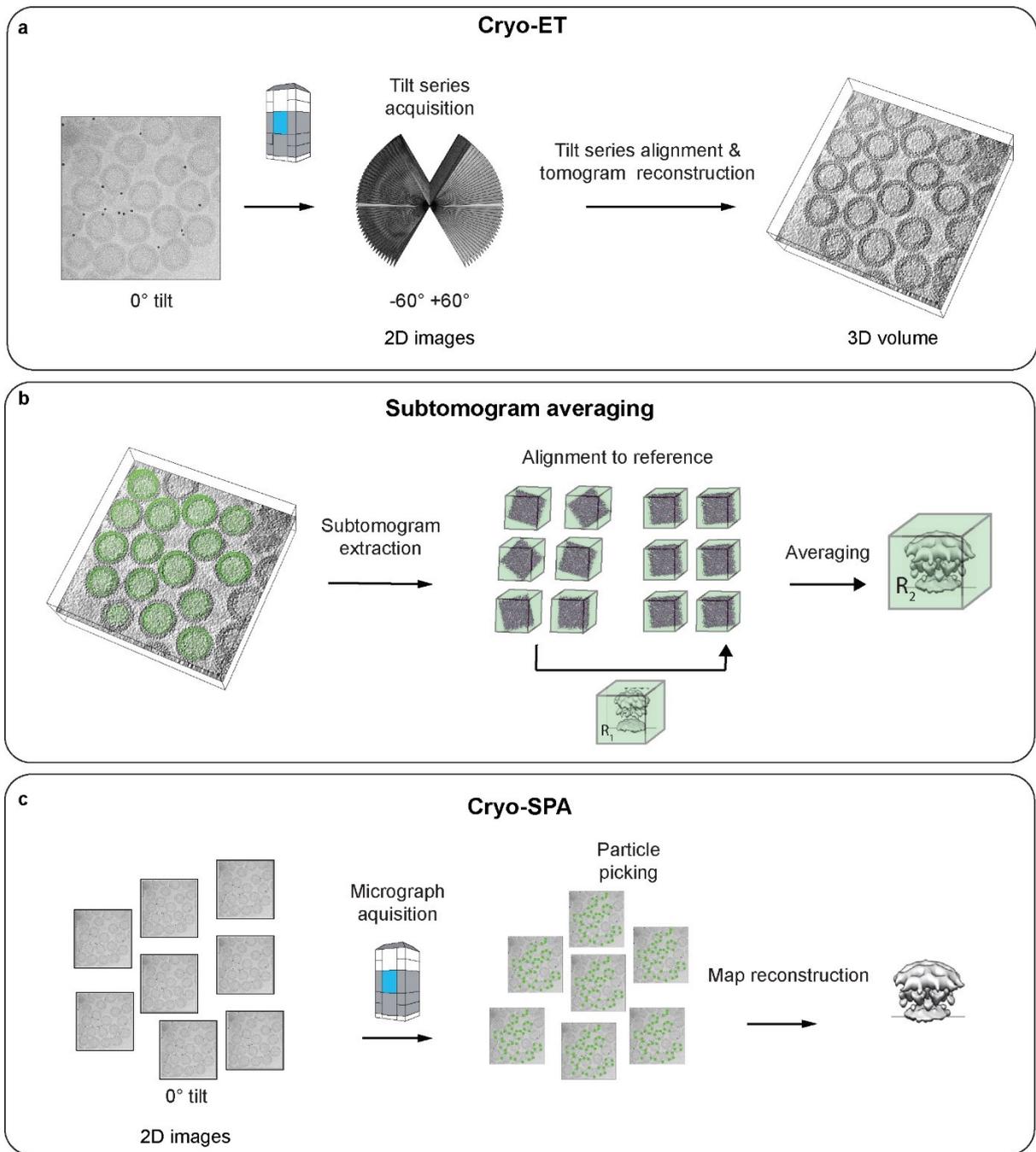
Figure 5: Different cryo-EM techniques and their workflow

(a) In cryo-ET tilt series are acquired, which are aligned and reconstructed into a tomogram. (b) Particles are identified and picked within the tomogram and subtomograms are extracted. Subtomogram averaging aligns and averages them to a reference, resulting in a 3D structure. (c) Cryo-SPA acquires 2D images of particles embedded in a thin vitreous layer in various orientations. After particles are picked their computational alignment generates a 3D structure.

Image source: adapted from (Obr and Schur 2019), License: 5912070018902

When comparing room-temperature thin-section EM (RT-EM) (for both 2D acquisition and tomography), cryo-ET, and cryo-SPA, each method offers distinct advantages and limitations. Within the overall sample, RT-EM and cryo-ET provide better contextualization and preserve the spatial relationships of proteins and structures. However, cryo-ET maintains the highest level of nativity, as samples are usually preserved in their natural environment. Sample preparation poses unique challenges for each technique. RT-EM involves chemical fixation and dehydration steps that can alter sample structure, while cryo-ET demands thin samples in a vitrified state to maintain native-like conditions. For cryo-SPA, achieving a high concentration of the protein of interest without preferred orientations is critical to ensure successful data acquisition.

The image processing capabilities vary tremendously as RT-EM is primarily limited to segmentation, whereas cryo-ET and cryo-SPA benefit from a wide range of advanced software tools for image alignment, reconstruction, and analysis. In summary, each technique plays a vital role in unraveling the structures of pleomorphic viruses, and combining these methods can provide a comprehensive understanding by leveraging their respective strengths. (Table 1)

Table 1 Comparison of different electron microscopy techniques

	Room-temperature thin section EM	Cryo-ET	Cryo-SPA
Context	✓	✓	x
Nativity	x	✓	~
Resolution	nm	nm/Å	Å
Sample preparation	Complex	Complex	Complex
Image processing	Limited	Advanced	Advanced

1.3.2 Artificial Intelligence (AI) based approaches to study poxviruses

No experimental structures are available for most poxvirus proteins, particularly the putative core proteins A10, 23K, A3, A4, and L4. Additionally, the lack of homologues makes homology modeling unfeasible, leaving a gap in structural information. AI-based approaches like AlphaFold are good opportunities to computationally predict protein structures and interpret experimental data.

AlphaFold - new Technologies for structural studies

Determining the three-dimensional structure of proteins is essential to elucidate their function and interactions within biological processes. X-ray Crystallography, Nuclear magnetic resonance (NMR) spectroscopy and cryo-EM are standard methods to determine protein structure. However, these techniques require specialized instrumentation and high-quality samples, often presenting a significant bottleneck. Therefore, the invention of AI-based programs for protein structure prediction is valuable for interpreting experimental data when the data lacks sufficient resolution to construct a *de novo* model.

The field of protein structure prediction already started in 1961 when the 'Anfinsen Dogma' suggested that the 3D structure of a protein is only determined by its amino acids (Anfinsen et al. 1961). From that moment, scientists tried to develop algorithms to predict 3D protein structures from amino acid sequences accurately. Since 1994, a competition organized by the Critical Assessment of Structural Prediction (CASP) community has taken place biennially where an amino acid sequence of a known but not yet published protein is released, and scientists can test their 3D structure that was predicted by their developed algorithm objectively (Lupas et al. 2021). Since then, many algorithms have been developed, and they can be classified into homology modeling, *de novo* modeling and machine learning (ML) based modeling. The most promising ML modeling in recent years is a deep learning algorithm that is the base for programs like RoseTTAFold (Baek et al. 2021), AphaFold (Jumper et al. 2021) and ESMFold (Lin et al. 2023). In 2020, the 14th CASP competition was won by the machine learning-based algorithm AlphaFold 2.0 (AF2) developed by DeepMind (Jumper et al. 2021, Kryshtafovych et al. 2021).

AF2 (Jumper et al. 2021) uses protein databases to compare amino acid sequences and identify conserved regions crucial for the protein structure. The program uses multiple sequence alignments (MSA) of homologous proteins and highlights their similarities and differences. A specialized neural network architecture called "Evoformer" processes the MSA and the pairwise amino acid interactions to capture relationships and build contextual information on how the protein should be folded. This information is combined in a structure module that refines the 3D coordinates and iteratively adjusts them to match the structural constraints. In the end, each region of the structure gets a confidence score, with the predicted local distance difference test (pLDDT) assigned, indicating the reliability of the structure. AF2 can also fold single and multi-protein complexes where multiple amino acid sequences of the same or a different protein are provided to build a multimer complex (Yang et al. 2023).

As of 2024, over 214 million proteins were already predicted by AF2 and significantly advanced research in the fields of biology and medicine (Varadi et al. 2024). The predictions are accessible through the AlphaFold Protein Structure Database, a collaboration between DeepMind and the European Bioinformatics Institute (EMBL-BI). One outstanding example of combining AF2 with experimental and computational techniques is the detailed model of the human nuclear pore complex architecture (Mosalaganti et al. 2022). Another example is the structure of VACV DNA helicase D5, where cryo-EM and AF2 predictions were integrated to construct a comprehensive structural model (Hutin et al. 2022).

The limitations of AI-based structural prediction include their dependency on available experimental databases and their reliance on established patterns. AF2 does not consider physical and chemical factors such as pH, temperature, and stoichiometry. While the algorithm can predict protein structures, it does not provide insights into protein function, dynamics, or how proteins behave within a complex cellular environment. The reliance on evolutionary data, like MSA, for accurate predictions is less effective for proteins with few homologous sequences, which is often true for novel or species-specific proteins. Furthermore, flexible and disordered regions are frequently predicted inaccurately.

The recently published AlphaFold 3 has addressed many previous limitations. This algorithm can now predict the joint structures of complexes that include proteins, small ligands, RNA, DNA, ions, and modified residues. Additionally, ligand binding predictions have been enhanced, allowing tasks that previously required multiple programs to be consolidated into a single tool. This improves the prediction of protein-ligand interaction, protein-nucleic acid interaction, and antibody-antigen prediction, which is valuable for a wide field of applications in biology, medicine and drug design. (Abramson et al. 2024)

Neural-network assisted restoration of missing information in cryo-ET

The missing wedge problem along with its anisotropy represents a significant challenge in the visualization and interpretation of tomograms. The tilt range of a tomogram usually does not exceed +60°/- 60° as higher tilts result in thicker samples and a poor SNR. This incomplete angle sampling creates a "missing wedge" that causes anisotropy along the optical axis in real space, resulting in missing information in Fourier space (Figure 6). Several strategies exist to address this issue. One method involves using a dual tilt axis acquisition approach, where the stage can also be rotated in-plane by 90° to acquire a second tilt series, resulting in a "missing pyramid" (Mastronarde 1997) This feature requires a special stage that can be integrated into some cryo-EM microscopes. However, it requires the acquisition of more tilt angles, which results in spending the same or a higher amount of electrons on a larger number of tilts, increasing the risk of sample damage and poor SNR.

Algorithms that use prior assumptions, such as density positivity or solvent flatness, to explain the structural features of reconstructed tomograms are another option for compensating for the missing wedge. Such assumptions have limited information and often do not succeed in recovering the missing wedge, as it is a complex biological system (Deng et al. 2016, Yan et al. 2019, Zhai et al. 2020). An recent development is the AI-based deep learning software IsoNet that is able to restore the missing wedge information and improves the SNR of electron tomograms (Liu et al. 2022).
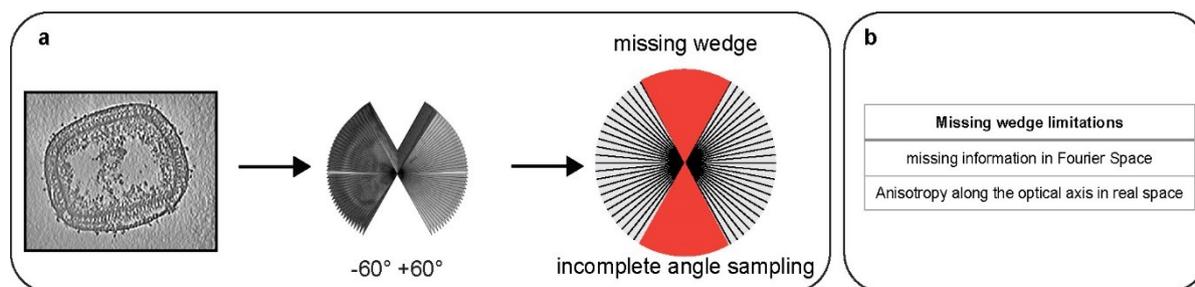
Figure 6: The missing wedge problem.

(a) Tomograms are acquired by collecting a tilt series of images, which are then computationally reconstructed into 3D volumes. The incomplete angular sampling at high tilts results in a region of missing data, known as the "missing wedge". (b) Limitations of the "missing wedge".

IsoNet employs a deep learning approach, utilizing a convolutional neural network (CNN) that iteratively learns from raw tomograms to restore the missing information. The software analyzes the provided tomograms to predict features that might fill the missing wedge, performing exceptionally well with spherical objects and those with uniform features across all orientations, such as viruses. Since side views are similar to top and bottom views due to the sample's uniformity, similar features can be expected to be found in the missing wedge region (Liu et al. 2022).

To benchmark the technique, IsoNet was tested on different biological samples. The correction of tomograms of HIV using IsoNet revealed that the missing wedge information could be effectively restored, resulting in a more spherical appearance of the virus with significantly improved contrast. In the case of *Trypanosoma brucei* flagella, the isotropic correction provided enhanced contrast in all three dimensions, allowing for a detailed visualization of the complex structure of the cellular organelle. A more complex and more extensive example involved tomograms of synapses in hippocampal neurons. Because of the increased thickness of the sample and its very low SNR, isotropic reconstruction proved to be more challenging. Nevertheless, the resulting tomograms displayed good structural integrity and substantially higher contrast.(Liu et al. 2022)

IsoNet is a useful tool for enhancing contrast and correcting for the missing wedge in cryo-ET. The program helps visualize tomograms in all three dimensions, with lower SNR and a higher contrast, thereby improving our understanding of complex biological structures. Additionally, IsoNet facilitates further processing by allowing researchers to select the features of interest for subtomogram averaging. However, it's important to approach the results with caution, as their accuracy depends heavily on the quality of the input training data.

AF2 enables protein structure prediction, and IsoNet enhances the resolution and interpretability of cryo-ET data. Together, these AI-based approaches are helpful tools to analyze and interpret experimental data and can help to gain valuable insights into the complex core architecture of poxviruses.

2 Aims and Objectives

Obtaining more knowledge on poxvirus morphogenesis could reveal a deeper insight into assembly mechanisms of these viruses and other large nucleocytoplasmic DNA viruses in general. As outlined in the introduction, the structural assembly of MVs is still widely unknown despite considerable efforts. Fundamental open questions include:

- Which structural proteins comprise the core layer, and where are they located in the core wall? Specifically which protein forms the palisade layer and the inner core wall?

- Which structural features comprise the viral core? Do viral core pores exist and what is the molecular structure and the functional state of these pores?

- Why does the virus core have a dumbbell-shaped form, and why is it so versatile throughout maturation of the virus? What defines the shape of the virus core?

For this thesis, I reason that these questions can be effectively addressed by revisiting VACV MVs using the latest cryo-EM and image processing developments, combined with optimized virus purification strategies to obtain viruses with high yields and purity.

This should allow me to achieve my **main aim, in which I want to substantially extend the available structural description of MVs to provide a complete MV core layer model.**

To this end, the work of this thesis is defined via three objectives:

Objective 1 - To establish purification protocols for high-resolution structural analysis of complete VACV virus particles and isolated virus cores

Using the VACV WR strain allows to draw general conclusions about poxvirus assembly features. Performing structural analysis via cryo-EM or cryo-ET requires purifying intact VACV MVs in high concentrations. Developing an optimized virus purification protocol to obtain high titers of intact viruses and an optimized vitrification protocol of VACV MVs, will ensure an optimal concentration, density and overall quality for subsequent cryo-EM and cryo-ET studies. Additionally, establishing a virus core purification protocol, where the outer membrane and the lateral bodies are stripped off the virus, will allow more detailed insights into viral core composition. A previous paper from Dubochet (Dubochet et al. 1994) showed that purified VACV cores are clean, and many structural details are visible in cryo-EM. Furthermore, this paper showed that single proteins are shed off purified viral cores, potentially making them amenable to SPA cryo-EM.

Objective 2- To elucidate VACV core morphology and the structural determination of core layers using cryo-ET

In this objective, optimizing a cryo-ET and subtomogram averaging workflow will allow in-depth structural analysis and give a contextual overview of the core morphology and the viral core proteins within intact VACV MVs. Structures such as the palisade layer, the inner core wall and the viral core pore as well as interactions between different structural entities will be visualized.

Objective 3 – Assigning molecular identities to the poxvirus core layer using multimodal cryo-EM and an AI assisted workflow

Identification of the major core proteins of VACV MVs and their exact location within the virus core wall will be performed with a multimodal approach of cryo-ET, cryo-EM with an AI assisted workflow. Cryo-ET and cryo-SPA of purified VACV MVs and their isolated cores as established in objective 2 should allow obtaining a high-resolution structure of the palisade layer, the inner core wall and potential viral core pores, and contextual knowledge about their location. Putative structural core proteins are already known from the literature, and an AI assisted workflow with Alphafold will be used to predict the protein structures. Subsequently, the predicted structures will be compared with experimentally derived structural densities to accurately identify the viral core proteins and reveal the correct core composition model. Finally, the structure will be compared with other Poxviridae family members to better understand protein conservation. Furthermore, interaction analysis will elucidate how different proteins interact with each other to obtain a more comprehensive understanding.

3 Material and Methods

The material and methods section has been described in the paper "Multi-modal cryo-EM reveals trimers of protein A10 to form the palisade layer in poxvirus cores" (Datler et al. 2024) and is reproduced here in adapted form by myself. The original draft has been initially written by myself, except for Chapter 3.6.6 and Chapter 3.8, which was originally drafted by Jesse Hansen, and Chapter 3.9, which was entirely written by Nicholas Armel.

3.1 VACV propagation

VACV Western Reserve (WR) sample was kindly received from Andreas Bergthaler (Medical University of Vienna & Center of Molecular Medicine) and propagated in HeLa cells that were kindly provided by Michael Sixt (Institute of Science and Technology Austria) and originally obtained from ATCC.

HeLa cells were cultured in DMEM GlutaMAX (Thermo Fisher Scientific, no. 31966047), supplemented with 10% (vol/vol) Fetal Bovine Serum (FBS) (Thermo Fisher Scientific, no. 10270106) and 1% (vol/vol) penicillin-streptomycin (Thermo Fisher Scientific, no. 15070063) in T-175 cell culture flasks (Corning, no. CLS431080) until they reached 80% confluency. Before infection, VACV was trypsinized with a 1:1 dilution of 0.25% Trypsin (Thermo Fisher Scientific, no. 25200056) for 30 minutes at 37 °C and then diluted in 2 ml infection medium containing DMEM, 2.5% FBS and 1% penicillin-streptomycin at an MOI of 1.5.

HeLa cells were washed once with PBS, by aspirating the culture medium and adding PBS. PBS was aspirated and the previously prepared 2 ml infection medium containing the virus at an MOI of 1.5 was added for 2 h at 37 °C and 5% CO_2. Culture flasks were manually tilted every 30 minutes to ensure equal distribution of the virus. 2 h after infection 33 ml of infection medium was added and cells were incubated for 3 days at 37 °C and 5% CO_2 until a cytopathic effect (CPE) was visible. The cells were harvested by scraping and the cell suspension was centrifuged at 1,200g for 10 minutes at 4 °C and the cell pellet was resuspended in 500 µl 10 mM Tris-HCl, pH 9 (Carl Roth, no. 9090.3) buffer. Cell pellets were stored at -80 °C until further use and underwent three times a freeze and thaw cycle and were vortexed after each thawing process to manually shear the cells. The cell suspension was centrifuged for 5 minutes at 300g and 4 °C and the supernatant was kept for further use. The cell pellet was resuspended again in 500 µl 10 mM Tris-HCl, pH 9 and centrifuged 5 minutes, 300g and 4 °C. The supernatants were pooled and were either aliquoted in 50 µl batches and stored at -80 °C until further use for virus propagation, or used to purify MVs for electron microscopy as described below.

3.2 VACV WR purification

The pooled supernatant was applied on a 6 ml sucrose cushion containing 36% Sucrose (Sigma-Aldrich, no. 84100) in 10 mM Tris-HCl buffer, pH 9 in small centrifuge tubes (Thermo Fisher Scientific, Thin-Walled WX, no. 03699). The tubes were filled

up with 10 mM Tris-HCl buffer and centrifuged at 32,900g and 4 °C for 80 minutes in an ultracentrifuge (Sorvall, WX100+, Rotor TH-641). The supernatant was aspirated and the pellet containing the MVs was resuspended in 500 µl 1 mM Tris-HCl, pH 9 and applied on a sucrose gradient (40%, 36%, 32%, 28%, 24% sucrose in 1 mM Tris-HCl buffer (2 ml each)) and centrifuged with 26,000g and 4 °C for 50 minutes.

The band containing the MVs, visible as a milky band, was collected and applied to a fresh tube with 10 ml 1 mM Tris-HCl buffer and centrifuged at 15,000g and 4 °C for 30 minutes to sediment the virus. The pellet was dissolved in 100 µl 1 mM Tris-HCl buffer, aliquoted, and frozen at -80 °C until further use.

3.3   Purification of VACV core

Samples were treated as described above until the last centrifugation step. The cell pellet was dissolved in 500 µl core stripping buffer containing 0.1% NP-40 (Thermo Fisher Scientific, no.85124), 50 mM DTT (Carl Roth, no. 69083), 50 mM Tris-HCl pH 9 and 2U DNAse (Promega, no. M6101) to shed the outer membrane of VACV, and incubated for 10 minutes at room temperature. The solution was applied on a 2 ml 24% sucrose cushion in 1 mM Tris-HCl buffer, pH 9, and filled up with 1 mM Tris-HCl buffer, pH 9. The sample was centrifuged at 20,000g and 4 °C for 30 minutes and the pellet was collected in 500 µl 1 mM Tris-HCl, pH 9. For sedimentation of the viral cores, the sample was centrifuged again at 15,000g and 4 °C for 30 minutes. The pellet was resuspended in 50 µl 1 mM Tris-HCl, pH 9, and stored in aliquots at -80 °C until further use. Cryo-SPA samples were prepared with four times more virus and 3 M KCl was added to a final concentration of 300 mM before storing at -80 °C.

3.4   Virus cryo-EM sample preparation and fixation

3.4.1   Preparation of purified whole virus for cryo-EM

Frozen aliquots of purified MVs were thawed on ice and sonicated three times 30 seconds in a sonication bath at 4 °C with the sweeping option. Trypsin (0.25%) was added in a 1:1 dilution and incubated for 30 minutes at 37 °C. The sample was mixed 1:1 with 4% Paraformaldehyde (PFA) (Merck, no. P6148) in 1 mM Tris-HCl buffer and incubated for 30 minutes at room temperature and 30 minutes at 37 °C (final concentration 2% PFA). Samples were frozen at -80 °C. The samples were thawed on ice and then sonicated three times for 30 seconds in a sonication bath at 4 °C with the sweeping option before vitrification.

3.4.2   Preparation of isolated virus cores for cryo-EM

Frozen aliquots of purified VACV cores were thawed on ice. For cryo-SPA samples, 3 M KCl was added after thawing, resulting in a final concentration of 210 mM after all dilution steps. All samples except the ones for cryo-SPA were sonicated three times for 30 seconds in a sonication bath at 4 °C with the sweeping option. Trypsin (0.25%) was added in a 1:1 dilution and incubated for 30 minutes at 37 °C. The samples were

fixed with a 1:1 dilution with 4% PFA in 1 mM Tris-HCl buffer, pH 9 and incubated for 30 minutes at room temperature and 30 minutes at 37 °C (final concentration 2%PFA). Samples were frozen again at -80 °C and then sonicated three times for 30 seconds in a sonication bath at 4 °C with the sweeping option.

3.4.3   Vitrification conditions

Samples were thawed on ice and cryo-ET samples were mixed with a 1:10 dilution of BSA-Gold (10 nm, Aurion Immuno Gold Reagents, no. 410.011) in PBS. Cryo-ET samples were applied on Quantifoil 300 mesh holey carbon grids (Quantifoil Micro Tools, R2/2 X-103-Cu300) and samples for cryo-SPA were applied to 200 mesh holey carbon grids (Quantifoil Micro Tool, R272X-103-Cu200). The grids were first glow discharged with an ELMO glow discharge unit (Cordouan Technologies) for 2.5 minutes and then vitrified with a Leica GP2 plunger (Leica Microsystems). 2.5 µl sample was applied on both sides of the grid and then back-side blotted in chamber conditions of 80% humidity at 4 °C. The grids were vitrified in liquid ethane at -185 °C and stored in liquid nitrogen until imaging.

3.5   Cryo-EM data acquisition

Screening of the vitrified grids to confirm the quality of the sample was done under cryogenic conditions on a Glacios cryo-TEM 200kV. Data sets were acquired under cryogenic conditions on a Thermo Fisher Scientific Titan Krios G3i TEM microscope (Thermo Fisher Scientific) operating at 300kV and equipped with a Bioquantum postcolumn energy filter and a Gatan K3 direct detector.

3.5.1   Cryo-ET data acquisition of VACV MVs and isolated viral cores

The cryo-electron tomography data was acquired using the SerialEM software package version 3.8 (Mastronarde 2005). Prior to data acquisition, gain reference images were obtained. The filter and microscope tuning were conducted using DigitalMicrograph 3.4.3, integrated into the Gatan Microscopy Suite v3.3 (Gatan), and SerialEM. Tilt series were collected with a filter slit width of 10 eV, utilizing a dose-symmetric tilt scheme (Hagen et al. 2017) covering a range from -66° to 66° with a 3° increment. The nominal defocus range was configured to -1.5 to -8 µm for whole VACV mature virions and -1.5 to -5 µm for isolated viral cores. The nominal magnification was set to 64,000×, resulting in a pixel size of 1.381 Å. Tilt images were taken as 5,760 × 4,092 pixel movies of ten frames with a cumulative dose of 165 e/$Å^2$. For detailed data acquisition settings, please refer to Table 2.

3.5.2   Cryo-SPA data acquisition of VACV cores

The automated collection process for the isolated viral core SPA dataset was set up using EPU 2.13 (Thermo Fisher Scientific) along with AFIS-corrected beam-shift image-shift acquisition. The soluble fraction purified from isolated cores was obtained via SerialEM 4.0 (Mastronarde 2005) with active beam tilt and astigmatism compensation. SPA micrographs were collected in counting mode, with a filter slit

width of 20 eV and using a 4-shot per hole data collection method. The nominal defocus was set to -1.25 to -3.0 μm for isolated cores and -1.5 to -2.2 μm for the soluble fraction. The nominal magnification was set to 81,000×, resulting in a pixel size of 1.06 Å. The isolated viral core dataset was collected from 5,760 × 4,092 pixel movies of 34 frames with a cumulative dose of 53.06 e/Å^2. The soluble fraction dataset was collected with a tilted stage of 25 degrees, resulting in 5,760 × 4,092 pixel movies of 54 frames with a cumulative dose of 80.20 e/Å^2. The decision to obtain tilted data was based on results from the isolated core dataset, indicating a preferred orientation in several classes. Details for data acquisition can be found in Table 2.

Table 2 Cryo-EM data acquisition details

	STA of A10 trimer from intact VACV virus (EMDB-17411) (EMDB-17413)	Cryo-ET VACV isolated cores (EMDB-17414)	A10 Trimer (residues 1-599) SPA (EMDB-17410) (PDB 8P4K)	Flower-shaped pore SPA (EMDB-17412)	A3 inner-core wall SPA (EMDB-18452)	SPA VACV soluble fraction of isolated cores
Data collection and processing						
Magnification	64,000	64,000	81,000	81,000	81,000	81,000
Voltage (kV)	300	300	300	300	300	300
Electron exposure (e–/Å^2)	165	165	53.06	53.06	53.06	80.0
Dose rate (eps)	18.59	22.74	24.458	24.458	24.458	23.43
Defocus range (μm)	-1.5 to -8.0	-1.5 to -5.0	-1.25 to -3.0	-1.25 to -3.0	-1.25 to -3.0	-1.5 to – 2.2
Pixel size (Å)	1.381	1.381	1.06	1.06	1.06	1.06
Acquistion scheme / tilt	-66/66°, 3°	-66/66°, 3°	0°	0°	0°	25°
Frame number	10	10	34	34	34	54
Symmetry imposed	C3	-	C3	C6	C1	-
Tomograms for STA/ micrographs for SPA	15	n/a	9,264	9,264	9,264	11,621
Initial particle images (no.)	-	-	224,331	26,127	211,924	-
Final particle images (no.)	27,922	-	24,943	14,330	18,452	-
Map resolution (Å)	13.1	-	3.8	7.2	20.7	-
FSC threshold	0.143	-	-	0.143	0.143	-
FSCref threshold		-	0.5	-	-	-
Refinement						
Model resolution (Å)	-	-	3.5	-	-	-
FSC threshold	-	-	0.143	-	-	-
Model resolution range (Å)	-	-	Infinity to 3.8	-	-	-
Map sharpening *B* factor (Å^2)	-2100	-	-	-424.39	-	-
Model composition						
Non-hydrogen atoms	-	-	14,568	-	-	-
Protein residues	-	-	599	-	-	-
B factors (Å^2)						
Protein	-	-	99.79	-	-	-
R.m.s. deviations						
Bond lengths (Å)	-	-	0.020	-	-	-
Bond angles (°)	-	-	2.618	-	-	-
Validation						
MolProbity score	-	-	0.79	-	-	-
Clashscore	-	-	1.00	-	-	-
Poor rotamers (%)	-	-	0.00	-	-	-
Ramachandran plot						
Favored (%)	-	-	98.99	-	-	-
Allowed (%)	-	-	1.01	-	-	-
Disallowed (%)	-	-	0.00	-	-	-

3.6   Image processing

3.6.1   Tomogram reconstruction

Initial reconstruction of the tomograms was performed with Tomoman (Wan 2020) to sort and create tilt stacks. The defocus was estimated using CTFFIND 4.1.14 (Rohou and Grigorieff 2015). IMOD 4.9.12 (Kremer et al. 1996) was utilized to align tilt series and create 8x binned tomograms using weighted-back projection (WBP). The reconstruction of the full tomograms was performed with NovaCTF (Turoňová et al. 2017) with simultaneous 3D CTF correction. The phase flip algorithm was used with a slab thickness of 15nm.

3.6.2 IsoNet

To enhance the quality of tomograms, IsoNet (Liu et al. 2022) was used to boost the SNR and to restore the missing wedge information. The tomograms were sorted by defocus with a group size of 5 to 10 tomograms per training. These raw bin8 tomograms underwent deconvolution using IsoNet, serving as the training data for the neural network over 50 iterations. The training involved using a mask with a patch size of 6 and subtomograms with a cube size of 64.

3.6.3 Template matching of the palisade layer of whole MV particles

Subtomogram averaging requires picked particles of the desired structural feature. Therefore, template matching was tested within the Dynamo software version 1.1.333 (Castaño-Díez 2017). To this end, initially particles were manually picked to create a starting reference for template matching in bin8. The software calculates cross-correlation of the template against the tomogram to generate a cross-correlation map, showing for each pixel the similarity of the neighborhood to the template. Various settings were tested, adjusting cone range, cone sampling, in-plane range, and in-plane sampling. After examining the cross-correlation maps, different thresholds were tried to identify suitable particles. The program mostly detected side views of the viral core palisade layer, as could be seen when the particles were back-projected to the whole tomogram, while the top and bottom views were missing.

3.6.4 Meshpicking of the palisade layer

Another approach within the Dynamo software (Castaño-Díez 2017) is to pick particles via defining surface meshes in the tomograms via a graphical user interface (gui). Individual IsoNet-corrected tomograms were loaded into the dtm slice gui, and then a new membrane model was created. The initial points (user points) were manually picked in z and y planes, and the virus's interior was selected by setting a center point to define the particles' angles (exterior facing). After initial picking of the points around the virus core, the program calculated control points, which are equi-spaced points distributed along user points to represent a continuous surface. The spacing of the generated points of the mesh was then recalculated to obtain 10,000 to 12,000 crop points. These positions were exported to a table, which was used to crop the particles from the tomograms. These cropped particles represent around fourfold of the expected particles, as it is recommended to use oversampling in the initial alignments of subtomogram averaging.

3.6.5 Subtomogram Averaging of the palisade layer

The subtomogram averaging workflow is illustrated in Figure 15. First, subtomogram averaging of IsoNet-corrected tomograms was used to get a good starting reference (and to also enable better segmentation of the virus core surface for subtomogram extraction, as explained above) and which was further used to drive the alignment of the WBP data.

In detail, Bin8 tomograms were IsoNet-corrected, as explained in Chapter 3.6.2, to compensate for the missing wedge and to obtain a better contrast to allow better visualization of the viral core (Figure 18). The definition of the subtomogram averaging particle picking is explained in Chapter 3.6.5, and all subsequent subtomogram averaging steps were performed in Dynamo version 1.1.333 (Castaño-Díez 2017). To generate a *de novo* reference, subtomograms were extracted with a cubic size of 464 Å³ from the bin8 tomograms and subjected to five rounds of alignments without applying symmetry. The first initial alignment reference was created by simply averaging all particles together and used for the following alignment projects.

The obtained *de novo* reference already displayed a hexamer-of-trimers arrangement. Initial mesh positions were extracted from the original WBP tomograms for further alignment projects and used the reference created from the Isonet corrected tomograms and started a new alignment with the WBP particles. After the C3 symmetry of the structure became apparent in the bin8 iterations, C3 symmetry was applied. The alignments were subsequently refined from bin8 over bin4 (subtomogram cubic size 464 Å³) to bin2 (subtomogram cubic size 398 Å³). The low-pass filter was gradually advanced, and the Euler angle scanning step and range decreased.

After the first two alignments in bin8, lattices were distance cleaned (6 pixels) to avoid duplicates and cc-threshold cleaned to remove subvolumes that did not align to the core. When reaching the bin2 stage, the dataset was split into even/odd half sets, and from this point onward, the even/odd datasets were treated independently. The low-pass filter never extended beyond 25 Å up to this point. The final half set averages were multiplied with a Gaussian-filtered cylindrical mask, and the resolution was determined by mask-corrected Fourier-shell correlation (Chen et al. 2013) after the final bin2 iteration. The final map was sharpened with an empirically determined B factor of -2,100 A2 and filtered to its measured resolution at the 0.143 FSC criterion (Rosenthal and Henderson 2003).

3.6.6   Single Particle Analysis

The processing details are briefly outlined here, with comprehensive information available in Figure 21 for the trimer processing and Figure 32 for the processing of the flower-shaped pore.

The movies from the dataset containing intact cores underwent motion correction with dose-weighting using the RELION 4.0-beta2 (Scheres 2012) implementation of MotionCorr2 with a patch size of 7 × 5. Subsequent processing of the Motion-corrected micrographs was performed using Cryosparc 4.0.0 (Punjani et al. 2017).

First the initial CTF parameters were estimated using CryoSPARC patch CTF. Initial particles were picked with a blob picker and extracted with a large box size (636 Å, bin4) to ensure that both large and small protein populations were captured during 2D classification. Through iterative 2D classification with varying mask sizes, particles could be sorted into different protein species with distinct classes. The following processing steps included re-extracting particles using more appropriate box sizes: 340 Å for trimers, 545 Å for the flower-shaped pore, and 636 Å for the side views of the

core wall. In all cases, initial 3D volumes were generated using CryoSPARC *ab initio* without symmetry, followed by the imposition of symmetry during 3D auto-refinement.

The 2D classes containing particles for the flower-shaped pore were carefully reconstructed using CryoSPARC non-uniform refinement. Afterward, they were locally sharpened and filtered within CryoSPARC. The particles of the A10 trimer were exported to RELION for further processing, which included 2D classification, 3D refinement using a mask containing the full trimer density, Bayesian polishing, defocus refinement, and focused 3D classification of A10 monomers. To improve the map quality, particle views were manually balanced for the trimer by removing over-represented top views during 2D classification. Finally, the resulting map underwent density modification (Terwilliger et al. 2020) using Phenix version 1.20-dev-4224 (Adams et al. 2010), with input of two half maps and the mask used during refinement. The 3DFSC calculations were conducted using the remote 3DFSC Processing Server (https://3dfsc.salk.edu).

The A10 trimer map's local resolution was calculated with RELION's implementation of ResMap (Kucukelbir et al. 2014) while cryoSPARC's local resolution estimate tool was employed for the flower-shaped pore. Additionally, the V4 tool was utilized in EMAN/1.9 (Ludtke et al. 1999) to generate projections of the final flower-shaped pore. The volume for the lower density of the core wall was produced in cryoSPARC from 3,795 side-view particles selected during 2D classification. An ab initio 2x binned initial model was generated without symmetry and 3D non-uniform refinement without symmetry resulted in a final map at a global resolution of approximately 20.7 Å. Subsequently, rigid-body fit of the highest-ranking Alphafold2 model of the A3 dimer into the density was done using the UCSF Chimera (1.17.1) 'fit in map' tool. During this process, a simulated map at a resolution of 20 Å (estimated resolution from cryoSPARC) and optimized by correlation was used to set options for real-time correlation. The optimal orientation of the model was determined based on the option that best satisfied the density, as measured by UCSF Chimera's output correlation score of each fit.

Following the same procedure as the previous dataset, the dataset collected at a 25° stage tilt containing the soluble fraction underwent motion correction using RELION (Scheres 2012). Data processing was done separately from the dataset containing intact cores. CryoSPARC (Punjani et al. 2017) patch CTF was utilized to estimate CTF and account for stage tilt. Particle picking was conducted using a blob picker, as previously described. The particles were extracted at bin2 with a box size of 340 Å and then subjected to iterative 2D classification with varying mask sizes and 250 classes.

3.7 Structure of putative structural core proteins predicted with AlphaFold

The initial structures of potential core proteins were predicted as monomers using AlphaFold 2.3.2. (Jumper et al. 2021) and Colabfold 1.5.2 (Mirdita et al. 2022). For orthopoxvirus proteins, five seeds were generated per model, while for parapoxvirus and entomopoxvirus models, four seeds were used, with one prediction per seed. All five models underwent relaxation using Amber relaxation and were then inspected manually. The model with the highest pLDDT score was chosen for further analysis.

The following proteins were folded as monomers: A10 (UniProt P16715, residues 1–614), A3 (UniProt P06440, residues 62–643), 23K (UniProt P16715, residues 698–891), A4 (UniProt P29191), and L4 (UniProt P03295, residues 33–251). Additionally, orf virus p4A (GenBank ID AY386264.1), AmEPV AMV139 (GenBank ID NP_064921.1 putative core protein), and MsEPV (GenBank ID AF063866.1 putative core protein P4a homolog) were also folded as monomers (Extended Data Fig. 1). Multimer predictions were made using the same settings as for monomer folds. The A10 trimer (UniProt. P16715, residues 1–614) and the A3 dimer (UniProt P06440, residues 62–643) were predicted as multimers. The predictions for the comparison of A10 trimer of variola virus (GenBank ID ABF23487.1, residues 1–615) and orf virus (GenBank ID AY386264.1, residues 1–905) to VACV WR (UniProt P16715, residues 1–614) were also made, as well as the multimer prediction of the A3 dimer. These predictions were conducted using AlphaFold version 2.3.2 (Jumper et al. 2021) and Colabfold (Mirdita et al. 2022).

3.8   Structure fitting and analysis

3.8.1   Modeling of the A10 trimer

The highest-ranking A10 model (residues 1–599) was fitted three times into the trimer map shown in Figure 22. The A10 trimer was refined and adjusted into the trimer map with imposed non-crystallographic C3 symmetry using Rosetta/3.13 (20220812 build) (Fleishman et al. 2011, Khatib et al. 2011, Maguire et al. 2021). Important refinement parameters included setting the resolution to our estimated value of 3.8 Å and setting the elec_dens_fast density weight to a value of 3.5, as recommended for maps at this resolution. The model was then examined and refined manually on a single chain using Coot (ver. 0.8.9.1 EL) (Emsley and Cowtan 2004), and it was copied to other C3 symmetry positions using UCSF Chimera (Pettersen et al. 2004). Afterward, the model was refined again using Rosetta, with non-crystallographic symmetry (NCS) applied. Model statistics were calculated using Molprobity in Phenix (Williams et al. 2018). Pairwise root-means-square-deviation (r.m.s.d.) calculations were done using the matchmaker tool in UCSF ChimeraX (version 1.5) (Pettersen et al. 2021). Finally, the final trimer model was fitted into the outer density of the flower-shaped pore and symmetrized sixfold using UCSF ChimeraX (Pettersen et al. 2021).

3.8.2   ConSurf analysis of the A10 trimer

The ConSurf analysis was conducted using the web server https://consurf.tau.ac.il/ with the default parameters. To start, the amino acid sequence for each of the proteins being analyzed was obtained by uploading the PDB file of the AlphaFold-predicted protein to the webserver. Then, PSI-PLAST (E-cutoff = 0.001) was used to identify similar sequences. For example, in the case of A10, this led to the identification of 31 sequences for comparison, including various members from the Chordopoxvirinae subfamily (such as orthopoxvirus, parapoxvirus, yatapoxvirus, molluscipoxvirus, among others), allowing for a comprehensive conservation analysis.

3.9 Proteomics of VACV cores

This entire chapter was written by Armel Nicholas and it is published in the paper "Multi-modal cryo-EM reveals trimers of protein A10 to form the palisade layer in poxvirus cores" (Datler et al. 2024).

3.9.1 Sample preparation

"To the soluble fraction from isolated cores (prepared as described above), 25 mM TCEP (Gold Biotechnology, no. 51805-45-9) and 4% SDS (Carl Roth, no. 8029,1) was added and boiled 10 min at 95 °C and was first cleaned up by SP3 using a commercial kit (PreOmics, 100 mg of beads per sample), then processed using the iST kit (PreOmics), according to the manufacturer's instructions. Tryptic digestion was stopped after 1 h, and samples were vacuum dried and then re-dissolved in the iST kit's LC LOAD buffer with 10 min sonication."(Datler et al. 2024)

3.9.2 LC–MS/MS analysis

"The sample was analyzed by LC–MS/MS on an Ultimate 3000 RSLC_Nano nano-HPLC (Thermo Fisher Scientific) coupled with a Q Exactive HF (Thermo Fisher Scientific), concentrated over an Acclaim PepMap C18 pre-column (5 µm particle size, 0.3 mm ID × 5 mm length, Thermo Fisher Scientific), then bound to an EasySpray C18 column (2 µm particle size, 75 µm ID × 50 cm length, Thermo Fisher Scientific) and eluted over the following 60 min gradient: solvent A, MS-grade H_2O + 0.1% formic acid; solvent B, 80% acetonitrile in H_2O + 0.08% formic acid; constant 300 nl min–1 flow; B percentage: 5 min, 1%; 45 min, 31%; 65 min, 44%.

Mass spectra were acquired in positive mode with a data independent acquisition method: FWHM 8 s, MS1 parameters: centroid mode, 1 microscan, 120,000 resolution, AGC target 3 × 106, 50 ms maximum IT, 400 to 1,005 m/z; DIA scans: 24 MS2 scans per cycle, 57 windows of 11.0 m/z width per cycle covering the range from 394.9319 to 1,022.21204 m/z (−0.005 m/z non-covered gap between adjacent windows), spectra acquired in Profile mode, with 1 microscan, at 30,000 resolution; AGC target 1 × 106, 60 ms maximum IT, NCE 27." (Datler et al. 2024)

3.10 Data visualization and figure preparation

The cryo-electron tomograms of entire VACV WR mature viruses and isolated cores, along with the lattice maps, were visualized using IMOD (Kremer et al. 1996), UCSF Chimera (Pettersen et al. 2004), and UCSF ChimeraX (Pettersen et al. 2021). The EM densities were visualized in UCSF ChimeraX. The figures were created using Adobe Illustrator 2023. The topology diagrams shown in Figure 10 were generated using Pro-Origami (Stivala et al. 2011).

3.11 Data availability

"The electron microscopy density maps of the A10 trimer and the hexameric flower-shaped pore, the subtomogram average of the palisade layer and representative tomograms for complete viruses, as well as isolated cores, have been deposited in the Electron Microscopy Data Bank under accession codes: EMD-17410, EMD-17411, EMD-17412, EMD-17413, EMD-17414 and EMD-18452.

The refined model of the A10 trimer has been deposited in the Protein Data Bank accession code: PDB 8P4K.

The UniProt codes of VACV core proteins used for structure prediction are: A10 (P16715), A3 (P06440), A4 (P29191) and L4 (P03295).

GenBank Protein IDs of variola virus A10 (ABF23487.1), monkeypox virus A10 (YP_010377118.1), rabbitpox virus A10 (AAS49831.1), cowpox virus A10 (ADZ29251.1), ectromelia virus A10 (NP_671631.1), orf virus P4a (AY386264.1), *Amsacta moorei entomopoxvirus* AMV139 (NP_064921.1) and *Melanoplus sanguinipes entomopoxvirus* putative core protein P4a homolog (AF063866.1) were used for protein sequence alignment." (Datler et al. 2024)

3.12 Vaccinia virus actin tail formation

This entire paragraph is from the paper "ArpC5 isoforms regulate Arp2/3 complex–dependent protrusion through differential Ena/VASP positioning" (Fäßler et al. 2023) that I co-authored. This part from the Material & Methods was initially drafted by myself.

"In the VACV infection assays, we followed a previously published protocol (Pfanzelter et al. 2018). The cells were infected with VACV (Western reserve strain, provided by A. Bergthaler, Medical University of Vienna) for 1 hour, washed three times with PBS and finally incubated in culture medium for an additional 7 hours. To ensure the inactivation of the virus, the VACV-infected cells were fixed with prewarmed 4% paraformaldehyde (Merck, no. P6148) in PBS for 1 hour. Following this, the cells were washed three times with PBS and then subjected to staining with Phalloidin-Atto594 and a previously characterized monoclonal anti-B5 antibody (Aldaz-Carroll et al. 2005, 2005)(63, 64) before being embedded in ProLong Gold Antifade Mountant (Thermo Fisher Scientific, no. P36934). The VACV-infected samples were visualized using a Zeiss Axio Observer equipped with 405-, 488-, 561-, and 640-nm laser lines and a Plan-APOCHROMAT 63×/1.4 Imm Corr DIC Oil M27 objective. Subsequent tasks such as converting image formats, creating merges/overlays, and measuring intensity and distance were carried out in Fiji v1.52p." (Fäßler et al. 2023)

4 Results

The results in the following chapter have been part of the publication "Multi-modal cryo-EM reveals trimers of protein A10 to form the palisade layer in poxvirus cores" (Datler et al. 2024), and were reproduced and adapted for this thesis. The draft was originally written by myself, except chapter 4.6, chapter 4.7, chapter 4.8 and chapter 4.10, which were originally drafted by both Jesse Hansen and myself.

4.1 Optimization of VACV purification

The VACV WR wildtype strain is commonly used in poxvirus research as it is the ideal role model for poxviruses, widely conserved within the *Orthopoxviridae* and can be handled under biosafety level 2 (BSL2) conditions. Purification of VACV MVs and MV cores was previously described (Easterbrook 1966, Dubochet et al. 1994, Moussatche and Condit 2015). However, I needed to optimize these protocols to purify MVs in such a quality and quantity that they could be used for high-resolution cryo-EM and cryo-ET studies.

4.1.1 Purification of whole VACV MVs for cryo-ET

Cryo-ET acquisition of the entire MV particle provides structural context and allows to draw general conclusions about poxvirus assembly features. Cryo-ET acquisition requires purified MVs in a high concentration with a good spread and density across holey carbon cryo-EM grids to prevent viruses from superimposing each other at high tilts during data collection and to avoid too thick vitreous ice. Hence, I optimized previously published purification protocols based on (Easterbrook 1966), trying different mammalian cells, buffers, pH values and sonication steps to achieve high virus titers and to prevent virus particles from clumping. In the first trials, BSC40 (Buffalo Green Monkey Kidney cells, ATCC) cells were used to propagate the virus, but the titers I obtained were lower than expected. After switching to HeLa cells, the titers increased by approximately 100-fold to 1000-fold. I tried several different buffers used for virus purification with different pH values such as HEPES buffer pH 7.5 and pH 9 and Tris-HCl buffer at pH 7 and pH 9. Ultimately, the 1mM and 10mM Tris-HCl buffers, pH 9 proved to be the most effective, yielding the highest amount of VACV. These buffers were used to isolate VACV with a subsequent sucrose gradient purification. After purification, the virus particles aggregated into big clumps, which were unusable for cryo-ET acquisition. To address this, the MVs were separated using trypsinization and sonication. After multiple trials with varying sonication durations and intensities, the optimal protocol was established and is described in detail in chapter 3.2..

VACV whole virus purification for cryo-ET

concentration	density and spread of virus	quality
low magnification	medium magnification	high magnification
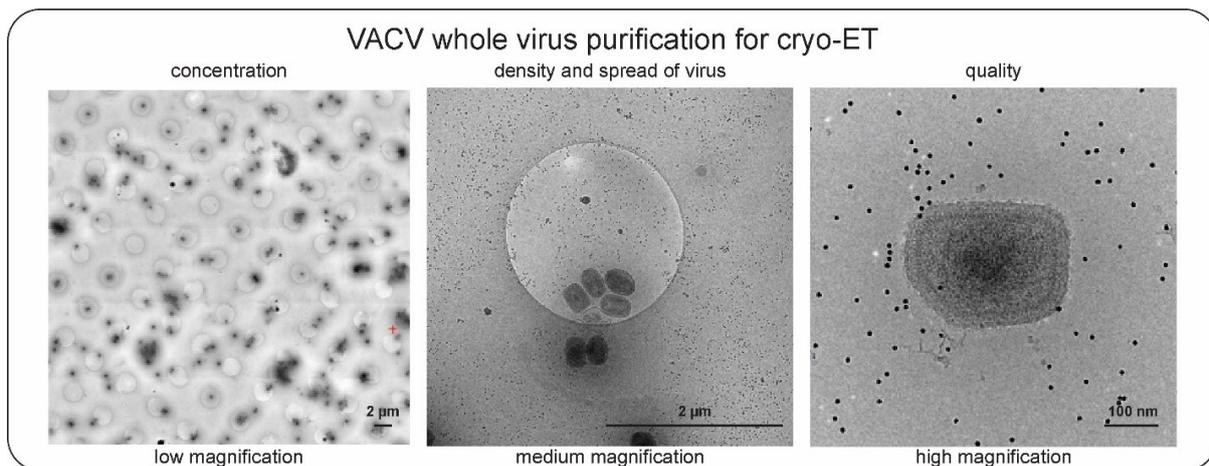
Figure 7: Quality of VACV whole virus purification

Cryo-EM images of purified VACV MVs on holey-carbon grids shown at low, medium and high magnification to display the concentration of the virions on the grid square, the density and spread of the virus in holes and the quality of the purified virus sample to estimate the quality for high-resolution acquisition, respectively.

Quality control of the virus samples was done via cryo-EM using a ThermoFisher Scientific Glacios TEM operated at 200kV. Low, medium and high magnification micrographs were used to estimate the concentration on the grid square, the density and spread of the virus in the holes of holey carbon films and the overall quality of the virus as is shown in Figure 7. At high magnification, the virus particles appear as expected, with an ideal distribution of gold particles. The dimensions of the VACV align with the reported literature (Hernandez-Gonzalez et al. 2023), and the virus particles display an intact envelope with a clearly visible outer membrane.

4.1.2 Purification of VACV MV cores for cryo-ET and cryo-SPA

The additional establishment of a virus core purification protocol allows more profound insights into the viral core composition. A previous paper demonstrated that isolated cores and fine structural details are visible after stripping the MVs of their outer membrane (Dubochet et al. 1994). Furthermore, single proteins are shed off the viral cores during purification, which are visible in micrographs (Dubochet et al. 1994). Using isolated viral cores for cryo-EM gives us more structural details and potentially allows obtaining high-resolution structures.

I used previously described core stripping protocols (Dubochet et al. 1994, Moussatche and Condit 2015) and optimized them by trying different concentrations of the detergent NP-40 and redox reagent DTT. NP-40 solubilizes the viral envelope, and DTT acts as reducing agent, breaking up disulfide bonds and preventing aggregation of the viral cores. DNAse was added to the stripping buffer to avoid DNA contamination. As shown in Figure 8a, analysis by cryo-TEM revealed the virus core concentration to be less abundant when compared to the whole virus purification. Still, the amount, spread, and the quality of cores was sufficient for cryo-ET acquisition. A detailed description of the viral core purification is provided in chapter 3.3.

However, a higher concentration of viral cores on the grid squares was necessary for cryo-SPA acquisition to obtain a sufficiently large dataset for subsequent image processing. The amount of starting virus used for the VACV purification was multiplied by four and the above described viral core purification protocol was adjusted according to previous purification protocols (Dubochet et al. 1994) where single proteins can already be seen on the micrographs. As suggested by Dubochet et al., high concentrations of KCl were added throughout the purification process to separate the viral cores and to shed parts of the palisade layer of the viral cores, as described in chapter 3.3. As for the purification of intact virus particles, quality control was again performed using a ThermoFisher Scientific Glacios TEM operated at 200kV at low, medium and high magnification. The results are shown in Figure 8b. The distribution and density of the viral cores was deemed sufficient, with some cores appearing intact while others were fragmented. The cores exhibit reduced density in the center, indicating the absence of the genome, likely due to perforations or structural defects in the core wall. The micrographs reveal that the viral cores are disrupted, with visible free DNA present in the surrounding buffer.
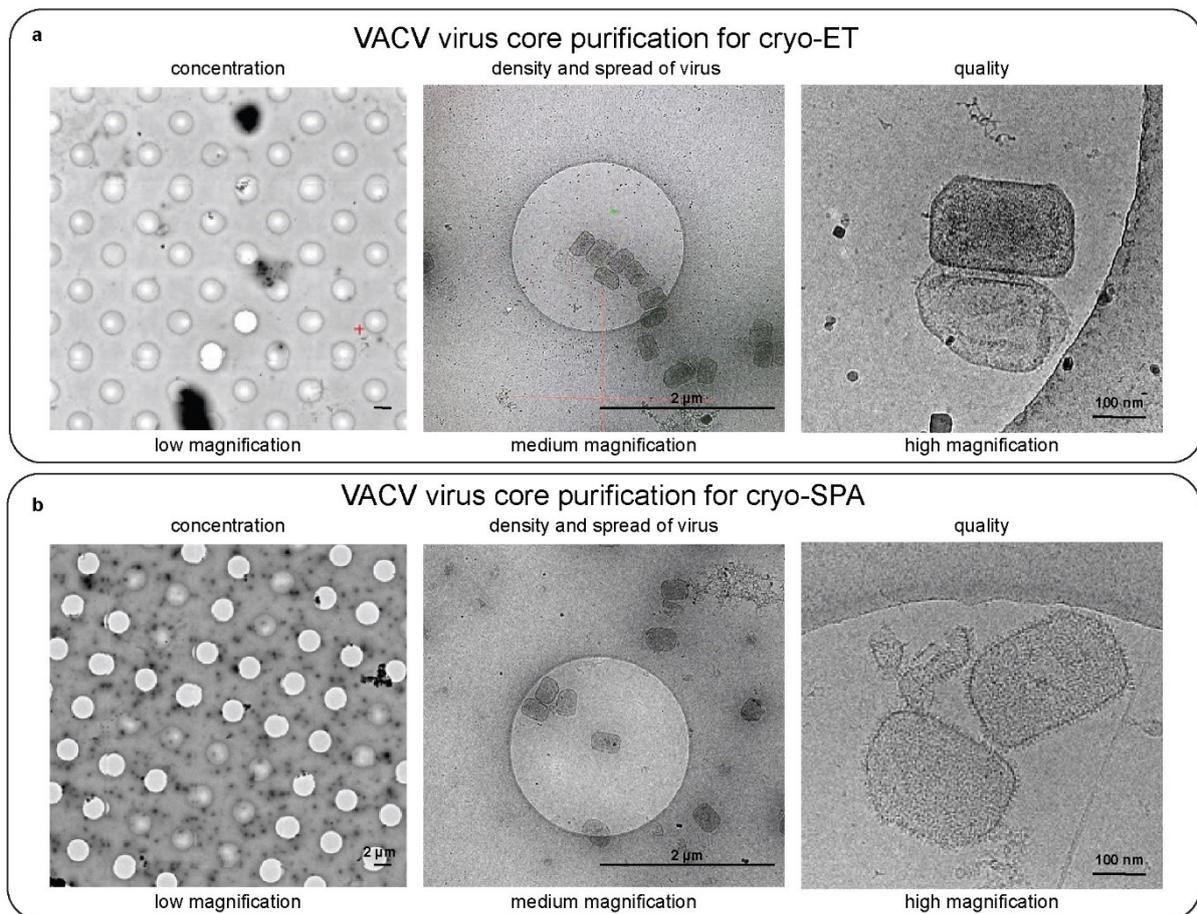
Figure 8: Quality of VACV core purification

Cryo-EM images of purified VACV MV cores on holey-carbon grids shown at low, medium and high magnification to display the concentration of the viral cores on the grid square, the density and spread of the viral cores in holes and the quality of the purified cores to estimate the quality for high-resolution acquisition, respectively. Purified VACV cores for cryo-ET (a), purified VACV cores for cryo-SPA with additional KCl treatment (b)

4.2   The protein composition of the isolated MV cores

According to literature VACV cores contain the putative structural core proteins A10, 23K, A4, A3, and L4 (Dubochet et al. 1994, Vanslyke and Hruby 1994, Jensen et al. 1996, Roos et al. 1996, 1996, Risco et al. 1999, Pedersen et al. 2000, Heljasvaara et al. 2001, Chung et al. 2006, Rodriguez et al. 2006, Jesus et al. 2014, 2015, Moussatche and Condit 2015, Mirzakhanyan and Gershon 2019, Bidgood et al. 2022).

To ensure that purified VACV cores retain these structural proteins, we used mass spectrometry to confirm the presence of these proteins in our isolated core sample. Specifically, we focused on the soluble fraction of the purified core sample that we used for cryo-SPA.

4.2.1   Proteomics

Mass spectrometry analysis of the soluble fraction of the isolated core sample revealed the presence of all the expected core proteins. Among the most abundant VACV proteins identified in the sample (listed in Table 3), the precursor of protein A10, also referred to as the major core protein 4a precursor (Uniprot ID P16715), was found to be the most prominent, followed by core protein L4, also known as core protein VP8 (Uniprot ID P03295). Furthermore, core protein A4 (Uniprot ID P29191) and A3 (Uniprot ID P06440) were identified within the top 10 entries, alongside other proteins known to be associated with the viral core and involved in the transcription and translation of the viral genome.

The comprehensive mass spectrometry results, including the host proteins, can be located in Appendix I. Among the most predominant proteins detected were a transmembrane serine protease, histones, mRNA capping proteins, cytochrome C, heat shock proteins, Annexin, and Keratin. Previous studies have also observed many of those host proteins in VACV virus purifications (Chung et al. 2006, Resch et al. 2007). Upon infection with VACV, the host cell rapidly halts protein synthesis, triggering significant cellular stress. This stress response is part of the cell's defense mechanism to combat viral invasion while maintaining homeostasis during the disruption of protein production pathways. Notably, stress-induced proteins such as heat shock proteins and Annexin are upregulated in response to cellular stress. Due to their high abundance, the detected proteins may remain in the sample post-purification. Additionally, proteins may be encapsulated within the viral core, which would require further investigation to evaluate the role of these in virus assembly and infectivity.

Table 3: Mass spectrometry data from soluble fraction of isolated cores

This list comprises mass spectrometry findings from the soluble fraction core sample, filtered to exclusively display proteins encoded by VACV sorted by log 10 expression. The full list, including host proteins can be found in Appendix I.

Table source: adapted from (Datler et al. 2024) , License: CC BY 4.0

UniProt ID	Common Names	Genes	Sequence Coverage %	Peptide Counts	Log 10 expression
P16715	*Major core protein A10 (4a) precursor*	*VACWR129*	60.5	69	9.5060
P03295	*Core protein L4 (VP8)*	*VACWR091*	54.2	16	9.1462
P07617	*Cap-specific mRNA (nucleoside-2'-O-)-methyltransferase*	*PAPS*	65.5	22	9.0373
P26669	*Cu-Zn superoxide dismutase-like protein A45R*	*VACWR171*	91.2	9	8.8556
P04195	*Cell surface-binding protein*	*VACWR113*	53.9	21	8.7152
P07396	*Phosphoprotein F17*	*VACWR056*	32.7	6	8.6891
P29191	*A4 (39kDa) core protein*	*VACWR123*	17.8	4	8.2534
P04298	*mRNA-capping enzyme catalytic subunit*	*VACWR106*	62.8	56	8.2447
P68692	*Glutaredoxin-1*	*VACWR069*	24.1	4	8.2045
P06440	*Major core protein A3 (4b)*	*VACWR122*	34.4	17	8.1866
P07392	*DNA-directed RNA polymerase 147 kDa polypeptide*	*PREPL;RPO147;*	55.5	69	8.0926
P24757	*DNA-directed RNA polymerase 35 kDa subunit*	*RPO35*	53.8	17	8.0675
P68428	*RNA polymerase-associated transcription-specificity factor RAP94*	*GFOD1;RAP94*	49.2	38	8.0537
Q76ZN5	*Profilin*	*VACWR167*	39.8	6	8.0401
P20636	*Early transcription factor 82 kDa subunit*	*VETFL*	55.2	34	8.0382
P68609	*DNA-directed RNA polymerase 22 kDa subunit*	*RPO22*	45.9	9	8.0209
P12926	*Core protease I7*	*VACWR076*	24.3	11	7.8991
P07614	*Protein L3*	*VACWR090*	32.9	12	7.8896
P04318	*mRNA-capping enzyme regulatory subunit*	*VACWR117*	72.1	19	7.8743
P07242	*Late transcription elongation factor H5*	*VACWR103*	44.8	10	7.8030
Q76ZP7	*DNA-directed RNA polymerase 133 kDa polypeptide*	*RPO132*	43.5	46	7.7817
P24758	*Protein A26*	*VACWR149*	35	15	7.7764
P07616	*Protein J1*	*VACWR093*	19.6	4	7.7718

4.2.2   AlphaFold prediction of major structural core proteins

Experimentally derived structures for putative core proteins A10, 23K, A3, A4 and L4 were unavailable. Therefore, we utilized AF2 (Jumper et al. 2021) and Colabfold (Mirdita et al. 2022) to accurately predict 3D protein structures of our core protein candidates (Figure 9). We intended to use these predicted models to interpret any of the structures obtained in our subsequent experiments. Figure 9 displays the predicted structural models as monomers and their prediction quality using a predicted local distance difference test (pLDDT) confidence score. Low confidence is shown in red, and high confidence in blue. All predicted proteins showed high pLDDT scores except protein A4, which displayed a disorganized structure. 23K exhibited a distinct triple helix confirmation and L4 demonstrated a particularly globular architecture.

AlphaFold-predicted models of major structural core proteins

A10
(1-614 Aa)

A3
(62-643 Aa)

23K
(698-891 Aa)

A4
(1-281 Aa)

L4
(33-251 Aa)

pLDDT score
0   50 70 90 100
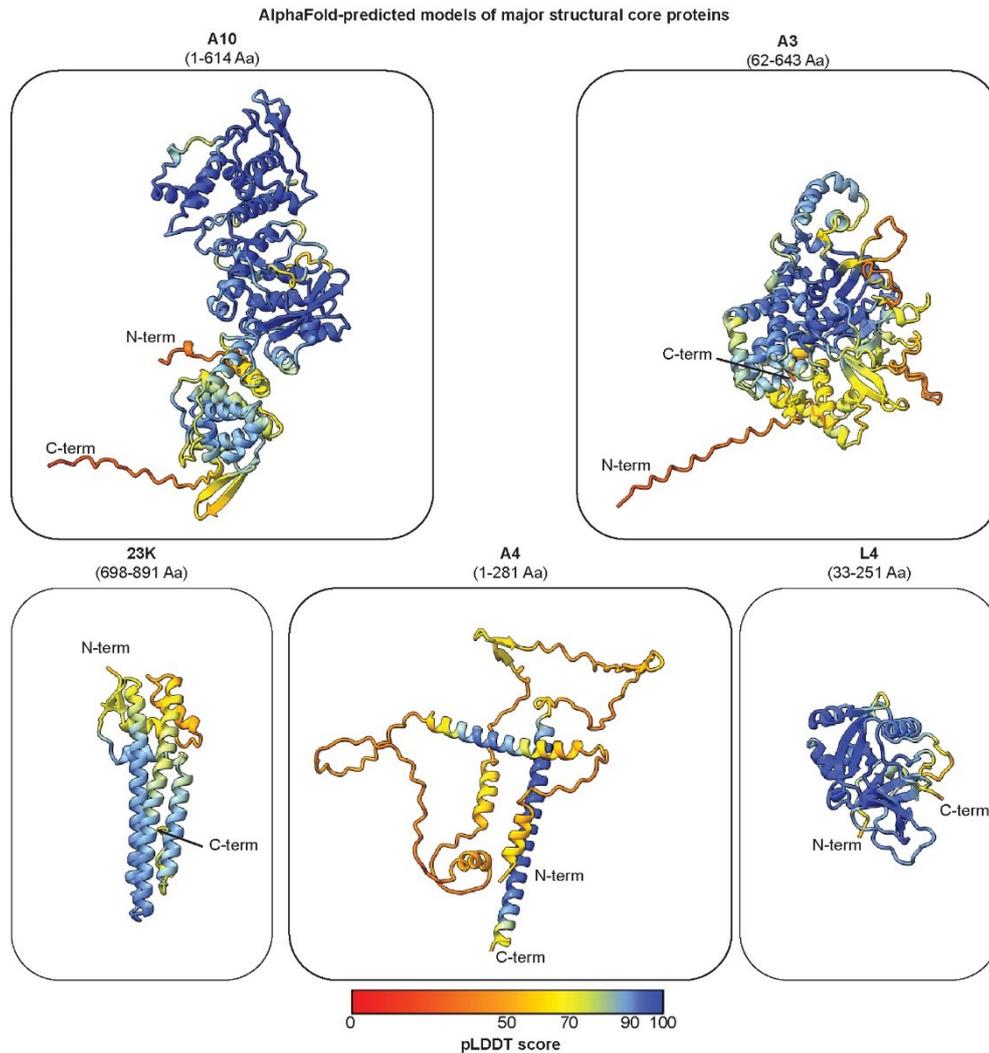
Figure 9: Alphafold predicted structures of putative core proteins

AlphaFold has predicted models for A10, A3, 23K, A4, and L4. The model's coloring indicates the pLDDT confidence score, with red for low confidence and blue for high confidence. The modeled residues are shown in brackets, and the N- and C-termini of the proteins are annotated.

Image source: (Datler et al. 2024) , License: CC BY 4.0

To better understand the structural characteristics of A10 and A3, we utilized Pro-Origami (Stivala et al. 2011) to create a topology map for the proteins (Figure 10). The topology analysis of A10 revealed an intricate fold with the amino terminus centrally positioned between two domains and a predominantly alpha-helical structure at the top. The protein folds back towards the center resulting in the C-terminus being in close proximity to the N-terminus (Figure 10b). Conversely, protein A3 exhibited a compact shape, with both the N- and C-terminal ends positioned close to each other on the same side of the protein structure (Figure 10d).
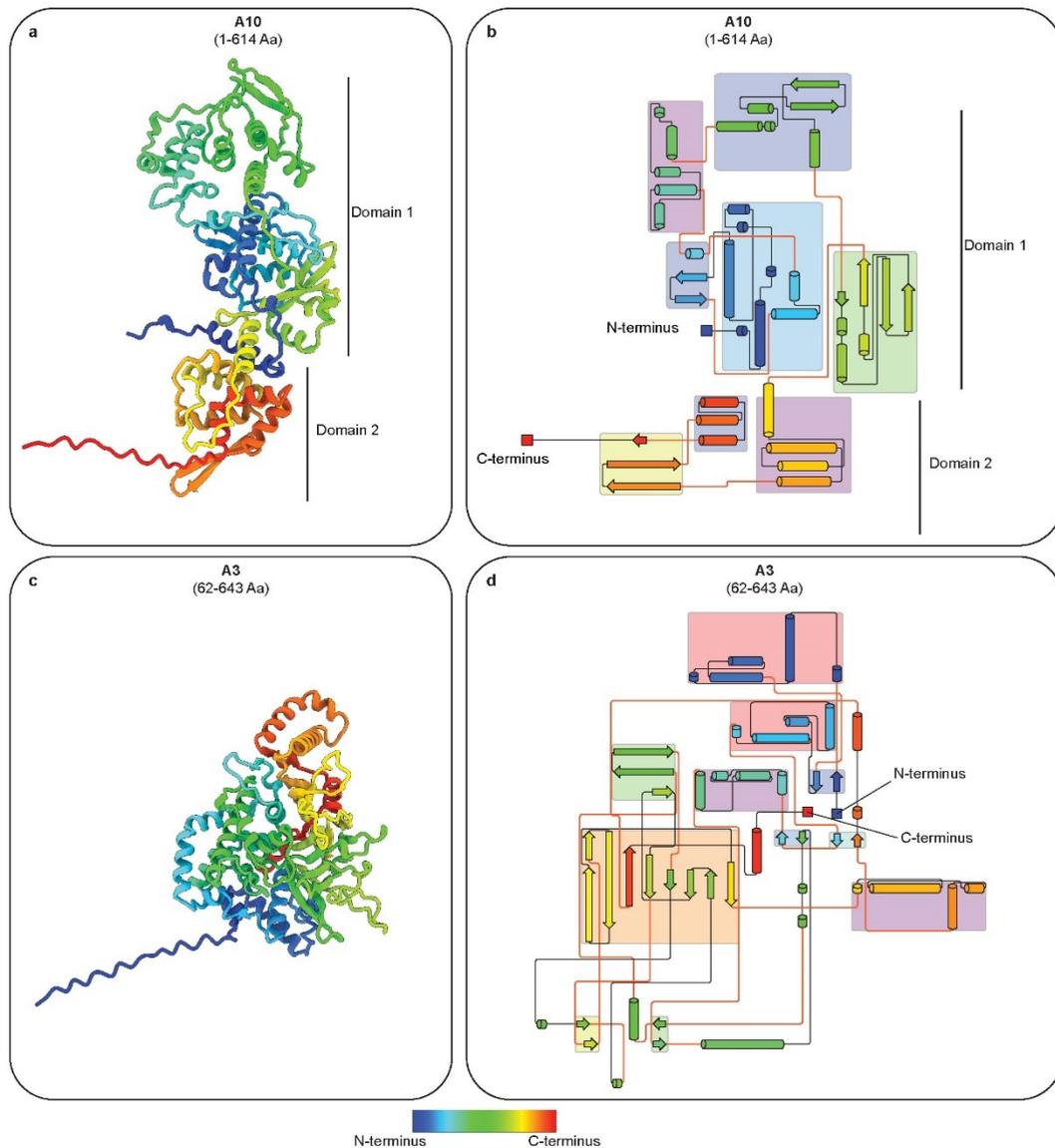
Figure 10: Illustration of protein topology and arrangement for A10 (panel a-b) and A3 (panel c-d)

"The coloring of the models and topology diagrams is according to residue positioning from the N-terminus (blue) to the C-terminus (red). (a)Cartoon ribbon representation of the Alphafold-predicted model of A10, colored from N-terminus to C-terminus. The two sub-domains (SD1 and SD2) of the A10 fold are annotated. b) Protein topology diagram of A10, showing the positioning of protein regions, such as specific secondary structures with respect to each other. Beta-strands and alpha-helices are shown as arrows and cylinders, respectively. Their length is proportional to their number of residues. Colored boxes represent motifs which are positioned nearby relative to one another and therefore are likely to form a structural group. (c) Cartoon ribbon representation of the Alphafold-predicted A3 colored from N-terminus to C-terminus.d) Protein topology diagram of A3, in the same depiction style as in (b). The topology diagrams shown in (b) and (d) were generated using Pro-Origami. The diagram in (b) was manually adapted to further improve visualization of the complex protein architecture. Orange line colors in (b) are used to highlight connections between structural groups. Orange lines in (d) are used to improve visualization but do not necessarily indicate connection between structural groups." (Datler et al. 2024)

Image source: (Datler et al. 2024) , License: CC BY 4.0

We further utilized Foldseek (van Kempen et al. 2024) and Dali (Holm 2022) analysis, which perform structural alignments and structural comparison in 3D to reveal similarities to other proteins. Surprisingly, protein A10 showed no similarities to other cellular or viral proteins. Protein A3 had its highest similarity to deubiquinating proteins, as recently suggested (Mutz et al. 2023). 23K, A4 and L4 proteins showed a very low probability in the Dali and Foldseek search that show that the protein folds seem to be unique to the poxvirus family (and that no related structures have been determined or predicted in the AF-database).

ConSurf Scoring of major structural core proteins
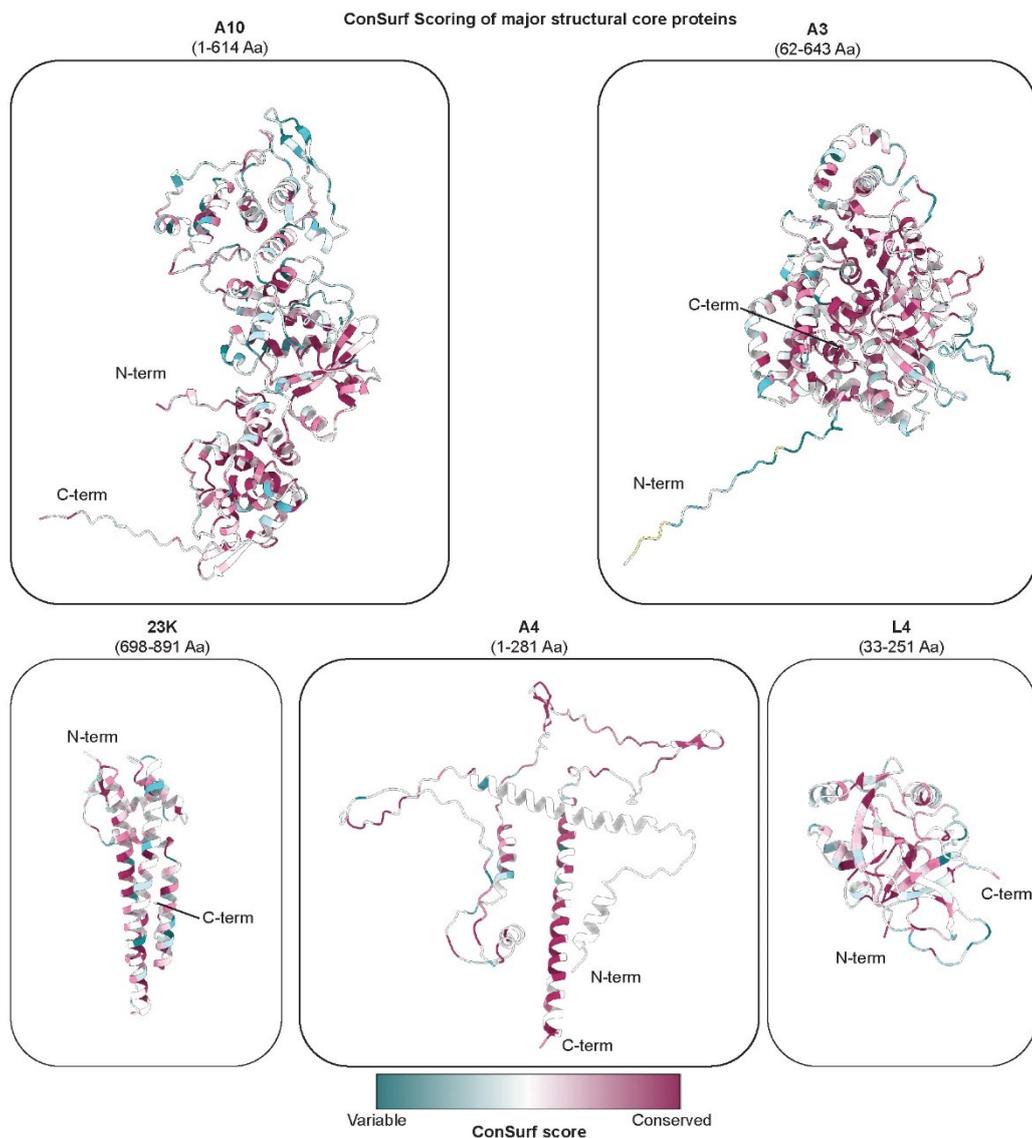
Figure 11: Consurf Analysis of putative structural core proteins Alphafold predicted models of A10, A3, 23K, A4 and L4

Proteins are displayed in the same orientation as shown in Figure 9 The color model is based on the ConSurf score, which ranges from variable (blue) to conserved (red).

Image source: (Datler et al. 2024) , License: CC BY 4.0

To better comprehend structural conservation of the proteins within the poxvirus family such as orthopoxvirus, parapoxvirus and yarapoxvirus among others, we performed a Consurf analysis (Yariv et al. 2023) to assess and visualize evolutionary conservation in the macromolecules of these five major structural proteins (Figure 11). The Consurf score of proteins is shown from variable (blue) to conserved areas (pink) in Figure 11. Proteins A10, A3, 23K and L4 exhibit a high Consurf score, indicating that the amino acid residues are highly conserved across the poxvirus family. In contrast, protein A4 displays a lower Consurf score throughout the protein sequence,

4.3   Cryo-ET of intact VACV mature virions

In the first attempt to study the intact structure of VACV MVs, I encountered substantial challenges. After optimization of the purification protocol for intact MVs (described in Chapter 3.4.1), the intact virions were used for cryo-ET. I acquired tilt series on a Thermo Fisher Scientific Titan Krios TEM operated at 300 kV. The initial processing steps involved reconstructing the tomograms using weighted-back projection (WBP) and correcting for the contrast transfer function (CTF) with novaCTF (Turoňová et al. 2017). Observing the structural detail of the virus particles was not trivial due to the low SNR.
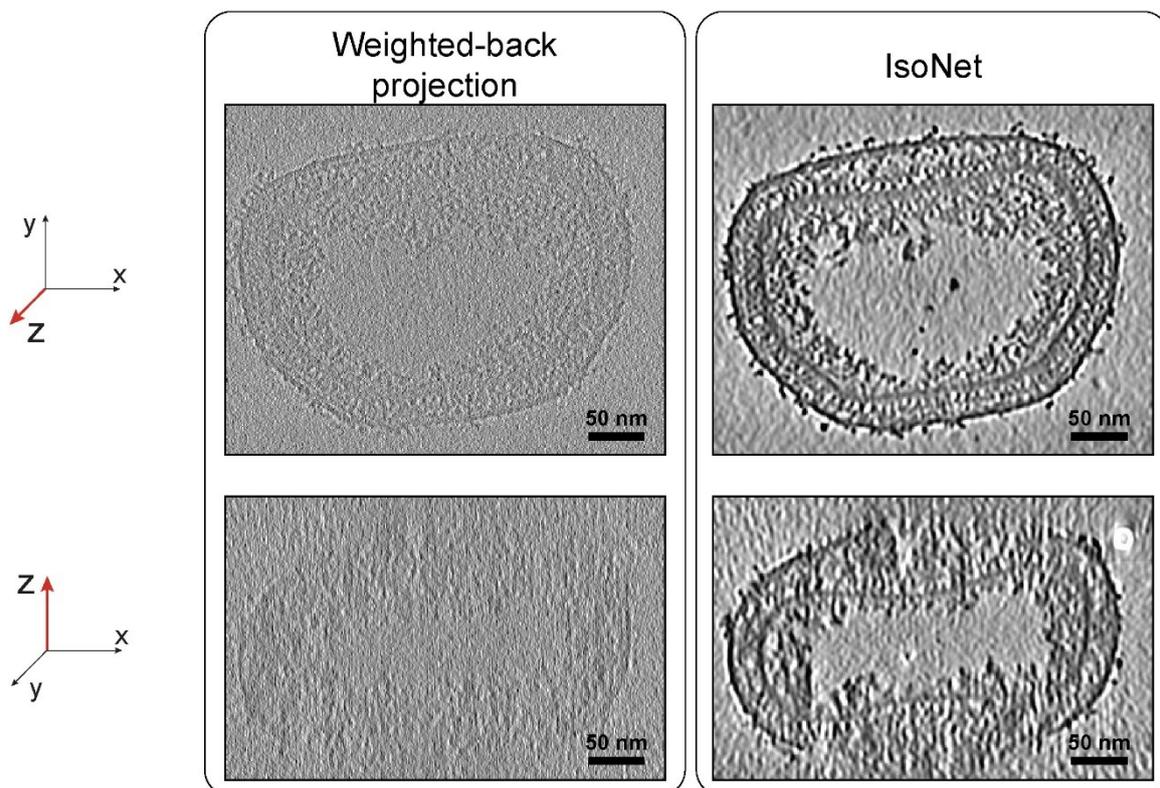
Figure 12: Comparison of the weighted-back projection (WBP) tomogram and IsoNet corrected tomogram of intact VACV MVs

This tomogram is a representative for 15 tomograms, and the axis view is referenced with a coordinate system displayed on the left. The upper panel shows a topview in xy projection, and the lower panel displays a side view in xz projection. The tomograms are displayed in the exact same position in both the WBP and the IsoNet corrected tomogram, ensuring a comprehensive comparison.

To address this, we experimented with various techniques, such as the Simultaneous Iterative Reconstruction Technique (SIRT, as implemented in IMOD (Kremer et al. 1996)) and a deconvolution filter (implemented in IsoNet (Liu et al. 2022)), to improve the contrast and visualize the structural features better. Details of the data acquisition and the data processing are described in chapter 3.5.1

Another major challenge was overcoming the anisotropy caused by the missing wedge. To address this problem, I employed the deep learning image processing software IsoNet (Liu et al. 2022) and successfully reconstructed the missing wedge information, and improved the SNR (Figure 12). Although VACV mature virions are relatively large, measuring approximately 360 x 250 x 220 nm (Cyrklaff et al. 2005, Hernandez-Gonzalez et al. 2023), we obtained tomograms that revealed detailed structural features.

In Figure 13, an intact mature virion of VACV is observed in the top view (xy) shown in the upper panel and side view (xz) shown in the lower panel, displaying a brick-shaped morphology containing a dumbbell-shaped core. The inner core wall (annotated in pink) is clearly visible, and the core surface is covered with spikes of the palisade layer, consistent with findings of previous studies (Cyrklaff et al. 2005, Hernandez-Gonzalez et al. 2023).
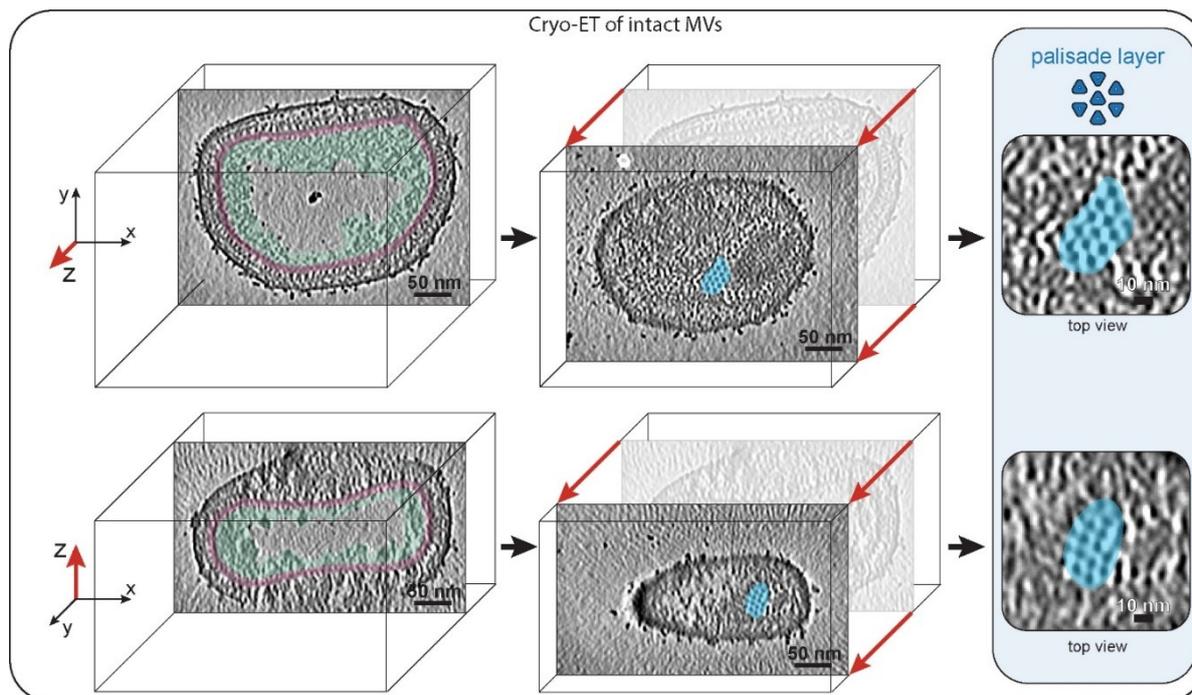
Figure 13: Cryo-ET tomogram of intact VACV MVs

The tomogram displays an intact mature virion in topview xy (upper panel) and sideview xz (lower panel); the coordinate system is displayed on the left. The tomogram, one of 15 aquired tomograms, is representative and is missing wedge corrected with IsoNet. The tomogram is shown in computational slices of 1.1nm thickness. The center panel shows the palisade layer in the tomogram, and a magnified view can be seen in the left panel. Structural features such as the inner core wall (pink), the DNA (green), and the hexameric arrangement of the palisade layer (blue) are annotated.

Image source: adapted from (Datler et al. 2024) , License: CC BY 4.0

The core interior is almost empty except for the condensed viral genome (annotated in green) just beneath the core wall. The central panel of Figure 13 highlights the palisade layer and its locally ordered arrangement within the tomogram, with a more detailed view in the right panel, where an ordered hexameric arrangement of the spikes can be seen.

4.3.1 Utilization of template matching to pick proteins of the palisade layer

To perform subtomogram averaging of the palisade layer, particles of this structural feature must be selected from the MVs in the tomograms, extracted into subvolumes, aligned and averaged. Different procedures for particle selection are possible. For instance, methods such as manual particle selection, mesh depiction of the core surface, or a template-matching approach can be utilized.

Manual selection of features is a laborious task and not sustainable for a large amount of data. Therefore, only a small number of particles were selected and used to create an initial reference with subtomogram averaging. This reference served as a template in a template-matching approach, utilizing the software Dynamo (Castaño-Díez 2017).

The software uses systematic cross-correlation of the template against the tomogram, producing a cross-correlation map. Each pixel in the map represents a score that shows the neighborhood's similarity to the template. Despite experimenting with different cross-correlation values, cone rotation angles and in-plane rotation angles, the software could not detect all spikes on the palisade layer surface. The selected particles mainly represented the top and bottom of the virus core wall; therefore, the side views were completely missing. This created an anisotropic effect because the particles were predominantly from specific orientations, while other critical orientations such as side views were missing. In subtomogram averaging, particles from diverse orientations are essential for constructing a high-resolution 3D structure. After multiple unsuccessful attempts, I decided to explore a different approach to select the particle positions.

4.3.2   Mesh depiction of the core wall of intact mature virions

After the template matching failed to produce the desired results, I used another feature of the Dynamo software (Castaño-Díez 2017) to pick particles in 3D. Initially, I used the program's graphical user interface (GUI) to manually define points of a grid following the surface of the core in xy and xz planes to overcome the issue of anisotropy caused by missing side views (see Figure 14). As mentioned earlier, selecting points in the xz plane was only possible after correcting the tomograms using IsoNet (Liu et al. 2022). In the GUI, the center of the mesh was defined to assign an outward-facing angle to the selected particles. Upon manual point selection, the program generated initial control points and a depiction mesh that could be refined to achieve a specific number of cropping points (Figure 14b). We estimated the expected number of particles on one viral core by calculating the average surface of the core and dividing it by the average size of the spikes, including the gaps (approx. 2.500 particles). To ensure that all of the spikes of the palisade layer were included, we oversampled them by 3, resulting in the cropping of approximately 7.500 particles per viral core.
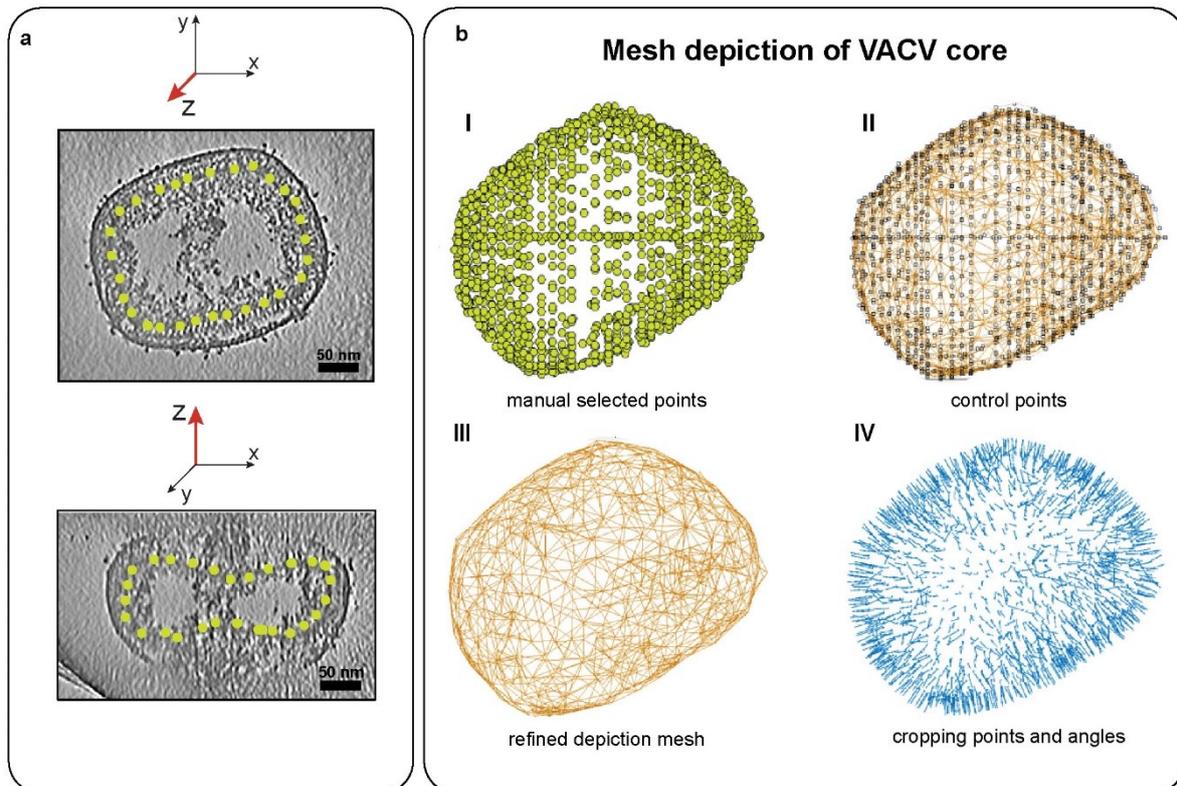
Figure 14: Mesh depiction of the VACV core wall of intact mature virions

(a) The Isonet corrected tomogram is displayed in xy and xz planes as indicated with the arrows, and annotated with manually selected cropping points. (b) The mesh depiction workflow illustrates the process of manually selecting points (I), creating computational control points (II), refining the depiction mesh (III), and determining the final cropping points and angle (IV). The mesh depiction was carried out using the Dynamo software.

4.3.3 Subtomogram averaging of the VACV core palisade layer in intact MVs

The obtained particles of the spikes of the palisade layer, selected with mesh picking, were initially cropped from the IsoNet-corrected tomograms and were subjected to subtomogram averaging alignments to obtain an initial reference. Then, the positions defined by the mesh were cropped from the bin8 tomograms and the structure derived from IsoNet was used as a reference for further alignment projects. Subsequently, alignments were performed in bin4, followed by alignment in bin2 using half-datasets (Figure 15a). During the binning process, the alignment masks were progressively reduced in size to focus more precisely on the region of interest. This procedure improves the alignment and the resolution of the palisade layer spikes. For more details on the processing, please refer to chapter 3.6.5 in Material and Methods.
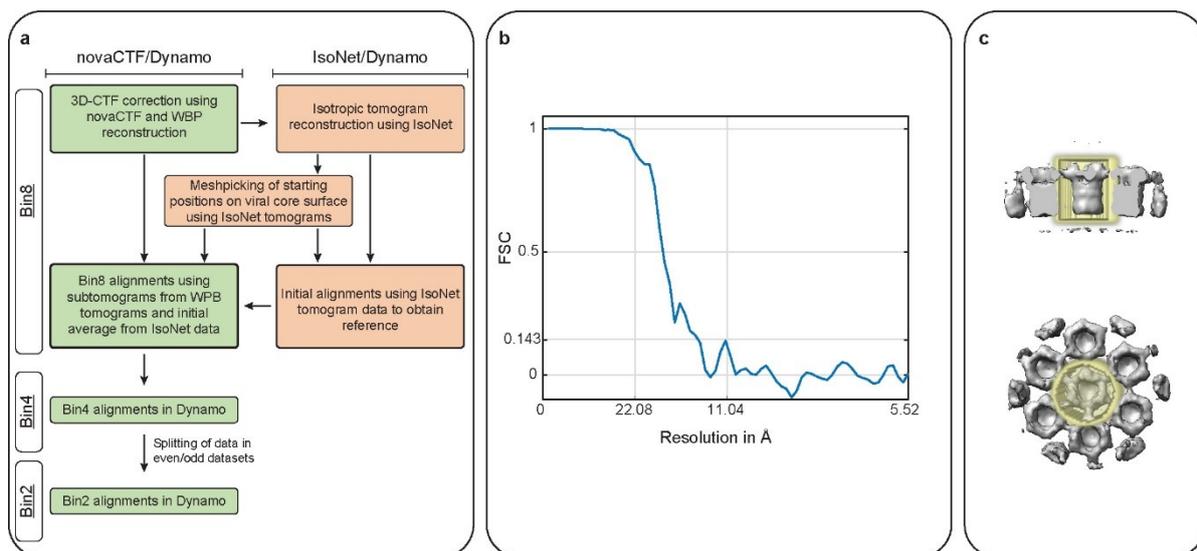
Figure 15: Workflow of cryo-ET processing and FSC curve of the VACV core wall in intact MVs

(a*) The workflow of subtomogram averaging of the core in intact MVs using novaCTF, mesh picking in IsoNet corrected tomograms and creation of an initial reference in Dynamo. (b) Fourier Shell correlation (FSC) created of half datasets of the final structure of the palisade layer (Figure 16) in bin2 shows a measured resolution of 13.1 Angstrom at the 0.143 criteria. (c) FSC mask used for FSC curve measurement.*

Image source: adapted from (Datler et al. 2024) , License: CC BY 4.0
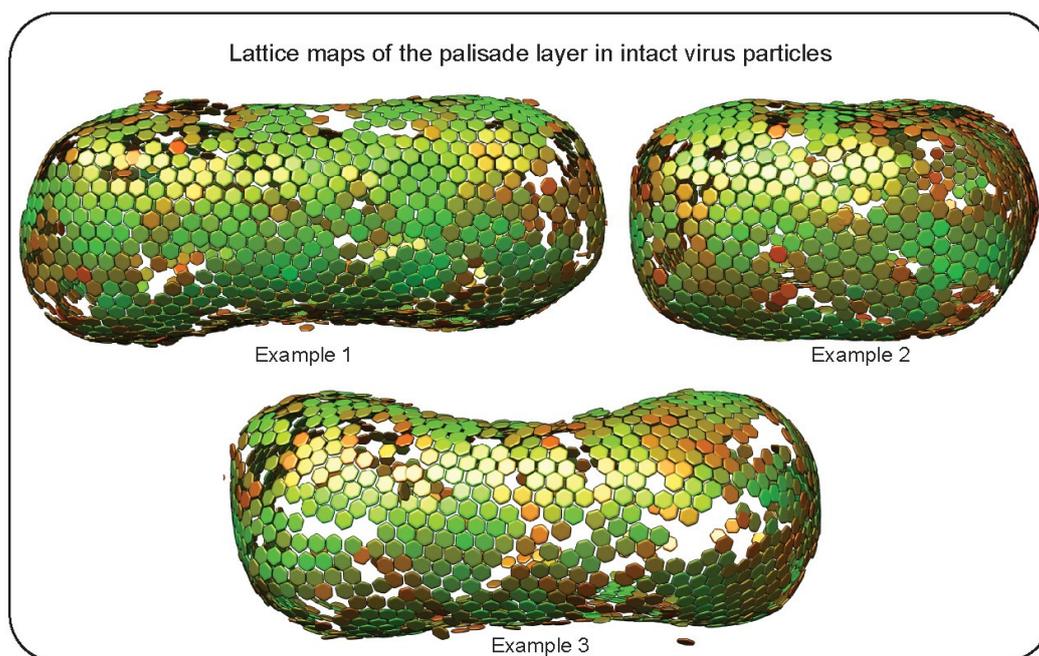
Figure 16: Lattice maps of the palisade layer in intact virus particles*.*

Three examples of the viral core lattice after final alignments are shown. The particles are represented as hexamers, with each hexamer indicating the central trimer of a hexagon. The particles are color-coded based on their cross-correlation coefficient (CCC), ranging from green (high CCC) to red (low CCC).

Image source: adapted from (Datler et al. 2024) , License: CC BY 4.0

42

The final structure at a global resolution of 13.1 Angstrom at the palisade layer revealed it to be composed of hexamers of trimers, consistent with previous suggestions (Cyrklaff et al. 2005, Hernandez-Gonzalez et al. 2023). Despite visualizing the overall arrangement, the achieved resolution was inadequate to identify the proteins forming the trimer in the palisade layer or the proteins of the inner core wall (Figure 15b). The mask used for FSC measurement is displayed in Figure 15c. The inner core wall is also visible in the cryo-EM density but does not exhibit an ordered arrangement. This suggests that the inner core wall's symmetry differs from the palisade layer's. The envelope of the MVs was visible in the structure, but only poorly defined, probably due to the varying distances between the core and the envelope, influenced by the dumbbell shape of the viral core (Figure 17).
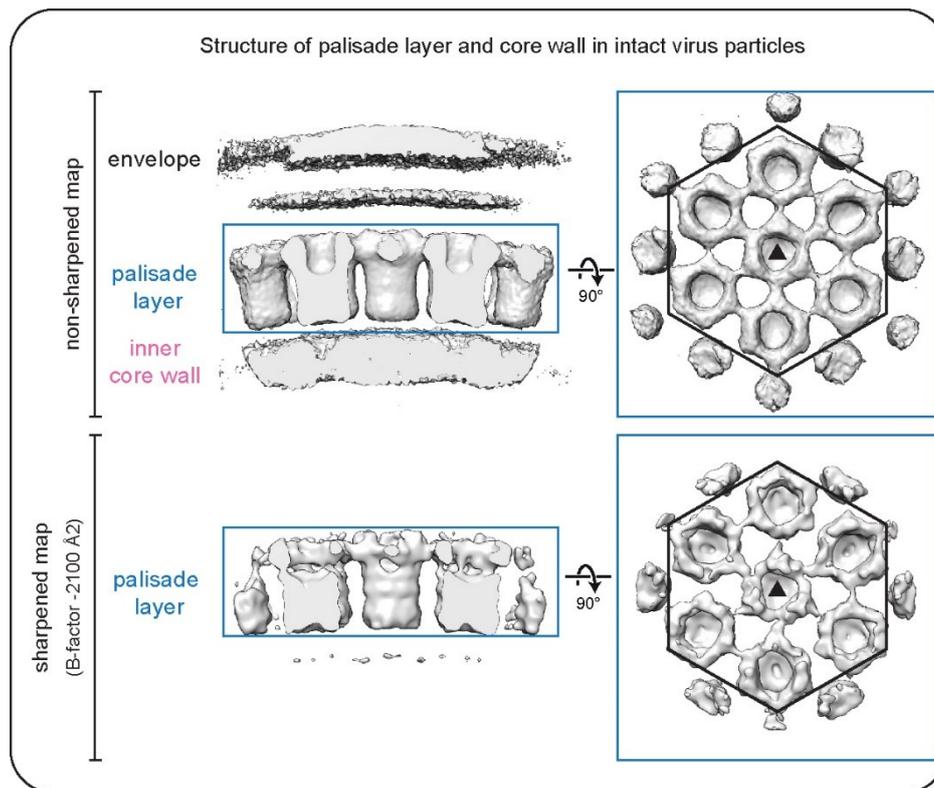
Figure 17: Structure of the palisade layer and core wall in intact virus particles

The non-sharpened EM density map in the upper panel displays the envelope, the palisade layer (blue) and the inner core wall (pink). The lower panel shows the sharpened EM density map of the palisade layer (blue). On the right side, turned in a 90° rotation, the top view of the palisade layer of the unsharpened (top) and sharpened map (bottom) is displayed annotated with a hexamer, indicating the hexamers of trimers and the triangle to indicate the central trimer.

Image source: adapted from (Datler et al. 2024) , License: CC BY 4.0

When back-projected on the original tomograms, the positions of the trimers showed a semi-continuous lattice around the viral cores, with gaps, cracks and local symmetry patches (Figure 16). We observed an average of 2280 trimers per viral core (s.d. +-309, n=15), not accounting for gaps and cracks. Similar to other viruses with incomplete hexagonal lattices like Gag protein in immature retroviruses (Mattei et al. 2016) and poxvirus D13 pseudohexagonal lattice (Hyun et al. 2011, 2022), pentamer formation that could allow for complete closure of the hexameric lattice was not found.

4.4   Cryo-ET of isolated MV cores reveals their complexity

The resolution achieved with subtomogram averaging of the intact VACV virions was insufficient to identify the proteins constituting the different layers of the core. Therefore, we adapted the sample preparation to render the specimen less complex for improved data acquisition. Following a method described by others (Dubochet et al. 1994, Moussatche and Condit 2015), we removed the viral envelope using DTT and NP40.
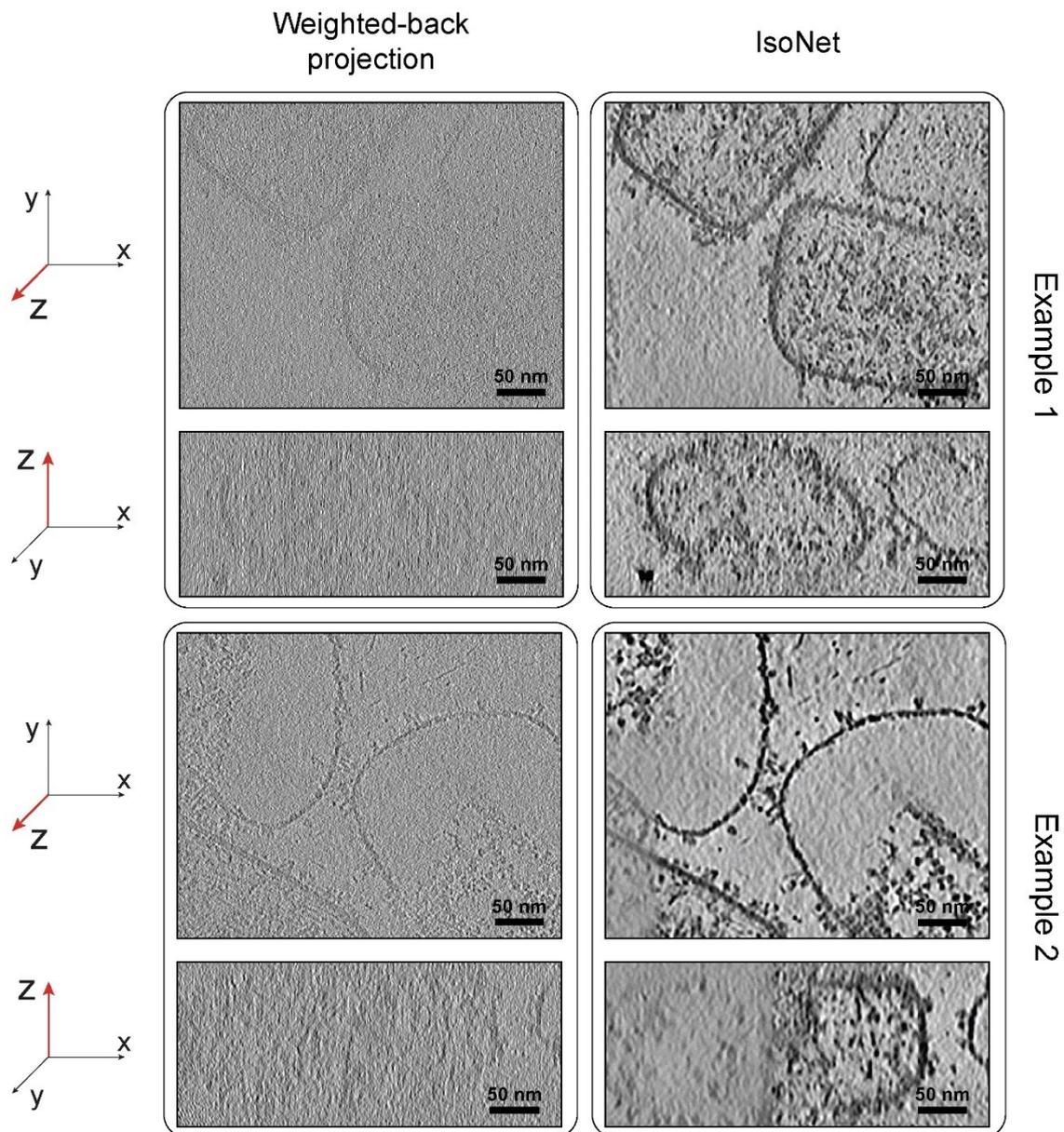
Figure 18: Comparison of the weighted-back projection (WBP) tomogram and IsoNet corrected tomogram of isolated VACV core MVs

Two different examples of VACV MV cores are shown with the WBP tomogram (left) and the IsoNet corrected tomogram (right). The tomograms are displayed in the exact same position in both the WBP and the IsoNet corrected tomogram, ensuring a comprehensive comparison. The axis view is referenced with a coordinate system displayed on the left.

44

After optimizing this viral core purification protocol (described in chapter 3.4.2), we acquired tilt series using a Thermo Fisher Scientific Titan Krios. The tomograms were reconstructed and CTF-corrected with novaCTF (Turoňová et al. 2017). We used IsoNet (Liu et al. 2022) to increase the SNR and compensate for the missing wedge to better characterize the individual structural features. Figure 18 shows two examples of isolated viral core tomograms with WBP tomogram (left) and IsoNet-corrected tomogram (right). The coordinate system on the left indicates the orientation of the tomogram. The viral cores appeared to be partially collapsed (Figure 18, Example 2), and in the intact cores, the DNA was not condensed underneath the core wall as seen in the intact MVs. Instead, the core structure appeared relaxed, and the DNA was visible throughout the core (Figure 18, Example 1).

The isolated and partially broken viral cores allowed us to observe distinct structural features in the tomograms (Figure 19). The three-dimensional (3D) nature of the tomograms also enabled us to identify the location of the structures within the viral core wall. We observed a pseudo-hexameric arrangement in the palisade layer, similar to what we had seen in the intact MVs. As reported earlier (Dubochet et al. 1994), due to the core isolation protocol, the proteins of the palisade layer were partially shed off, providing the opportunity to observe side views where single proteins were still attached to the core and looked like spikes. In the top view, these spikes exhibited a more trimeric formation, suggesting the protein to be a trimer. We were also able to identify the regular arrangement of the inner core wall in both side and top views. In the top view, the inner core wall appeared to consist of square-like shapes measuring 7.4 nm x 7.4 nm with at least a 2-fold symmetry arranged in a regular pattern. This confirmed our earlier conclusion from intact MVs that the inner core wall follows a different symmetry compared to the palisade layer.
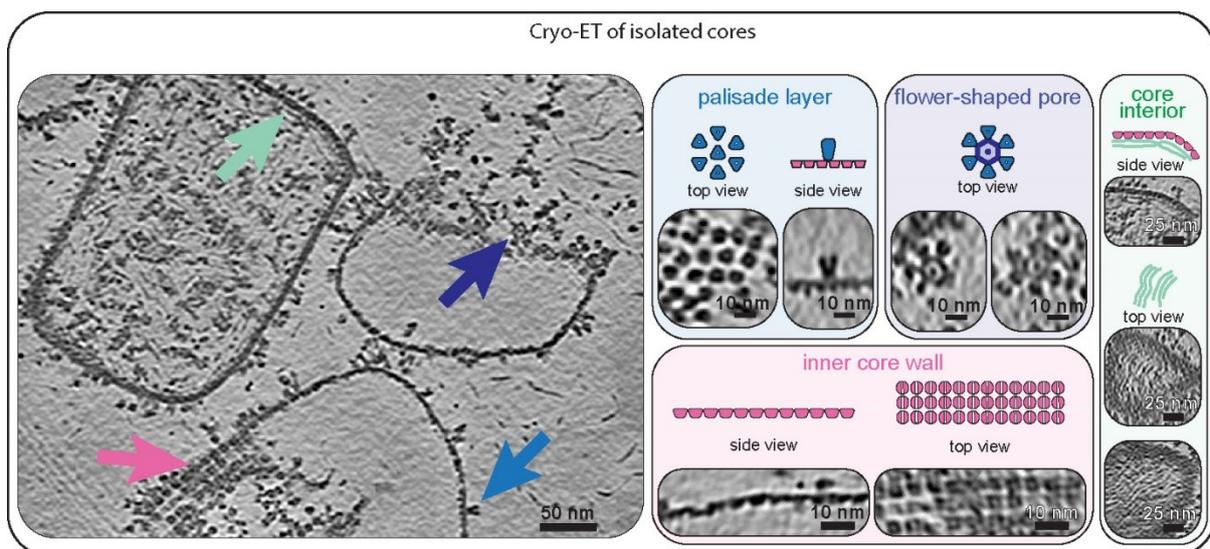
Figure 19: Cryo-ET of isolated MV cores

Representative IsoNet corrected tomogram of isolated MV cores shown as computational slices with 1.1 nm thickness. Annotated notable structural features include the palisade layer (blue) with a hexameric arrangement, the symmetric pattern observed in the inner core wall (pink), a flower-shaped pore (purple), and visible DNA (green) within the core interior.

Image source: adapted from (Datler et al. 2024) , License: CC BY 4.0

Additionally, embedded in the palisade layer, we observed flower-shaped pores with a 29 nm diameter, consisting of up to 6 petals, and an inner circular lumen with an 11 nm diameter and a density inside. These pores could be the core wall pores previously reported in negative stain EM images (Pedersen et al. 2000) and low-resolution cryo-ET (Cyrklaff et al. 2005, Moussatche and Condit 2015, Hernandez-Gonzalez et al. 2023). In side views, we also observed DNA inside the viral core underlying the inner core wall. Due to most of the viral cores being partially broken or deformed, we deemed performing subtomogram averaging by drawing meshes around the viral cores as done with the intact virus not to be the most promising approach.

4.5   SPA of isolated VACV cores reveals a treasure chest of structures

The obtained tomograms of isolated cores were already of a very high quality, but subtomogram averaging of the palisade layer was not possible due to the disruption of the viral cores and the inability to pick positions around them because of their distortion. Therefore, we aimed to use the viral core purification sample for cryo-SPA to obtain high-resolution structures. Cryo-SPA averages 2D images of the proteins in different orientations, reducing noise and improving 3D accuracy. Unlike cryo-ET, it also avoids the missing wedge artifact. We increased the input virus fourfold before purifying viral cores to ensure enough particles for acquisition. Additionally, we added KCl to better separate the cores from each other and to shed off individual proteins from the palisade layer (described in detail in chapter 3.4.2). The high concentration of KCl also reduces electrostatic interactions, stabilizes the virus particles by balancing the ionic environment and increasing the viral cores' solubility. We acquired two SPA datasets on the Thermo Fisher Scientific Titan Krios operated at 300 kV, one of the isolated viral cores and one of the soluble fraction of the purification that was subjected to filtration upon vitrification.

In the micrographs, flower-shaped pores were already visible to the eye, and the trimers of the palisade layer could be seen on the core wall (Figure 20a), as well as proteins in the soluble fraction (Figure 20b). The micrographs were motion-corrected with RELION 4.0 (Scheres 2012) and the CTF estimation and subsequent particle picking, extraction, 2-dimensional (2D) classification and 3D-refinement were performed in CryoSparc 4.0.0 (Punjani et al. 2017) (Figure 21a). The results of the 2D classification of the datasets revealed a structural treasure chest of the VACV core. We identified trimers of the palisade layer in the side and top views as individual proteins in the soluble fraction and several trimers as multimeric assemblies as part of the viral core. The classes also exhibited the trimer still attached to the inner core wall, with the core wall being less defined than the trimers. This observation again suggests that the inner core wall adheres to a different symmetric order than the palisade layer. Interestingly, the inner core wall classes displayed a consistent dimeric pattern in the side and top views.

The flower-shaped pore flanked by the petals consisting of trimers of the palisade layer was exclusively found in the top or slightly tilted view. Side views of the pore might be present, but we were not able to find them in our sample. In addition to the assigned structures, we found soluble tetrameric and pentameric structures that could originate

from the viral cores or as a co-purified contaminant. The identity of these structures could not be assigned (Figure 20c).
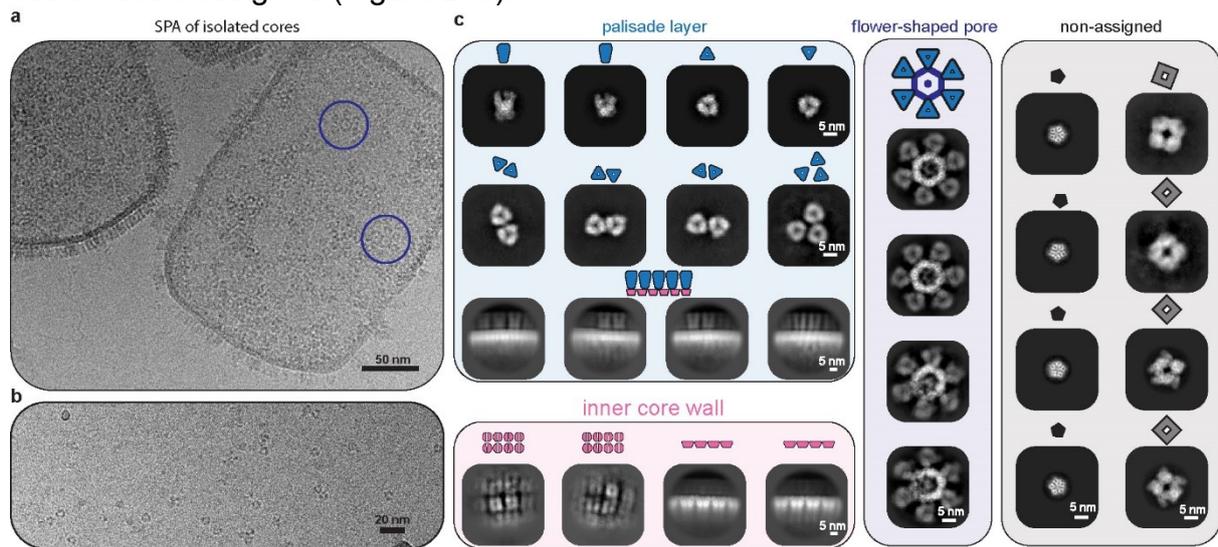
Figure 20: Treasure chest of the VACV core

Cryo-SPA acquired micrographs (representative of 9,264) that show the isolated cores (a) and the soluble core proteins (b) The flower-shaped pore is encircled in purple. (c) Gallery of different 2D classes obtained for cryo-SPA processing. The trimers of the palisade layer (blue) the inner core wall (pink), the flower shaped pore (purple) and non assigned structures (grey).

Image source: adapted from (Datler et al. 2024) , License: CC BY 4.0

4.6   The palisade layer is composed of trimers of A10

The dominant feature in our SPA dataset was the trimer of the palisade layer, for which we obtained high-quality classes from various angles in the 2D classification. We first utilized 3D refinement in RELION (Scheres 2012) and obtained a global resolution of 4.2 Å. To achieve this, we iteratively removed duplicates, carried out multiple 2D classifications, and employed 3D masked auto-refine with C3 symmetry. The detailed processing steps can be found in Figure 21a.

Subsequent processing with Phenix density modification (Terwilliger et al. 2020) improved the global resolution to 3.8 Å, enabling the identification of bulky sidechains and the alpha-helical pitch, as depicted in Figure 22.

The EM density was of sufficient quality to enable fitting of our previously generated AlphaFold models of the putative structural core proteins. This confirmed that A10 (residues 1- 599) fits into the experimentally derived density as a trimer. These results align with a recently published study that suggested the trimerization of protein A10 using XL-MS and modeling (Mirzakhanyan et al. 2023).

47
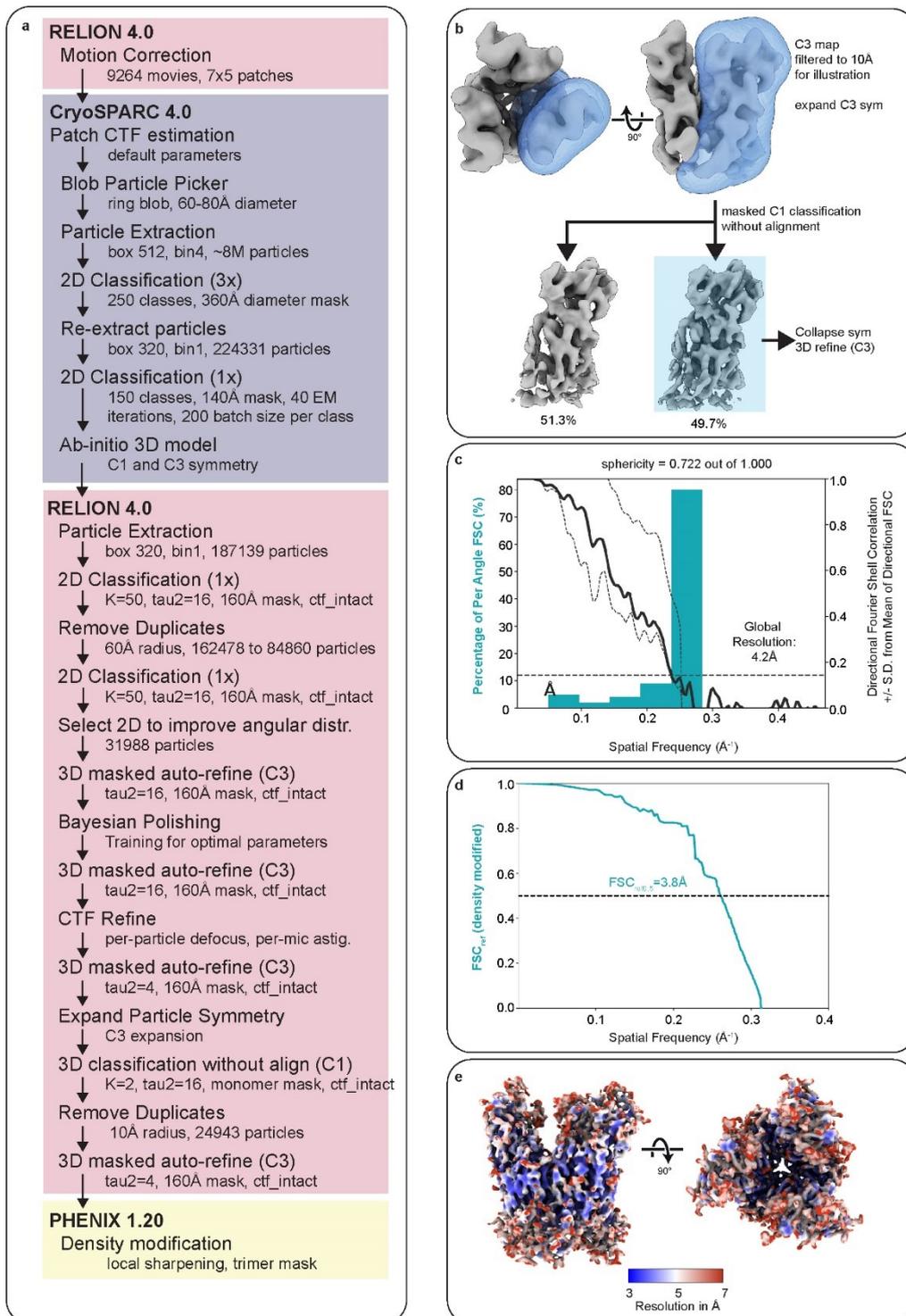
Figure 21: Cryo-EM SPA workflow for the A10 Trimer

(a) Detailed A10 SPA processing workflow. (b) The 3D classification scheme was employed to eliminate particles with low-quality asymmetric units within the A10 trimer. A blue mask was utilized for classification, and the particles designated within the blue box were chosen for the final refinement. (c) The 3D FSC calculations were performed on the masked Relion half maps. In the cyan histogram, the fraction of particles reaching the corresponding resolution can be seen, while the black curves display the global FSC with extensive angular sampling, along with the FSC +/− SD. The global resolution indicated is at the FSC 0.143 cutoff. (d) The FSCRef calculation of the Phenix density modified map with a cutoff value 0.5 of the estimated resolution. (e) A10 trimer density map showing the local resolution.
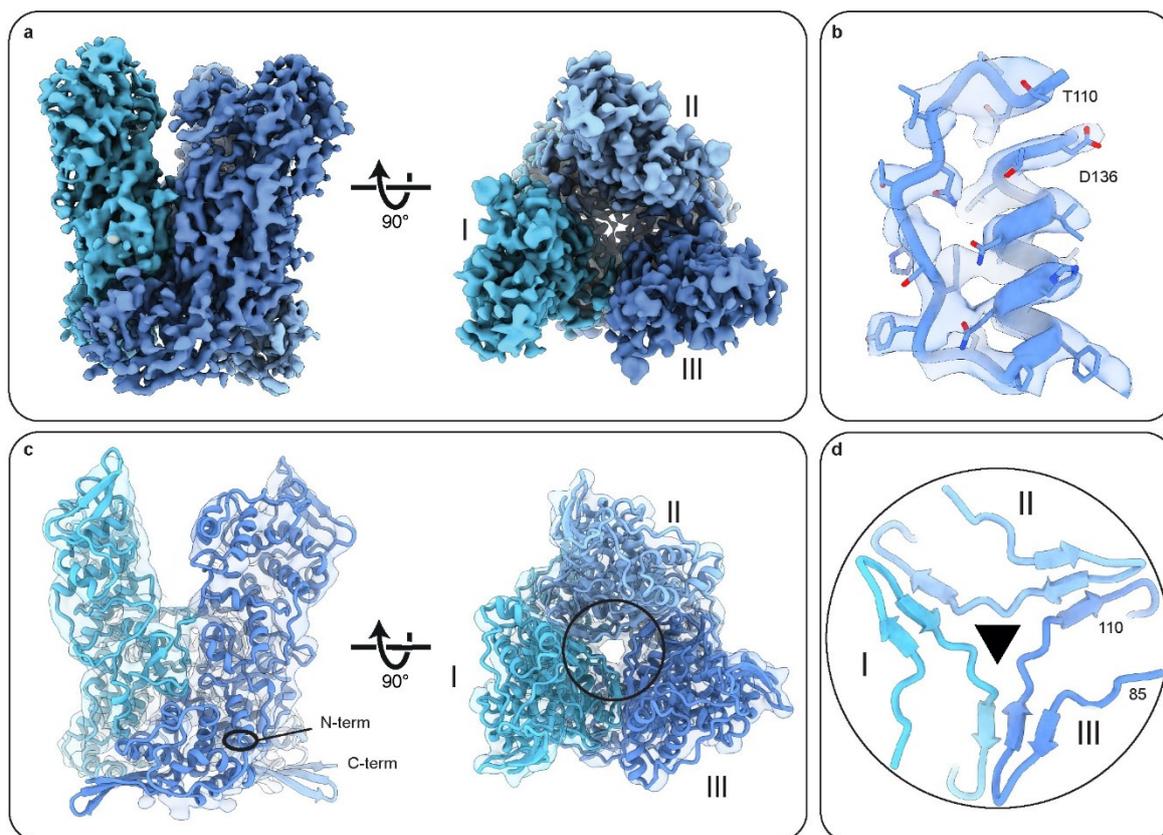
Figure 22: Cryo-SPA structure of the A10 trimer

(a*) C3 symmetrized cryo-EM density map with a 3.8 Å resolution shown as side view and 90° tilted as top view. (b) The primary protein sequence in the core region of the A10 trimer inserted into the density map shows fitting side chains that verify protein identity. (c) Cryo-EM density map of the A10 trimer with fitted refined model. In (d), the central trimer contacts from (c) (shown as a circle) are illustrated, indicating residues 85–110. These residues are involved in hetero-oligomer beta-sheet interactions with adjacent monomers. The N terminus and C terminus of the displayed protein region are labeled with their respective residue numbers. Individual monomers of the A10 protein are annotated in different shades of blue and are marked with the numbers I, II and III. (a-d)*

Image source: (Datler et al. 2024) , License: CC BY 4.0

4.7   Extensive interactions stabilize the A10 trimer

Next, we analyzed contacts and their properties within the A10 trimer to understand the stabilizing interactions that form this key unit of the palisade layer. To this end we used the online resource PDBePISA server (Krissinel and Henrick 2007) to compute properties and pinpoint potential key contacts at the A10 trimer interface (Figure 23). Due to the resolution of around 4 Å, specific structural characteristics, such as small or negatively charged side chains, may not be clearly discernible, thus limiting interpretability. Our investigation indicated that hydrophobic interactions dominate the 2,104-Å2 buried surface area per protomer pair (Figure 23). Additionally, there are several conserved inter-chain salt bridges binding central alpha helices within the trimer (Figure 24a). Notably, each pair of protomers constructs a heterodimeric three-stranded beta-sheet (residues 85–110, Figure 22d), comprising two strands from one monomer and one strand from the neighboring monomer. Furthermore, a hydrogen bond network on the outward surface of the beta-sheet fortifies this interaction, while

the opposite side of the beta-sheet tightly accommodates underlying hydrophobic side chains (Figure 24b).
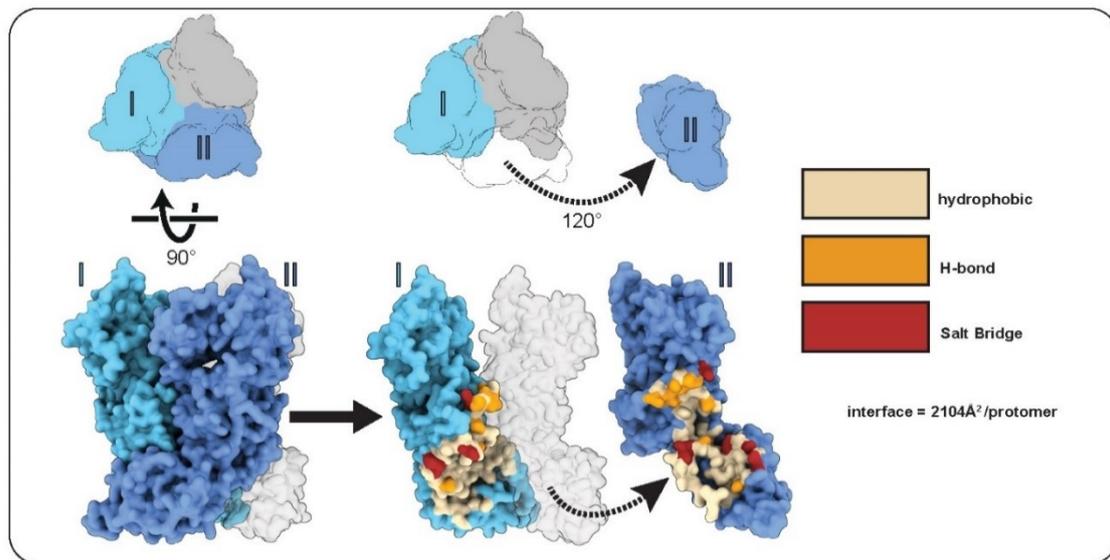
Figure 23: A10 trimer is stabilized through extensive interactions

The model illustrates the crucial inter-chain contacts at the oligomerization interface between two A10 monomers labeled as I and II. The model shows monomer II pulled away by 120° to reveal the underlying interactions. Hydrophobic interactions are depicted in yellow, hydrogen bonds in orange, and salt bridges in red.

Image source: adapted from (Datler et al. 2024) . License: CC BY 4.0
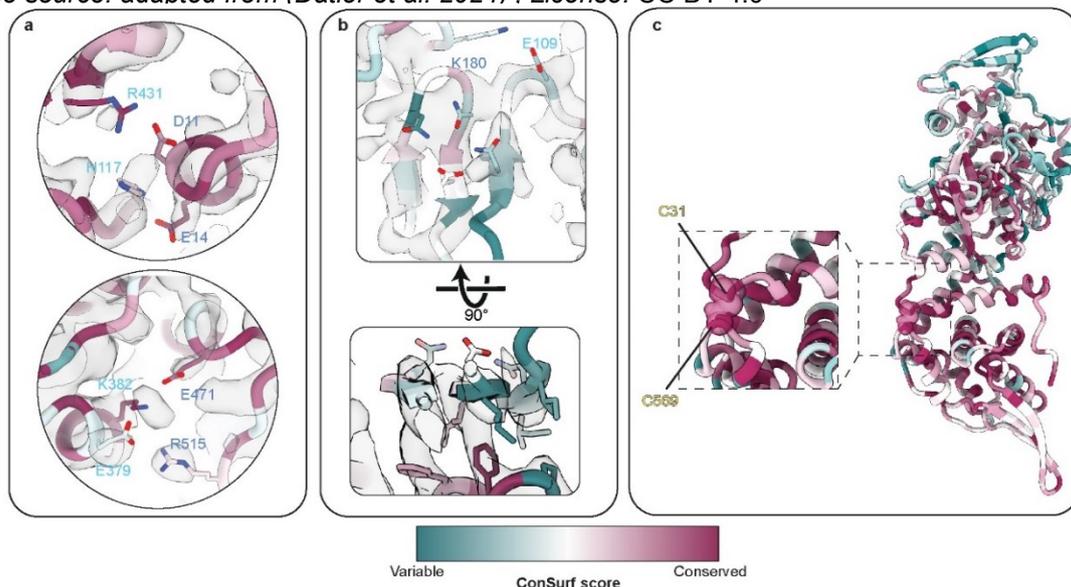
Figure 24: Inter- and intramolecular interactions of the A10 trimer

(a) Detailed view of salt bridges that were already shown in Figure 23, annotated with residue numbers. (b) The top view displays a salt bridge and a hydrogen bond network within the inter-subunit beta-sheet located at the core of the trimer. The bottom view illustrates the hydrophilic packing underneath the beta-sheet. EM – density is shown for all images. (c) C31 and C589 are potentially interacting cysteines; the orientation of the monomer is identical to that shown in Figure 23. The atoms of the cysteine residues are displayed as semitransparent spheres. The Color code according to conservation of the residues analysed with Consurf analysis from green (variable) to red (conserved)

Image source: adapted from (Datler et al. 2024) , License: CC BY 4.0

50

Earlier findings suggested that VACV core proteins form disulfide bonds within MVs, enhancing the stability of released virus particles (Locker and Griffiths 1999). Upon examining our A10 model, we discovered two highly conserved cysteine residues (C31 and C569) positioned closely together (Figure 24c). These residues potentially link the N-terminal and C-terminal domains of the A10 monomer together, helping to stabilize its overall configuration.
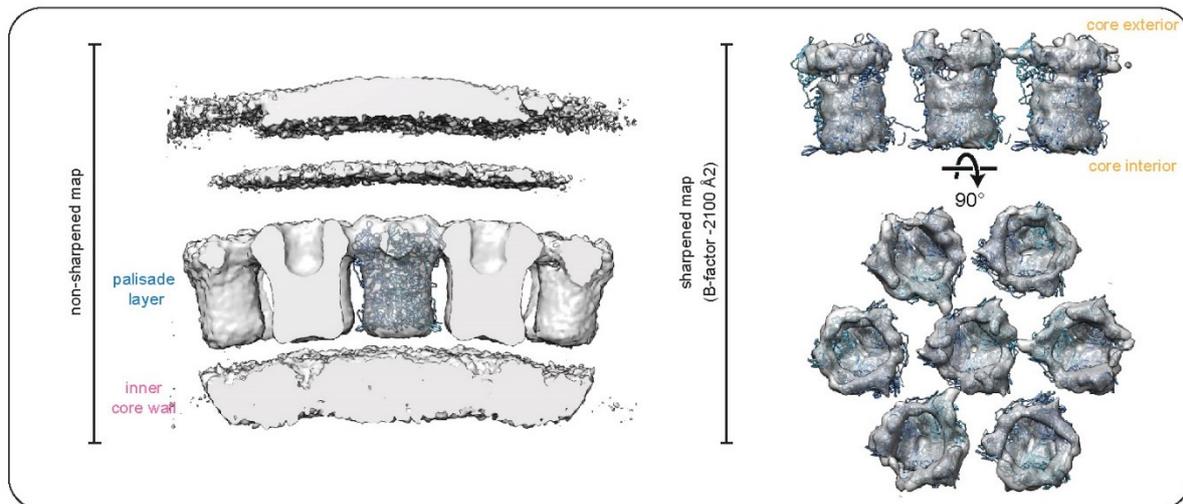
Figure 25: A10 trimer fits into the STA result of whole MVs

The A10 structure was fitted into the subtomogram averaging results of the palisade layer of whole MVs using a rigid body fit. On the left side, there is a side view of the non-sharpened subtomogram averaging map, with the trimer structure fitted into the center of the palisade layer. On the right side, the sharpened subtomogram averaging map is displayed, with the fitted A10 structures shown in side and top views. The density of the inner core wall and the envelope was removed for better visibility.

Image source: adapted from (Datler et al. 2024) , License: CC BY 4.0

4.8 The palisade layer forms weak interactions between trimers

To gain a more thorough understanding of the positioning of the A10 trimer within the VACV core wall, we attempted to fit the A10 trimer model into the subtomogram averaging density previously obtained from the cryo-ET dataset of the intact MV viruses (Chapter 4.3.3) (Figure). At the resolution of our structure, the trimer model occupied the density without any noticeable unoccupied areas (Figure 25). This strongly indicates that most of the palisade layer is composed of A10 trimers. The A10 trimers are arranged in a pseudohexagonal pattern. However, as observed in the subtomogram averaging results of the cryo-ET data (Figure), the trimers show substantial spacings between them, suggesting weak lateral interactions. This is in line with the lattice maps of the intact MV cores shown in Figure 26a, which reveal gaps and cracks and show different positionings of the trimers with respect to each other.

The 2D classification from the cryo-SPA dataset displays several classes with trimers oriented differently to each other (Figure 26b). This suggests that the core's stability relies not only on the palisade layer but also on the inner core wall, or other proteins that form interactions above the trimer.
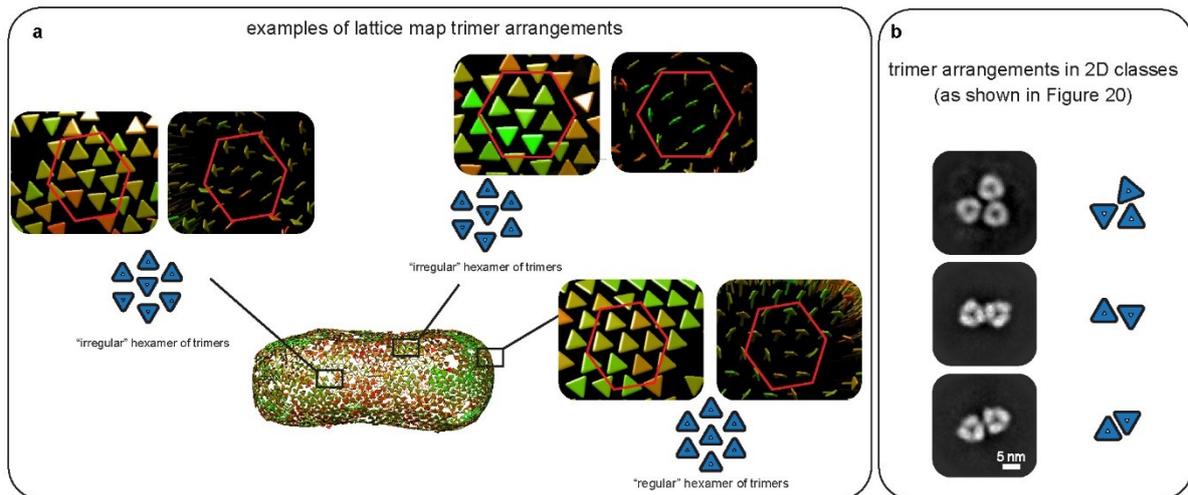
Figure 26: A10 trimer arrangement is variable within the palisade layer

(a) *A lattice map of the viral core's palisade layer, obtained from cryo-ET data of whole MV virions, shows the particles' final positions after subtomogram averaging at different locations around the viral core. The trimer arrangement within the hexamer varies in position relative to each other but is not dependent on the lattice curvature. Trimers are depicted as triangles, shown in z and y orientation. The color of the triangles indicate the cross-correlation coefficient (CCC) of the alignment, ranging from red (low CCC) to green (high CCC). (b) Furthermore, different 2D classes of cryo-SPA show the trimer arrangement of trimers released from the core wall, as depicted in Figure 20 These trimers exhibit different orientations and the schematic representation illustrates them.*

Image source: adapted from (Datler et al. 2024) , License: CC BY 4.0

4.9   A10 trimers are likely conserved in ortho- and parapoxviruses

Protein A10 is highly conserved among members of the *Orthopoxvirus* genus, but shows less conservation in more distantly related poxviruses such as parapoxviruses and entomopoxviruses. The A10 protein of VACV WR shares more than 97% sequence identity with other viruses of the *Orthopoxvirus* genus, such as variola virus, monkeypox virus, rabbit poxvirus and cowpox virus. Other members of the poxvirus family, for example, the P4a protein of Orf virus, which belongs to the parapoxvirus genus, shares 40% sequence identity with an alignment length of 511 base pairs. entomopoxviruses exhibit even lower sequence conservation than parapoxviruses, with *Amsacta moorei entomopoxvirus* (AmEPV) and *Melanoplus sanguinipes entomopoxvirus* (MsEPV) showing only about 22% sequence identity. The alignment spans 114 or 145 residues out of 1,149 for the putative core protein AMV139 of AmEPV, and 293 residues out of 1,306 for the putative core protein homolog P4a of MsEPV (Table 4).

Table 4 Conservation of A10 protein within the poxvirus family

Sequence identity of different poxvirus strains compared to VACV Western Reserve.

Table source: adapted from (Datler et al. 2024) , License: CC BY 4.0

Taxonomy	Genus	Protein	GenBank Protein ID	full length protein	alignment length	Sequence identity %
Variola virus	Orthopoxvirus	A10	ABF23487.1	892	615	96.6
Monkeypox virus	Orthopoxvirus	A10	YP_010377118.1	891	614	96.6
Rabbitpox virus	Orthopoxvirus	A10	AAS49831.1	892	615	99.2
Cowpox virus	Orthopoxvirus	A10	ADZ29251.1	891	614	97.6
Ectromelia virus	Orthopoxvirus	A10	NP_671631.1	891	614	96.3
Orf virus	Parapoxvirus	P4a (A10)	NP_957863.1	905	511	39.5
Amsacta moorei entomopoxvirus	Entomopoxvirus	AMV139 (putative core protein)	NP_064921.1	1149	145/114	22.1
Melanoplus sanguinipes entomopoxvirus	Entomopoxvirus	Putative core protein P4a homolog	AAC97675.1	1306	293	21.2

The AlphaFold-predicted structures of orthopoxvirus A10, show remarkable similarity when computationally predicted using AF2 (Figure 27a). Due to lower sequence identity, AlphaFold predictions of *Entomopoxviruses*, such as AmEPV and MsEPV, have an overall more different fold, as can be seen in Figure 27b. A similar fold to the A10 protein of VACV can be exhibited if only the residues that align with sequence identity are compared. Therefore, AmEPV residues 618-850 and MsEPV residues 784-1014 show a similar structure to VACV A10 residues (370-599). For AmEPV, the structural loop between the halves of the fold shows a different angle that leads to a slightly different orientation(Figure 27c).

Based on these analyses, it is apparent that trimer interaction and palisade formation share similarities in *Orthopoxviruses* and *Parapoxviruses*. However, the constitution of the palisade layer could differ in *Entomopoxviruses*. Nevertheless, the base of the palisade layer might be responsible for the conserved interactions with the underlying core layers in all poxviruses, as only the upper part might differ.
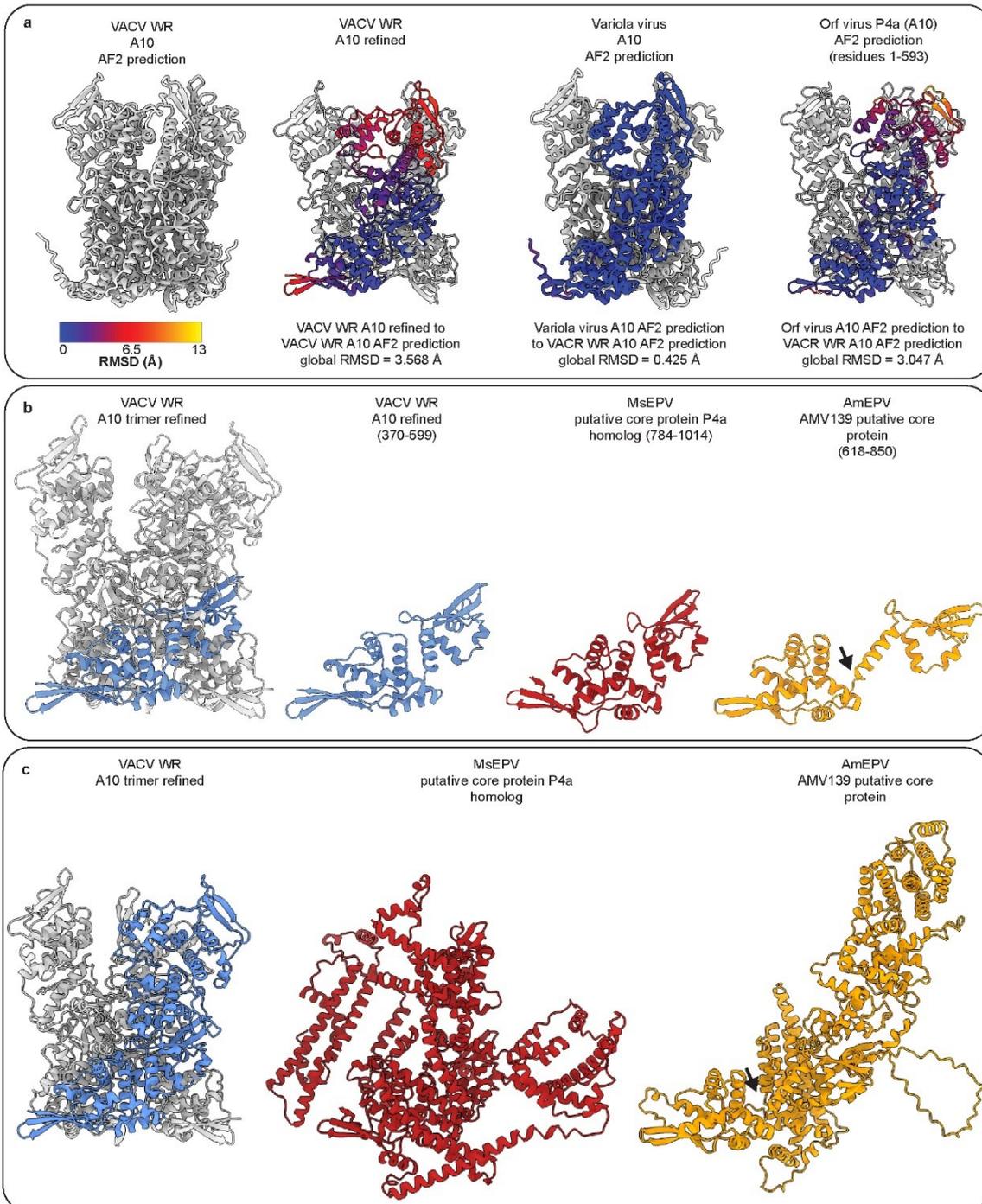
Figure 27: Structural conservation of the A10 protein within the poxvirus family

(a) *The comparison of the original VACV WR A10 AlphaFold prediction structure with the refined VACV WR structure and the AlphaFold prediction of variola virus A10 protein and the parapoxvirus orf virus P4a (A10) protein (residues 1-593). The color code represents the root mean square deviation (RMSD) value, ranging from low to high (blue to yellow). A lower value indicates higher similarity. (b) Comparison of the VACV WR refined A10 structure (residues 370-599) with similar parts of the putative core protein P4a homolog of MsEPV (residues 784-1014) and putative core protein of AmEPV (residues 618-850). (c) Comparison of the proteolytic cleaved monomer of VACV WR refined A10 structure with the whole monomer of the putative core protein P4a homolog of MsEPV and the whole monomer of the putative core protein of AmEPV. In b – c VACV A10 is shown in blue, MSEPV P4a in red and AmEPV AMV139 in yellow for better visualization.*

Image source: adapted from (Datler et al. 2024) , License: CC BY 4.0

4.10 Positioning of core wall proteins in respect to the trimer

To provide a more detailed model of the MV core wall, we mapped the results of the previously published XL-MS of VACV core proteins onto our A10 trimer map shown in Figure 28 (Mirzakhanyan and Gershon 2019). The contact sites between proteins A10 and 23K (represented by yellow spheres) and A10 and the L4 protein (depicted by blue spheres) are located throughout the trimer model without any specific location. However, interactions between A10 and A3 (shown as red spheres) are exclusively found on the side of the trimer model which faces the core interior. Interestingly, the connections between A10 and A4 (indicated by pink spheres) are mainly on the exterior side of the trimer (Mirzakhanyan and Gershon 2019).

To more comprehensively understand the protein-protein interactions, we examined the surface charge maps of our A10 trimer structure and the AlphaFold-predicted structures of the putative core proteins A3, A4, 23K, and L4. The surface charge maps of the trimer show a negatively charged area on the exterior side of the trimer model (annotated with an arrow) and a positively charged side of the A4 protein that is not clearly visible in our surface charging map (Figure 29). Together with the MS cross-correlation results, this could indicate that A4 interacts with A10 on the exterior site. Previous immunogold labeling experiments of purified viral cores detected A10 and A4 at the surface of the core wall, which supports this finding. (Pedersen et al. 2000, Moussatche and Condit 2015)
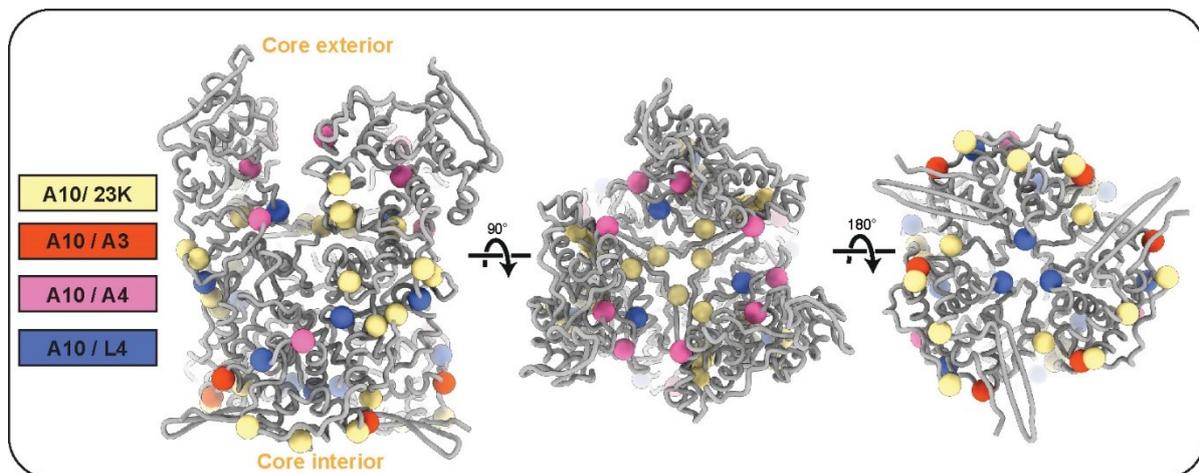
Figure 28: Visualization of XL-MS interactions of the A10 trimer

Contact sites between proteins A10 and 23K are shown as yellow spheres, A10 and A3 as red spheres, A10 and A4 as pink spheres and A10 and L4 as blue spheres. Display of previous results of XL-MS correlations of A10 with structural core proteins. (Mirzakhanyan and Gershon 2019)

Image source: adapted from (Datler et al. 2024) , License: CC BY 4.0

The bottom of the A10 trimer is positively charged, and the predicted A3 protein's surface charge model shows negative charges on one side (Figure 29). This finding is in line with the XL-MS result, which also indicates an interaction site between A10 and A3 at the base of A10 (Mirzakhanyan and Gershon 2019).
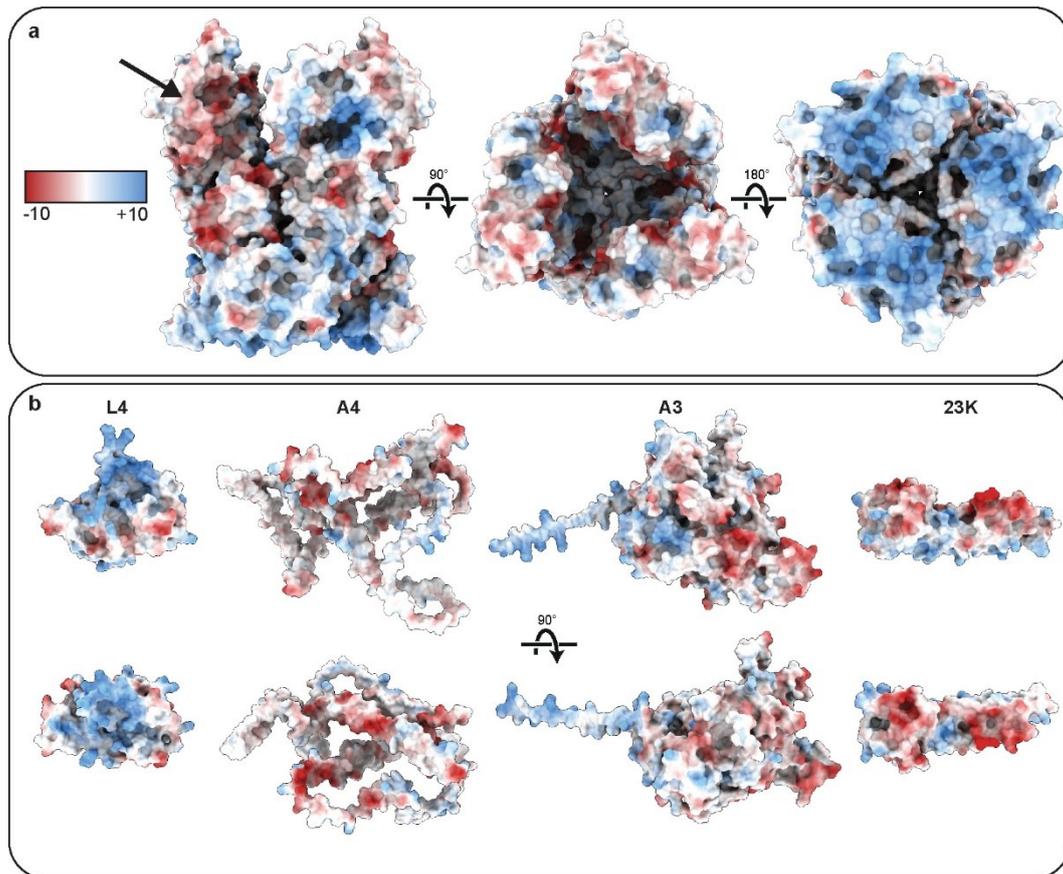
Figure 29: Surface charge map interpretation of VACV core proteins

(a) *Surface charge analysis of the A10 trimer map. Suggested binding site of A4 is annotated with an arrow. (b) Surface charging analysis of AlphaFold predictions of proteins L4, A4, A3 and 23K. a-b.) Negative charge is shown in red and positive charge shown in blue*

Image source: adapted from (Datler et al. 2024) , License: CC BY 4.0

4.10.1   The inner core wall potentially consists of A3 proteins

In the cryo-SPA data, classes were identified that clearly display the inner core wall of VACV (Figure 20, Figure 30a). To better understand the protein interactions in the inner core wall and identify the proteins, we re-examined the inner core wall structure to achieve a high-resolution structure. Due to anisotropy and preferred orientation, a resolution of only 20 Å was reached that was insufficient to identify the protein (or proteins) that make up the inner core wall (Figure 30b).

To identify potential candidates, we used information from indirect observations. The base of the A10 protein, which forms the palisade layer, carries a positive charge. Consequently, the underlying inner core wall must consist of a protein that has a negative charge. AF2 predictions of the putative core proteins showed that protein A3 exhibits a negative charge on one side. Previous studies utilizing immunogold labeling suggested its location is on the inner side of the core wall (Moussatche and Condit 2015). However, upon analyzing the cryo-EM density of the inner core wall, it became evident that A3 must be at least a dimer to fit. Therefore, AlphaFold (Jumper et al.

56

2021) was utilized to predict the dimer structure of protein A3. We further examined the structure using a surface charge map, and the negatively charged patch on one side of the dimer provided insight into the protein's orientation towards A10 (Figure 30c). The dimer of A3 fitted well into the density map, with its N-terminus oriented toward the core interior, as illustrated in Figure 30d. Furthermore, previous XLS-MS results show that the residues crosslinking with A10 face the core exterior (Mirzakhanyan and Gershon 2019). However, this fit remains speculative and will require further experimental validation for confirmation.
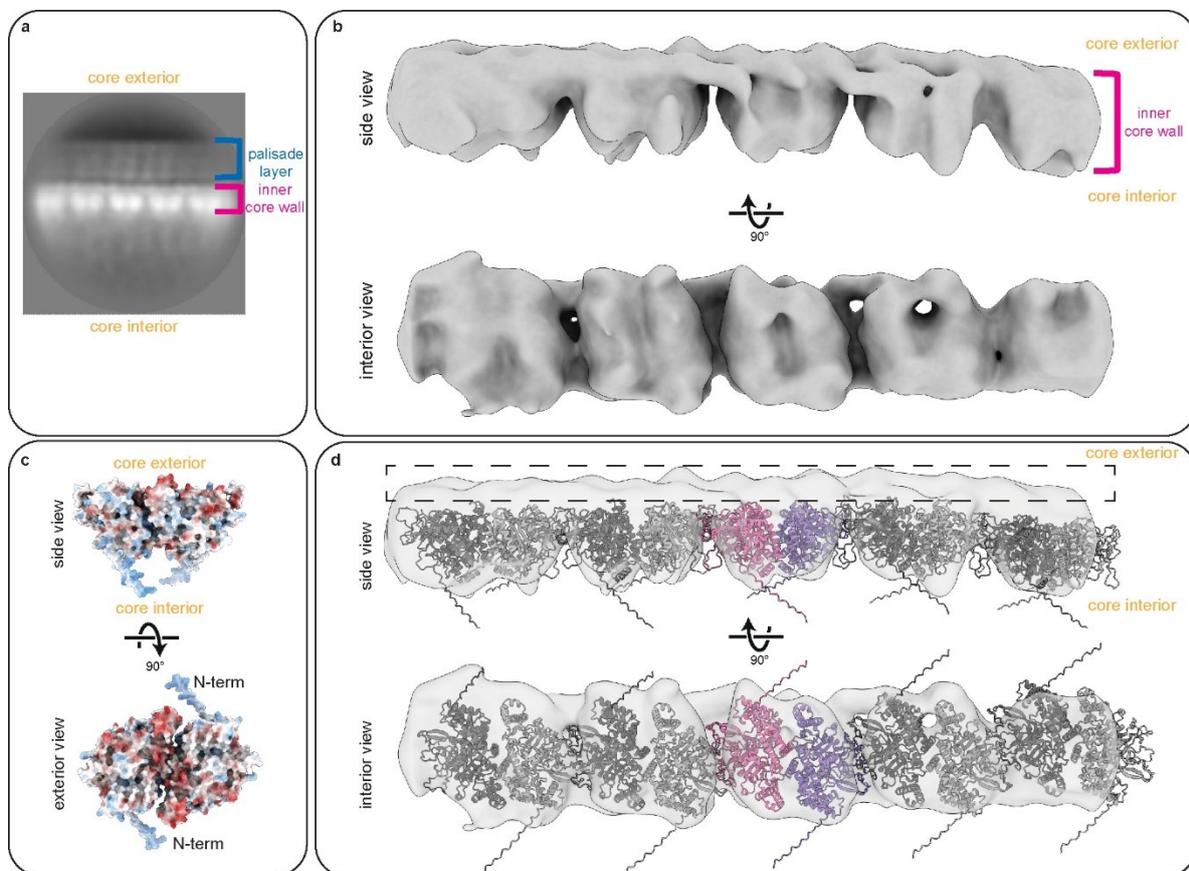
Figure 30: Cryo-EM density of the inner core wall with a fitted dimer of A3

(a) *2D classification of cryo-SPA data with annotated palisade layer and inner core wall. (b) low-resolution cryo-EM density map of the inner core wall in side view (top) and top view (bottom) (c) Alphafold predicted structure of the dimer of A3 protein with surface charging annotation. Same color scheme as inFigure 29. (d) A3 dimer fitted in cryo-EM density map of the inner core wall in side view (top) and interior view (bottom). Unoccupied are shown in dashed lines.*

Image source: (Datler et al. 2024) , License: CC BY 4.0

4.10.2   The flower-shaped pore of the MV core wall

Next, we focused on the hexameric flower-shaped pore, another structure revealed by the 2D classes of the VACV treasure chest (Chapter 4.5). This structure was already evident in the cryo-ET dataset of isolated viral cores, where we identified that they are embedded in the core layer at the same height as the trimers, with the inner hexamer slightly further extended towards the MV envelope as previously described (Hernandez-Gonzalez et al. 2023) (Figure 31).
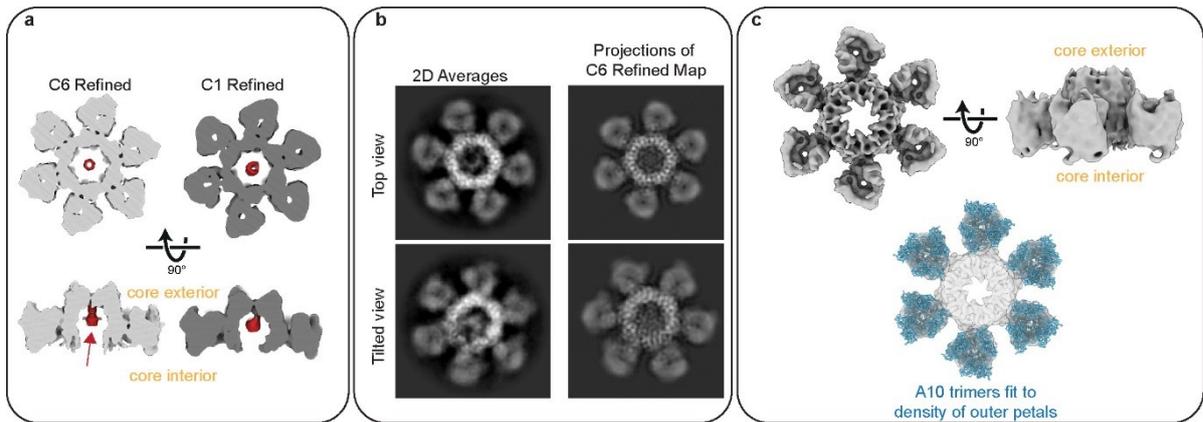
Figure 31: The flower-shaped pore of the VACV core wall

(a) *Slices through the cryo-EM density of the flower-shaped pore C1 and C6 refined in top and side view with annotated core interior and exterior. The unidentified donut-shaped central density is depicted in red (b) 2D class averages of the flower-shaped pore in top and bottom view (left) compared to the projections of the C6 symmetrized refined map. (c) C6 symmetrized cryo-EM density map of the flower-shaped pore (top) with fitted A10 trimers in the petals (bottom). The central hexamer density could not be identified and, therefore, remains empty.*

Image source: (Datler et al. 2024) , License:
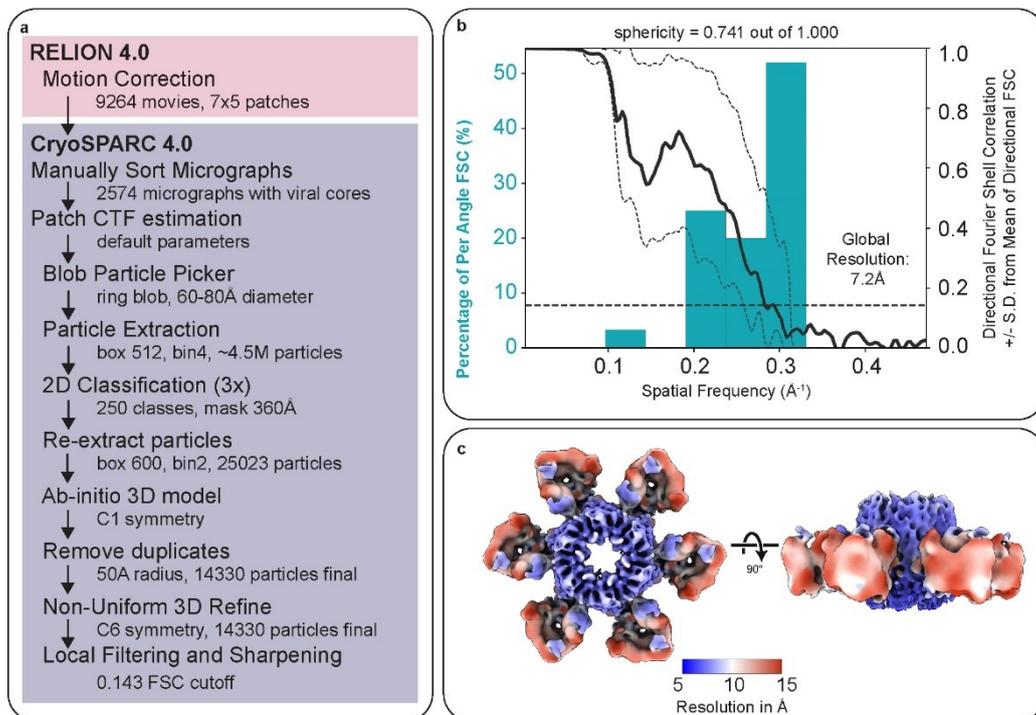
Figure 32: "SPA processing workflow of the flower-shaped pore

(a*) Summary of processing steps used in the SPA workflow. (b) 3D FSC calculations of the masked CryoSPARC half maps. The cyan histogram depicts the fraction of particles that reach the corresponding resolution, and the black curve shows the global FSC +/- SD of FSCs calculated with extensive angular sampling. The global resolution indicated is at FSC 0.143 cutoff. (c) Local resolution of the flower-shaped pore cryo-EM density map." (Datler et al. 2024)*

Image source: (Datler et al. 2024) , License:

We achieved a resolution of 7.2 Å that was sufficient to identify the outer proteins that form the flower's petals, and we could fit six of the A10 trimers (Figure 31c). Most of the flower-shaped pore particles were top views and slightly tilted views, causing a strong anisotropy. Therefore, the cryo-EM map exhibited insufficient resolution or quality for detailed structural interpretation. Despite trying different protein candidates we could not fit a model into the central hexameric density. The detailed processing steps are shown in Figure 32.

Overall, these findings offer valuable insights into the proteins incorporated within the viral core. However, many proteins remain that we observed in our structural "treasure chest" of 2D classes, for which we have not yet determined high-resolution structures. Further exploration and elucidation of these components will provide a deeper understanding and enhance our comprehensive knowledge of the viral core architecture.

5 Discussion

The discussion in the following chapter has been part of the publication "Multi-modal cryo-EM reveals trimers of protein A10 to form the palisade layer in poxvirus cores" (Datler et al. 2024), and were reproduced and adapted for this thesis. The draft was originally written by myself.

Our research has identified A10 as the key protein forming the palisade layer of VACV and revealed additional novel structural features of the core, particularly extending our understanding of the virus core (Figure 33). We have used a combination of cryo-ET, which provided the overall context of the protein interactions, and cryo-SPA to identify a high-resolution structure of the trimer protein A10. This allowed us to substantially revise the core wall morphology, facilitating a more comprehensive and detailed model by integrating prior descriptions of protein interactions and spatial arrangements within the core wall.

Structural core proteins of the VACV core wall

Our observations show that trimers of A10 are building the palisade layer of the poxvirus core. This finding corrects earlier suggestions based on immunolabelling proposing that the spike protein of the palisade layer consists of A4 (Cudmore et al. 1996, Roos et al. 1996, Pedersen et al. 2000). These previous studies suggested that A4 is located on the exterior side of the viral core, presumably interacting with A10. The A4 protein is relatively abundant and has been shown to constitute 13.8% of the
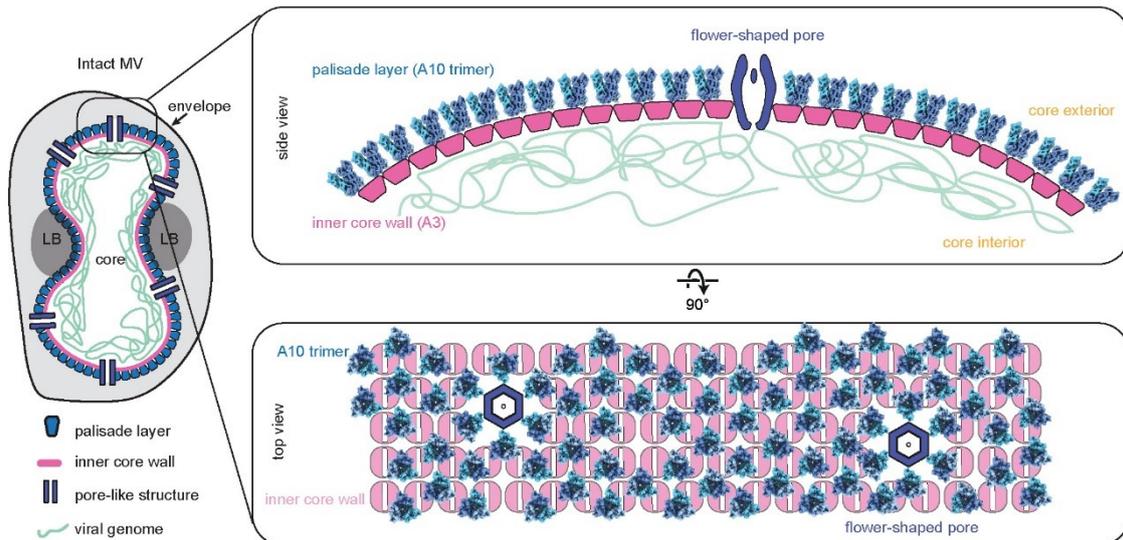
Figure 33: Revised structural model of the VACV core wall

A schematic overview of the revised model for the VACV core wall suggests that protein A10 constitutes the palisade layer, positioned above an inner core wall with distinct symmetry. Protein A4 is likely situated on the outer surface of the palisade layer. The inner core wall appears to be primarily formed by A3, with L4 potentially acting as a DNA-binding protein that anchors the viral genome to the core wall. The core structure is punctuated by flower-shaped pores, which are unevenly distributed across its surface.

Image source: (Datler et al. 2024) , License: CC BY 4.0

mass of the virion core, supporting its role in decorating the entire core surface when forming the palisade layer (Holowczak and Joklik 1967, Sarov and Joklik 1972). Additionally, immunoprecipitation studies revealed that A10 and A4 build a stable complex even prior to the proteolytic cleavage of the precursor protein (Risco et al. 1999). Our determination of the A10 trimer structure, combined with insights from XL-MS experiments, allow us to provide the context that interaction sites for A4 are primarily located on the exterior of the A10 trimer (Mirzakhanyan and Gershon 2019) (Figure 28). Furthermore, A4 partially disassociates after purification of the cores with NP40 and DTT, while A10 remains incorporated in the viral core. This reinforces the hypothesis that the protein is located on the exterior side of the palisade layer (Cudmore et al. 1996, Moussatche and Condit 2015).

Moreover, the extended shape of A4 resembles minor coat proteins, known as cement proteins, as for example found in adenoviruses, which help assemble and stabilize the viral shell (Dai et al. 2017, Gallardo et al. 2021). Additionally, A4 has been described as a matrix-like protein, potentially linking the core to surrounding membranes through interactions with other viral membrane proteins (Cudmore et al. 1996). In line with this, a recent study demonstrated that the A10 trimer interacts closely with an N-terminal alpha helix of A4. This interaction suggests that while A10 forms the palisade layer, A4 functions as a linker to stabilize intra-trimer connections and links to the sourrounding membranes, contributing to the structural integrity and stability of the viral core (Hernandez-Gonzalez et al. 2024).

Previous results of XL-MS suggested that protein A10 directly interacts with A3 (Mirzakhanyan and Gershon 2019). This implies that A3 is part of the inner core wall as the linkages to A3 are exclusively positioned at the bottom side of the trimer (Figure 28). Additional evidence supporting this placement comes from immunogold labeling studies, which showed that A3 becomes detectable only after the core is disrupted by hypertonic shock and protease treatment (Moussatche and Condit 2015). Our low-resolution structural map of the inner core wall further indicates that a dimer of A3 fits well within the observed density (Figure 30). XL-MS data also suggest that a small unoccupied region at the top of the inner core wall could potentially be occupied by 23K, given its close proximity to A10 (Mirzakhanyan and Gershon 2019). While these findings point to possible interactions among A10, 23K, and A3, further experimental work would be necessary to define their positions and interactions within the core wall accurately. To achieve this goal, we could revisit the purified viral cores to gather additional cryo-EM data on the inner core wall. Alternatively, we could disrupt the viral core to display individual proteins and their protein interactions. An alternative approach could include cryo-EM of *in vitro* assemblies of purified proteins.

L4 is a major DNA binding protein (Bayliss and Smith 1997, Jesus et al. 2014) located in the viral core's interior (Pedersen et al. 2000, Moussatche and Condit 2015). The predicted interaction between L4 and A10 is inconclusive when analyzing XL-MS (Mirzakhanyan and Gershon 2019). L4 is a relatively small protein of about 25 kDa and could be responsible for DNA recruitment to the inner core wall. Our cryo-ET data reveal condensed DNA adhering to the inner surface of the core wall, which is in contradiction to previous studies where viral DNA was observed in tubule formations within the core (Condit et al. 2006). Notably, while VACV shows condensed viral DNA, MPXV does not exhibit this morphology (Hong et al. 2024). To further investigate if L4 is responsible for the different DNA morphology in VACV and MPXV, we could further

compare the L4 sequences, revisit our cryo-EM 2D classes and conduct mutation studies with L4 knockouts. However, it is important to consider that L4 may have additional functions beyond DNA binding, which could result in a virus that is unable to reproduce.

Building on these findings, we propose the following structural arrangement of major core proteins within the viral core. A4 is positioned on the exterior surface of the palisade layer, A10 forms the palisade layer itself, as shown in this study and A3 comprises the inner core wall. Given the absence of unoccupied density within the palisade layer, 23K may be either beneath the trimer or within the viral core pore. The precise localization of L4, whether functioning as DNA-binding protein inside the core, situated within the viral core pore or contributing to its central density, requires further investigation.

Comparison to previous poxvirus cryo-EM studies

VACV MVs are large pleomorphic viruses with size dimensions of about 360 nm × 250 nm × 220 nm. These dimensions are reported across multiple studies (Cyrklaff et al. 2005, Hernandez-Gonzalez et al. 2023, 2024, Datler et al. 2024, Liu et al. 2024), often with slight variations, which can be attributed to factors such as the tilt of the virus within the ice layer influencing measurement, inaccuracies associated with manual measurement techniques, and differences in sample preparation and purification protocols. One of the main factors is still the intrinsic pleomorphicity of the virions, which might also depend on the cell type in which they are produced.

A recent structural study of monkeypox virus (MPXV) revealed that MVs exhibit dimensions of approximately 313 × 267 × 236 nm, indicating a notably shorter long axis compared to VACV MVs (Hong et al. 2024). In this study MPXV isolated from patients was propagated in Vero cells, which are derived from African green monkey cells. This atypical morphology is similarly observed in MVA when propagated in HeLa cells, where the long axis only measures 296 nm. MVA has undergone over 500 passages in chicken embryo fibroblasts and can no longer produce infectious particles in human cells (Gallego-Gómez et al. 2003). These findings suggest that atypical morphological changes may be characteristic of zoonotic infections where the virus has not fully adapted to its new host.

The role of A10 in the viral core lattice

The viral lattice of the poxvirus core is composed of several key components, as described above. The role of each layer in the viral lattice's stability and shape determination remains unclear. Given the observation that the palisade layer of VACV consists of A10 trimers, the question of their function and whether they stabilize the viral core lattice remains.

The spacing between the trimers is substantial, including gaps and cracks, and accordingly, the observed lateral interactions are not extensive. In the 2D classes of the cryo-SPA data we observe trimers with different orientations to each other. This suggests that the regularity of the palisade layer might dissolve upon entering a new cell, suggesting the palisade layer is not the key stabilizer of the viral core. In the lattice maps of intact VACV MVs, we can also observe different trimer orientations

underlining the flexibility of the viral core lattice (Figure 26). However, this indicates that additional proteins are necessary to account for the tightly packed organization of the viral lattice. Conditional mutation studies showed that without A10 the correct assembly to MVs is not possible. Instead, the inner core wall builds stacks that are supposedly labeled by L4, F17 and E8 antibodies (Heljasvaara et al. 2001, Rodriguez et al. 2006). The integrity of the viral core must be achieved by another layer, such as the inner core wall or protein A4, that could link the trimers on exterior side of the viral core.

Therefore, A10 and A4 may help maintain the palisade layer in a semi-regular lattice to act as the shape and curvature-defining unit of the viral core, providing the force and scaffold needed to condense the core into a dumbbell-shaped form. Further suggesting that the inner core wall acts as a stabilizer with strong protein-protein interactions. This is evident in isolated cores, where trimers can detach from the inner core wall during the purification while the core itself remains structurally intact.

Functional implications of A10 trimer stability and variability in the viral life cycle

Our AF2 predicted A10 structure fitted precisely in our refined cryo-EM structure, that was obtained from the soluble fraction of purified viral cores, and could even accommodate bulky side chains and apha-helical pitch (Figure 22). The intra-trimer stability is maintained through hydrophobic interactions, conserved interchain salt bridges, and a heterodimeric beta-sheet that closely binds the monomers of the trimers together. Additionally, a disulfide bond within A10 monomers seems to stabilize the protein fold, by linking conserved cysteine residues at the N- and C-terminal ends. This is in line with a report that found disulfide formation important for virus stability (Locker and Griffiths 1999),

A recent study by Liu & Corroyer-Dulmot et al. also demonstrates that A10 assembles into trimers within the palisade layer (Liu et al. 2024). The structure was obtained from trimers still attached to the surface of isolated cores, and the received classes indicate increased flexibility, in addition to our observations of variable interactions between trimers. These classes indicate that some trimers exhibit an open, hollow central region, and some show a closed state. Furthermore, the researchers propose that the open state of the trimer is physiologically important, as it has been observed in detergent-stripped virions as well as *in situ* within virions after infection. Another study reached the same conclusion, suggesting that A4 is released from the palisade layer after viral membrane loss or viral core purification (Hernandez-Gonzalez et al. 2024). The release of the membrane could trigger a physiological release of A4, which could lead to a more flexible and relaxed palisade layer that might be necessary for viral infection and early transcription of viral proteins.

Flower shaped core pore

The flower-shaped core pore that we identified in our cryo-EM datasets seems to have an abundance of approximately 20 flowers per core, presumably more, as our data did not allow us to determine the exact number quantitatively. Cryo-EM studies already described this pore-like structure and their function was postulated to be directly involved in the mRNA extrusion into the cytoplasm of infected cells (Cyrklaff et al.

2005, Moussatche and Condit 2015). Our cryo-EM density reveals an additional density in the center of the flower lumen (Figure 31). Theoretically, the lumen would be large enough to accommodate DNA, but further experimental proof will be required. We identified the six flower petals as A10 trimers; however, the proteins constituting the interior region of this structure remain unidentified. A recent cryo-EM study suggests that the structure may correspond to protein E6, which plays a critical role in viral assembly and the release of viral mRNA from the core during the early stages of infection (Hernandez-Gonzalez et al. 2024). Another cryo-ET study proposes that the interior region could also consist of protein A10 that rearranges and causes the ring-like structure during early infection to facilitate viral DNA release (Liu et al. 2024). Further experimental proof is needed to identify the viral core pore's role and constitution.

Cryo-SPA reveals the treasure chest of the VACV core

Cryo-SPA has already been shown to work on simpler viruses such as HIV-2 and pleomorphic in vitro virus-like particles (VLPs) (Talledge et al. 2023). Our study further shows that cryo-SPA is also feasible and applicable to more complex viruses like poxvirus. Working with isolated viral cores, which encapsulate all components necessary for the virus to infect new cells, enabled us to visualize several structural entities exhibiting different symmetries (Figure 20). Given their prevalence, they most likely play a relevant role in the viral lifecycle. Notably, the pentamer was the most frequently observed structure, and the tetramer structure was also highly abundant. Despite considerable efforts we were not able to obtain a higher resolution structure, of the pentameric and tetrameric classes. Additionally, due to preferred orientation problems, caused by working with isolated core particles which are still large in size, we encountered difficulties with certain structures, such as the flower-shaped core pore. Modifications to the core purification and vitrification protocols may address these issues in the future.

A10 protein conservation within poxviruses

The structural conservation of the A10 protein in poxviruses is relatively high among orthopoxviruses and parapoxviruses, but shows less conservation in other families, such as entomopoxviruses. A recent study on MPXV, a member of the *Ortopoxviridae*, reveals a similar cryo-EM density structure for the trimer, reinforcing this observation (Hong et al. 2024).

However, the palisade layer and the structure of the A10 trimer are very similar to VACV. This raises the question of whether the less conserved A10 proteins in other viruses, such as entomopoxvirus, are associated with a different viral core and palisade layer morphology. The A10 protein of AmEPV is up to 250 AA longer than the A10 of VACV and the sequence identity is only 22 %. Based on the AF2 structure, the base of the trimer appears similar (Figure 27). The top of the entomopoxvirus trimer shows substantial differences to the orthopoxvirus trimer, which could heavily influence the palisade layer morphology. Early TEM studies of ultrathin sections from entomopoxvirus-infected cockroach cells reveal virions that closely resemble those of VACV (Radek and Fabel 2000). Nevertheless, the resolution is insufficient to determine if they possess a similar palisade layer. Therefore, it would be an interesting experimental approach to explore whether virions from viruses with less conserved

A10 proteins develop a palisade layer and whether it serves a similar function as in *Orthopoxviridae*.

Future work

Future efforts should focus on refining purification protocols to achieve higher-resolution structures of additional components, enabling detailed visualization of lateral bodies and their contents and elucidating the structures of the pentamers and tetramers observed in our cryo-SPA data. Addressing the challenge of preferred orientation will also be essential to obtain more accurate structural characterizations of these core elements.

In vitro studies, similar to those conducted on VACV D13, could be extended to the unidentified core proteins. D13 has been extensively characterized using both x-ray crystallography and cryo-EM, demonstrating its trimeric formation *in vitro* and *in vivo*, resulting in a hexameric lattice found in the immature virus (Bahar et al. 2011, Hyun et al. 2011, 2022). *In vitro*, reconstitution of A10 trimers or other oligomers would be an interesting approach for reconstituting the VACV core *in vitro*. This could be challenging due to the limited biochemistry that has been previously described.

Given the recent reemergence of poxviruses causing global epidemics, it is crucial to better understand the different stages of the viral lifecycle. In summary, our findings allow us to extend the model of the core wall assembly substantially.

References

Abella JVG, Galloni C, Pernier J, Barry DJ, Kjær S, Carlier M-F, Way M. 2016. Isoform diversity in the Arp2/3 complex determines actin filament dynamics. Nature Cell Biology 18: 76–86.

Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, Bodenstein SW, Evans DA, Hung C-C, O'Neill M, Reiman D, Tunyasuvunakool K, Wu Z, Žemgulytė A, Arvaniti E, Beattie C, Bertolli O, Bridgland A, Cherepanov A, Congreve M, Cowen-Rivers AI, Cowie A, Figurnov M, Fuchs FB, Gladman H, Jain R, Khan YA, Low CMR, Perlin K, Potapenko A, Savy P, Singh S, Stecula A, Thillaisundaram A, Tong C, Yakneen S, Zhong ED, Zielinski M, Žídek A, Bapst V, Kohli P, Jaderberg M, Hassabis D, Jumper JM. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630: 493–500.

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D: Biological Crystallography 66: 213–221.

Aldaz-Carroll L, Whitbeck JC, Ponce de Leon M, Lou H, Hirao L, Isaacs SN, Moss B, Eisenberg RJ, Cohen GH. 2005. Epitope-mapping studies define two major neutralization sites on the vaccinia virus extracellular enveloped virus glycoprotein B5R. Journal of Virology 79: 6260–6271.

Alvarez DE, Agaisse H. 2013. The formin FHOD1 and the small GTPase Rac1 promote vaccinia virus actin–based motility. Journal of Cell Biology 202: 1075–1090.

Anfinsen CB, Haber E, Sela M, White FH. 1961. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. Proceedings of the National Academy of Sciences of the United States of America 47: 1309–1314.

Ansarah-Sobrinho C, Moss B. 2004. Role of the I7 Protein in Proteolytic Processing of Vaccinia Virus Membrane and Core Components. Journal of Virology 78: 6335–6343.

Arakawa Y, Cordeiro JV, Schleich S, Newsome TP, Way M. 2007. The release of vaccinia virus from infected cells requires RhoA-mDia modulation of cortical actin. Cell Host & Microbe 1: 227–240.

Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, Wang J, Cong Q, Kinch LN, Schaeffer RD, Millán C, Park H, Adams C, Glassman CR, DeGiovanni A, Pereira JH, Rodrigues AV, van Dijk AA, Ebrecht AC, Opperman DJ, Sagmeister T, Buhlheller C, Pavkov-Keller T, Rathinaswamy MK, Dalwadi U, Yip CK, Burke JE, Garcia KC, Grishin NV, Adams PD, Read RJ, Baker D.

2021. Accurate prediction of protein structures and interactions using a three-track neural network. Science (New York, N.Y.) 373: 871–876.

Bahar MW, Graham SC, Stuart DI, Grimes JM. 2011. Insights into the Evolution of a Complex Virus from the Crystal Structure of Vaccinia Virus D13. Structure(London, England:1993) 19: 1011–1020.

Bayliss CD, Smith GL. 1997. Vaccinia virion protein VP8, the 25 kDa product of the L4R gene, binds single-stranded DNA and RNA with similar affinity. Nucleic Acids Research 25: 3984–3990.

Belongia EA, Naleway AL. 2003. Smallpox Vaccine: The Good, the Bad, and the Ugly. Clinical Medicine and Research 1: 87.

Bidgood SR, Samolej J, Novy K, Collopy A, Albrecht D, Krause M, Burden JJ, Wollscheid B, Mercer J. 2022. Poxviruses package viral redox proteins in lateral bodies and modulate the host oxidative response. PLoS Pathogens 18: e1010614.

Bisht H, Weisberg AS, Szajner P, Moss B. 2009. Assembly and Disassembly of the Capsid-Like External Scaffold of Immature Virions during Vaccinia Virus Morphogenesis. Journal of Virology 83: 9140–9150.

Blanchard TJ, Alcami A, Andrea P, Smith GL. 1998. Modified vaccinia virus Ankara undergoes limited replication in human cells and lacks several immunomodulatory proteins: implications for use as a human vaccine. Journal of General Virology, 79: 1159–1167.

Carter GC, Law M, Hollinshead M, Smith GL. 2005. Entry of the vaccinia virus intracellular mature virion and its interactions with glycosaminoglycans. The Journal of General Virology 86: 1279–1290.

Castaño-Díez D. 2017. The Dynamo package for tomography and subtomogram averaging: components for MATLAB, GPU computing and EC2 Amazon Web Services. Acta Crystallographica. Section D, Structural Biology 73: 478–487.

Chen S, McMullan G, Faruqi AR, Murshudov GN, Short JM, Scheres SHW, Henderson R. 2013. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. Ultramicroscopy 135: 24–35.

Chichón FJ, Rodríguez MJ, Risco C, Fraile-Ramos A, Fernández JJ, Esteban M, Carrascosa JL. 2009. Membrane remodelling during vaccinia virus morphogenesis. Biology of the Cell 101: 401–414.

Chung C-S, Chen C-H, Ho M-Y, Huang C-Y, Liao C-L, Chang W. 2006. Vaccinia Virus Proteome: Identification of Proteins in Vaccinia Virus Intracellular Mature Virion Particles. Journal of Virology 80: 2127–2140.

Chung CS, Hsiao JC, Chang YS, Chang W. 1998. A27L protein mediates vaccinia virus interaction with cell surface heparan sulfate. Journal of Virology 72: 1577–1585.

Cifuentes-Munoz N, El Najjar F, Dutch RE. 2020. Chapter Three - Viral cell-to-cell spread: Conventional and non-conventional ways. Pages 85–125 in Kielian M, Mettenleiter TC, and Roossinck MJ, eds. Advances in Virus Research, vol. 108. Academic Press.

Condit RC, Moussatche N, Traktman P. 2006. In a nutshell: structure and assembly of the vaccinia virion. Advances in Virus Research 66: 31–124.

Cudmore S, Blasco R, Vincentelli R, Esteban M, Sodeik B, Griffiths G, Krijnse Locker J. 1996. A vaccinia virus core protein, p39, is membrane associated. Journal of Virology 70: 6909–6921.

Cudmore S, Cossart P, Griffiths G, Way M. 1995. Actin-based motility of vaccinia virus. Nature 378: 636.

Cyrklaff M, Risco C, Fernández JJ, Jiménez MV, Estéban M, Baumeister W, Carrascosa JL. 2005. Cryo-electron tomography of vaccinia virus. Proceedings of the National Academy of Sciences 102: 2772–2777.

Dai X, Wu L, Sun R, Zhou ZH. 2017. Atomic Structures of Minor Proteins VI and VII in Human Adenovirus. Journal of Virology 91: e00850.

Dales S. 1963. The Uptake and Development of Vaccinia Virus in Strain L Cells Followed with Labeled Viral Deoxyribonucleic Acid. The Journal of Cell Biology 18: 51–72.

Dales S, Mosbach EH. 1968. Vaccinia as a model for membrane biogenesis. Virology 35: 564–583.

Damon IK, Damaso CR, McFadden G. 2014. Are We There Yet? The Smallpox Research Agenda Using Variola Virus. PLOS Pathogens 10: e1004108.

Datler J, Hansen JM, Thader A, Schlögl A, Bauer LW, Hodirnau V-V, Schur FKM. 2024. Multi-modal cryo-EM reveals trimers of protein A10 to form the palisade layer in poxvirus cores. Nature Structural & Molecular Biology 1–10.

Dehaven BC, Gupta K, Isaacs SN. 2011. The vaccinia virus A56 protein: a multifunctional transmembrane glycoprotein that anchors two secreted viral proteins. The Journal of General Virology 92: 1971–1980.

Deng Y, Chen Y, Zhang Y, Wang S, Zhang F, Sun F. 2016. ICON: 3D reconstruction with 'missing-information' restoration in biological electron tomography. Journal of Structural Biology 195: 100–112.

Donnelly SK, Weisswange I, Zettl M, Way M. 2013. WIP Provides an Essential Link between Nck and N-WASP during Arp2/3-Dependent Actin Polymerization. Current Biology 23: 999–1006.

Dubochet J, Adrian M, Richter K, Garces J, Wittek R. 1994. Structure of intracellular mature vaccinia virus observed by cryoelectron microscopy. Journal of Virology 68: 1935–1941.

Easterbrook KB. 1966. Controlled degradation of vaccinia virions in vitro: an electron microscopic study. Journal of Ultrastructure Research 14: 484–496.

Edward Jenner. 1802. An inquiry into the causes and effects of the variolæ vaccinæ, a disease discovered in some of the western counties of England, particulary Gloucestershire, and known by the name of the cow pox. 2nd ed. Springfield [Mass.] : Re-printed for Dr. Samuel Cooley, by Ashley & Brewer, 1802.

van Eijl H, Hollinshead M, Smith GL. 2000. The vaccinia virus A36R protein is a type Ib membrane protein present on intracellular but not extracellular enveloped virus particles. Virology 271: 26–36.

Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. Acta Crystallographica Section D: Biological Crystallography 60: 2126–2132.

Ericsson M, Cudmore S, Shuman S, Condit RC, Griffiths G, Locker JK. 1995. Characterization of ts 16, a temperature-sensitive mutant of vaccinia virus. Journal of Virology 69: 7072–7086.

Ericsson M, Sodeik B, Locker JK, Griffiths G. 1997. In VitroReconstitution of an Intermediate Assembly Stage of Vaccinia Virus. Virology 235: 218–227.

Fäßler F, Javoor MG, Datler J, Döring H, Hofer FW, Dimchev G, Hodirnau V-V, Faix J, Rottner K, Schur FKM. 2023. ArpC5 isoforms regulate Arp2/3 complex–dependent protrusion through differential Ena/VASP positioning. Science Advances 9: eadd6495.

Fleishman SJ, Leaver-Fay A, Corn JE, Strauch E-M, Khare SD, Koga N, Ashworth J, Murphy P, Richter F, Lemmon G, Meiler J, Baker D. 2011. RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. PloS One 6: e20161.

Frischknecht F, Moreau V, Röttger S, Gonfloni S, Reckmann I, Superti-Furga G, Way M. 1999. Actin-based motility of vaccinia virus mimics receptor tyrosine kinase signalling. Nature 401: 926–929.

Gallardo J, Pérez-Illana M, Martín-González N, San Martín C. 2021. Adenovirus Structure: What Is New? International Journal of Molecular Sciences 22: 5240.

Gallego-Gómez JC, Risco C, Rodríguez D, Cabezas P, Guerra S, Carrascosa JL, Esteban M. 2003. Differences in virus-induced cell morphology and in virus maturation between MVA and other strains (WR, Ankara, and NYCBH) of vaccinia virus in infected human cells. Journal of Virology 77: 10606–10622.

Greseth MD, Traktman P. 2022. The Life Cycle of the Vaccinia Virus Genome. Annual Review of Virology 9: 239–259.

Griffiths G, Wepf R, Wendt T, Locker JK, Cyrklaff M, Roos N. 2001. Structure and Assembly of Intracellular Mature Vaccinia Virus: Isolated-Particle Analysis. Journal of Virology 75: 11034–11055.

Hagen WJH, Wan W, Briggs JAG. 2017. Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. Journal of Structural Biology 197: 191–198.

Hebben M, Brants J, Birck C, Samama J-P, Wasylyk B, Spehner D, Pradeau K, Domi A, Moss B, Schultz P, Drillien R. 2007. High level protein expression in mammalian cells using a safe viral vector: modified vaccinia virus Ankara. Protein Expression and Purification 56: 269–278.

Heljasvaara R, Rodríguez D, Risco C, Carrascosa JL, Esteban M, Rodríguez JR. 2001. The Major Core Protein P4a (A10L Gene) of Vaccinia Virus Is Essential for Correct Assembly of Viral DNA into the Nucleoprotein Complex To Form Immature Viral Particles. Journal of Virology 75: 5778–5795.

Hernandez-Gonzalez M, Calcraft T, Nans A, Rosenthal PB, Way M. 2023. A succession of two viral lattices drives vaccinia virus assembly. PLoS biology 21: e3002005.

Hernandez-Gonzalez M, Calcraft T, Nans A, Rosenthal PB, Way M. 2024. Palisade structure in intact vaccinia virions. mBio 15: e03134-23.

Heuser J. 2005. Deep-etch EM reveals that the early poxvirus envelope is a single membrane bilayer stabilized by a geodetic 'honeycomb' surface coat. The Journal of Cell Biology 169: 269–283.

Holm L. 2022. Dali server: structural unification of protein families. Nucleic Acids Research 50: W210–W215.

Holowczak JA, Joklik WK. 1967. Studies on the structural proteins of vaccinia virus: I. Structural proteins of virions and cores. Virology 33: 717–725.

Hong Y, Huang B, Zhang J, Peng C, Kong W, Tan W, Li S. 2024. Molecular architecture of monkeypox mature virus. Cell Discovery 10: 1–3.

Hutin S, Ling WL, Tarbouriech N, Schoehn G, Grimm C, Fischer U, Burmeister WP. 2022. The Vaccinia Virus DNA Helicase Structure from Combined Single-Particle Cryo-Electron Microscopy and AlphaFold2 Prediction. Viruses 14: 2206.

Hyun J, Matsunami H, Kim TG, Wolf M. 2022. Assembly mechanism of the pleomorphic immature poxvirus scaffold. Nature Communications 13: 1704.

Hyun J-K, Accurso C, Hijnen M, Schult P, Pettikiriarachchi A, Mitra AK, Coulibaly F. 2011. Membrane Remodeling by the Double-Barrel Scaffolding Protein of Poxvirus. PLOS Pathogens 7: e1002239.

Ichihashi Y, Oie M, Tsuruhara T. 1984. Location of DNA-binding proteins and disulfide-linked proteins in vaccinia virus structural elements. Journal of Virology 50: 929–938.

Ichihashi Y, Takahashi T, Oie M. 1994. Identification of a vaccinia virus penetration protein. Virology 202: 834–843.

Jensen ON, Houthaeve T, Shevchenko A, Cudmore S, Ashford T, Mann M, Griffiths G, Krijnse Locker J. 1996. Identification of the major membrane and core proteins of vaccinia virus by two-dimensional electrophoresis. Journal of Virology 70: 7485–7497.

Jesus DM, Moussatche N, Condit RC. 2014. Vaccinia virus mutations in the L4R gene encoding a virion structural protein produce abnormal mature particles lacking a nucleocapsid. Journal of Virology 88: 14017–14029.

Jesus DM, Moussatche N, McFadden BD, Nielsen CP, D'Costa SM, Condit RC. 2015. Vaccinia virus protein A3 is required for the production of normal immature virions and for the encapsidation of the nucleocapsid protein L4. Virology 481: 1–12.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596: 583–589.

Katsafanas GC, Moss B. 2007. Colocalization of Transcription and Translation within Cytoplasmic Poxvirus Factories Coordinates Viral Expression and Subjugates Host Functions. Cell Host & Microbe 2: 221–228.

van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J, Gilchrist CLM, Söding J, Steinegger M. 2024. Fast and accurate protein structure search with Foldseek. Nature Biotechnology 42: 243–246.

Khatib F, Cooper S, Tyka MD, Xu K, Makedon I, Popovic Z, Baker D, Players F. 2011. Algorithm discovery by protein folding game players. Proceedings of the National Academy of Sciences of the United States of America 108: 18949–18953.

Kremer JR, Mastronarde DN, McIntosh JR. 1996. Computer Visualization of Three-Dimensional Image Data Using IMOD. Journal of Structural Biology 116: 71–76.

Krissinel E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. Journal of Molecular Biology 372: 774–797.

Kryshtafovych A, Schwede T, Topf M, Fidelis K, Moult J. 2021. Critical assessment of methods of protein structure prediction (CASP)-Round XIV. Proteins 89: 1607–1617.

Kucukelbir A, Sigworth FJ, Tagare HD. 2014. Quantifying the local resolution of cryo-EM density maps. Nature Methods 11: 63–65.

Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, Smetanin N, Verkuil R, Kabeli O, Shmueli Y, dos Santos Costa A, Fazel-Zarandi M, Sercu T, Candido S, Rives A. 2023.

Evolutionary-scale prediction of atomic-level protein structure with a language model. Science 379: 1123–1130.

Liu J, Corroyer-Dulmont S, Pražák V, Khusainov I, Bahrami K, Welsch S, Vasishtan D, Obarska-Kosińska A, Thorkelsson SR, Grünewald K, Quemin ERJ, Turoňová B, Locker JK. 2024. The palisade layer of the poxvirus core is composed of flexible A10 trimers. Nature Structural & Molecular Biology 31: 1105–1113.

Liu Y-T, Zhang H, Wang H, Tao C-L, Bi G-Q, Zhou ZH. 2022. Isotropic reconstruction for electron tomography with deep learning. Nature Communications 13: 6482.

Locker JK, Griffiths G. 1999. An Unconventional Role for Cytoplasmic Disulfide Bonds in Vaccinia Virus Proteins. The Journal of Cell Biology 144: 267–279.

Ludtke SJ, Baldwin PR, Chiu W. 1999. EMAN: semiautomated software for high-resolution single-particle reconstructions. Journal of Structural Biology 128: 82–97.

Lupas AN, Pereira J, Alva V, Merino F, Coles M, Hartmann MD. 2021. The breakthrough in protein structure prediction. The Biochemical Journal 478: 1885–1890.

Maguire JB, Haddox HK, Strickland D, Halabiya SF, Coventry B, Griffin JR, Pulavarti SVSRK, Cummins M, Thieker DF, Klavins E, Szyperski T, DiMaio F, Baker D, Kuhlman B. 2021. Perturbing the energy landscape for improved packing during computational protein design. Proteins 89: 436–449.

Malkin AJ, McPherson A, Gershon PD. 2003. Structure of Intracellular Mature Vaccinia Virus Visualized by In Situ Atomic Force Microscopy. Journal of Virology 77: 6332–6340.

Mallardo M, Leithe E, Schleich S, Roos N, Doglio L, Krijnse Locker J. 2002. Relationship between vaccinia virus intracellular cores, early mRNAs, and DNA replication sites. Journal of Virology 76: 5167–5183.

Mallardo M, Schleich S, Locker JK. 2001. Microtubule-dependent Organization of Vaccinia Virus Core–derivd Early mRNAs into Distinct Cytoplasmic Structures. Molecular Biology of the Cell 12: 3875.

Maruri-Avidal L, Domi A, Weisberg AS, Moss B. 2011. Participation of Vaccinia Virus L2 Protein in the Formation of Crescent Membranes and Immature Virions. Journal of Virology 85: 2504–2511.

Mastronarde DN. 1997. Dual-axis tomography: an approach with alignment methods that preserve resolution. Journal of Structural Biology 120: 343–352.

Mastronarde DN. 2005. Automated electron microscope tomography using robust prediction of specimen movements. Journal of Structural Biology 152: 36–51.

Mattei S, Schur FK, Briggs JA. 2016. Retrovirus maturation—an extraordinary structural transformation. Current Opinion in Virology 18: 27–35.

McKelvey TA, Andrews SC, Miller SE, Ray CA, Pickup DJ. 2002. Identification of the orthopoxvirus p4c gene, which encodes a structural protein that directs intracellular mature virus particles into A-type inclusions. Journal of Virology 76: 11216–11225.

Mercer J, Helenius A. 2008. Vaccinia Virus Uses Macropinocytosis and Apoptotic Mimicry to Enter Host Cells. Science 320: 531–535.

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold: making protein folding accessible to all. Nature Methods 19: 679–682.

Mirzakhanyan Y, Gershon P. 2019. The Vaccinia virion: Filling the gap between atomic and ultrastructure. PLOS Pathogens 15: e1007508.

Molinie N, Gautreau A. 2018. The Arp2/3 Regulatory System and Its Deregulation in Cancer. Physiological Reviews 98: 215–238.

Mosalaganti S, Obarska-Kosinska A, Siggel M, Taniguchi R, Turoňová B, Zimmerli CE, Buczak K, Schmidt FH, Margiotta E, Mackmull M-T, Hagen WJH, Hummer G, Kosinski J, Beck M. 2022. AI-based structure prediction empowers integrative structural analysis of human nuclear pores. Science 376: eabm9506.

Moussatche N, Condit RC. 2015. Fine structure of the vaccinia virion determined by controlled degradation and immunolocalization. Virology 475: 204–218.

Mutz M, W R, G F, Tg S, Ev K, B M. 2023. Exaptation of Inactivated Host Enzymes for Structural Roles in Orthopoxviruses and Novel Folds of Virus Proteins Revealed by Protein Structure Modeling. mBio 14.

Newsome TP, Weisswange I, Frischknecht F, Way M. 2006. Abl collaborates with Src family kinases to stimulate actin-based motility of vaccinia virus. Cellular Microbiology 8: 233–241.

Obr M, Schur FKM. 2019. Chapter Five - Structural analysis of pleomorphic and asymmetric viruses using cryo-electron tomography and subtomogram averaging. Pages 117–159 in. Advances in Virus Research, vol. 105. Academic Press.

Pedersen K, Snijder EJ, Schleich S, Roos N, Griffiths G, Locker JK. 2000. Characterization of Vaccinia Virus Intracellular Cores: Implications for Viral Uncoating and Core Structure. Journal of Virology 74: 3525–3536.

Peters D, Mueller G. 1963. THE FINE STRUCTURE OF THE DNA-CONTAINING CORE OF VACCINIA VIRUS. Virology 21: 267–269.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. Journal of Computational Chemistry 25: 1605–1612.

Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE. 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Science: A Publication of the Protein Society 30: 70–82.

Peukes J, Xiong X, Erlendsson S, Qu K, Wan W, Calder LJ, Schraidt O, Kummer S, Freund SMV, Kräusslich H-G, Briggs JAG. 2020. The native structure of the assembled matrix protein 1 of influenza A virus. Nature 587: 495–498.

Pfanzelter J, Mostowy S, Way M. 2018. Septins suppress the release of vaccinia virus from infected cells. The Journal of Cell Biology 217: 2911–2929.

Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nature Methods 14: 290–296.

Radek R, Fabel P. 2000. A New Entomopoxvirus from a Cockroach: Light and Electron Microscopy. Journal of Invertebrate Pathology 75: 19–27.

Resch W, Hixson KK, Moore RJ, Lipton MS, Moss B. 2007. Protein composition of the vaccinia virus mature virion. Virology 358: 233–247.

Risco C, Rodríguez JR, Demkowicz W, Heljasvaara R, Carrascosa JL, Esteban M, Rodríguez D. 1999. The vaccinia virus 39-kDa protein forms a stable complex with the p4a/4a major core protein early in morphogenesis. Virology 265: 375–386.

Risco C, Rodríguez JR, López-Iglesias C, Carrascosa JL, Esteban M, Rodríguez D. 2002. Endoplasmic Reticulum-Golgi Intermediate Compartment Membranes and Vimentin Filaments Participate in Vaccinia Virus Assembly. Journal of Virology 76: 1839–1855.

Rodriguez D, Bárcena M, Möbius W, Schleich S, Esteban M, Geerts WJC, Koster AJ, Griffiths G, Locker JK. 2006. A vaccinia virus lacking A10L: viral core proteins accumulate on structures derived from the endoplasmic reticulum. Cellular Microbiology 8: 427–437.

Rodríguez JR, Risco C, Carrascosa JL, Esteban M, Rodríguez D. 1998. Vaccinia virus 15-kilodalton (A14L) protein is essential for assembly and attachment of viral crescents to virosomes. Journal of Virology 72: 1287–1296.

Rohou A, Grigorieff N. 2015. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. Journal of Structural Biology 192: 216–221.

Roos N, Cyrklaff M, Cudmore S, Blasco R, Krijnse-Locker J, Griffiths G. 1996. A novel immunogold cryoelectron microscopic approach to investigate the structure of the intracellular and extracellular forms of vaccinia virus. The EMBO Journal 15: 2343–2355.

Rosenthal PB, Henderson R. 2003. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. Journal of Molecular Biology 333: 721–745.

Sanderson CM, Hollinshead M, Smith GL. 2000. The vaccinia virus A27L protein is needed for the microtubule-dependent transport of intracellular mature virus particles. The Journal of General Virology 81: 47–58.

Sarov I, Joklik WK. 1972. Studies on the nature and location of the capsid polypeptides of vaccinia virions. Virology 50: 579–592.

Scaplehorn N, Holmström A, Moreau V, Frischknecht F, Reckmann I, Way M. 2002. Grb2 and Nck act cooperatively to promote actin-based motility of vaccinia virus. Current biology: CB 12: 740–745.

Scheres SHW. 2012. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. Journal of Structural Biology 180: 519–530.

Schur FKM, Obr M, Hagen WJH, Wan W, Jakobi AJ, Kirkpatrick JM, Sachse C, Kräusslich H-G, Briggs JAG. 2016. An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. Science 353: 506–508.

Senkevich TG, White CL, Koonin EV, Moss B. 2002. Complete pathway for protein disulfide bond formation encoded by poxviruses. Proceedings of the National Academy of Sciences 99: 6667–6672.

Silva DCM, Moreira-Silva EA dos S, Gomes J de AS, Fonseca FG da, Correa-Oliveira R. 2010. Clinical signs, diagnosis, and case reports of Vaccinia virus infections. Brazilian Journal of Infectious Diseases 14: 129–134.

Smith GL, Mackett M, Moss B. 1984. Recombinant Vaccinia Viruses as New Live Vaccines. Biotechnology and Genetic Engineering Reviews 2: 383–407.

Sodeik B, Doms RW, Ericsson M, Hiller G, Machamer CE, Hof W van 't, Meer G van, Moss B, Griffiths G. 1993. Assembly of vaccinia virus: role of the intermediate compartment between the endoplasmic reticulum and the Golgi stacks. The Journal of Cell Biology 121: 521–541.

Stivala A, Wybrow M, Wirth A, Whisstock JC, Stuckey PJ. 2011. Automatic generation of protein structure cartoons with Pro-origami. Bioinformatics (Oxford, England) 27: 3315–3316.

Stokes GV. 1976. High-voltage electron microscope study of the release of vaccinia virus from whole cells. Journal of Virology 18: 636–643.

Strassburg MA. 1982. The global eradication of smallpox. American Journal of Infection Control 10: 53–59.

Szajner P, Weisberg AS, Lebowitz J, Heuser J, Moss B. 2005. External scaffold of spherical immature poxvirus particles is made of protein trimers, forming a honeycomb lattice. Journal of Cell Biology 170: 971–981.

Talledge N, Yang H, Shi K, Coray R, Yu G, Arndt WG, Meng S, Mendonça LM, Castaño-Díez D, Aihara H, Mansky LM, Zhang W. 2023. HIV-2 Immature Particle Morphology Provides Insights into Gag Lattice Stability and Virus Maturation. 2022.02.01.478508.

Taylor MP, Koyuncu OO, Enquist LW. 2011. Subversion of the actin cytoskeleton during viral infection. Nature Reviews Microbiology 9: 427–439.

Terwilliger TC, Ludtke SJ, Read RJ, Adams PD, Afonine PV. 2020. Improvement of cryo-EM maps by density modification. Nature methods 17: 923–927.

Townsley AC, Weisberg AS, Wagenaar TR, Moss B. 2006. Vaccinia virus entry into cells via a low-pH-dependent endosomal pathway. Journal of Virology 80: 8899–8908.

Turoňová B, Schur FKM, Wan W, Briggs JAG. 2017. Efficient 3D-CTF correction for cryo-electron tomography using NovaCTF improves subtomogram averaging resolution to 3.4 Å. Journal of Structural Biology 199: 187–195.

Vanslyke JK, Hruby DE. 1994. Immunolocalization of Vaccinia Virus Structural Proteins during Virion Formation. Virology 198: 624–635.

Varadi M, Bertoni D, Magana P, Paramval U, Pidruchna I, Radhakrishnan M, Tsenkov M, Nair S, Mirdita M, Yeo J, Kovalevskiy O, Tunyasuvunakool K, Laydon A, Žídek A, Tomlinson H, Hariharan D, Abrahamson J, Green T, Jumper J, Birney E, Steinegger M, Hassabis D, Velankar S. 2024. AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. Nucleic Acids Research 52: D368–D375.

Wan W. 2020. williamnwan/TOMOMAN: TOMOMAN 08042020 (https://zenodo.org/record/4110737).

Wan W, Kolesnikova L, Clarke M, Koehler A, Noda T, Becker S, Briggs JAG. 2017. Structure and assembly of the Ebola virus nucleocapsid. Nature 551: 394–397.

Watanabe R, Zyla D, Parekh D, Hong C, Jones Y, Schendel SL, Wan W, Castillon G, Saphire EO. 2024. Intracellular Ebola virus nucleocapsid assembly revealed by in situ cryo-electron tomography. Cell 187: 5587-5603.e19.

Weisberg AS, Maruri-Avidal L, Bisht H, Hansen BT, Schwartz CL, Fischer ER, Meng X, Xiang Y, Moss B. 2017. Enigmatic origin of the poxvirus membrane from the endoplasmic reticulum shown by 3D imaging of vaccinia virus assembly mutants. Proceedings of the National Academy of Sciences 114: E11001–E11009.

Welch MD, Way M. 2013. Arp2/3-mediated actin-based motility: a tail of pathogen abuse. Cell host & microbe 14: 242–255.

Westwood JCN, Harris WJ, Zwartouw HT, Titmuss DHJ, Appleyard G. 1964. Studies on the Structure of Vaccinia Virus. Microbiology, 34: 67–78.

Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB, Snoeyink J, Adams PD, Lovell SC, Richardson JS, Richardson DC. 2018. MolProbity: More and better reference data for improved all-atom structure validation. Protein Science: A Publication of the Protein Society 27: 293–315.

Wilton S, Mohandas AR, Dales S. 1995. Organization of the Vaccinia Envelope and Relationship to the Structure of Intracellular Mature Virions. Virology 214: 503–511.

Yan R, Venkatakrishnan SV, Liu J, Bouman CA, Jiang W. 2019. MBIR: A cryo-ET 3D reconstruction method that effectively minimizes missing wedge artifacts and restores missing information. Journal of Structural Biology 206: 183–192.

Yang Z, Zeng X, Zhao Y, Chen R. 2023. AlphaFold2 and its applications in the fields of biology and medicine. Signal Transduction and Targeted Therapy 8: 1–14.

Yariv B, Yariv E, Kessel A, Masrati G, Chorin AB, Martz E, Mayrose I, Pupko T, Ben-Tal N. 2023. Using evolutionary data to make sense of macromolecules with a "face-lifted" ConSurf. Protein Science : A Publication of the Protein Society 32: e4582.

Zhai X, Lei D, Zhang M, Liu J, Wu H, Yu Y, Zhang L, Ren G. 2020. LoTToR: An Algorithm for Missing-Wedge Correction of the Low-Tilt Tomographic 3D Reconstruction of a Single-Molecule Structure. Scientific Reports 10: 10489.

A. Appendix 1 -Contribution to the paper "ArpC5 isoforms regulate Arp2/3 complex–dependent protrusion through differential Ena/VASP positioning"

The results in the following chapter have been part of the publication "ArpC5 isoforms regulate Arp2/3 complex–dependent protrusion through differential Ena/VASP positioning" (Fäßler et al. 2023), and were reproduced, adapted and summarized for this thesis. The text regarding VACV experiments was originally written by Florian Fäßler and myself and adapted for this thesis.

The Arp2/3 complex is essential for the formation of branched actin networks, which are critical for various cellular functions, including cell migration (Molinie and Gautreau 2018). The Arp2/3 complex consists of seven subunits: Arp2, Arp3 (with isoforms Arp3 and Arp3B), ArpC1 (with isoforms Arp1a and ArpC1b), ArpC2-5 (with isoforms ArpC5 and ARPC5L), which together facilitate the nucleation and stabilization of branched actin filament networks critical for cellular dynamics (Abella et al. 2016). Specifically, the ArpC5 and ArpC5L isoforms within the complex play critical roles in modulating its activity, initiating actin branching, and stabilizing branch junctions. In the study titled "ArpC5 isoforms regulate Arp2/3 complex–dependent protrusion through differential Ena/VASP positioning" by Florian Fäßler (Fäßler et al. 2023), which I contributed to as a co-author, we employed reverse genetics and cellular structural biology to explore how ArpC5 and ArpC5L differentially influence cell migration. ArpC5 and ArcpC5L knock out (KO) cells have quite distinct phenotypes. ArpC5 KO have reduced motility, narrow lamellipodia formation, changes in actin architecture, and reduced recruitment of EnA/VASP actin elongators, suggesting that ArpC5 acts as a regulator of actin polymerization efficiency in lamellipodia. ArpC5L KO cells showed in comparison faster ,migration, wider lammelipodia and actin filaments more parallel to the direciton of pertrusion.

VACV hijacks the host actin cytoskeleton and recruits the Arp2/3 complex to facilitate its movement and cell-to-cell spread. The VACV tail model was used to investigate how actin tails behave in the ArpC5 and ArpC5L KO cells and to assess whether the actin tail formation behavior aligns with previously reported ArpC5 isoform knockdown results (Abella et al. 2016). Abella *et al.*, proposed that VACV actin tails become shorter in ArpC5L knockdown cells and longer in ArpC5 knockdown cells. This observation does not fit our findings regarding the lamellipodium, a broad, sheet-like protrusion at the leading edge of a migrating cell, in B16-F1 melanoma and RAT2 fibroblast KO cells, which exhibit differing phenotypes in lamellipodial width. Therefore, we wanted to determine if VACV tails behave differently than lamellipodia.

Our finding in VACV tail formation in B16-F1 cells showed reduced VACV tail length formation for both KOs compared to WT cells, with ArpC5L KO tails being longer than ArpC5 KO tails. A similar pattern was observed in Rat2 cells, but here ArpC5 KO tails were shorter than the WT and ArpC5L KO tails were longer (Figure 34) just as observed in lamellipodia of our KO cells. Therefore, while our finding of VACV tail lengths in ArpC5/5L KO cells resembles the lamellipodial phenotypes, the extent of these effects differs between Rat2 and B16-F1 cells.
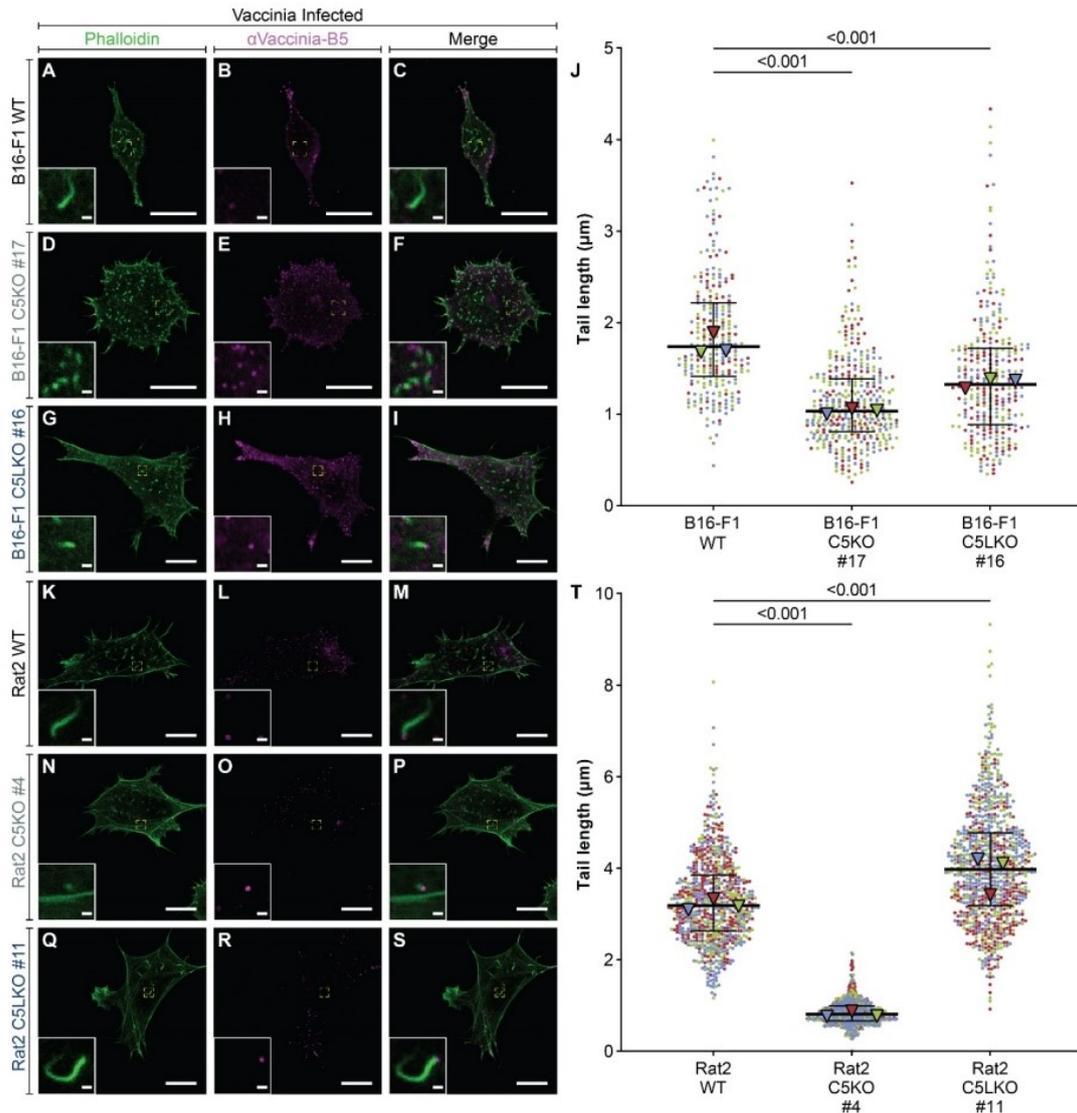
Figure 34: Effects on Vaccinia virus tails after loss of ArpC5 isoform

"(A-I) Representative epifluorescence micrographs of Vaccinia virus-infected B16-F1 wildtype (A, B, C), C5KO (D, E, F), and C5LKO (G, H, I) cells visualizing the actin cytoskeleton and thus the Vaccinia tails using fluorescent phalloidin (A, D, G,) and the viruses using Vaccinia-B5 antibody (B, E, H). Overlays of the signals are shown in (C, F, I). Insets show individual viruses and their tails. Dashed yellow rectangles indicate their position in the micrographs. (J) Vaccinia tail lengths were measured in infected B16-F1 wildtype, C5KO, and C5LKO cells. Kruskal-Wallis test combined with Dunn's multiple comparison test on pooled data from 3 independent experiments, n=263, 391, 292, p values shown in the chart. Black lines indicate overall medians (1.739µm, 1.033µm, 1.,324µm) and quartile ranges. (K-S) Representative epifluorescence micrographs of Vaccinia virus-infected Rat2 wildtype (K, L, M), C5KO (N, O, P), and C5LKO (Q, R, S) cells visualizing the actin cytoskeleton and thus Vaccinia tails using the same staining as in (A-I). Insets show individual viruses and their tails. Their position in the micrographs is indicated by dashed yellow rectangles. (T) Vaccinia tail lengths were measured in infected Rat2 wildtype, C5KO, and C5LKO cells. KruskalWallis test combined with Dunn's multiple comparison test on pooled data from 3 independent experiments, n=738, 728, 846, p values shown in the chart. Black lines indicate overall medians (3.183µm, 0.8095µm, 3.974µm) and quartile ranges. Scale bars, 20µm in standard panels and 1µm in enlarged insets. Data points are color-coded according to the individual experiments. Triangles indicate the medians of the respective experiments."

79
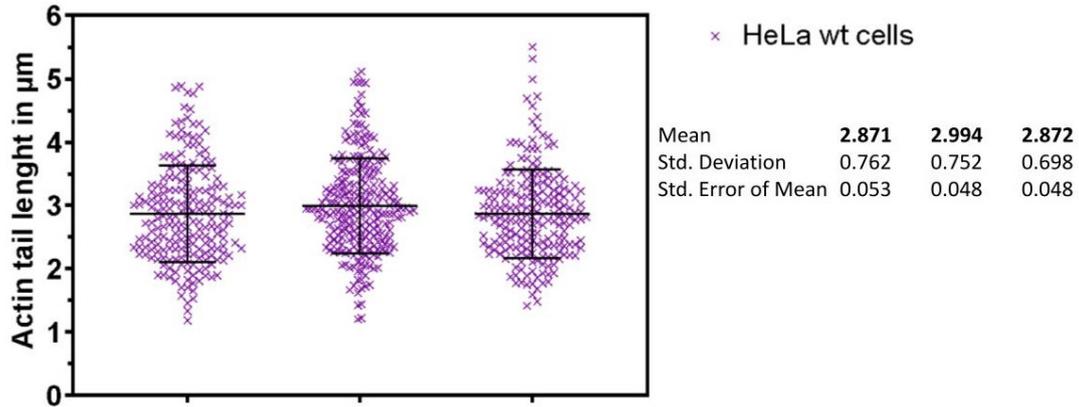
Figure 35: VACV tail length in HeLa cells

Measurement of VACV actin tail length in HeLa wt cells was replicated in three independent experiments. Each dataset represents an individual experiment and each point refers to an individual actin tail measurement.

To assess the robustness of our VACV tail model compared to previous studies, we measured the actin tail length in VACV infected HeLa wt cells (Figure 35). We did this due to the fact that in HeLa cells, the ArpC5/5L isoform-dependent actin tail lengths were the opposite of those in B16-F1 and Rat2 cells (Abella et al. 2016). Our VACV tail measurement aligned with the reported actin tail length of VACV infected HeLa wt cells as described by Abella et al. 2016. This indicates the VACV tail system, I established provided reproducible results and also implied that any differences between KO and knockdown cells might not be due to differences in the infection protocol. Instead the observed phenotypic differences between those two studies might be due to the differences in genetic manipulation, specifically knockdown versus complete KO cells. While we observed varying ArpC5 isoform expression levels across cell lines, overexpression of either isoform did not reverse isoform-specific KO effects (Figure 2, Fäßler et al. 2023). This suggests that ArpC5 isoforms clearly impact VACV actin tail formation differently but do so in a cell type-specific manner. This implies additional regulatory factors are involved in isoform specific Arp2/3 complex regulations in both different cell types when VACV tails are formed.

B. Appendix 2

Mass spectrometry data from soluble fraction of isolated cores. This list comprises mass spectrometry findings from the soluble fraction core sample, sorted by log 10 expression and shows the 76 most abundant proteins of 1019 proteins that were detected by mass spectrometry. The full list can be found in the supplementary information of the paper "Multi-modal cryo-EM reveals trimers of protein A10 to form the palisade layer in poxvirus cores" (Datler et al. 2024). The color code from green to red annotates high to low abundance. VACV proteins are highlighted in yellow. The shortened list, filtered for VACV proteins, can be found in **Table 3**.

UniProt ID	Common Names	Genes	Sequence Coverage %	Peptide Counts	Log 10 expression
CON__cRAP112P00761	;Protease serine 1;;Serine protease 1;Serine protease 1;Transmembrane protease serine 13;Transmembrane protease serine 13, Isoform 2;Transmembrane protease serine 13, Isoform 3;Transmembrane protease serine 13, Isoform 6	PRSS1;TMPRSS13	61	13	10.2062
P16715	Major core protein A10 (4a) precursor	VACWR129	60.5	69	9.5060
P62805	Histone H4	H4C16	58.3	13	9.3992
P03295	Core protein L4 (VP8)	VACWR091	54.2	16	9.1462
P07617	Cap-specific mRNA (nucleoside-2'-O-)-methyltransferase	PAPS	65.5	22	9.0373
CON__cRAP023P00767			5.7	2	9.0345
P26669	Cu-Zn superoxide dismutase-like protein A45R	VACWR171	91.2	9	8.8556
P04195	Cell surface-binding protein	VACWR113	53.9	21	8.7152
P07396	Phosphoprotein F17	VACWR056	32.7	6	8.6891
A0A590UJJ6;B4DEB1;K7EK07;K7EMV3;K7ES00;P68431;P84243;Q16695;Q71DI3	Histone H3;Histone H3;Histone H3 (Fragment);Histone H3;Histone H3.3 (Fragment);Histone H3.1;Histone H3.3;Histone H3.1t;Histone H3.2	H3-3A;H3-3B;H3-4;H3-5;H3C12;H3C13	12.5	3	8.6196
O60814;P57053;P58876;P62807;Q5QNW6;Q5QNW6-2;Q93079;Q99877;Q99879;Q99880;U3KQK0	Histone H2B type 1-K;Histone H2B type F-S;Histone H2B type 1-D;Histone H2B type 1-C/E/F/G/I;Histone H2B type 2-F;Histone H2B type 2-F, Isoform 2;Histone H2B type 1-H;Histone H2B type 1-N;Histone H2B type 1-M;Histone H2B type 1-L;Histone H2B	H2BC1;H2BC10;H2BC12;H2BC12L;H2BC13;H2BC14;H2BC15;H2BC18;H2BC5;H2BC9	47.6	9	8.4889
P59942	Mitochondrial coiled-coil domain protein 1	MCCD1	7.6	1	8.3465
P12532	Creatine kinase U-type, mitochondrial	CKMT1A;CKMT1B;CKMT2	51.3	19	8.3168
C9JFR7;CON__cRAP026P99999;CON__P99999	Cytochrome c (Fragment);;Cytochrome c	CYCS	53.5	6	8.3058
P29191	A4 (39kDa) core protein	VACWR123	17.8	4	8.2534
P04298	mRNA-capping enzyme catalytic subunit	VACWR106	62.8	56	8.2447
P11142	Heat shock cognate 71 kDa protein	HSPA8	48.6	30	8.2155
P68692	Glutaredoxin-1	VACWR069	24.1	4	8.2045
P06440	Major core protein A3 (4b)	VACWR122	34.4	17	8.1866
CON__cRAP054P04264;CON__P04264	;Keratin, type II cytoskeletal 1	KRT1	41	35	8.1639
A0A0U1RR32;A0A0U1RRH7;P04908;P0C0S8;P20671;Q16777;Q6FI13;Q7L0L0;Q93077;Q96KK5;Q99878;Q9BTM1	Histone H2A;Histone H2A;Histone H2A type 1-B/E;Histone H2A type 1;Histone H2A type 1-D;Histone H2A type 2-C;Histone H2A type 2-A;Histone H2A type 3;Histone H2A type 1-C;Histone H2A type 1-H;Histone H2A type 1-J;Histone H2A.J	;H2AC12;H2AC14;H2AC17;H2AC19;H2AC20;H2AC25;H2AC6;H2AC7;H2AC8;H2AJ;H2AZ2;HCG_2039566	24.9	5	8.1621

P08758	*Annexin A5*	*ANXA5*	61.6	21	8.1436
E7EWB4;Q9BPU6	*Dihydropyrimidinase-related protein 5 (Fragment);Dihydropyrimidinase-related protein 5*	*DPYSL5*	3.7	1	8.1410
A0A2R8Y7R2;CON__cRAP036P68871;CON__P68871;E9PFT6;P02042	*Hemoglobin subunit beta;;Hemoglobin subunit beta;Hemoglobin subunit delta;Hemoglobin subunit delta*	*;HBB;HBD;HBE1;HBG1;HBG2*	17.1	2	8.1050
P07392	*DNA-directed RNA polymerase 147 kDa polypeptide*	*PREPL;RPO147*	55.5	69	8.0926
P24757	*DNA-directed RNA polymerase 35 kDa subunit*	*RPO35*	53.8	17	8.0675
F5H265	Polyubiquitin-C (Fragment)	RPS27A;UBA52;UBB;UBC	69.8	7	8.0647
P68438	*RNA polymerase-associated transcription-specificity factor RAP94*	*GFOD1;RAP94*	49.2	38	8.0537
Q76ZN5	*Profilin*	*VACWR167*	39.8	6	8.0401
P20636	*Early transcription factor 82 kDa subunit*	*VETFL*	55.2	34	8.0382
P16104;Q96QV6	Histone H2AX;Histone H2A type 1-A	;H2AC1;H2AC21;H2AJ;H2AX;H2AZ2	32.2	5	8.0268
P68609	*DNA-directed RNA polymerase 22 kDa subunit*	*RPO22*	45.9	9	8.0209
G3V595;P01011;P01011-2	*Alpha-1-antichymotrypsin (Fragment);Alpha-1-antichymotrypsin;Alpha-1-antichymotrypsin, Isoform 2*	*SERPINA3*	2.3	1	8.0110
J3QRG3;Q86WV1;Q86WV1-2	*Src kinase-associated phosphoprotein 1 (Fragment);Src kinase-associated phosphoprotein 1;Src kinase-associated phosphoprotein 1, Isoform 2*	*SKAP1*	11.1	1	7.9939
CON__cRAP094P62937;CON__P62937	*;Peptidyl-prolyl cis-trans isomerase A*	PPIA;PPIAL4A;PPIAL4C;PPIAL4D;PPIAL4E;PPIAL4F;PPIAL4G;PPIAL4H	48.5	8	7.9045
P12926	*Core protease I7*	*VACWR076*	24.3	11	7.8991
P07614	*Protein L3*	*VACWR090*	32.9	12	7.8896
P04318	*mRNA-capping enzyme regulatory subunit*	*VACWR117*	72.1	19	7.8743
A0A0S2Z377	Annexin (Fragment)	ANXA6	55.9	39	7.8672
A0A7P0Z497;P23284	*Peptidyl-prolyl cis-trans isomerase;Peptidyl-prolyl cis-trans isomerase B*	*PPIB*	56.9	12	7.8554
CON__cRAP041P35527;CON__P35527	*;Keratin, type I cytoskeletal 9*	*KRT12;KRT9*	47.4	25	7.8128
P07242	*Late transcription elongation factor H5*	*VACWR103*	44.8	10	7.8030
P09525;P09525-3	*Annexin A4;Annexin A4, Isoform 3*	*ANXA4*	64.3	22	7.7942
Q04837	*Single-stranded DNA-binding protein, mitochondrial*	*SSBP1*	49.3	6	7.7905
Q76ZP7	*DNA-directed RNA polymerase 133 kDa polypeptide*	*RPO132*	43.5	46	7.7817
P24758	*Protein A26*	*VACWR149*	35	15	7.7764
D6R9L0;D6R9Z1;D6RAU2;D6RFX4;D6RFZ9;D6RHH4;E9PD14;H0Y8W2;H0YAF8;J3KPE3;P63244	*Receptor of-activated protein C kinase 1 (Fragment);Receptor of-activated protein C kinase 1 (Fragment);Receptor of-activated protein C kinase 1;Receptor of-activated protein C kinase 1 (Fragment);Receptor of-activated protein C kinase 1 (Fragment);Receptor of-*	*RACK1*	7.7	2	7.7759

	activated protein C kinase 1;Receptor of-activated protein C kinase 1;Receptor of-activated protein C kinase 1 (Fragment);Receptor of-activated protein C kinase 1 (Fragment);Receptor of-P35908activated protein C kinase 1;Receptor of activated protein C kinase 1				
P07616	Protein J1	VACWR093	19.6	4	7.7718
CON__cRAP053P35908;CON__	;Keratin, type II cytoskeletal 2 epidermal	KRT2;KRT72;KRT77;KRT79;NEFH	49.6	33	7.7517
P17066	Heat shock 70 kDa protein 6	HSPA6;HSPA7	15.4	9	7.7506
P68698	DNA topoisomerase 1B	TOP1	59.9	14	7.7456
Q8N257	Histone H2B type 3-B	H2BC19P;H2BC20P;H2BC26;H2BK1	47.6	9	7.7433
P21603;P21603-2	DNA-directed RNA polymerase 30 kDa polypeptide;DNA-directed RNA polymerase 30 kDa polypeptide, Isoform Short	RPO30	50.6	13	7.7432
A0A6Q8PG15;A0A7I2PK44;O95831;O95831-3	Apoptosis-inducing factor 1, mitochondrial;Apoptosis-inducing factor 1, mitochondrial;Apoptosis-inducing factor 1, mitochondrial;Apoptosis-inducing factor 1, mitochondrial, Isoform 3	AIFM1	36.9	19	7.7394
A0A1B0GVI3;CON__cRAP039P13645;CON__P13645	Keratin, type I cytoskeletal 10;;Keratin, type I cytoskeletal 10	KRT10;KRT12;KRT24;KRT25;KRT26;KRT27;KRT28;KRT35;KRT37	36.2	22	7.7256
P18377	Phospholipase-D-like protein K4	VACWR035	30.4	15	7.7231
E5RGN0;E5RIU4;E5RIY8;E5RJM2;Q00536;Q00536-2;Q00536-3	Cyclin-dependent kinase 16 (Fragment);Cyclin-dependent kinase 16 (Fragment);Cyclin-dependent kinase 16 (Fragment);Cyclin-dependent kinase 16 (Fragment);Cyclin-dependent kinase 16;Cyclin-dependent kinase 16, Isoform 2;Cyclin-dependent kinase 16, Isoform 3	CDK16	2	1	7.6915
P04406;P04406-2	Glyceraldehyde-3-phosphate dehydrogenase;Glyceraldehyde-3-phosphate dehydrogenase, Isoform 2	GAPDH	40.3	9	7.6748
P04308	Early transcription factor 70 kDa subunit	VETFS	30.6	16	7.6658
P68623	Protein L5	VACWR092	24.2	4	7.6414
Q9UL15;Q9UL15-2	BAG family molecular chaperone regulator 5;BAG family molecular chaperone regulator 5, Isoform 2	BAG5	1.6	1	7.6364
Q5TEC6	Histone H3-7	H3-3B;H3-5;H3-7	14	3	7.6216
P05807	Nucleoside triphosphatase I	NPH1	46.9	25	7.6173
P12927	RNA helicase NPH-II	NPH2	47.8	31	7.5841
A0A0A0MT40;H3BUD4;H7C276;Q15111;Q15111-2;Q9UPR0;Q9UPR0-2;Q9UPR0-3	Phosphoinositide phospholipase C;Phosphoinositide phospholipase C (Fragment);Phosphoinositide phospholipase C (Fragment);Inactive phospholipase C-like protein 1;Inactive phospholipase C-like protein 1, Isoform 2;Inactive phospholipase C-	PLCL1;PLCL2	0.9	1	7.5660

	like protein 2;Inactive phospholipase C-like protein 2, Isoform 2;Inactive phospholipase C-like protein 2, Isoform 3				
P16233	*Pancreatic triacylglycerol lipase*	*PNLIP*	4.7	2	7.5570
Q14008;Q14008-2;Q14008-3	*Cytoskeleton-associated protein 5;Cytoskeleton-associated protein 5, Isoform 2;Cytoskeleton-associated protein 5, Isoform 3*	*CKAP5*	10.1	18	7.5416
P68613	*Late transcription factor 1*	*VLTF1*	8.1	2	7.5125
P07239	*Dual specificity protein phosphatase H1*	*H1L*	43.9	6	7.4970
A0A8I5KZ38;P35580;P35580-2;P35580-3;P35580-4;P35580-5	*Myosin-10;Myosin-10;Myosin-10, Isoform 2;Myosin-10, Isoform 3;Myosin-10, Isoform 4;Myosin-10, Isoform 5*	*MYH10*	0.4	1	7.4937
P68611	*DNA-directed RNA polymerase 19 kDa subunit*	*RPO19*	37.8	6	7.4932
A0A0G2JIW1;P0DMV8	*Heat shock 70 kDa protein 1B;Heat shock 70 kDa protein 1A*	*HSPA1A;HSPA1B*	48.8	26	7.4866
Q2KHM9;Q2KHM9-2	*Protein moonraker;Protein moonraker, Isoform 2*	*KIAA0753*	0.8	1	7.4825
CON__cRAP119P00772		*CELA1*	26.7	4	7.4555
P07611	*Myristoylated protein G9*	*VACWR087*	26.5	8	7.4532
P68710	*DNA polymerase processivity factor component A20*	*VACWR141*	1.6	1	7.4508